REVISTA DE FILOSOFÍA HOMÓNIMA

UNIVERSIDAD ANDRÉS BELLO

REVISTA DE MÉTODO, INTERDISCIPLINA Y METAFILOSOFÍA

Volumen 1, número 1, 2026

REVISTA DE FILOSOFÍA

ISSN: 3087-260X

Universidad Andrés Bello | Universidad Andrés Bello | Facultad de Educación y Ciencias Sociales

www.revista.homonima.com/
homonima.revista@gmail.com
Facultad de Educación y Ciencias Sociales - UNAB
Fernández Concha 700, 7591538 Las Condes, Región Metropolitana

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026)
ISSN 3087-260X
Santiago de Chile

TABLA DE CONTENIDOS

Editorial *1*

Artículos *6*

Sobre la carencia de análisis (meta)filosófico sobre él término conciencia y autoconciencia en la contemporaneidad. Javier Castillo 6

La verdadera red social: la amistad. Lucas Peñaylillo 48

Un naturalismo minimalista. Constantino Contreras 79

Ontología del color y tesis de la continuidad: relacionalismo adverbial como fusión de estilos de razonamiento. Rodolfo Aldea 112

El filósofo de la ciencia como pensador público: estándares profesionales para la práctica filosófica. Gabriel Donoso 128

¿Qué es leer y escribir filosofía?. Oscar Morales 162

Profundidad, estética y la actual filosofía de las matemáticas: un caso de estudio del teorema de Picard-Lindelöf. Facundo Camargo; Unai León Rico 178

La superación de los desacuerdos profundos: Revisión de bisagras y racionalidad dinámica. Leandro Lema 214

Traducciones *237*

Modelos en Ciencia. Por Roman Frigg & Stephan Hartmann. Traducción por Nibaldo Lorca Améstica 237

Una apología de los conflictos entre Metafísica y Ciencia en la Metafísica Naturalizada. Por Rasmus Jaksland. Traducido por Gabriel Donoso 310

Scanlon sobre la Ontología de la relaciones de razón. Por John Skorupski. Traducción por Cristóbal Andrade 355

Entrevista *379*

Enseñanza, método y práctica de la filosofía. Entrevista al Dr. Luis Placencia / Entrevistador Kuminak Lefio 379

Sección de Lógica y Matemática *439*

Métodos lógicos y matemáticos en filosofía: perspectivas y motivaciones. Francisco Castro Richter 439

Algunas observaciones sobre gráficos peircianos modales. Felipe Morales Carbonell. Comentarios por Kuminak Lefio 453

Evitando malinterpretar el mítico teorema de incompletitud de Gödel: qué es lo que el teorema dice en realidad. Sebastián Castillo 473

Ensayos de Lectura, Escritura y Metodología Filosófica *485*

Breve guía de cómo escribir un ensayo filosófico. Kuminak Lefio; Oscar Morales; Alan Martin; Manuel Rodriguez 485

Reseñas *504*

Lenguaje, mente y cultura: Reseña a "Cartografías Lingüísticas: un abordaje desde y hacia la interdisciplinariedad". Rodrigo Manríquez 504

HOMÓNIMA
Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 1-5
ISSN 3087-260X
Santiago de Chile

EDITORIAL

Carla Román

Kuminak Lefio

La Revista de Filosofía Homónima nace en junio de 2025, *ad portas* de la consumación de un entorno político hostil en Chile, y se edifica como el intento de revisitar el tema de "la importancia de las humanidades", su aplicabilidad fuera de los espacios habituales y el rol de ellas en nuestra sociedad. En nuestro caso, el rol de la filosofía en nuestra nación. Consideramos indispensable preguntarnos qué es el ejercicio que nos merecemos, si vale la pena, para qué nos sirve (si es el caso), qué alcanzas reales tiene y qué creemos que podemos hacer con ella. A priori de cualquier empresa cognoscitiva de un filósofo/a, para encausar adecuadamente cualquier intento de pregunta y respuesta filosófica. Todas preguntas legítimas y necesarias de ser respondidas por y para cualquier filósofo/a. Nuestro mayor interés es ofrecer un espacio de reflexión, en lengua hispana, tanto para novatos/as como *seniors/doctores* sobre nuestro ejercicio, acerca de por qué la tarea que hacemos es valiosa, y de qué forma lo es para cada uno, desde todas nuestras perspectivas, inconmensurables a veces, pero *orientativas* y plurales.

A esto le colocamos un nombre, "metafilosofía", y tiene una razón en particular. Un día un profesor nos dijo: "la reflexión de los métodos y el ejercicio filosófico es algo que tiene que hacer cualquier buena

filosofía". En nuestro caso, la reflexión "metafilosófica" no es algo que se da por sentado o que nos conformemos con un tratamiento secundario de la misma. Nosotros tenemos la convicción de que el ejercicio reflexivo sobre nuestra propia disciplina debe pensarse y sistematizarse, y para eso debe nombrarse, en este caso como "metafilosofía" para entenderla como una disciplina valiosa por sí sola. Inseparable de la filosofía misma, y de sus vínculos con otras áreas de la filosofía, pero que distingue su propia nobleza al descubrirla y hablar de ella, sin atemorizarnos por su gran alcance cogitabundo.

La pregunta "¿qué es la filosofía?" es tan antigua como el primer libro y, como señaló un profesor, con respuestas "tan grandes como el océano", la pregunta de qué es la filosofía sigue siendo tan importante hoy como nunca antes. Tan importante es no por una cuestión curricular, sino por que define la empresa que cada uno de nosotros desempeña día a día, cuando uno se dedica insaciablemente a estas tareas, y que impregna en cada una de las discusiones que tenemos con nuestros pares. Lo que consideramos como un ejercicio filosófico, también lo consideramos por su valor o bondad, en decir, por ser un buen ejercicio filosófico. Lo que no es buena filosofía, posiblemente sea muy difícil para nosotros considerarlo filosofía. Y lo que consideramos de valor filosófico será decisivo para el tipo de compromisos filosóficos que tomaremos como indiscutibles para todas nuestras decisiones filosóficas.

Es cierto que tenemos muchos problemas disponibles a ser discutidos en filosofía, en nuestra alma mater disciplinaria, sin embargo, algunos de nuestros mayores problemas son qué metodologías usamos para responderlas, qué es una pregunta filosófica y cómo lo hacemos para ponernos de acuerdo. Sin duda, el cómo hacer que diferentes corrientes filosóficas, ya no que se adecuen, sino solo que se comprendan, es de las mayores batallas que tenemos como disciplina. Es aquello que nos caracteriza. Somos así: "parte de una disciplina que se encarga de tratar el desacuerdo", que "opera bajo la lucha de los sistemas", que "lidia bajo un entorno hostil", sin ella nunca consolidar nada, así diferenciándose de nuestras "hermanas más disciplinadas y dominadas por la práctica". En

medio de vaguedades, malentendidos y desacuerdos profundos, el responder al cómo, qué hacemos y bajo que finalidad nos orienta en nuestras posibilidades del ejercicio filosófico. Como dijo un profesor: "¿qué llamaría la atención de estas cosas rarísimas que estudiamos los filósofos para otros?"; "¿es trasmisible esta disciplina para otras disciplinas?"; y "¿qué le trasmitiríamos de valor a ellas, que nosotros los filósofos damos por aludido, pero que no sabemos explicar?"

Nuestro interés es saber explicar por qué nuestra disciplina es valiosa tanto para nosotros, como para otros; saber entender cómo es posible orientar el conocimiento filosófico para que cale y circule dentro de nuestras sociedades, y así volverlas más solidarias y responsables. Este camino nos compromete con una perspectiva interdisciplinaria reflexiva y un vínculo mutuo con nuestro contexto y las sociedades que habitamos.

En la importancia de la reflexión filosófica, encontramos la importancia de contraer acuerdos y desacuerdos con otros y de trabajar en conjunto a nuestros pares, de manera colaborativa y facilitadora, para mejorar la comprensión de la tarea que llevamos diariamente. Entre estudiantes y académicos, la importancia del diálogo académico, el buen ejercicio y la confianza investigativa, creemos, que son indispensables para cualquier empresa cognoscitiva como la nuestra.

Lo que, como directores de la Revista Homónima, queremos consolidar es una revista académica de filosofía, tanto nacional como internacional, dirigida, por igual, a estudiantes de pregrado y posgrado en filosofía, investigadores jóvenes y académicos consolidados, cuyo propósito sea fomentar el pensamiento original y la colaboración interdisciplinaria desde las primeras etapas de la formación filosófica. Homónima impulsa la metafilosofía, la cuestión del método en filosofía y la interdisciplinariedad, trabajando en conjunto con otras disciplinas para abordar temas comunes y reflexionar sobre sus fundamentos y supuestos internos.

Queremos agradecer especialmente al Dr. Luis Placencia, por siempre motivar nuestros proyectos, por preocuparse por su consumación, por ayudarnos cuando otros no lo hacían. Pues, sin su

apoyo, Homónima no podría haberse alcanzado fácilmente. Igualmente, a la Dra. Ruth Espinoza y la Universidad Andrés Bello, que, gracias a su cálida acogida, pudimos darle un espacio institucional a la Revista. También nos gustaría agradecer a los profesores y académicos Alan Martín, Manuel Rodríguez, Felipe Morales y Sebastián Castillo. Pues, sin su colaboración y su apertura a preguntarse sobre qué es y cómo enseñar nuestra disciplina y hacerla mejor para los estudiantes, no habrían sido posible nuestras secciones de "Lectura y Escritura Filosófica" y "Lógica y Matemática". Y a todo nuestro equipo editorial, desde sus variadas contribuciones en la revisión de textos y las tareas especializadas —por ejemplo, en la traducción o creación de ensayos —, les ofrecemos nuestro eterno aprecio y la feliz proyección de seguir colaborando con ellos para hacer de la academia filosófica aquello que más nos gustaría que llegase a ser.

También nos gustaría agradecer a Nicholas Rescher[1] (1928–2024) †, por inspirarnos, formarnos desde la distancia, por acompañarnos en nuestro proceso y comprensión filosófica desde nuestros inicios. Sin su aporte, nunca habría sido posible la idealización de este proyecto.

Esta edición se caracteriza por su pluralidad temática; este número se abre para abordar temáticas tales como los desacuerdos filosóficos y científicos, el rol público del filósofo, la semiótica, el naturalismo, la conciencia y autoconciencia, entre otros. Nada de aquello pudo haber sido posible sin los estándares que cada uno aplicó en sus investigaciones y a nivel personal, así como el reflejo de la importancia de la buena ética de investigación y traducción, la confianza académica e investigativa con nuestros pares y la virtud de compartir el conocimiento de forma libre y accesible.

[1] Filósofo analítico-pragmatista, germanoestadounidense de la Universidad de Pittsburgh. Rescher se desempeñó como presidente de la American G.W. Sociedad Leibniz, American Metaphysical Society, American Philosophical Association, y C.S. Peirce Society. Además, fue el fundador de *American Philosophical Quarterly*, *History of Philosophy Quarterly*, y *Public Affairs Quarterly*.

Como revista creemos que en el compromiso de tener un comportamiento honesto, y valores como la confianza, equidad, respeto, responsabilidad y valentía en cada uno de nuestros caminos investigativos, es fundamental. Agradecemos a cada uno de los editores, revisores y autores que trabajaron bajo estos compromisos y normas, siempre respetando la pluralidad de ideas de cada uno y ofreciendo soluciones a los problemas que tuvimos en el proceso de la revista, haciendo este trabajo editorial posible.

La revista Homónima agradece a todos quienes hicieron posible esta publicación del número 1 de la revista, que ahora queda en manos de sus gentiles lectores.

Carla Román
Kuminak Lefio
Santiago, Chile.
27 / 01 / 2026

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 6-47
ISSN 3087-260X
Santiago de Chile

SOBRE LA CARENCIA DE ANÁLISIS (META)FILOSÓFICO DE LOS TÉRMINOS "CONCIENCIA" Y "AUTOCONCIENCIA" EN LA CONTEMPORANEIDAD[1]

Javier Castillo[2]
Universidad de Chile

Resumen

Este artículo señala una carencia en la filosofía de la mente contemporánea: la ausencia de un análisis adecuado del concepto de "conciencia". Siguiendo a Balibar (2013), se sostiene que su actual polisemia tiene un origen lockeano. Al trazar una genealogía del

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Castillo, J. (2026). Sobre la carencia de análisis (meta)filosófico de los términos "conciencia" y "autoconciencia" en la contemporaneidad. *Revista de Filosofía Homónima, 1*(1), pp. 6-47. https://doi.org/xxxxxxx[Por designar]

[2] Javier Castillo es Licenciado en Filosofía, Universidad de Chile. Magíster en Filosofía, Universidad de Chile. Actualmente me encuentro cursando el programa de Doctorado en Filosofía en la Universidad de Chile, bajo la Beca de Doctorado Nacional ANID (año académico 2022, folio 21221670). Email: javier.e.castillo.v@gmail.com. ORCID: https://orcid.org/0000-0002-4893-5425

concepto, es posible identificar una bifurcación histórica no suficientemente examinada entre una tradición constitutiva (de Descartes a Kant) y otra psicológico-descriptiva (iniciada con Locke). Frente a este panorama, la tradición que va de Descartes a Kant —reinterpretada por autores como Placencia (2017) y Hennig (2007), y articulada mediante el "modelo de propiedad" (apoyado en Castillo & Fuentes, 2024)— ofrece una concepción que podría esclarecer ciertas controversias vigentes. Finalmente, se propone que la estrategia (meta)filosófica, aplicada a esta genealogía, puede revelar la contingencia de dichos problemas, mostrando que surgen precisamente de la adopción acrítica del paradigma lockeano.

Conceptos clave: *Paradigma constitutivo, paradigma psicológico-descriptivo, disolución, modelo de propiedad, bifurcación conceptual.*

On the lack of (meta)philosophical analysis of the terms "consciousness" and "self-consciousness" in contemporary thought

Abstract

The present paper points out a deficiency in contemporary philosophy of the mind: the absence of an adequate analysis of the concept of "consciousness". Following Balibar (2013) it is argued that the current polisemy of the term has a Lockean origin. By tracing the genealogy of the concept, it is possible to identify a historical bifurcation that has not been sufficiently examined, between a constitutive tradition (from Descartes to Kant) and a psychological-descriptive tradition (that began with Locke). Given this scenario, the tradition that goes from Descartes to Kant —reinterpreted by authors such as Placencia (2017) and Hennig (2007), and articulated through the "property model" (supported by Castillo & Fuentes, 2024)— offers a perspective that

could clarify certain current controversies. Finally, it is proposed that the (meta)philosophical strategy, applied to this genealogy, can reveal the contingency of such problems, showing that they arise precisely from the uncritical adoption of the Lockean paradigm.

Key words: *Constitutive paradigm, psychological-descriptive paradigm, dissolution, property model, conceptual bifurcation.*

1. Introducción. La Metafilosofía como constitutiva de la práctica filosófica

Toda empresa filosófica que aspire al rigor se encuentra, en su mismo umbral, con una pregunta inevitable y fundacional: ¿qué es la filosofía? Esta interrogante, lejos de ser un preámbulo accesorio o una mera cuestión metodológica preliminar, constituye el gesto autorreflexivo que, desde sus orígenes socráticos, ha definido la peculiar naturaleza del quehacer filosófico. Como señala Castillo, el pensar filosófico se distingue por un componente reflexivo esencial, poseyendo un doble objeto: un fenómeno determinado y la investigación misma sobre dicho fenómeno (2021, p. 22). Esta capacidad de volverse sobre sí misma no es un atributo secundario, sino la característica que, de manera primordial, diferencia a la filosofía de otras disciplinas, como las ciencias positivas, las cuales operan típicamente sobre objetos externos a su propia práctica sin cuestionar de manera permanente sus fundamentos.

Es en este contexto donde adquiere relevancia el término "metafilosofía". En un sentido estricto, siguiendo a Nudler (2012), este concepto designa la reflexión sistemática que la filosofía ejerce sobre su propia naturaleza, sus métodos, sus fines, sus límites y su lugar dentro del entramado cultural e intelectual. Sin embargo, desde el principio, esta definición encierra una paradoja constitutiva. La expresión "filosofía de la filosofía" parece, por un lado, establecer una distinción de niveles —un metalenguaje sobre el lenguaje objeto de la filosofía— pero, por otro,

inevitablemente colapsa en una identidad: reflexionar filosóficamente sobre la filosofía es, en sí mismo, un acto filosófico. Esta autorreferencialidad no es una patología del pensamiento, sino su condición de posibilidad. A diferencia de disciplinas como la física o la biología, que cuando desean examinar sus fundamentos deben recurrir a una mirada externa (la filosofía de la ciencia, la historia de la ciencia), la filosofía lleva a cabo esta tarea desde su propio interior; "sigue habitando su propia casa" (Nudler, 2012, p. 20). El "tomarse a sí misma como problema formaría parte de la naturaleza misma de la filosofía" (Nudler, 2012, p. 20).

Esta íntima relación suscita una cuestión fundamental: ¿es la metafilosofía un dominio separable de la filosofía misma? Aunque es posible —y frecuente— realizar filosofía sin tematizar explícitamente la pregunta por su esencia, toda práctica filosófica conlleva, de modo implícito, una concepción de lo que está haciendo. La elección de problemas, los métodos empleados, el lenguaje utilizado y la actitud hacia la historia de la disciplina revelan, "por activa o por pasiva, la idea de filosofía que lo inspira" (Nudler, 2012, p. 21). La metafilosofía, por tanto, es esencial no porque todo filósofo deba escribir un tratado sobre ella, sino porque opera como un sustrato inevitable que configura y dirige todo pensamiento filosófico[3].

Esta reflexión conduce a una conclusión radical que este artículo adopta como punto de partida: la filosofía es, en su mismo núcleo, (meta)filosófica[4]. Esta tesis encuentra un respaldo en la tradición. Ya

[3] Las críticas a la metafilosofía, como la de autores que la ven como un síntoma de agotamiento o un ensimismamiento estéril (Esposito y Nancy, 2006, p. 163), caen en una paradoja performativa: la negación misma de la metafilosofía es un acto metafilosófico (Nudler, 2012, p. 21). Negar su relevancia implica, necesariamente, ejercerla. Por ello, lejos de ser una práctica superflua, la metafilosofía es la autoconciencia de la tensión constitutiva de la filosofía: su permanente apertura a lo problemático y su irreductible vocación por dar un sentido comprensivo no solo al mundo, sino también a su propio lugar en él (Nudler, 2012, p. 27).

[4] Aunque el tratamiento detallado de esta cuestión excede los límites propuestos para el presente artículo, conviene advertir al lector sobre el uso de la expresión "(meta)filosofía". Esta elección se justifica por dos motivos principales. En primer lugar,

Sócrates, a partir del oráculo délfico, inauguraba la práctica filosófica no como la posesión de un saber positivo, sino como la conciencia de la propia ignorancia, un "saber en grado sumo" que consistía precisamente en "ser consciente de ignorar algo" (Abalo, 2015, p. 12; cf. Platón, *Apología*, 21d). Se trata de un conocimiento de segundo orden, un "conocimiento del conocimiento" que sella el vínculo indisoluble entre filosofía y autoconciencia. Kant, a su vez, en la *Crítica de la razón pura*, identificaba la tarea de la razón madura como el "conocimiento de sí", un tribunal que debe determinar "tanto sus fuentes, como del alcance y de los límites" de su propio conocimiento (KrV, A xi-xii). Para Kant, este autoconocimiento trascendental es una tarea necesaria para evitar el error metafísico que surge de la ignorancia respecto de los límites de nuestras capacidades cognitivas (Placencia, 2015, pp. 548-549). En último término, la filosofía se revela como una forma de autoconocimiento, una actividad en la que, al ocuparnos de objetos, nos hacemos objetos a nosotros mismos mediante esa investigación (Castillo, 2021).

2. Hacia un análisis (meta)filosófico de la conciencia: La carencia contemporánea y la necesidad de una genealogía

Partiendo de la comprensión de la filosofía como una práctica intrínsecamente autorreflexiva, este artículo aborda una carencia significativa en el panorama filosófico contemporáneo: la ausencia de un análisis (meta)filosófico y genealógico de los términos centrales "conciencia" y "autoconciencia" dentro de la filosofía de la mente. El

como se argumentó, la filosofía parece poseer de modo connatural un "doble objeto", característica que el prefijo "meta" busca reflejar. No obstante, es también preciso señalar que el término "filosofía", por sí mismo, podría considerarse suficiente para transmitir este sentido. En segundo lugar, y en atención a los lineamientos temáticos de la revista, se ha optado por emplear la denominación "metafilosofía", ya que este vocablo se ha consolidado en el contexto contemporáneo para destacar dicha particularidad disciplinar. Su uso aquí pretende, por tanto, ajustarse a dicha convención y evitar posibles confusiones terminológicas.

objetivo central es argumentar la necesidad y fecundidad de dicho enfoque, proponiendo que muchos desacuerdos actuales en torno a la conciencia podrían disolverse mediante un rastreo de las condiciones históricas y conceptuales que configuraron su uso predominante.

La metodología propuesta es doble. En primer lugar, se adopta una *perspectiva genealógica* destinada a rastrear el sentido actual del concepto de "conciencia" mediante el examen de sus antecedentes históricos más relevantes, identificando aquel punto de inflexión en el que este concepto adquirió los rasgos que definen su empleo contemporáneo. En segundo lugar, se aplica un *enfoque (meta)filosófico crítico*, es decir, una reflexión sobre el propio marco conceptual en el que dichos términos se formulan y discuten.

En este caso en particular, aplicar dicha estrategia a ciertos problemas contemporáneos sobre la conciencia consiste en llevar a cabo un análisis que exponga tanto su momento de inflexión histórica en Locke como el marco conceptual desde el cual surge su uso restringido. De este modo, podría demostrarse que algunos debates actuales no constituyen necesariamente indicios de un misterio metafísico insondable, sino síntomas de un uso categorial no examinado.

El núcleo de la genealogía aquí esbozada sigue la interpretación de Balibar (2013), quien identifica en Locke el punto de inflexión decisivo. Según Balibar, es en Locke donde la "conciencia" (*consciousness*) se consolida como un concepto central y adquiere su carácter moderno, psicológico y epistemológico, hasta el punto de poder considerarse una "invención" del concepto moderno (p. 73 ss.). Mientras en borradores de la década de 1670 del *Ensayo sobre el Entendimiento Humano* el término estaba notablemente ausente, en la edición de 1690, y especialmente con la adición del capítulo sobre la identidad personal (II.xxvii) en 1694, se convierte en el fundamento mismo del *sí mismo* (*self*) (Balibar, 2013, p. 100). Esta tradición, que concibe la conciencia predominantemente como un fenómeno psicológico, un campo de interioridad y la base de la identidad personal, ha marginalizado otras dimensiones esenciales.

Efectivamente, la herencia lockeana se manifiesta en la polisemia radical y a menudo no examinada que caracteriza el uso contemporáneo del término. Como advierte Van Gulick (2014), los múltiples sentidos de "consciente" y "conciencia" distan de ser unívocos, y el fracaso en distinguirlos genera confusiones y desacuerdos más verbales que sustanciales. Usos como "conciencia fenoménica" en el contexto de la "brecha explicativa" (Levine, 1983) o "autoconciencia" en teorías de orden superior son herederos directos de esta psicologización (Rosenthal, 2005), la cual sentó las bases para su posterior naturalización.

No obstante, la tradición lockeana no es la única disponible en el panorama moderno. Otras caracterizaciones de la conciencia, que han recibido considerablemente menos atención en los debates contemporáneos, se desarrollaron en paralelo. Una línea de pensamiento que va de Descartes a Kant, por ejemplo, consolidó una comprensión de la conciencia como certeza inmediata (el *cogito*) y como fuente de normatividad constitutiva. En particular, Kant despliega una riqueza de sentidos que la filosofía posterior ha tendido a simplificar. De su obra, en este artículo se mencionarán dos acepciones cruciales: 1) la autoconciencia como apercepción trascendental (*Selbstbewusstsein*); 2) la conciencia como *Gewissen* (conciencia moral); ambas bajo la interpretación de un modelo de propiedad (Castillo & Fuentes, 2024). Estas dimensiones han sido, en el mejor de los casos, relegadas o reinterpretadas de forma reduccionista frente al paradigma psicológico lockeano.

Cabe señalar que una disolución completa mediante una genealogía exhaustiva excede los límites de este trabajo. Este artículo se concibe como un primer paso programático, cuyo objetivo inmediato es diagnosticar la carencia, esbozar la metodología y trazar el mapa de la bifurcación conceptual entre la tradición lockeana-psicológica y otras tradiciones marginadas. La ejecución completa de este programa, que exigiría un análisis textual detallado de las tradiciones aristotélica, cartesiana, lockeana y kantiana, así como de su recepción contemporánea, se plantea como una tarea pendiente para futuras investigaciones.

3. Locke: La invención de la conciencia psicológica y el *Self*

Como se anticipó, de acuerdo con la tesis central de Étienne Balibar (2013), desarrollada en *Identity and Difference: John Locke and the Invention of Consciousness*, es en la obra de John Locke donde el concepto de "conciencia" (*consciousness*) experimenta una transformación decisiva que puede describirse como una verdadera "invención" del concepto moderno de conciencia (Balibar, 2013, p. 73)[5]. Según Balibar, Locke no solo consolida la conciencia como un concepto filosófico central, sino que la establece como el fundamento de una nueva concepción del sujeto individual moderno, un sujeto que "secretamente se llama a sí mismo *My Self*" (Balibar, 2013, p. 121). Este giro es tan radical que configura gran parte del debate filosófico posterior sobre la subjetividad.

A. El giro anti-lingüístico y el aislamiento de lo mental

Locke lleva a cabo lo que Balibar denomina un "giro anti-lingüístico" o la instauración de un "psicologismo" en filosofía (Balibar, 2013, p. 104). Este giro se fundamenta en una separación radical entre el dominio de lo mental (pensamiento, ideas) y lo verbal (lenguaje, palabras). En el Libro III del *Ensayo*, Locke argumenta que "las palabras, en su significación primaria o inmediata, no representan nada más que las ideas en la mente de quien las usa" (Locke, *Ensayo*, III.ii.2, citado en Balibar, 2013, p. 103). Las palabras son, por tanto, "signos de concepciones internas" y "marcas de las ideas dentro de su propia mente" (Balibar, 2013, p. 103).

Esta separación permite a Locke postular la existencia de "proposiciones mentales" y "proposiciones verbales", distinguiendo así entre una "verdad mental" (*vérité mentale*) y una "verdad verbal" (*vérité*

[5] Para profundizar en la historia del término 'conscience' y 'consciousness' en los predecesores de Locke se puede revisar a Thiel (1991, 1994) y Davies (1990).

verbale) (Balibar, 2013, p. 103). Según Balibar, "el nacimiento del psicologismo en filosofía" es "menos el resultado de una crítica de la idea de una verdad necesaria existente en sí misma que una descalificación del lenguaje como elemento originario del pensamiento (un 'giro anti-lingüístico' por así decirlo)" (2013, p. 104). Este psicologismo no es el efecto del nacimiento de una psicología, sino su condición de posibilidad, "el programa que los psicólogos buscarán llevar a cabo" (Balibar, 2013, p. 104).

B. La conciencia como principio lógico-psicológico de identidad

La refutación lockeana de la doctrina de las ideas innatas en el Libro I del *Ensayo* constituye el momento fundacional de su teoría de la conciencia. Locke relativiza los principios lógicos universales (como "lo que es, es") como creencias adquiridas, no innatas, para luego redescubrirlos inscritos en la estructura misma de la mente. La tesis fundamental que emerge es que "es imposible que el hombre no piense que piensa, o que piense sin pensar" (Balibar, 2013, p. 108). Un "pensamiento no consciente" es, por tanto, una contradicción en los términos, un "no-pensamiento".

La conciencia se convierte así en el principio de no-contradicción y de auto-identidad de la mente. Balibar argumenta que Locke redescubre aquí las funciones de la certeza que caracterizaban al "cogito" cartesiano, pero distribuyéndolas a lo largo de una constitución teórica del sujeto, en lugar de concentrarlas en el enunciado puro del 'yo' (Balibar, 2013, p. 106). La singularidad del sujeto (el *self*) puede describirse sin necesidad de que este se enuncie inmediatamente en primera persona, como en la meditación cartesiana. La conciencia es, en esencia, "la presencia para sí, como 'percepción', de una acción que sucede" (Balibar, 2013, p. 119), y este "percatarse de lo que pasa en la propia mente" constituye el principio de identidad en trabajo.

C. La topografía del pliegue: Conciencia como sentido interno y reflexión

Locke rompe con la concepción de la conciencia como una entidad sustancial (el alma como *res cogitans*). En su lugar, la concibe como un sistema o *modo de operación*. Balibar utiliza el término "topografía" para describir la concepción lockeana: un "mapa de las relaciones entre interioridad y exterioridad" (2013, p. 123). La conciencia no es una cosa que se posee, sino un campo de relaciones (interior/exterior, sensación/reflexión, palabra/idea) que constituyen lo mental.

En el Libro II, Locke postula un "sentido interno" (*internal sense*) o reflexión, como la segunda fuente de ideas, junto con la sensación. Balibar desarrolla una detallada "topografía de la conciencia" lockeana para explicar esta estructura. La mente se vuelve sobre sí misma en un movimiento de "reflexión" que crea un "pliegue" (*repli*). Este pliegue define una nueva interioridad, internalizando la exterioridad de la sensación y, a su vez, desplegando "una cuasi-exterioridad en el corazón de la interioridad" (Balibar, 2013, p. 116) [6]. Desde esta perspectiva, el "pliegue" descrito por Balibar no sería un movimiento homogéneo, sino

[6] Cabe aclarar que con esta posición Balibar no afirma que el sentido "reflexivo" sea el único posible. La metáfora del "pliegue" podría sugerir una reflexión omnipresente y constante, lo que entraría en conflicto con los pasajes donde Locke señala que los niños y los adultos no atentos no reflexionan, aunque sí son conscientes. Weinberg resuelve esto al separar conceptualmente los términos (2021, p. 132). A partir de esta lectura, se podría criticar la posición de Balibar en la medida en que entraría en conflicto con los pasajes donde Locke señala que los niños y los adultos no atentos no reflexionan, aunque sí son conscientes. Weinberg resuelve esto al separar conceptualmente los términos (2021, p. 132).
Sin embargo, no creo que la precisión de Weinberg contradiga la posición de Balibar, en cambio, pienso que esta lectura topográfica puede ser enriquecida con el análisis funcional de Weinberg (2021). Para Weinberg, la reflexión es específicamente un acto voluntario y de segundo orden que genera ideas (ideas de reflexión), mientras que la conciencia es un elemento intrínseco y no-ideacional de toda percepción (pp. 131-132). Weinberg distingue ambos conceptos para resolver problemas de coherencia en el sistema lockeano, como el regreso al infinito y la violación del principio empirista. La conciencia, como elemento intrínseco, no es una fuente de ideas (pp. 128-129).

que contiene una distinción operativa crucial: la vivencia inmediata (conciencia) y la generación de conocimiento objetivo sobre ella (reflexión).

Locke utiliza la metáfora del "escenario" (*stage*) interno, donde los pensamientos son actores que desfilan ante la mirada de la mente. La conciencia es, entonces, "la percepción de lo que pasa en nuestra propia mente" (Locke, *Ensayo*, II.i.19), una percepción que duplica las operaciones mentales y las convierte en objeto de su propia contemplación. Weinberg (2021) argumentaría que esta "mirada" tiene dos modalidades radicalmente distintas: una es la conciencia misma (la percepción de "mismo orden" que acompaña ineludiblemente a cada acto mental), y la otra es la reflexión propiamente dicha (el acto de atender y formar ideas a partir de esa conciencia) (p. 132). La primera es pasiva e inevitable; la segunda, activa y contingente. Así, la estructura de duplicación y pliegue es lo que permite la autoconciencia y el autoconocimiento, pero su mecánica interna es más compleja de lo que la metáfora espacial sugiere por sí sola.

En cualquier caso, lo interesante es la aparición de la noción de "interioridad". La interioridad no es simplemente lo opuesto a lo exterior, sino un espacio complejo donde lo exterior es interiorizado y reproyectado, "un interior que idealmente contiene tanto a sí mismo como a su otro" (Balibar, 2013, p. 123). En definitiva, mientras la topografía de Balibar describe la arquitectura del espacio mental lockeano, el análisis de Weinberg detalla los *mecanismos generativos* que operan dentro de él, mostrando que la posibilidad de una psicología como ciencia se funda no en el mero "pliegue", sino en la capacidad distintiva de la reflexión para transformar la experiencia vivida en ideas objetivables[7].

[7] Esta síntesis sugiere que las lecturas de Balibar (2013) y Weinberg (2021) son complementarias más que antagónicas. Balibar ofrece el marco histórico-estructural de la "invención" de la conciencia moderna, mientras que Weinberg proporciona la microfísica de su funcionamiento, asegurando su coherencia lógica interna. Juntas, brindan una mejor comprensión de la revolución lockeana.

D. La conciencia como fundamento del *Self* temporal y responsable

La culminación de este sistema es la teoría lockeana de la identidad personal, desarrollada en el célebre capítulo añadido en 1694 (II.xxvii). Para Locke, la identidad personal no reside en la sustancia (ya sea material o inmaterial), sino exclusivamente en la autoconciencia (*self-consciousness*) y en la continuidad de la conciencia extendida en el tiempo a través de la memoria.

Balibar describe este como el "Cogito" lockeano, fundamentalmente diferente del cartesiano. Mientras que Descartes se sitúa en la certeza puntual del "yo pienso", Locke sitúa la identidad del *self* en "el elemento 'histórico' de la memoria". La conciencia implica intrínsecamente una temporalidad, una "retención" del pasado y una "protención" del futuro (Balibar, 2013, p. 119). Es esta conciencia extendida en el tiempo la que "une existencias y acciones muy remotas en el tiempo en la misma persona" (Locke, *Ensayo*, II.xxvii.16).

Este *self* temporal es, ante todo, un sujeto moral y jurídico. La identidad personal es un "término forense" (*Forensick Term*), que "apropia acciones y su mérito" (Balibar, 2013). La persona es responsable de todas aquellas acciones de las que puede apropiarse mediante la conciencia, y será llamada a rendir cuentas por ellas, incluso en el Juicio Final. Balibar subraya que la retención del pasado se une a la conciencia presente "en virtud del horizonte del juicio por venir en el que está siempre ya inscrita" (2013, p. 120). Así, la unidad del sujeto lockeano es "indisolublemente lógica (auto-idéntica), moral y jurídica (responsabilidad, apropiación) y psicológica (interioridad y temporalidad)" (Balibar, 2013, p. 121).

E. Conclusión: La invención del sujeto moderno

Para Balibar, la invención lockeana de la conciencia no sacrifica las significaciones simbólicas tradicionales asociadas al individuo, pero las seculariza al historicizar e inscribir las marcas de la trascendencia en la inmanencia de la relación de la mente consigo misma. Locke construye una topografía compleja donde la conciencia se define por una serie de articulaciones problemáticas (interior/exterior, sensación/reflexión, palabra/idea, pasividad/actividad, afección/percepción) que prescribirán de antemano los emplazamientos teóricos para gran parte de los problemas futuros de la filosofía de la conciencia, incluida la hipótesis de un "inconsciente".

De este modo, la conciencia lockeana, como principio lógico, sentido interno y fundamento de un *self* temporal y responsable, constituye la primera gran doctrina moderna del sujeto individual, un sujeto que "secretamente se llama a sí mismo *My Self*" y cuya invención, según Balibar, "sigue siendo interminable" (2013).

4. La taxonomía contemporánea como herencia del paradigma lockeano

La taxonomía contemporánea de los sentidos de "conciencia", herramienta fundamental para delimitar los problemas en filosofía de la mente, es en gran medida la sistematización y radicalización del giro psicológico e internalista inaugurado por John Locke. Como señala Berger (2020), "un obstáculo mayor para el estudio de la conciencia es que las personas usan el término 'conciencia' y expresiones relacionadas como 'consciencia' de muchas maneras diferentes" (p. 252). La organización tripartita propuesta por Van Gulick (2014) –*creature consciousness*, *state consciousness* y *consciousness as an entity*–, junto con sus subcategorías, puede entenderse como la articulación moderna de este paradigma, donde la conciencia se aborda predominantemente como un fenómeno naturalizable, interior y analíticamente desagregable.

En lugar de un análisis pormenorizado de cada subcategoría, es más ilustrativo ofrecer una caracterización general de cada tipo y mostrar

cómo, en conjunto, son herederos de la redefinición lockeana. Este enfoque permite concentrarse en el origen lockeano de dos caracterizaciones cruciales para los debates actuales: la "conciencia fenoménica" (en el contexto de la llamada "brecha explicativa") y la "autoconciencia" tal como es teorizada en los modelos de orden superior. Ambas son descendientes directos de la psicologización lockeana.

4.1. Caracterización general y herencia lockeana

(1) **Conciencia de criatura (*Creature Consciousness*):** Se refiere a la atribución de conciencia a un organismo en su totalidad. Incluye sentidos graduales como la sensibilidad (*sentience*), la vigilia (*wakefulness*), la autoconciencia (*self-consciousness*) y la noción de "lo que se siente ser" (*what-it-is-like*) (Van Gulick, 2014, §2.1). El núcleo de esta herencia radica en el giro lockeano que establece la conciencia no como un atributo binario de un alma sustancial, sino como un conjunto de capacidades y estados cognitivos graduales, dependientes de la actualización de la reflexión. Locke funda un criterio funcional: la conciencia requiere la "percepción de lo que pasa en la mente de un hombre" (Locke, *Ensayo*, II.i.19), anticipando así la distinción entre mera vida y vida consciente.

(2) **Conciencia de estado (*State Consciousness*):** Atribuye conciencia a estados mentales particulares. Incluye variantes como los estados fenoménicos (*phenomenal*), los estados de los que se es consciente (*conscious-of*), la conciencia de acceso (*access consciousness*) y la conciencia narrativa (Van Gulick, 2014, §2.2). La correspondencia aquí con la "topografía" lockeana es directa. Al definir la conciencia como percepción interna (reflexión), Locke convierte cualquier estado mental en potencial objeto de una conciencia de segundo orden. Su arquitectura, que distingue entre el contenido de la sensación y su percepción reflexiva,

prefigura las distinciones contemporáneas entre contenido representacional y su cualidad fenoménica o su disponibilidad para sistemas cognitivos superiores (acceso, narrativa).

(3) **Conciencia como entidad (*Consciousness as an Entity*):** Considera la conciencia como una "cosa" o sustancia en sí misma. Van Gulick (2014, §2.3) nota que esta visión es minoritaria hoy, pues predomina entender la conciencia como una propiedad o proceso. Locke es pivotal aquí: su teoría desplaza la conciencia del estatus de *res cogitans* cartesiana a un principio de operación o "campo de relaciones" (Balibar, 2013, p. 123). Al hacer de la conciencia el principio lógico-psicológico de identidad personal fundado en la memoria, no en una sustancia, Locke anticipa los modelos relacionales y funcionales contemporáneos.

El origen lockeano de esta taxonomía es, por tanto, *estructural*. Su "giro anti-lingüístico" (Balibar, 2013, p. 104), que aísla lo mental como un dominio interno investigable, crea el espacio conceptual para desagregar analíticamente la conciencia en facultades, estados y funciones. Esto hace posible la pregunta descriptiva (*what?*) que organiza la taxonomía (Van Gulick, 2014, §3). Sin embargo, dos herencias específicas merecen atención detallada por su centralidad en los debates explicativos.

4.2. Herencias específicas: Conciencia fenoménica y autoconciencia

a) Conciencia Fenoménica: De la "idea del pensamiento" a la brecha explicativa.

La noción lockeana de conciencia como *percepción interna* sienta las bases para la distinción moderna entre el *carácter cualitativo* y la *estructura fenoménica*, dos aspectos cruciales de la "pregunta descriptiva" (*what?*) (Van Gulick, 2014, §3).

1. **Carácter Cualitativo ("*Qualitative Character*"):** Van Gulick (2014, §4.2) lo equipara a los *qualia* o "sensaciones crudas" (*raw feels*), como la rojez experimentada al ver un tomate. Locke, al

hacer de la reflexión una fuente de "ideas simples" análoga a la sensación, genera justamente la "idea del pensamiento mismo" (Balibar, 2013, p. 112). Esta idea es el correlato introspectivo de la vivencia, el primer registro filosófico de que los actos mentales tienen una cualidad experiencial propia y accesible. Su descripción de la conciencia como aprehensión de "lo que pasa" (*what passes*) captura la intuición nuclear detrás de la pregunta "¿qué se siente ser?" (*what-it-is-like*) (Nagel, 1974). Sin embargo, Locke también anticipa la complejidad del "espacio cualitativo" (*qualia space*). Su empirismo, al derivar todas las ideas de la sensación o la reflexión, sugiere que las cualidades no son propiedades simples y *sui generis*, sino que se organizan en redes de semejanza y diferencia (p.ej., el espectro de colores) que podrían ofrecer "ganchos" (*hooks*) para conexiones psicofísicas inteligibles (Van Gulick, 2014, §4.2).

2. **Estructura Fenoménica ("*Phenomenal Structure*"):** Van Gulick (2014, §4.3) advierte que lo fenoménico va más allá de los *qualia* e incluye la organización intencional rica de la experiencia: representaciones de tiempo, espacio, causalidad, yo y mundo. La "topografía" lockeana del pliegue (Balibar, 2013) es una cartografía pionera de esta estructura. Al duplicar la operación mental en un segundo nivel de percepción, Locke no solo aísla una cualidad, sino que instaura una *relación intencional reflexiva* que es constitutiva de lo mental. Su análisis de la identidad personal, fundada en la memoria que teje una continuidad temporal, es una descripción primigenia de la *estructura narrativa* y diacrónica de la conciencia fenoménica. Esto refleja la visión "densa" (*thick view*) de la fenomenología, según la cual la experiencia consciente está imbuida de contenido cognitivo y no se limita a sensaciones "puras" (Van Gulick, 2014, §4.3).

El legado decisivo aquí es que, al hacer de la subjetividad fenoménica el dato primario de la psicología, Locke crea las condiciones para la "brecha

explicativa" (*explanatory gap*). La pregunta explicativa (*how?*) de Van Gulick (2014, §5) busca entender cómo estados no conscientes (neurales, físicos) pueden dar lugar a la conciencia. El núcleo duro de este problema es la P-conciencia o conciencia fenoménica (Van Gulick, 2014, §5.1). La definición lockeana, al establecer la conciencia como un dominio de percepción interno irreducible a lo meramente físico-funcional, dibuja los límites de esa brecha. La dificultad para derivar *a priori* la "idea del pensamiento" a partir de descripciones cerebrales, o para hacer inteligible el vínculo (*intelligible link*) entre ellas (Van Gulick, 2014, §5.1, §5.2), es el heredero directo de haber fundado la mente en la autotransparencia de la reflexión.

b) Autoconciencia y el modelo de orden superior:

La teoría lockeana de la identidad personal provee la arquitectura básica para las teorías de orden superior (*Higher-Order theories*) que dominan la explicación contemporánea de la autoconciencia.

Esta posición puede describirse como un "modelo de duplicación reflexiva". Para Locke, ser consciente de un deseo o un recuerdo implica un acto de reflexión (percepción interna) dirigido a ese estado[8]. Esta es exactamente la estructura central de las teorías HO: un estado mental M es consciente si y solo si está acompañado por un estado meta-mental (de orden superior) cuyo contenido es que uno está en M (Van Gulick, 2014, §9.1). La "pregunta forense" lockeana –¿de qué acciones puedo apropiarme mediante la conciencia?– opera bajo este mismo principio: la responsabilidad moral requiere ese acto de auto-atribución reflexiva. Esto resuelve, en su origen, el "problema de la generalidad" (*generality problem*) esbozado por Van Gulick (2014, §9.1): ¿por qué pensar en un

[8] Cabe aclarar que, en este contexto, se aborda únicamente uno de los sentidos del término "conciencia" en el pensamiento de Locke. Como señala Weinberg (2016), es posible distinguir en Locke al menos dos acepciones: la conciencia como (1) un estado mental inseparable del acto de percepción que nos hace conscientes de nosotros mismos como seres perceptivos, y (2) el yo continuo del que somos conscientes en dichos estados conscientes (p. 153). Esta sección se centrará en desarrollar esta segunda acepción propuesta por la autora.

estado mental lo hace consciente, pero pensar en una papa no? Para Locke, la respuesta está en la naturaleza misma de lo mental como aquello que *es* percepción de sí; la reflexión no añade conciencia a algo exterior, sino que actualiza la autotransparencia constitutiva de la operación mental.

En síntesis, la psicologización lockeana no solo suministra el léxico inicial para hablar de la conciencia fenoménica y la autoconciencia, sino que diseña el mismo *campo de batalla* filosófico. Al definir la conciencia como un espacio interno de cualidades accesibles y relaciones reflexivas, Locke plantea de manera inaugural el desafío de naturalizar lo fenoménico (la brecha) y ofrece el esquema estructural (la duplicación reflexiva) sobre el que se construirán las principales teorías explicativas de la autoconciencia. Su legado es, por tanto, tanto descriptivo como programático.

5. Críticas de Bennett y Hacker al modelo lockeano

Las críticas de Bennett y Hacker (2022) al uso contemporáneo de la noción de conciencia, desarrolladas en su análisis de la neurociencia y la filosofía de la mente, constituyen una diagnosis profunda de los problemas conceptuales generados por la predominancia acrítica del paradigma lockeano. Aunque sus críticas se dirigen explícitamente a las prácticas y afirmaciones de neurocientíficos y filósofos contemporáneos[9] y no explícitamente al uso moderno del término, me parece que el marco que desmontan es precisamente el heredado de la "invención" lockeana de la conciencia psicológica descrita en la sección 3. Locke redefinió la conciencia (*consciousness*) como un "sentido interno" o reflexión, un principio lógico-psicológico de identidad personal, y un espacio interno de cualidades accesibles. La taxonomía contemporánea analizada en la sección 4 —especialmente las nociones de "conciencia fenoménica" y

[9] Los autores y obras a los que Bennett y Hacker principalmente dirigen su crítica son: Crick (1995); Crick y Koch (1998); Edelman y Tononi (2000); Tononi (2005); Searle (1984); Dehaene, Lau y Kouider (2017); McGinn (1987); y Gazzaniga (1997).

"autoconciencia" reflexiva— es la sistematización de este legado. Las críticas de Bennett y Hacker, por tanto, aunque formuladas en el contexto de debates actuales, son perfectamente aplicables a las características fundamentales del modelo lockeano, revelando los vicios conceptuales que este introdujo y que la neurociencia cognitiva ha naturalizado. En última instancia, su obra es una crítica radical al modelo psicológico y reflexivo del uso del término "conciencia", tal como fue institucionalizado por Locke y amplificado por la ciencia cognitiva moderna.

5.1. La falacia mereológica: La interiorización lockeana y su localización cerebral errónea

Una de las críticas más fundamentales de Bennett y Hacker es lo que denominan "la falacia mereológica en la neurociencia": el error lógico de atribuir a una parte de un animal (específicamente, el cerebro) propiedades o capacidades que solo tienen sentido atribuir al animal como un todo (Bennett & Hacker, 2022, p. 266). Los autores documentan cómo neurocientíficos prominentes (Edelman, Tononi, Gazzaniga, Dehaene) y filósofos (Searle, McGinn) afirman rutinariamente que "la conciencia surge en el cerebro", que es "una propiedad de una red neural", o que "el cerebro causa la conciencia" (Bennett & Hacker, 2022, pp. 265-267). Esta falacia, argumentan, es un caso particular de atribuir a una parte (el cerebro) un atributo que solo se predica legítimamente del organismo completo.

Esta falacia es una consecuencia directa y casi inevitable de la interiorización lockeana de la conciencia. Al redefinirla como la "percepción de lo que pasa en nuestra propia mente" (Locke, citado en Balibar, 2013), Locke desplaza el *locus* de la conciencia de la relación entre el animal y su entorno (propia de Aristóteles) o de la autotransparencia constitutiva del pensamiento (propia de Descartes), hacia un espacio interno, un "escenario" o "teatro" mental (Balibar, 2013,

p. 116). Este espacio, aunque inicialmente no se identifica con un órgano físico, establece una topografía de interioridad que invita a su localización material. Si la conciencia es algo que "pasa dentro", y si el cerebro es el órgano central de la vida mental, entonces es un paso corto —pero conceptualmente fatal— concluir que la conciencia "pasa" *en el cerebro.* Como señalan Bennett y Hacker, "no es el cerebro el que está consciente o inconsciente, sino la persona cuyo cerebro es" (2022, p. 269). El error lockeano, por tanto, no está meramente en señalar un "interior", sino en concebirlo como un contenedor de eventos psicológicos que luego pueden ser mal-asignados a una de sus partes materiales. La falacia mereológica es, así, el síntoma de una psicologización que ha perdido de vista al sujeto vivo y unitario.

5.2. El mito de la privacidad y el acceso privilegiado como dogmas lockeanos naturalizados

Bennett y Hacker critican lo que consideran uno de los dogmas centrales de la filosofía de la mente contemporánea, heredado directamente de la epistemología lockeana: la idea de que la conciencia es un fenómeno esencialmente privado, al que solo el sujeto tiene un acceso privilegiado e infalible (Bennett & Hacker, 2022, p. 267). Neurocientíficos como Damasio, Edelman y Tononi describen la conciencia como un "fenómeno enteramente privado, de primera persona" (Bennett & Hacker, 2022, p. 267). Filósofos como Searle y Dennett argumentan que mi estado de conciencia me es accesible de una manera radicalmente diferente y superior a como lo es para cualquier otro (Bennett & Hacker, 2022, p. 267).

Este dogma es la herencia directa de la psicologización lockeana. Al definir la conciencia como un "sentido interno", Locke la constituye como un espacio de observación al que, por definición, solo el propio sujeto puede acceder directamente. La reflexión es, en su modelo, un ojo interior privado que solo yo puedo usar para mirar mis propias ideas. Esta

es la raíz de la privacidad psicológica moderna. Sin embargo, Bennett y Hacker argumentan que esta concepción es un *error conceptual*. No hay nada "intrínsecamente privado" en la conciencia intransitiva: que alguien recupere la conciencia después de una operación es un hecho públicamente observable en su comportamiento (se sienta, mira alrededor, habla) (Bennett & Hacker, 2022, pp. 269-270). Tampoco en la transitiva: que alguien sea consciente de la presencia de otra persona se manifiesta en su conducta (una sonrisa de reconocimiento, una reacción, un comentario) (Bennett & Hacker, 2022, p. 269).

Más aún, la afirmación "sé que estoy consciente" no es el reporte de una observación interna privilegiada. No hay criterios de evidencia para tal afirmación; uno no puede equivocarse al hacerla. Es más bien una señal análoga a decir "hola" o preguntar la hora, que muestra que uno ha recobrado la conciencia (Bennett & Hacker, 2022, p. 277). El error lockeano es transformar lo que es una condición para la experiencia (estar consciente) en un *objeto* de una experiencia interna especial (la reflexión). Esta transformación genera la ilusión de un reino privado, inescrutable para los demás, y por tanto, misterioso.

5.3. Crítica contemporánea de los conceptos de conciencia fenoménica y autoconciencia de orden superior

La crítica más incisiva de Bennett y Hacker se dirige a desmantelar los dos conceptos centrales que la taxonomía contemporánea (sección 4.2) identifica como herederos directos del giro lockeano: la *conciencia fenoménica* (con su núcleo en los *qualia* y la "brecha explicativa") y la autoconciencia teorizada según *modelos de orden superior*. Su análisis muestra que estos conceptos no son descubrimientos filosóficos, sino artefactos conceptuales generados por los presupuestos erróneos del modelo psicológico-reflexivo instaurado por Locke.

a) Crítica a la "conciencia fenoménica" y el problema de los qualia como ilusión lockeana

La noción contemporánea de "conciencia fenoménica" (P-conciencia), definida por su "carácter cualitativo" subjetivo (*what-it-is-like*), es para Bennett y Hacker el resultado de una inflación y confusión conceptual con raíces en el "sentido interno" lockeano.

1. **El origen lockeano del *quale*:** Como se señaló en 4.2, Locke, al hacer de la reflexión una fuente de "ideas simples" análoga a la sensación, genera la "idea del pensamiento mismo" (Balibar, 2013, p. 112). Este movimiento es fundacional: transforma la *conciencia de* algo (transitiva) en la *percepción interna de* una cualidad de la propia vivencia. Bennett y Hacker argumentan que este es el error germinal. La "idea del pensamiento" no es un *quale* privado, sino una confusión gramatical. Cuando Locke describe la conciencia como aprehensión de "lo que pasa", sienta las bases para la pregunta moderna "¿qué se siente ser?" (Nagel). Pero esta pregunta, según Bennett y Hacker, está mal formulada. "Preguntar '¿qué se siente ver rojo?' solo es inteligible si se pregunta por la actitud de la persona hacia su experiencia (si le resulta agradable, llamativa, etc.), no por una cualidad intrínseca e inefable de la experiencia misma" (Bennett & Hacker, 2022, p. 295). Para la mayoría de las experiencias perceptuales neutrales (ver un poste, oír un ruido de fondo) *no hay un quale distintivo*. La búsqueda de uno es un sinsentido inducido por el modelo que exige que toda "idea" (lockeana) o "estado consciente" (contemporáneo) tenga un contenido cualitativo.
2. **Problematización de la "inefabilidad" y la "brecha explicativa":** La supuesta inefabilidad de los qualia —su resistencia a la descripción pública y su papel en la "brecha explicativa"— es, para Bennett y Hacker, un síntoma de un error más profundo sobre el lenguaje y la mente (Bennett & Hacker,

> 2022, pp. 302-307). El modelo lockeano, al internalizar la cualidad en la "idea" privada, hace parecer que conceptos como "rojo" se definen por referencia a una muestra mental a la que solo yo accedo. Esto, argumentan, es incoherente. Los conceptos perceptuales se aprenden y definen mediante muestras públicas, no por ostensión privada (Bennett & Hacker, 2022, p. 306). La dificultad para "explicar la rojez del rojo" (Crick) o para hacer "inteligible el vínculo" entre cerebro y experiencia (Van Gulick) nacería de un pseudoproblema. No hay una "rojez fenoménica" *además de* la propiedad pública de los objetos rojos que personas con visión normal pueden discriminar. La "brecha explicativa" sería, por tanto, una *brecha categorial ilusoria*, creada por una mala caracterización de lo que hay que explicar. Lo que requiere explicación científica no es cómo el cerebro "genera qualia", sino cómo sustenta las capacidades de discriminación sensorial, atención y conocimiento que constituyen la percepción.

La consecuencia de esta posición sería la disolución del problema en cuestión. Al desmontar la noción de *qualia* como un artefacto del modelo de percepción interna, Bennett y Hacker desactivan el núcleo del llamado "problema duro" de la conciencia (Chalmers, 1996) o la supuesta brecha identificada en la “brecha explicativa” (Levine, 1983). Si no hay *qualia* inefables que surjan misteriosamente de la actividad neural, entonces la tarea de la neurociencia se redimensiona. Su labor es investigar los mecanismos neurales que subyacen a las capacidades cognitivas y perceptuales de los organismos (discernir colores, sentir dolor, atender a estímulos), no "correlacionar" o "producir" fantasmas fenoménicos (Bennett & Hacker, 2022, p. 307). Así, la "conciencia fenoménica" deja de ser un *explanandum* especial.

b) Crítica a los modelos de autoconciencia de orden superior como duplicación reflexiva lockeana

Los modelos de orden superior (HO) para la autoconciencia, en los que un estado mental se vuelve consciente al ser objeto de un segundo estado meta-mental, son la formalización contemporánea de la "arquitectura de duplicación lockeana" descrita en 4.2. Bennett y Hacker atacan esta arquitectura en su raíz.

1. **El duplicado interno:** Los modelos HO reproducen exactamente la estructura del "sentido interno" lockeano: para que un dolor sea consciente, debe ser "percibido" por un acto interno de reflexión o monitoreo. Bennett y Hacker argumentan que esto conduce a un regreso al infinito o a una circularidad absurda (Bennett & Hacker, 2022, pp. 277, 319). Más fundamentalmente, comete un error categorial: trata la conciencia transitiva (darse cuenta de algo) como si fuera una especie de percepción interna dirigida a un objeto mental. Pero "ser consciente de mi dolor" no es percibir un objeto interno llamado "dolor"; es, simplemente, tener un dolor que, en virtud de su intensidad o contexto, ocupa mi atención o mi pensamiento. No requiere un segundo estado mental. La autoconciencia, en su sentido ordinario, no es un mecanismo de escaneo interno, sino la capacidad de un sujeto de articular reflexivamente sus propios estados, una capacidad que se ejerce en el lenguaje y la acción, no en un teatro de la mente.
2. **Crítica a la "metacognición" y el "acceso privilegiado":** Las teorías HO suelen vincularse a una idea de "metacognición" o "monitoreo" que otorga un acceso privilegiado a los propios estados. Bennett y Hacker ya han desmontado el mito del acceso privilegiado (sección 5.3). Extienden esta crítica a los modelos HO: la relación que tengo con mi propio dolor no es epistémica (no *sé* que tengo dolor en base a evidencia interna), sino *gramatical.* Mi enunciado "me duele" es una expresión

> adverbial del dolor, no un informe basado en una observación (Bennett & Hacker, 2022, p. 277). Los modelos HO, al modelar la autoconciencia como una forma de conocimiento interno, perpetúan el error lockeano de convertir la condición de la experiencia (la subjetividad) en un objeto de conocimiento especial dentro de la mente.

La consecuencia de esta posición consiste en que la autoconciencia no es un mecanismo de orden superior. Para Bennett y Hacker, la autoconciencia no es un mecanismo psicológico adicional que se active sobre estados mentales preexistentes. Es, más bien, una capacidad de orden personal ligada al uso del lenguaje y a la vida en una comunidad social. Un ser autoconsciente no es aquel cuyo cerebro ejecuta operaciones de meta-representación, sino aquel que puede decir "yo", que puede dar razones de sus actos, que puede reflexionar sobre su pasado y planificar su futuro. Esta capacidad depende, ciertamente, de un cerebro en funcionamiento, pero no es *idéntica* a ningún proceso cerebral de "monitoreo". El modelo HO, al buscar internalizar esta capacidad en un mecanismo causal, cae nuevamente en la falacia mereológica y en la psicologización extrema de un concepto que es, en esencia, relacional y social.

La alternativa de Bennett y Hacker es regresar a una caracterización de los fenómenos mentales como atributos de la persona en su totalidad, situada en un mundo público y dotada de capacidades (perceptuales, cognitivas, lingüísticas) cuya base neural puede ser estudiada sin necesidad de postular reinos fenoménicos misteriosos o homúnculos metacognitivos. Justamente esta solución se conecta un uso del término "conciencia" soslayado por la filosofía contemporánea y que puede apreciarse, al menos, en el pensamiento de Descartes o Kant.

6. Modelo constitutivo o de propiedad

Las críticas de Bennett y Hacker al modelo psicológico-reflexivo de la conciencia (sección 5) no solo exponen los vicios conceptuales del

paradigma lockeano, sino que, de manera implícita, apuntan hacia una comprensión alternativa de la conciencia que escapa a dichos vicios. Esta comprensión alternativa puede identificarse en lo que denominamos *modelo constitutivo*, ejemplificado de manera paradigmática, aunque con diferencias sustanciales, en el uso que Descartes y Kant hacen de los términos *conscientia*, *Bewusstsein* y *Gewissen* (conciencia moral), respectivamente. Mientras que el modelo lockeano y sus herederos contemporáneos conciben la conciencia como un *campo de interioridad psicológica* susceptible de estudio naturalista, el modelo constitutivo la entiende como una *condición de posibilidad* o un *principio de imputación* de carácter fundamentalmente normativo. Las críticas de Bennett y Hacker —a la falacia mereológica, al mito de la privacidad, a la inflación conceptual y a los artefactos de la conciencia fenoménica y la autoconciencia reflexiva— son aplicables al primer modelo, pero no tocan de la misma manera al segundo, pues Descartes y Kant, cada uno a su modo, evitan precisamente las confusiones categoriales que aquel introduce.

La continuidad entre la sección 5 y la presente radica en que la crítica contemporánea de Bennett y Hacker, al desmantelar el modelo lockeano, redescubre por contraste los contornos de un modelo de conciencia que nunca fue psicológico en primer lugar. Lo que Bennett y Hacker defienden —la atribución de estados conscientes a la persona como un todo en un contexto público de conducta y lenguaje— comparte un aire de familia, pese a las distancias históricas y sistemáticas, con la concepción cartesiana de la *conscientia* como saber constitutivo del *cogito*, con la concepción kantiana del *Gewissen* como tribunal interno de la razón práctica así como también con un cierto sentido de *Selbstbewusstsein* en su filosofía teórica. En ambos casos, la "conciencia" no nombra un estado o proceso interno observable, sino una relación normativa que constituye al sujeto como tal. Explorar este modelo no solo proporciona una salida al callejón sin salida lockeano, sino que también muestra que las críticas de Bennett y Hacker, lejos de ser meramente

destructivas, ayudan a recuperar una comprensión de la conciencia más rica y filosóficamente sólida.

6.1. Descartes: *Conscientia* como certeza inmediata y saber constitutivo

En Descartes, el concepto de *conscientia* se erige como un pilar metafísico-epistemológico que garantiza la certeza inmediata y posee una normatividad constitutiva para lo mental. Contrariamente a la interpretación que lo considera el fundador de una "filosofía de la conciencia" subjetivista —una interpretación que el modelo lockeano posterior podría asumir—, un examen detenido de sus textos, guiado por la investigación modélica de Boris Hennig, revela que Descartes se atiene en gran medida al significado tradicional del término, integrándolo en un nuevo marco sistemático (Hennig, 2006, 2007; Placencia, 2017).

La primera evidencia de esta continuidad es la notable economía léxica con la que Descartes emplea el sustantivo *conscientia*. Como señala Placencia (2017), "el sustantivo *conscientia* no es empleado más de 20 veces" en toda la edición de Adam y Tannery, y está completamente ausente en el texto central de las *Meditaciones Metafísicas* (p. 283). Esta escasez sugiere que no se trata de un concepto técnico novedoso destinado a nombrar un nuevo dominio psicológico, sino de un término cuyo significado, heredado de la tradición estoica, escolástica, patrística y del derecho romano, se da por sentado (Hennig, 2007, p. 469).

La tesis fundamental cartesiana es la inseparabilidad entre pensamiento y conciencia. Descartes es categórico: "Empleo el término *pensamiento* para todo aquello que está en nosotros y de lo que somos conscientes de modo inmediato" (Descartes, AT VII, 160, 7-8). Y en otra definición crucial: "Entiendo bajo el nombre *pensamiento* todo aquello de lo que somos conscientes como ocurriendo en nosotros, en cuanto hay en nosotros consciencia de ello" (Descartes, AT VIII A, 7, 20-22). Esto establece que la *conscientia* no es un acompañamiento contingente, sino la condición misma que constituye un evento

mental *como* pensamiento. Esta concepción es inmunes a la crítica de la "falacia mereológica". Para Descartes, no es el cerebro ni una parte del hombre lo que es consciente, sino el *yo* pensante (*res cogitans*) en su totalidad indubitable. La *conscientia* es el modo de ser del *cogito*, no una propiedad de un órgano o parte de algo.

Frente a interpretaciones que ven en la *conscientia* un acto de segundo orden o reflexivo (un *actus reflexus*) [10] —precisamente el germen del modelo lockeano—, Hennig (2007) propone recuperar el sentido tradicional y etimológico del término. *Conscientia* deriva de *cum-scire* (saber-con), y en su origen designaba un "saber compartido", como el de un cómplice o testigo de un acto, particularmente en contextos forenses (pp. 469-470). Hennig traslada este modelo al ámbito cartesiano: la *conscientia* no es un pensamiento que se vuelve sobre otro (lo que generaría el regreso al infinito que Bennett y Hacker critican en los modelos de orden superior), sino el "saber-con" que constituye el ser cogitativo de un evento mental. Un proceso (incluso uno físico) no es un "pensamiento" hasta que es elevado a la esfera de lo evaluable según criterios de verdad, coherencia y significado por la *conscientia* (Hennig, 2007, pp. 479-480; Placencia, 2017, p. 286-287). La fórmula clave es: el pensador piensa algo *en cuanto (quatenus)* es consciente de él (Hennig, 2007, p. 467).

[10] La tesis fundamental cartesiana es que no puede haber pensamiento sin un conocimiento inmediato del mismo. Descartes es categórico: "no podemos tener ningún pensamiento del que no seamos conscientes en el mismo momento en que está en nosotros" (Descartes, 1641 AT VII, 246; CSM II, 171, como se cita en Lähteenmäki, 2008, p. 182). La *conscientia* en Descartes no debe entenderse primordialmente como un *actus reflexus* (un acto de segundo orden)
Como también destaca Placencia (2017, p. 284n), la interpretación de Cottingham (1978) es una "forma de interpretar el sentido de "conciencia" es, en rigor, un tipo específico de interpretación que puede subsumirse bajo una clase más amplia que corresponde a lo que podría Barth denominar "interpretaciones del metaacto actual". Se trata de aquellas concepciones según las cuales la conciencia consiste en la ejecución de un acto de segundo orden que está dirigido a uno de primer orden, del que deviene consciente el sujeto del acto de segundo orden a través, justamente, de ese acto (Barth 2011, p. 154)".

Esta perspectiva de evaluación, que es inherente a la *conscientia*, en última instancia remite a un observador ideal: Dios. Por lo tanto, el *cogito* no es un acto de pura interioridad autorreferencial y privada —el "mito de la privacidad" que Bennett y Hacker atacan—, sino un acto que se realiza bajo la mirada de un testigo omnisciente, internalizando así la estructura del "saber compartido" (Hennig, 2007, p. 484). La *conscientia* es, pues, un saber práctico y activo –un heredero directo de su sentido clásico– que, en el marco cartesiano, se internaliza sin perder su carácter normativo y constitutivo. No es que yo me observe a mí mismo pensando, sino que mi pensar es, desde su origen, un actuar sometido a una norma de claridad y distinción cuya garantía última es divina. En este sentido, la *conscientia* en Descartes es fundamentalmente un *saber constitutivo* que, lejos de encerrar al sujeto en sí mismo, lo abre a una dimensión de validez trascendente. Esta estructura normativa la distancia radicalmente del "sentido interno" lockeano, que es precisamente lo que Bennett y Hacker diagnostican como la fuente de la psicologización y la posterior naturalización errónea.

6.2. Kant: La conciencia moral (*Gewissen*) como tribunal interno

Como demuestra Placencia (2017), el punto de contacto sistemático más profundo entre Kant y la tradición de la *conscientia* no se encuentra en su filosofía teórica –donde busca distanciarse de la sustancialización cartesiana del *cogito*–, sino en el ámbito de la filosofía práctica, a través de su concepto de conciencia moral (*Gewissen*). Aquí, Kant retoma y transforma radicalmente la herencia normativa de la *conscientia* dentro de su marco crítico, desplazando el foco de la certeza teórica a la imputación moral (Placencia, 2017).

Para Kant, el *Gewissen* no es una facultad cognitiva más, ni un sentimiento empírico, sino un "hecho de la razón" práctico que actúa como instancia última de juicio. Es descrito de manera constante como un "tribunal interno" (*innerer Gerichtshof*), una instancia judicial que

cada ser racional lleva consigo mismo (Kant, *MS*, 06:438; Placencia, 2017, p. 290). Esta metáfora forense, que Kant emplea desde sus primeras lecciones de ética, subraya su función de juzgar, sancionar o exculpar al agente respecto de sus propias acciones y máximas. Esta concepción evita por completo la "inflación conceptual" y la creación de "qualia". El *Gewissen* no es una cualidad fenoménica de la experiencia, sino un principio de imputación que opera en el foro interno de la razón.

Placencia (2017) destaca tres características centrales del *Gewissen* kantiano que revelan profundas analogías estructurales con la *conscientia* cartesiana interpretada por Hennig (2007):

1. *Es un instinto natural:* En sus lecciones, Kant caracteriza al *Gewissen* como un "instinto" para destacar que su operación no es voluntaria (*nach Belieben*), sino que es una disposición constitutiva del agente moral. Sin su posesión, el ser humano sería "moralmente ciego" y no podría ser sujeto de imputación. Esta "naturalización" señala su carácter de *condición de posibilidad* de la moralidad, no su origen empírico.
2. *Es infalible en su dominio propio:* Kant sostiene una tesis fuerte sobre la infalibilidad del *Gewissen*. Esto no significa que no podamos equivocarnos en el juicio objetivo sobre si una acción es un deber, sino que no podemos errar en el juicio subjetivo sobre si hemos sometido escrupulosamente nuestra acción a ese juicio objetivo. "Una conciencia moral que yerra es un absurdo", afirma Kant (1797, *MS*, 06:401), porque el *Gewissen* es precisamente la conciencia de este acto de examinación moral.
3. *Es constitutivo de la moralidad de la acción:* Al igual que la *conscientia* en Descartes constituye un evento como pensamiento, el *Gewissen* en Kant constituye la acción como moral, no en cuanto a su contenido material, sino en cuanto a la actitud formal con la que fue emprendida. El *Gewissen* "causa la acción respectiva en cuanto ella es sujeto a evaluación y en cuanto ella depende de la deliberación" (Hennig, 2007, p. 479). Su función es garantizar la *Gewissenhaftigkeit* (escrupulosidad), es

decir, que el agente haya procedido con la meticulosidad debida en la aplicación de la ley moral a su caso particular.

6.3. Autoconciencia (*Selbstbewusstsein*) bajo el modelo de propiedad

Para comprender plenamente la estructura no reflexiva de esta conciencia y su conexión con la autoconciencia teórica, es crucial incorporar la propuesta de Castillo y Fuentes (2024) sobre el "modelo de propiedad". Según estos autores, las nociones kantianas con el prefijo 'auto-' (como *Selbstbewusstsein* y *Autonomie*) han sido tradicionalmente interpretadas mediante un modelo reflexivo, que entiende el prefijo como una relación de algo consigo mismo (p. ej., el sujeto se convierte en objeto de sí mismo). Esta interpretación genera problemas insuperables, como regresiones al infinito y una hipostatización del yo (Castillo Vallez & Fuentes González, 2024, pp. 569-571), problemas idénticos a los que Bennett y Hacker señalan en los modelos de orden superior.

Frente a esto, proponen un "modelo de propiedad" donde el prefijo 'auto-' no indica una relación reflexiva, sino que el acto en cuestión es *propio del sujeto*, es decir, es realizado por el sujeto mismo. "Con la adición del prefijo 'auto-' no se está refiriendo a algo distinto de lo que designa la expresión sin el prefijo, sino al modo o al cómo del acto designado por esta última" (Castillo Vallez & Fuentes González, 2024, p. 569). Este modelo evita la hipostatización y se ajusta mejor al carácter formal y condicional de la filosofía trascendental kantiana.

La aplicación de este modelo ilumina la verdadera naturaleza del *Gewissen* y revela una profunda conexión estructural con la autoconciencia teórica (*Selbstbewusstsein*). En el ámbito teórico, la conciencia (*Bewusstsein*) para Kant no es un acto aislado, sino la condición formal que posibilita la experiencia[11]. Según el modelo de

[11] Si bien la propuesta de Castillo y Fuentes (2024) es fundamental para comprender el sentido primordial y no hipostasiado del prefijo 'auto-' en términos

propiedad, la conciencia no debe entenderse como una relación intencional con un objeto externo, sino como el modo en que el sujeto constituye sus representaciones como *propias*. Esto se vincula directamente con la unidad sintética de la apercepción, que Kant describe como el "supremo principio del entendimiento" (*KrV*, B 136). Así, la conciencia no es un acto que se añade a la representación, sino la forma en que toda representación es reconocida como propia por el sujeto. Este carácter propio no surge de una reflexión explícita, sino que es una condición trascendental de la experiencia (Castillo Vallez & Fuentes González, 2024, p. 573).

La autoconciencia (*Selbstbewusstsein*) ha sido tradicionalmente interpretada como una forma de autoconocimiento reflexivo, donde el

---

de propiedad o modo de constitución del sujeto, es crucial reconocer que esto no agota la semántica de "conciencia" en Kant. Junto a este sentido estructural y trascendental, convive de manera legítima un sentido reflexivo o cuasi-lockeano de la conciencia.

Kant distingue explícitamente entre una conciencia inmediata en la aprehensión y una conciencia mediata en la reflexión, donde esta última es definida precisamente como "la representación de otra representación que tenemos" (AA 09, 33). Este es un acto de segundo orden, de carácter psicológico-empírico, mediante el cual el sujeto tematiza sus propios estados internos como objetos de atención. Este "yo en cuanto objeto" del sentido interno (AA 07, 134 nota) corresponde al sentido introspectivo y reflexivo que Locke defendería y que la tradición asoció con la *conscientia*.

Por lo tanto, la gran contribución del modelo de propiedad no es la negación de este nivel reflexivo, sino la demostración de su dependencia y fundamentación en una estructura más originaria. La reflexión (el modelo reflexivo) opera sobre un contenido que ya ha sido previamente constituido como unitario y propio gracias a la apercepción trascendental (el modelo de propiedad). Como señala Kant, la "unidad analítica de la apercepción" (propia de la reflexión y la autoadscripción) solo es posible bajo la presuposición de una "unidad sintética" previa (B 133). En otras palabras, puedo tomarme *reflexivamente* a mí mismo como objeto de introspección solo porque mis representaciones están ya unificadas de manera *propia* y no reflexiva en la conciencia trascendental.

En conclusión, lejos de ser excluyentes, el modelo de propiedad (trascendental y constitutivo) y el modelo reflexivo (empírico e introspectivo) son niveles complementarios en la arquitectura kantiana. El primero explica la posibilidad misma de una experiencia unificada y de la autoconciencia; el segundo describe un ejercicio concreto y derivado que dicha experiencia posibilita. La genialidad de Kant reside en haber subordinado el segundo al primero, evitando así los problemas de la regresión al infinito y la hipostatización, sin por ello negar la fenomenología de la introspección.

sujeto se toma a sí mismo como objeto. Sin embargo, esta lectura choca con la advertencia kantiana de que el "yo pienso" no representa un objeto, sino una condición formal de la experiencia (*KrV*, A 346/B 404). El modelo reflexivo falla porque "busca legitimar el fenómeno de la autoconciencia mediante una relación de identidad, cuya única manera de volverse inteligible depende de que dicha relación refiera a un objeto" (Castillo Vallez & Fuentes González, 2024, p. 570).

Frente a esto, el modelo de propiedad propone que la autoconciencia no es una relación consigo mismo, sino la forma en que los actos cognitivos son realizados por el sujeto. No se trata de que el yo se conozca a sí mismo, sino de que el yo se constituye como tal en el ejercicio mismo de sus facultades. Como señalan los autores: "La autoconciencia sólo sería la contracara de la conciencia de objetos y no una particularidad única o esencial del sujeto que conoce" (Castillo Vallez & Fuentes González, 2024, p. 574). Esto se refleja en la unidad trascendental de la apercepción, que no es un objeto sustancial, sino la condición formal que unifica las representaciones bajo el "yo pienso". Este "yo" no es un *algo*, sino un *cómo*: el modo en que las representaciones son sintetizadas y reconocidas como propias.

7. Conclusiones en torno al modelo constitutivo o de propiedad

El análisis de esta sección 6 permite consolidar un resultado fundamental: el *modelo constitutivo o de propiedad* emerge no solo como una alternativa histórica al paradigma lockeano, sino como la única vía conceptual coherente para superar las aporías generadas por la psicologización de la conciencia. Este modelo, cuyos contornos se delinean en Descartes y alcanzan su formulación más sistemática en Kant, redefine la conciencia como una relación de apropiación que constituye al sujeto, en lugar de como un espacio interno de observación. Los resultados obtenidos se articulan en tres niveles interconectados: 1) la resolución de las críticas específicas al modelo lockeano, 2) la explicitación de la estructura no

reflexiva de la conciencia, y 3) la reconciliación genealógica de la polisemia del término.

7.1. El modelo constitutivo como antídoto a las críticas de Bennett y Hacker

La contrastación con las críticas desarrolladas en las secciones 3, 4 y 5 demuestra que el modelo constitutivo es inmune a los principales vicios diagnosticados en el paradigma lockeano-reflexivo.

- *Superación de la falacia mereológica*: Mientras el modelo lockeano, al concebir la conciencia como un "escenario" interno (Balibar, 2013, p. 116), allanó el camino para su errónea localización cerebral (Bennett & Hacker, 2022, p. 266), el modelo constitutivo evita por principio esta confusión categorial. Para Descartes, la *conscientia* es el modo de ser de la *res cogitans* en su indivisibilidad; no es una propiedad de una parte, sino la esencia misma del sujeto pensante como totalidad (Descartes, AT VII, 160). En Kant, el *Gewissen* es un "tribunal interno" (Kant, MS, 06:438; citado en Placencia, 2017, p. 290) que pertenece a la persona como agente racional, no a un órgano. Así, este modelo coincide con la insistencia de Bennett y Hacker en que la conciencia se predica de la persona íntegra en un contexto de conducta y lenguaje, no de sus partes.
- *Disolución del mito de la privacidad y el acceso privilegiado:* El modelo lockeano del "sentido interno" creó la ilusión de un reino privado e inefable. Bennett y Hacker desmontan este dogma, señalando que la conciencia se manifiesta en conductas públicas (2022, pp. 269-270). El modelo constitutivo anticipa esta crítica al rechazar la idea de un acceso introspectivo privilegiado a un objeto mental. La *conscientia* cartesiana no es una observación privada, sino un "saber-con" (*cum-scire*) que, en su estructura misma, remite a un testigo ideal (Dios), internalizando así una dimensión de publicidad y validez (Hennig, 2007, p. 484).

El *Gewissen* kantiano no juzga sobre cualidades fenoménicas privadas, sino sobre la conformidad de la máxima con la ley moral, un principio universalizable. La normatividad, aquí, reemplaza a la privacidad.

- *Desactivación de los sentidos reflexivos de la conciencia fenoménica y la autoconciencia:* Las nociones de *qualia* y la arquitectura de duplicación de los modelos de orden superior (HO) son, como muestra la sección 4, herederas directas del giro lockeano. Bennett y Hacker las identifican como artefactos conceptuales (2022, pp. 295, 319). El modelo constitutivo disuelve estos problemas en su raíz. Al entender la conciencia como propiedad constitutiva (el *cómo* del acto mental o moral) y no como cualidad observada o estado meta-cognitivo, elimina la necesidad de postular cualidades fenoménicas inefables o un homúnculo reflexivo. La "autoconciencia" (*Selbstbewusstsein*), bajo el modelo de propiedad, no es un acto reflexivo, sino "la contracara de la conciencia de objetos" (Castillo Vallez & Fuentes González, 2024, p. 574), es decir, el modo en que el sujeto se apropia de la síntesis de sus representaciones. Del mismo modo, la autonomía kantiana "no es cuestión de 'reflexión' sino de 'propiedad'" (Castillo Vallez & Fuentes González, 2024, p. 577).

7.2. La estructura de la conciencia como propiedad no-reflexiva

Un hallazgo de este artículo es la explicitación parcial de la estructura interna del modelo constitutivo, gracias a la aplicación del "modelo de propiedad" (Castillo & Fuentes, 2024). Este análisis revela que:

- La conciencia (*Bewusstsein*, *conscientia*) es el *modo formal* en que una representación o una máxima es reconocida como *propia* por el sujeto. No es un elemento añadido, sino la forma de la aprehensión o la deliberación.
- La autoconciencia (*Selbstbewusstsein*) y la conciencia moral (*Gewissen*) son, por tanto, *actualizaciones de esta relación de*

propiedad en los ámbitos teórico y práctico. No suponen que el yo se tome a sí mismo como objeto (modelo reflexivo), sino que el yo se constituye como unificador de lo múltiple (teoría) o como responsable ante la ley (práctica) en el ejercicio mismo de sus facultades.

- Esta estructura evita las *regresiones al infinito* y la *hipostatización del yo* que plagan al modelo lockeano y sus herederos HO, ya que no se requiere una serie de actos observantes, sino que se describe una condición transcendental o práctica de posibilidad.

7.3. La bifurcación como clave de la polisemia actual

El contraste realizado a lo largo de este trabajo confirma y profundiza la tesis de Balibar (2013) sobre la "invención" lockeana de la conciencia, revelando la existencia de una bifurcación en la modernidad:

- Línea constitutivo-de propiedad (Descartes-Kant): Preserva y transforma el legado de la *conscientia* como saber compartido y principio forense, definiendo la conciencia fundamentalmente como un principio de imputación y validez (lógica, moral, existencial).
- Línea psicológico-reflexiva (Locke y su legado): Inventa la conciencia como *campo de interioridad psicológica*, sentando las bases de la taxonomía contemporánea (sección 4) y constituyendo el blanco de las críticas de Bennett y Hacker (sección 5).

En consecuencia, la polisemia actual del término "conciencia" no es un accidente, sino el síntoma de esta bifurcación que podría rastrearse históricamente. El predominio del segundo sentido en la filosofía de la mente y la ciencia cognitiva corre el peligro de restringir tanto la formulación como la solución de los problemas que estas disciplinas se proponen.

La posición de Balibar (2013) es, por tanto, crucial para apreciar que el sentido de conciencia que domina la filosofía de la mente y las ciencias cognitivas tiene que ver principalmente con el nivel reflexivo y

psicológico atribuido a Locke. Así, mientras la ciencia cognitiva se ocupa de la "conciencia-accesibilidad" o la "conciencia-fenoménica" —herederas del problema lockeano de la interioridad y el "escenario" mental—, las éticas contemporáneas emplean de manera más abierta, a menudo implícitamente, con la herencia kantiana del *Gewissen* y su modelo de conciencia como tribunal normativo internalizado.

Comprender esta compleja genealogía permite desmontar la idea de una "filosofía de la conciencia" monolítica. La tensión moderna entre la conciencia como hecho psicológico (Locke) y la conciencia como instancia constitutiva (Descartes-Kant) no es una mera discrepancia histórica, sino una división conceptual que estructura los debates actuales. Reconocer este legado dual es el primer paso para abordar con mayor precisión la pregunta por lo que significa ser consciente, evitando reducir una dimensión a la otra y apreciando la riqueza de un problema que es, a la vez, metafísico, epistemológico, moral y psicológico.

En síntesis, los resultados de esta investigación demuestran que el modelo constitutivo o de propiedad no es una reliquia histórica, sino un marco conceptual que podría resolver los problemas generados por el paradigma psicológico. Al redefinir la conciencia como una relación de propiedad —donde el sujeto se constituye en el acto mismo de apropiarse de sus pensamientos y acciones bajo una ley de razón—, este modelo ofrece una salida a algunos callejones sin salida con los que se ha encontrado el ejercicio la naturalización reductiva, sin caer en el misterio dualista. Recuperar este sentido de la conciencia es esencial para una filosofía que aspire a comprenderlo a cabalidad.

Estos hallazgos tienen implicaciones para la investigación futura. Sugieren que la aparente obviedad de la conciencia como "objeto de investigación" natural es, en realidad, un artefacto histórico. Lo que se investiga bajo ese nombre en la ciencia cognitiva contemporánea no es un fenómeno suficientemente delimitado, sino un concepto históricamente constituido cuyos contornos responden a decisiones filosóficas específicas, predominantemente lockeanas. Por tanto, la estrategia de disolución (meta)filosófica que orienta este artículo implica

necesariamente realizar el análisis genealógico que permita problematizar esta naturalización acrítica.

Puestas las cosas de este modo, la vía hacia la disolución no pasa, entonces, por resolver el, por ejemplo, "problema difícil" de la conciencia en sus propios términos, sino en el ejercicio de reconfigurar el campo de indagación en cual se encuentra para luego reevaluar la condición de problema de dicho asunto. Al revelar la contingencia histórica de la concepción lockeana y contrastarla con otra tradición alternativa, se socava la pretensión de universalidad de los problemas que aquella genera. Si la conciencia psicológica es solo una de las posibles concepciones —y no la más fundamental—, entonces los desafíos que plantea (el "hard problem", los qualia, la privacidad) pierden su estatus de misterios metafísicos perennes y se revelan como dificultades generadas por un marco categorial particular y potencialmente deficiente.

En definitiva, el camino hacia una filosofía de la mente más lúcida y productiva requiere volver la mirada, (meta)filosóficamente, hacia los cimientos históricos y conceptuales de su vocabulario fundamental. Esta tarea de autoevaluación crítica no constituye un atributo secundario de la filosofía, sino su elemento distintivo fundamental, aquello que la diferencia de otras disciplinas (Castillo, 2021, p. 22).

REFERENCIA

Abalo, F. (2015). El problema de determinar la “sabiduría” en los primeros diálogos de Platón. *Revista de Filosofía, 71*, 9-21.

Anthony, M. V. (2002). Concepts of consciousness, kinds of consciousness, meanings of 'consciousness'. *Philosophical Studies, 109*(1), 1–16. https://doi.org/10.1023/A:1015711419870

Balibar, É. (2013). *Identity and difference: John Locke and the invention of consciousness*. Verso.

Barth, C. (2011). Bewusstsein bei Descartes. *Archiv für Geschichte der Philosophie, 93*(2), 162–194. https://doi.org/10.1515/agph.2011.007

Bennett, M. R., & Hacker, P. M. S. (2022). *Philosophical foundations of neuroscience* (2nd ed.). Wiley-Blackwell.

Berger, J. (2020). Kinds of consciousness. En J. Berger, *Philosophy of mind: The basics* (pp. 251-264). Routledge.

Castillo, J. (2021). La ciencia no piensa: Heidegger y la tarea de la filosofía como autoconsciencia. *Aporía. Revista Internacional de Investigaciones Filosóficas, 21*, 21-34.

Castillo Vallez, J., & Fuentes González, O. J. (2024). El modelo reflexivo y el modelo de propiedad: La interpretación del prefijo "auto-" en las nociones kantianas de autoconciencia y autonomía. *Revista de Filosofía, 79*(1), 567–582. https://doi.org/10.4067/S0718-43602024000100567

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford University Press.

Cottingham, J. (1978). 'Descartes on thought'. *The Philosophical Quarterly, 28*(112), 208–214. https://doi.org/10.2307/2219082

Crick, F. (1995). *The astonishing hypothesis: The scientific search for the soul*. Touchstone.

Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex, 8*(2), 97–107. https://doi.org/10.1093/cercor/8.2.97

Davies, C. G. (1990). *'Conscience' as consciousness: The idea of self-awareness in French philosophical writing from Descartes to Diderot*. The Voltaire Foundation.

Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness, and could machines have it? *Science, 358*(6362), 486–492. https://doi.org/10.1126/science.aan8874

Descartes, R. (1996). *Œuvres de Descartes* (C. Adam & P. Tannery, Eds.). J. Vrin. (Obra original publicada entre 1897-1913).

Edelman, G. M., & Tononi, G. (2000). *Consciousness: How matter becomes imagination*. Allen Lane, The Penguin Press.

Esposito, R., & Nancy, J.-L. (2006). Diálogo sobre la filosofía por venir. En S. Cabanchik (Ed.), *El abandono del mundo* (pp. 161-181). Grama.

Gazzaniga, M. S. (1997). Consciousness and the cerebral hemispheres. En M. S. Gazzaniga (Ed.), *The new cognitive neurosences* (pp. 1391–1400). MIT Press.

Hennig, B. (2007). Cartesian conscientia. *British Journal for the History of Philosophy, 15*(3), 455–484. https://doi.org/10.1080/09608780701444915

Kant, I. (1902-). *Kant's gesammelte Schriften*. Herausgegeben von der Preußischen (später Deutschen) Akademie der Wissenschaften.

Kant, I. (2010). *Crítica de la razón pura* (M. Caimi, Trad.). Fondo de Cultura Económica. (Obra original publicada en 1781/1787).

Lähteenmäki, V. (2008). Orders of consciousness and forms of reflexivity in Descartes. En S. Heinämaa, V. Lähteenmäki & P. Remes (Eds.), *Consciousness: From perception to reflection in the history of philosophy* (pp. 177-201). Springer.

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly, 64*(4), 354–361. https://doi.org/10.1111/j.1468-0114.1983.tb00207.x

Locke, J. (1975). *An essay concerning human understanding* (P. H. Nidditch, Ed.). Oxford University Press. (Obra original publicada en 1689). https://doi.org/10.1093/actrade/9780198243861.book.1

McGinn, C. (1987). Could a machine be conscious? En C. Blakemore & S. Greenfield (Eds.), *Mindwaves: Thoughts on intelligence, identity, and consciousness* (pp. 279–286). Basil Blackwell.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review, 83*(4), 435–450. https://doi.org/10.2307/2183914

Nudler, Ó. (2012). Los problemas de la filosofía de la filosofía. En Ó. Nudler (Ed.), *Desafíos filosóficos de la actualidad* (pp. 19-47). Editorial Trotta.

Placencia, L. (2015). Autoconocimiento, acción y autointerpretación: Opacidad y transparencia del saber de sí mismo en la filosofía práctica de Kant. *Anuario Filosófico, 48*(3), 543–563.

Placencia, L. (2017). Descartes y Kant sobre la conciencia: Observaciones sobre un olvidado punto de contacto sistemático entre Kant y Descartes. *CON-TEXTOS KANTIANOS. International Journal of Philosophy, 5*, 278–295. https://doi.org/10.5281/zenodo.805946

Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies, 49*(3), 329–359. https://doi.org/10.1007/BF00355521

Rosenthal, D. M. (2005). *Consciousness and mind.* Oxford University Press. https://doi.org/10.1093/oso/9780198236979.001.0001

Searle, J. R. (1984). *Minds, brains and science*. Harvard University Press.

Thiel, U. (1991). Cudworth and seventeenth-century theories of consciousness. En S. Gaukroger (Ed.), *The uses of antiquity* (pp. 79-99). Kluwer Academic Publishers.

Thiel, U. (1994). Hume's notions of consciousness and reflection in context. *British Journal for the History of Philosophy, 2*(2), 75–115. https://doi.org/10.1080/09608789408570905

Tononi, G. (2005). Consciousness, information integration, and the brain. En S. Laureys (Ed.), *Progress in brain research: Vol. 150. The boundaries of consciousness: Neurobiology and neuropathology* (pp. 109–126). Elsevier. https://doi.org/10.1016/S0079-6123(05)50009-8

Van Gulick, R. (2014). Consciousness. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2014 ed.). Recuperado de https://plato.stanford.edu/archives/spr2014/entries/consciousness/

Weinberg, S. (2016). *Consciousness in Locke.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198749011.001.0001

Weinberg, S. (2021). Locke on consciousness and reflection. En J. Gordon-Roth & S. Weinberg (Eds.), *The Lockean mind* (pp. 127–137). Routledge.

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 48-78
ISSN 3087-260X
Santiago de Chile

LA VERDADERA RED SOCIAL: LA AMISTAD[1]

Lucas Bastián Peñaylillo Morales[2]
Pontificia Universidad Católica de Chile

Resumen

Desde la distinción tomista de los cuatro tipos de órdenes con los que la razón se relaciona y una metodología filosófica propia de la *Ética Nicomáquea*, pretendo reflexionar, en filosofía de la tecnología, acerca del rol de las redes sociales contemporáneas en la felicidad de la persona humana. Defiendo que estas tecnologías contribuyen, en muchos casos, al impedimento de la verdadera amistad, fundamento de la felicidad que se completa en la contemplación del amigo como ser personal. Para ello, (1) antes del desarrollo argumental, expongo una reflexión metodológica acerca de cómo abordar un problema en

[1] Versión corregida de un ensayo más breve enviado a la Editorial Tanto Monta para su concurso Persona y Sociedad, el cual fue inicialmente el trabajo final para un seminario sobre filosofía de la tecnología con el profesor Federico Ponzoni (Facultad de Filosofía, Pontificia Universidad Católica de Chile, 2025-1). Agradezco a los revisores de la Revista Homónima por sus valiosos comentarios.
Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Peñaylillo, L. (2026). La verdadera red social: la amistad. *Revista de Filosofía Homónima, 1*(1), pp. 48-78. https://doi.org/xxxxxxx[Por designar]

[2] Lucas Peñaylillo es estudiante de Filosofía de la Pontificia Universidad Católica de Chile. Email: lpeaylillo@uc.cl. ORCID: https://orcid.org/0009-0001-4543-2232.

filosofía de la tecnología, fundada en Aristóteles, santo Tomás de Aquino, Federico Ponzoni y Antonio Millán-Puelles; (2) ya en el desarrollo, describo brevemente la tecnología de las redes sociales y algunos de sus efectos psicológicos y sociales; (3) explico la amistad y su rol en la realización del ser personal, desde una filosofía tomista; (4) analizo cómo la felicidad en la amistad se ve impedida, en alguna medida, por las redes sociales, a través de la perspectiva platónica y mcluhaniana de la tecnología. Finalmente, (5) concluyo que, en realidad, la bondad de esta tecnología queda supeditada al valor de la persona del amigo, desde lo cual se abren líneas de investigación, y reivindico el valor de la verdadera red social: la amistad.

Palabras clave: *metodología filosófica, tecnología, redes sociales, amistad, persona, tomismo, felicidad.*

The true social network: friendship

Abstract

Based on Thomas Aquinas' distinction between the four types of orders with which reason relates and a philosophical methodology characteristic of *Nicomachean Ethics*, I intend to reflect, in the philosophy of technology, on the role of contemporary social networks in human happiness. I argue that these technologies often contribute to the impediment of true friendship, the foundation of happiness that is completed in the contemplation of the friend as a personal being. To this end, (1) before developing my argument, I present a methodological reflection on how to approach a problem in philosophy of technology, based on Aristotle, St. Thomas Aquinas, Federico Ponzoni, and Antonio Millán-Puelles; (2) in the development, I briefly describe social media technology and some of its psychological and social effects; (3) I explain friendship and its role in the realization of

the personal being, from a Thomist philosophy; (4) I analyze how happiness in friendship is impeded, to some extent, by social media, through the Platonic and McLuhanian perspective of technology. Finally, (5) I conclude that, in fact, the goodness of this technology is contingent upon the value of the friend as a person, which opens up lines of research, and I vindicate the value of the true social network: friendship.

Keywords: *philosophical methodology, technology, social networks, friendship, person, Thomism, human flourishing.*

… hay que recordar dos verdades fundamentales. La primera es que el hombre está llamado a vivir en la verdad y en el amor. La segunda es que cada hombre se realiza mediante la entrega sincera de sí mismo.

—Juan Pablo II, Gratissimam Sane, 1994

Introducción
Prólogo metodológico

Antes de comenzar con el ensayo, conviene realizar un excurso metodológico. Este trabajo se enmarca, en alguna medida, en la filosofía de la tecnología. En esta parte de la filosofía, dada la variedad de posturas posibles y la novedad de la disciplina[3], no hay consenso en su objeto y método propio. En este trabajo utilizaré una metodología propia del

[3] Una novedad sistemática, pues ya en Platón existía una reflexión sobre la técnica en cuanto tal (*Fedro*, 274c-275d), y más tarde en Aristóteles con su *Ética Nicomáquea* y su *Retórica*.

aristotelismo, desarrollada más tarde por santo Tomás de Aquino y sus seguidores, especialmente la hoy llamada Nueva Teoría de la Ley Natural —NNLT—, no sin advertir que el modo de presentar este método queda sujeto a revisión para un mayor desarrollo.

Como dice Aristóteles en la *Ética Nicomáquea*, "[l]a exposición será satisfactoria si hacemos las aclaraciones que la materia admite, pues no hay que buscar del mismo modo la exactitud en todas las argumentaciones…" (I 3, 1094b10-15). En efecto, "… es propio del hombre instruido buscar en cada género tanta exactitud cuanta la naturaleza de la cosa consiente" (*E. N.*, I 3, 1094b20-25), entonces, si el objeto que analizamos está sujeto a una variabilidad tal que las verdades sobre él solo valen para la mayoría de los casos, a partir de ellas solo podremos concluir verdades así variables y, en definitiva, contingentes, pero generales (*E. N.*, I 3, 1094b15-25). Toda ciencia queda así determinada por su objeto, pues ante todo estudiamos la realidad y es ella la que nos impone el modo de conocerla (Giussani, 2008, pp. 19–20). Es que el estudio con el que nos encontramos versa sobre los objetos de la parte 'calculadora' de la razón, es decir, objetos 'que pueden ser de otra manera' (*E. N.*, VI 1, 1139a5-15), ya que los artefactos tienen su principio en el que los produce y no en sí mismos (*E. N.*, VI 4, 1140a1-16)[4]. Por tanto, la filosofía de los artefactos —arte y tecnología—, y en concreto de la tecnología, parece ser una ciencia práctica (*Met.*, VI 1, 1025b20-25), que estudia —hasta donde su objeto lo permite— las causas y lo universal y necesario en la tecnología (*Met.*, I 1, 981a1-b5; *An. Post.*, I 33, 88b30-89a10), aplicando los principios universales y simples de la ciencia a lo particular y compuesto (I *In Ethic*, lect. 3).

[4] Un ejemplo de la variabilidad de la tecnología, incluso entre culturas y momentos del tiempo, es dado por la obra de Yuk Hui: *La pregunta por la técnica en China*. Más allá de su propia filosofía de la tecnología, la reflexión de Hui (2024) parte de la base de esta variabilidad, desde los orígenes de la técnica y su relación con la humanidad y el universo, de modo que es comprendida distinta en, por ejemplo, la cultura asiática (p. 31).

Esta reflexión es ampliada más adelante por el esquema tomista. En su *Comentario a la Ética a Nicómaco*, S. Tomás indica —siguiendo la *Metafísica* de Aristóteles— que lo propio de la sabiduría, perfección de la razón, es conocer el orden de la realidad completa (I *In Ethic.*, lect. 1). Sin embargo, la razón se relaciona con el orden de cuatro modos: pasivamente, con un orden que ella no produce —el de la naturaleza—; con uno que sí produce de su misma operación —el de la reflexión racional—; uno que produce en la operación de la voluntad —el de la acción humana—; y uno que produce en los entes causados por ella —el de los artefactos— (I *In Ethic.*, lect. 1)[5]. Todos estos órdenes avanzan gradualmente, desde lo más íntimo del sujeto cognoscente y su pasividad, hasta la relación activa de la razón con lo extrínseco a ella[6]; desde el conocimiento que recibe de lo externo, pasando por la actividad propia de la razón y la voluntad que impera actos exteriores, hasta la producción racional de algo externo al sujeto.

De cada uno de estos órdenes reales, cognoscibles por la razón, surgen ciencias —filosóficas o particulares—, todas las cuales, como bien muestra Ponzoni (en prensa), interesan al análisis filosófico de la tecnología (pp. 5–9): las ciencias naturales —incluidas, con S. Tomás, la metafísica y la filosofía de la naturaleza—, la lógica —junto con la filosofía

[5] La teoría de los cuatro órdenes, propiamente tomista, es presentada contemporáneamente por los teóricos de la NNLT, especialmente Finnis (2011a, pp. 135–139, 2011b, pp. 36–38, 2019, pp. 25–28) y, más recientemente, por Tollefsen (2018) y Ponzoni (en prensa). Como indica Tollefsen (2018), la NNLT comprende los cuatro órdenes no solo como distintos entre sí, sino también como irreducibles en un sentido fuerte, es decir, que ninguno puede estar fundado en otro (pp. 249–250). Finnis (2011b, 2019), en cambio, sostiene la irreductibilidad, pero no esta no-fundación, pues, por ejemplo, el tercer orden se funda ontológicamente en el primero, aunque en el ámbito explicativo cada orden sea autosuficiente (pp. 36–38, 115–116). Con la no-fundación, así como la autosuficiencia explicativa, sin embargo, no concuerdo, dada la dependencia epistemológica de la filosofía práctica a la especulativa, en directa negación de la aplicación de la *ley de Hume* y la *falacia naturalista* a la filosofía tomista. Aunque en este trabajo no puedo sustentar mi desacuerdo —véase, por ejemplo, Widow (2020)—, considero que esto no desvía en lo esencial la teoría de los cuatro órdenes en su aplicación a la filosofía de la tecnología.

[6] Agradezco al profesor Ponzoni por esta aclaración.

de la ciencia— y la ética —incluyendo política y economía— (Finnis, 2011a, pp. 135–139; I *In Ethic.*, lect. 1). Aunque un análisis del cuarto orden, estudiado por las 'artes mecánicas' (I *In Ethic.*, lect. 1), se hace esencial para la filosofía de los artefactos, pero en aquello que interesa a la reflexión filosófica, en este caso, sobre la técnica: *el sentido último* del orden de la tecnología (Millán-Puelles, 2013, p. 52)[7].

Tenemos, por tanto, el objeto material y formal de la filosofía de la tecnología: los artefactos (tecnológicos) —objeto material— en cuanto a su sentido último —objeto formal—. Este es el que le da el carácter propiamente filosófico a esta disciplina, de lo contrario estaríamos ante una mera τέχνη, que conozca y produzca los artefactos particulares y no estudie sus causas últimas. El *sentido último* se dividirá precisamente en los cuatro ordenes, con lo cual se podrá estudiar tanto las causas (materiales, formales y eficientes) de los artefactos y los atributos que de suyo le pertenecen —primer, segundo y cuarto orden—[8], como la finalidad[9] que tienen, para la vida propiamente humana y en sí mismos —tercer y cuarto orden—[10]. Sin embargo, dada la particularidad propia de la tecnología, su sentido solo podrá ser estudiado, como en la ética, una vez repasadas sus concreciones[11].[12]

[7] Otro modo de explicar los cuatro ordenes está en Millán-Puelles (2013, pp. 51–53), al presentar la división de las disciplinas filosóficas: metafísica, filosofía de la naturaleza, lógica, ética y filosofía de la cultura (dividida en arte y técnica).

[8] Sirviéndonos, de modo general, de la metafísica y la filosofía de la naturaleza, en la filosofía; y de la física, química, ciencias computacionales, psicología, entre otras, en las ciencias particulares.

[9] Con ello tenemos las cuatro causas aristotélicas, esenciales para la ciencia (*An. Post.*, II 11, 94a20-25).

[10] Sirviéndonos de la ética —con el fin último del hombre— y de un análisis de la finalidad intrínseca del artefacto y sus efectos, como aquello propio de la filosofía de la tecnología, que puede servirse de la sociología y la psicología para un estudio cultural de estos.

[11] Las singularidades de la tecnología sirven de insumo inicial para el estudio de sus causas y fines, pero en sí mismas su estudio corresponde, no a la filosofía —que estudia lo universal—, sino que a las diversas técnicas que las produzcan.

[12] El presente esquema es propio, sujeto a desarrollo, y de ningún modo adscribible a las posturas de Finnis, Tollefsen o Ponzoni.

Lo que haré en este artículo, dada esta metodología, es ampliar la discusión en filosofía de la tecnología continuando la propuesta de Ponzoni (en prensa, p. 5) desde el análisis, como indica el nombre de la disciplina, de la tecnología —no del arte—: una entidad causada exteriormente por la razón que "… pretende alcanzar un fin que el ser humano imagina por la sola virtud del artefacto o procedimiento o técnica empleados…"[13] (Ponzoni, en prensa, pp. 6–7). El arte, por el contrario, aunque estudiado por la filosofía de los artefactos —o de la cultura, en terminología de Millán-Puelles—, tiene su propia disciplina, pues corresponde a un objeto causado por la razón que, a diferencia de la tecnología, "… pretende alcanzar un fin que el ser humano imagina … necesariamente obtenido a través de otra persona" (Ponzoni, en prensa, pp. 6–7), como ocurre con aquellas obras que solo pueden ser apreciadas si se comprende de antemano el arte —por ejemplo— musical que hay detrás (Ponzoni, en prensa, 7).

En palabras de Millán-Puelles (2013), hay artificios "… que persiguen la belleza [el arte], [y otros] que se ordenan a la utilidad o al placer [la tecnología]…" (p. 52). Aunque ambas distinciones se relacionan, no son exactamente las mismas. En efecto, tanto la belleza como la utilidad o el placer, en el sentido que parece desprenderse de este autor, son objetivos, en la medida en que estas propiedades existen en el artefacto mismo con independencia de la admiración de las personas. Por supuesto, la belleza será admirada por un sujeto, y la utilidad y el placer serán *para* un sujeto, pero estas cosas se dan porque el productor ha logrado un artificio de tales cualidades que logra ser bello en sí o susceptible de ser usado para ciertas finalidades, entre ellas el placer. En el sentido de Ponzoni, en cambio, el arte requiere, para lograr aquello que el productor buscaba, no solo sus cualidades intrínsecas —una belleza en sí—, sino el hecho mismo de que otros las reconozcan. Concuerdan ambos, no obstante, en que el arte se relaciona con la experiencia estética y la tecnología con la utilidad para un fin ulterior. Por mi parte, considero

[13] Esta y todas las traducciones de textos en inglés son propias.

la postura objetiva más adecuada, pues, incluso aunque el artista requiera de otros para apreciar su obra, él es el primero que la aprecia incluso desde que la comienza a producir —teniendo la idea en su mente—, precisamente porque "… por la sola virtud del artefacto o procedimiento o técnica empleados…" (Ponzoni, en prensa, pp. 6–7) se logra lo buscado: la belleza[14]. La anterior discusión, sin embargo, como todo este excurso metodológico, debe ser ampliada y está sujeta a debate, cuya materia es suficiente para un trabajo individual que lo aborde.

Ahora bien, podría objetarse que este marco teórico es obsoleto, pues se funda en Aristóteles y el Aquinate, autores lejanos a la tecnología como la conocemos hoy. Sin embargo, no solo Millán-Puelles y Ponzoni con la NNLT, autores actuales, defienden su uso, sino que las enseñanzas de Aristóteles y Tomás, por sí mismas, contienen verdades metafísicas y éticas que se mantienen a lo largo de la historia, precisamente por el realismo que las subyace: el objeto conocido determina mi conocimiento de él, de modo que el método queda determinado por aquel, y por ello las reflexiones que comiencen desde la evidencia de la verdad del objeto tienen un derecho epistémico perenne a ser reconocidas, mientras continúen respetando los límites del método, entre ellos la consideración contingente de la materia. Por ello, podemos aplicar esta filosofía a los problemas actuales, siempre teniendo presente sus límites y, por supuesto, asumiendo la verdad de este realismo, que desde la antigüedad ha sido asumido y desde la modernidad cuestionado.

[14] Dado el marco tomista de este trabajo, la belleza se comprende como una propiedad ontológica del ente que, según santo Tómas, es "… aquello cuya vista agrada" (*S. Th.*, I, q. 5, a. 4, ad 1). En ella confluyen los sentidos, pero especialmente el gozo de la voluntad ante un bien contemplado por el intelecto. Tal gozo de lo contemplado se funda en que el objeto posee proporción, perfección y claridad (*S. Th.*, I, q. 5, a. 4, ad 1; q. 39, a. 8c), lo que le da una dimensión objetiva al arte. Ello no excluye, sin embargo, que ciertas concepciones sociales de lo 'feo' puedan ser consideradas objeto del arte, por cuanto aquello podría ser contemplado desde este concepto metafísico de lo bello. Un análisis exhaustivo de este concepto, y del objeto del arte como disciplina, excede el presente trabajo.

Con esta metodología, analizaré la tecnología de las redes sociales y cómo afectan a la vida buena de amistad supuesta la dignidad de la persona humana. Tomaré en consideración el primer, tercer y cuarto orden, y las ciencias filosóficas respectivas, delimitadas y profundamente desarrolladas por S. Tomás. Todo ello en respuesta al llamado de Ponzoni (en prensa): "La razón tiene trabajo que hacer…" (p. 10), pues el auge de la tecnología hace necesario un análisis exhaustivo de su racionalidad intrínseca y sus implicancias éticas para enfrentar los problemas, como el acá presentado, que suscita. Este trabajo, por tanto, es una aplicación a un caso concreto de la reflexión sobre el sentido último de la tecnología en cuanto a su finalidad.

Intención, tesis y esquema argumental

Ante los efectos de las redes sociales que se mencionarán más adelante y, especialmente, mi experiencia en el uso de ellas, considero que hace falta, por lo menos, esbozar filosóficamente la relación que tiene esta tecnología con la *amistad.* ¿Qué es la amistad? Es algo por lo cual *vivir* en su sentido más pleno. Sin embargo, las redes *sociales* parecen arrastrarnos a la introspección, no filosófica —que ciertamente luego se comparte con los amigos—, sino que encerrada en la máquina digital. Las risas vacías, la soledad, la compulsión y el tiempo irrecuperable generan la necesidad de dilucidar qué hacer respecto de las redes sociales, para la vida buena, mía y de *mis* amigos. Dilucidación la cual debe hacerse, a mi parecer, desde una filosofía realista y sistemática, que abarque los cuatro órdenes de la *realidad*: el tomismo.

Por este motivo, desde un análisis del orden natural, práctico y técnico, sostengo que las redes sociales contemporáneas, en muchos casos, no contribuyen a la realización de la amistad —fundamento de la felicidad que se completa en la contemplación del amigo—, sino que nos alejan de la ordenación amorosa y racional al otro como ser personal —bien infinito, tal como fue creado por Dios—, ordenación requerida para un buen uso de las redes sociales. Sin embargo, debe distinguirse entre la 'amistad' fundada en redes sociales y las redes sociales fundadas en la

amistad. Los casos perjudiciales a los que remite nuestra tesis se dan por el espejismo que puede generar una red social no sustentada en una amistad verdadera, es decir, en la 'amistad' fundada en redes. Al contrario, cuando la amistad real está en el fondo o la red social contribuye a generarla, obtenemos un instrumento valioso para la sociabilidad humana.

Para demostrar esta tesis será necesario pasar por tres etapas. Primero, explicaré brevemente qué son las redes sociales contemporáneas y sus efectos negativos inmediatos, tales como adicción, psicopatologías y aislamiento ideológico —primer y cuarto orden—. Segundo, mostraré la importancia de la amistad como bien fundamental para la vida buena en orden a la felicidad, dado el valor infinito de la persona —primer y tercer orden—. Tercero, explicaré cómo las redes sociales podrían evitar el recto orden hacia el otro como amigo, lo que, en último término, es impedir la felicidad propiamente humana realizada en comunidad —primer, tercer y cuarto orden—. Finalmente, concluyo que, en realidad, la bondad de esta tecnología queda supeditada al valor de la persona del amigo, desde lo cual se abren líneas de investigación, y reivindico el valor de la verdadera red social: la amistad.

1. Las redes sociales y sus efectos

Ponzoni (en prensa) ya da cuenta del problema que intento presentar, dejándolo abierto a un mayor desarrollo (pp. 1–2). En efecto, afirma, "[l]a amistad [y] la política … deberían ser perseguidas por una persona racional buscando la felicidad porque ellas son aspectos esenciales del florecimiento humano" (en prensa, p. 2), pero la tecnología contemporánea "… no nos ha hecho más felices" (p. 1), por ir contra estos bienes (pp. 1–2). Específicamente, la tecnología de las redes sociales contemporánea ha generado una serie de efectos importantes en la vida social, como veremos a continuación en el marco del primer y cuarto orden de la razón.

La literatura científica sobre las redes sociales es vasta y un análisis pormenorizado de ella excede no solo el objeto de este trabajo, sino

también su misma disciplina. Bastará señalar, sucintamente, lo que la comunidad científica ha concluido en cuanto a los efectos psicológicos y sociales de esta tecnología, pero no sin antes esbozar una definición de ella. Las redes sociales se enmarcan en las llamadas 'plataformas digitales', que podemos comprender simplificadamente como estructuras compuestas de un elemento técnico —*software* y *hardware*— y uno organizacional, dentro de los cuales operan variedad de otros recursos tecnológicos (de Reuver et al., 2018, pp. 126–127). En estos recursos se encuentran las aplicaciones o *apps*, *softwares* específicos que son utilizados por los usuarios finales de las plataformas digitales (de Reuver et al., 2018, p. 126). Si bien la conceptualización de las plataformas es discutida, este modo amplio de entenderlas bastará para ir al funcionamiento de las redes sociales en cuanto tales, las cuales son *apps* dentro de las plataformas, por ejemplo: Facebook, Instagram, Snapchat, TikTok, YouTube, X (Twitter), entre otras. De hecho, los *smartphones* permiten acceso constante a Internet y diversas aplicaciones, de modo que el análisis de las redes es casi el mismo que el de estos dispositivos (Oberst et al., 2016, p. 53).

Las aplicaciones señaladas contienen las características que Haidt (2024) atribuye a las redes sociales desarrolladas desde la década del 2010: (1) perfiles de usuario —información personal e intereses públicos—; (2) contenido generado por usuarios —difusión de fotos, textos, videos, etc.—; (3) redes de contactos —conexión entre perfiles—; (4) interactividad entre usuarios —interacción mutua mediante texto y reacciones— (pp. 141–142). Estas redes, indica Haidt (2024), cambiaron el paradigma a la hora de relacionarse con los *smartphones* —especialmente en la juventud— y, así, con las demás personas en sociedad, en *La Gran Reconfiguración* (p. 18, 48–49, 141–142). Pero esto no es solo un fenómeno anglosajón, pues en Chile el 95,2% de la población adulta utiliza teléfonos inteligentes (Activa, 2022, p. 12), y el acceso a ellos por niños y adolescentes es evidente.

En cuanto a los efectos de las redes sociales[15], el mismo Haidt (2024) da cuenta de bastantes, sustentando su tesis de que "… esta Gran Reconfiguración de la infancia es la principal razón del tsunami de enfermedades mentales en los adolescentes…" (p. 59). Para este autor no es coincidencia que desde la misma década en que la juventud se desarrolló inmersa en plataformas digitales, *smartphones* y redes sociales, hayan aumentado exponencialmente las tasas de depresión grave, enfermedades mentales en general, ansiedad, autolesiones y suicidio en adolescentes (Haidt, 2024, pp. 33–49).

Su tesis, sin embargo, no solo es sustentada a lo largo de su libro, sino que tiene apoyo en otros autores de la literatura especializada:

1. El fenómeno 'miedo de perderse algo' —la necesidad de saber lo que otros hacen para participar de ello—, contribuye a un uso negativo de las redes sociales y, en consecuencia, a las psicopatologías relacionadas (Oberst et al., 2016, p. 57). A este se añaden muchos otros fenómenos estudiados (Treviño Benavides et al., 2023).
2. Las redes tienen un componente adictivo altamente estudiado, con individuos presentando las mismas características que con adicciones a sustancias (Kuss & Griffiths 2017, p. 6; Nowland et al., 2017, p. 76; Treviño Benavides et al., 2023, pp. 7–8).
3. Están relacionadas fuertemente con síntomas depresivos, de ansiedad, baja autoestima, poca calidad de sueño, entre otros (Kelly et al., 2018, pp. 62–63; Ivie et al., 2020, p. 167; Valkenburg et al., 2022).
4. Su uso se asocia con mayor soledad en los usuarios, en el caso de que la utilicen para reemplazar la falta de relaciones *offline*; al mismo tiempo, la soledad determina un uso negativo de las

[15] Dejo de lado, más allá de lo ya señalado, un análisis exhaustivo de la estructura intrínseca de esta tecnología, lo que daría lugar a tesis interesantes respecto a sus efectos y a sus posibles soluciones.

redes, contribuyendo a mantener la soledad *offline* (Nowland et al., 2017, pp. 78–80).

5. En vez de contribuir a la idea de 'esfera pública' habermasiana, en ellas se dan discusiones políticas menos respetuosas, más probabilidad de acoso digital y se interactúa más con gente que piensa como uno, esto es, hay *echo chambers* producidas por *filter bubbles*, lo que lleva a polarización política y la animadversión entre grupos (Kruse et al., 2018; Terren & Borge, 2021; Pariser, 2011; Kubin & Sikorski, 2021; Lorenz-Spreen et al., 2022, p. 80; Hobolt et al., 2024, pp. 1472–1476).

Dada la naturaleza estadística de los estudios, se debe tener presente que correlación no implica causalidad, y así se manifiesta en estas investigaciones: "Puede ser que las personas con ansiedad y depresión tiendan a usar más las redes sociales, o también puede ser que quienes usan intensivamente las redes sociales desarrollen una mayor depresión" (Oberst et al., 2016, p. 52). Así ocurre con todos los efectos que se relacionan con las redes sociales, por mucha evidencia que haya. Asimismo, los estudios no son unánimes, pues se ha cuestionado la *gravedad* —no la existencia— de los efectos negativos de las redes sociales (Odgers & Jensen 2020; Ivie et al., 2020) e incluso en cada estudio se presentan dudas metodológicas y de generalización de sus resultados. La muestra de estudios encontrados apoya en su mayoría una tesis negativa acerca de los efectos de las redes sociales, pero el fondo de la cuestión no es empírico, sino filosófico.

Por ello, dudas de resultados científicos no obstan a analizar filosóficamente tales efectos negativos, los cuales existen; sean necesarios o contingentes, causa o efecto, graves o leves en general o en casos aislados; por cuanto la correlación de estos con el uso de redes sociales, y la experiencia común de su uso, nos dan lugar a un análisis de la amistad que muestra que ella es necesaria para cualquier uso de un instrumento que pretenda apoyar la sociabilidad humana. Así, con ojos de filósofo moral y de la tecnología se presenta una problemática que los estudios científicos manifiestan en potencia: el daño individual y social que

efectivamente existe puede ser reparado, en parte, cuando existe una verdadera amistad en el fondo de la sociabilidad online, sin la cual las redes sociales no son verdaderamente sociales; y aunque no haya unanimidad ni generalidad absoluta, los daños causados o agravados por las redes se producen, justamente, por impedir la verdadera amistad, pues si ella se diera, esos daños no ocurrirían como efectivamente lo hacen, pues la amistad es fuente de felicidad humana. Las conclusiones, dada nuestra metodología realista, no serán categóricas, como tampoco lo son los estudios científicos, pero el fondo ético y metafísico de la persona y la amistad serán como una directriz permanente ante el auge de las tecnologías sociales. En cuanto a la red social en sí, mi tratamiento del asunto no es nada nuevo, pero lo interesante será analizarlo a la luz filosófica del valor de la amistad[16].

2. La amistad en la persona humana para la felicidad

2.1. La persona humana

Decía Canals que es propio de la persona, en cuanto se sigue de su misma esencia, ser "… sujeto y término de amor y amistad" (1989). Es una descripción especialmente precisa para el tema que nos ocupa, pues el valor de la amistad que defiendo se funda en una metafísica de la persona. No pretendo innovar en esta cuestión, pues grandes maestros han reflexionado sobre ello —en primer lugar S. Tomás—, sino solo subirme en sus hombros para lograr demostrar mi punto.

Una breve explicación de la clásica definición boeciana de persona, asumida por S. Tomás, será necesaria para comenzar. "… *persona es una sustancia individual de naturaleza racional*" (*S. Th.*, I, q. 29, a. 1, ad 1). Lo que se intenta predicar con *persona* no es una determinada naturaleza, sino que *lo singular* en la naturaleza racional, esto es, la sustancia misma de los entes racionales, que, a diferencia de los

[16] Más allá de los estudios científicos, piense usted en su propia experiencia con las redes y lo que ve en sus cercanos: ¿acaso no hay un uso absolutamente inútil de ellas, muchas veces espontáneo, que impide el mejor desarrollo personal, profesional y comunitario?

accidentes, se individualiza por sí y subsiste en sí misma y no en otro (*S. Th.*, I, q. 29, a. 1c; *De potentia*, q. 9, a. 1c). Pero no es cualquier individuo del género de las sustancias, sino que uno designado con un término especial por el modo dignísimo de subsistir propio de los racionales (*De potentia*, q. 9, a. 2, ad 1; *S. Th.*, I, q. 29, a. 3, ad 2).

Las personas son "… lo que es más perfecto en toda la naturaleza" (*S. Th.*, I, q. 29, a. 3c), lo más digno del universo, por la perfección de su naturaleza y la dignidad de su fin: son las únicas que dominan su propio acto, son libres; y son las únicas que realizan el fin último del universo, el conocimiento y amor de Dios (*C. G.*, III, c. 111, n. 1).

El primer aspecto es al que el Aquinate da énfasis para otorgar un nombre especial a los individuos de naturaleza racional, ya que "… tienen dominio sobre sus propios actos; y no son solo hechas actuar, como las otras; sino que pueden actuar por sí mismas…" (*S. Th.*, I, q. 29, a. 1c). La operación personal es algo originalísimo por el dominio del propio acto, que individualiza a la persona más que a cualquier otro ente (Echavarría, 2013, p. 283), lo cual solo se explica por el modo especial en que ella posee su propio ser (Amado, 2014, p. 52). También el hombre singular, compuesto de materia y forma e individualizado por su materia (*S. Th.*, I, q. 29, a. 2, ad 3; *De potentia*, q. 9, a. 1, ad 6; a. 4c), es persona humana, en cuanto con *persona* significamos aquello individualizado en una naturaleza racional (*De potentia*, q. 9, a. 1, ad 6; a. 4c). Pero no es la individuación de la materia la que nos provee la especial razón de persona, pues si así fuera sus actuaciones se circunscribirían a aquello que su materia específica permite (Amado, 2014, p. 51).

En efecto, "… el ser de una cosa es propio de ella y distinto del ser de cualquier otra…" (*De potentia*, q. 7, a. 3c) y el ser mismo "… es lo más íntimo en todas las cosas" (*S. Th.*, I, q. 105, a. 5c) y es lo formal en cada cosa (*S. Th.*, I, q. 4, a. 1, ad 3). Así, persona indica, ante todo, la individualidad propia del *esse* del ente en cuestión —"Porque el ser pertenece a la constitución misma de la persona" (*S. Th.*, III, q. 19, a. 1, ad 4)—, en cuya especial posesión somos capaces del dominio de una operación que trasciende la determinación material y da lugar a la

elección por la apertura intelectual al bien universal (Amado, 2014, p. 52; *De malo*, q. 6c; *S. Th.*, I, q. 105, a. 4c). La libertad, entonces, se sigue de un especial ser dado en un modo de ser racional, pues el obrar sigue al ser (Amado, 2014, p. 52).

Pero este ser especialísimo, dado en la naturaleza racional, es capaz del conocimiento y amor de Dios. El universo completo tiene por fin último a Dios (*S. Th.*, I, q. 44, a. 4c) y de ello no se desentiende la persona humana, que por su voluntad se inclina al bien universal —Dios— (*S. Th.*, I, q. 105, a. 4c). Las demás criaturas, en cuanto persiguen su propia perfección, buscan asemejarse a la bondad divina, pues nada es bueno si no en cuanto participa de Su semejanza (*S. Th.*, I, q. 44, a. 4c; ad 3), pero lo hacen solo por cierta imitación de su bondad, a diferencia de la persona, que alcanza el fin último "… conociendo y amando a Dios…" (*S. Th.*, I-II, q. 1, a. 8c).

Por su naturaleza, la felicidad de la persona creada se realiza en la contemplación de la verdad de lo más inteligible con el hábito de la sabiduría (X *In Ethic.*, lect. 10, n. 1478-1479; *S. Th.*, I-II, q. 3, a. 5c). Así, hay felicidad solo mediante una posesión intelectiva —seguida por el gozo de la voluntad— (*S. Th.*, I-II, q. 3, a. 4c; a. 5, ad 2), en que el cognoscente se asimila a lo conocido —la perfección de lo conocido se encuentra en el cognoscente (*De veritate*, q. 2, a. 2c)—, en cuya semejanza conoce la realidad (*De veritate*, q. 1, a. 1c; *S. Th.*, I, q. 85, a. 2; ad 1). Esta semejanza del racional solo es posible, al igual que la libertad, por la especial perfección del ser personal: un ser intrínsecamente inmaterial, que puede por tanto recibir la actualidad de todas las formas (Widow, 2014, pp. 61–63). Dios, lo más inteligible, colmaría suficientemente el apetito del bien en sí (*S. Th.*, I-II, q. 2, a. 8c; q. 5, a. 3c) y en Él conoceríamos el orden universal completo; pero en esta vida resulta imposible dada la limitación de nuestro conocimiento (*S. Th.*, I, q. 12, a. 4c). Sin embargo, la posesión cognoscitiva de Dios, lejos de nosotros, no es lo único que da plenitud al ser personal: hay otro objeto proporcionado a la contemplación.

Hay otro ser de algún modo infinito: la persona creada (Echavarría, 2013, pp. 295–299; Amado, 2014, p. 57). Aunque en ella no se contemple, como en la Verdad Suma, el orden universal, no por nada es "… lo que es más perfecto en toda la naturaleza" (*S. Th.*, I, q. 29, a. 3c). En efecto, es a ella hacia lo que se ordenan todas las partes del universo, porque la persona humana, por dignidad de fin, es la que más se asemeja a la perfección divina y 'es en cierto modo todas las cosas', en cuanto no está limitada a su esencia individual, sino que se abre a la perfección esencial de toda entidad, acercándose a la infinitud del ser por el conocimiento intelectual (*C. G.*, III, c. 112, n. 5; *S. Th.*, I, q. 14, a. 1c; *De veritate*, q. 2, a. 2c). Así, "… es posible que en una sola cosa exista la perfección de todo el universo" (*De veritate*, q. 2, a. 2c). La persona humana, aunque singular, es verdadero objeto de contemplación y felicidad, dada la perfección de un ser intelectual que realiza la perfección del todo universal.

2.2. Amistad: felicidad contemplativa en la conversación

La anterior explicación, propia del primer orden de la razón —metafísica—, con elementos del tercero —ética teleológica—, da lugar a que la persona humana sea fin último universal, objeto de contemplación, esto es, de felicidad. Estas no son palabras vacías: la persona humana, *el prójimo*, es el objeto propio de nuestra felicidad, la cual solo se realiza en amistad.

Es común experiencia humana que la amistad es un bien humano (Finnis, 2011a, p. 88), pero más que la amistad como una virtud o una relación —que lo es (*E. N.*, VIII 5, 1157b5-10; Finnis 2011a, 135; *S. Th.*, II-II, q. 23, a. 3, ad 1)—, se reconoce principalmente, en la experiencia práctica, el valor intrínseco del amigo mismo en su particularidad personal (Finnis, 2011a, p. 142). Ya decía Aristóteles que los amigos son cosa necesaria para la vida (*E. N.*, VIII 1, 1155a1-5), no solo por la utilidad o el placer que nos pueden dar, sino que principalmente por la bondad del vivir mismo del amigo virtuoso, que se ama por ser él mismo y no otro, siendo necesario para la felicidad (*E. N.*, IX 9, 1170a10-20; *E.*

E., VII 2, 1237a40-1237b5). Lo que la reflexión metafísica nos ha mostrado, por tanto, no es ajeno a la experiencia común y evidente de la amistad: la persona es un bien perfectísimo, necesario para nuestra plenitud.

La amistad es un tipo de amor, no el que busca algo meramente para sí —concupiscencia— y tampoco el mero deseo del bien al otro —benevolencia—, sino que es una comunicación y benevolencia *mutuas* (*S. Th.*, II-II, q. 23, a. 1c), en la que se reúnen cinco condiciones recíprocas, pues se quiere: que el amigo viva, que tenga bienes, hacerle el bien, convivir con él y coincidir en sus sentimientos (*S. Th.*, II-II, q. 25, a. 7c).

Justamente, las dos aristas de la dignidad humana, perfección de naturaleza —dominio de sí— y dignidad de fin —realiza en sí el fin último del universo—, son relacionadas directamente por el Aquinate en su tratamiento de la amistad:

> … el amor de amistad es doble: primero, está el amor por el amigo a quien ofrecemos nuestra amistad, segundo, el amor por aquellas cosas buenas que deseamos para nuestro amigo … La amistad … no puede tener por objeto a una criatura irracional: primero porque la amistad es hacia aquel a quien deseamos cosas buenas, mientras que, hablando propiamente, no podemos desear cosas buenas a una criatura irracional, porque no es capaz, propiamente hablando, de poseer el bien, siendo esto propio de la criatura racional la cual, a través de su libre albedrío, es dueña de disponer del bien que posee. … Segundo, porque toda amistad está fundada en alguna comunión de vida; pues *nada es tan propio de la amistad como vivir juntos*, como el Filósofo demuestra (*Ethic.* viii, 5). Ahora bien, las criaturas irracionales no pueden tener comunión en la vida humana que es regulada por la razón. Por tanto, la amistad con criaturas irracionales es imposible… (*S. Th.*, II-II, q. 25, a. 3c).

Como es evidente, solo la persona es sujeto y término de amor y de amistad, porque su especial acto de ser intelectual la hace capaz de poseer el bien libremente, así como comunicar su propia intimidad —*communicationem habere in vita humana*— mediante la razón. Lo esencial para la contemplación del otro como persona es, como se ve, la *communicatione vitae* mediante el pensamiento, que se da en la *convivencia* con el otro. "La amistad no puede existir sino hacia criaturas racionales, que son capaces de corresponder el amor, y de comunicarse uno con otro en las diversas tareas de la vida" (*S. Th.*, I, q. 20, a. 2, ad 3). Es el amor hacia el amigo, aprehendido como bien infinito, el que nos mueve a querer unirnos con él, pero es en la comunicación y convivencia donde se perfecciona nuestro ser intelectual, realizando tal unión realmente (*S. Th.*, I-II, q. 28, a. 1c; ad 2).

Porque, así como la contemplación de Dios consiste en amarlo y conocerlo, contemplar a la persona del amigo consiste, también, en cierto amor y conocimiento, pero hacia el ser mismo de ella. Sin embargo, su ser en sí es incomunicable, por pertenecerle íntima y singularísimamente (Martínez, 2006, pp. 84–85). El único modo de comunicar su perfección, dando lugar a la posesión intelectual de su ser, es mediante la operación propia del racional: la palabra y el amor (Martínez, 2006, p. 87). En la conversación, al manifestar los pensamientos, es donde especialmente se deleitan los amigos en la bondad amada del ser del otro, compartiendo el modo de vida propio de cada uno: en el juego, en lo cotidiano, en lo filosófico y, especialmente, en lo matrimonial (Letelier, 2024, p. 18). La importancia de la comunicación estriba en lo siguiente:

> Conversar con alguien implica llegar a saber, en un único acto, cómo es algo según alguien y quién es ese que habla de eso. El que habla, hablando de cualquier cosa, habla también de sí, pues 'lo conocido está en el cognoscente según el modo del cognoscente': quien expresa lo que entiende, expresa también su modo de entenderlo y a sí mismo como sujeto que entiende (Letelier, 20240, p. 23).

"Porque el árbol se conoce por su fruto. … de la abundancia del corazón habla la boca" (Mateo, 12:34)[17]. El corazón del otro, lo más íntimo de sí, se manifiesta mediante el pensamiento, que está, sin embargo, obstaculizado por el cuerpo y, posiblemente, por la negación de la voluntad (*S. Th.*, I, q. 107, a. 1, ad 1). El obstáculo corporal, no obstante, no implica el desdén a lo material. Al contrario, la persona humana, compuesta de forma y materia, existe por el acto de ser de su alma subsistente, que informa a la materia corporal y le comunica su propia actualidad para constituirse en sustancia (*S. Th.*, I, q. 76, a. 1, ad 5; a. 7). Así, el cuerpo es *humano* por la forma sustancial que le da la especie y el ser: el alma racional (*C. G.*, IV, c. 81, n. 7). Es de la esencia del hombre darse en un cuerpo que recibe el *esse* del alma; y entonces es propio de su individualidad personal el realizarse en tres dimensiones, sujeto al movimiento, al tiempo y al lugar (Echavarría, 2019, pp. 347–348).

Por ello, el modo de expresar lo entendido en la persona *humana*, en cuanto su ser es corporal, es una expresión, no de meros conceptos, sino que, de palabras habladas, gesticuladas, con movimientos, con emociones. La unión real con el amigo, la real *con-vivencia*, no puede realizarse sino en la presencia corporal, un estar-con locativo, del amigo, pues la persona humana *es* corporal, por lo que provee de signos sensibles para expresar lo íntimo de su pensamiento, en la medida en que así lo quiera (*S. Th.*, I, q. 107, a. 1, ad 1).

Esta corporeidad existente, desde la que se comunica lo íntimo de la persona humana, implica también la vida orgánica y animal, por la cual se poseen partes orgánicas funcionales asociadas a la distinción sexual, y al conocimiento y apetito sensibles, pues el alma intelectual está en todo el cuerpo y cada una de sus partes (*S. Th.*, I, q. 76, a. 8; q. 78, a. 1; Echavarría, 2019, pp. 348–355). Por ello, la contemplación intelectual del amigo pasa necesariamente por la comprensión profunda de y enfocada

[17] Agradezco a nuestro Señor por darme la idea de poner esta cita.

en las cualidades sensibles del amigo, su temperamento, sus deseos, sus modos de reaccionar, de hablar, de moverse, sus gestos, su tono, su vestimenta, su calor y *cercanía*, todo lo cual, y más, expresa la intimidad espiritual de su ser personal precisamente porque el *esse* corporal es el *esse* recibido del alma intelectual, que constituye la personalidad (Echavarría, 2019, pp. 363–364).

En suma, el amigo manifiesta su originalidad irrepetible en la conversación amistosa; con su presencia y unión real puede conocerse intelectual y sensiblemente su individualidad corpórea personal, especialmente cuando el bien que une a los amigos no es algo superficial, sino que es la misma singularidad del otro y su ordenación virtuosa a un bien común, el cual hará la amistad más perfecta y estable tanto más perfecto sea este bien —Dios—. En la comunicación de vida, entonces, "[l]o realmente valioso … será la subjetiva disposición del amigo respecto de ese tema, manifestativa de quién es él, y el gozo consiguiente de compartirla recíprocamente con el amigo" (Letelier, 2024, p. 24).

3. ¿Las redes sociales impiden la felicidad?

Heidegger (2021) ya veía, desde otro aparataje conceptual, que la tecnología contemporánea, la esencia de la técnica, no hacía otra cosa que mostrarnos la realidad, el ser, como un fondo permanente para utilizarse merced de objetivos utilitarios (pp. 15–17, 23). De ello no quedaba ajena la persona misma, que ahora era vista como un mero solicitador del fondo e incluso parte de este, impidiéndosele develar el mundo de un modo 'originario' ajeno a la utilidad (Heidegger, 2021, pp. 18, 25–26). No es que ocurra necesariamente esto con las redes sociales, pero lo que manifiesta la literatura científica, así como la experiencia cotidiana, es que en alguna medida nos alejan de la verdadera unión amorosa con la otra persona, como mostraré con un análisis propio del cuarto orden de la razón, desde la finalidad intrínseca de la tecnología.

Como decía Platón en el famoso mito de Theuth (*Fedro*, 274a-c), las γράμματα (*grámmata*) —representando a la tecnología en general— son un φάρμακον (*fármakon*), esto es, tanto un veneno como un

medicamento. Las letras, tecnología del hombre, tienen esta dualidad intrínseca que, aunque las intenciones de su productor sean la memoria y la sabiduría, inevitablemente conllevarán, en alguna medida, olvido y apariencia de saber (*Fedro*, 275a). Así ocurre con las redes sociales, a las que se puede aplicar el análisis del autodenominado tomista McLuhan (1969): "… el medio en sí es el mensaje…", de modo que aquello que queremos transmitir, efectuar y lograr con la tecnología está más determinado por el medio mismo en que y a través del cual se realiza (p. 5). Los medios modifican nuestro modo de percibir el mundo, pues amplifican nuestros sentidos, los cuales se adaptan a la tecnología y, simultáneamente, esta modifica la cultura dada su influencia en nuestra comprensión del mundo (McLuhan, 1969, p. 4).

La estructura de las redes sociales, como fue descrita al inicio de este ensayo, da lugar a una ampliación sensorial sin precedentes, que genera un espejismo de conectividad y sociabilidad amistosa, una ilusión de *communicationem habere in vita humana*, pues de hecho es lo que tienen por finalidad intrínseca. Sin embargo, en muchos casos el efecto es el contrario, como se manifiesta en las consecuencias estudiadas científicamente de las redes sociales: su uso, en realidad, parece alejar de la verdadera conexión humana necesaria para la felicidad, en cuanto favorecen la soledad, la adicción, la discusión irrespetuosa y el desarrollo de psicopatologías o, cuando menos, solo agravan las ya existentes. El φάρμακον está presente, pues no negamos, dada la contingencia de las cosas humanas —siguiendo a Aristóteles—, que esta tecnología pueda tener efectos positivos consiguiendo su finalidad, así como en los bienes de la vida pasa lo contrario, pues "… muchos han muerto a causa de la riqueza y otros a causa de la valentía" (*E. N.*, I 3, 1094b15-20). Así, las redes sociales, en muchos casos, pueden reunir personas, generar amistades, relaciones afectivas y ofrecer un sano entretenimiento, pero ellas solo son instrumentos de tales beneficios, con su posible contracara perjudicial, al igual que las γράμματα pretendían dar memoria y sabiduría al hombre, pero causaron también olvido y apariencia de saber.

Sin embargo, el principal perjuicio, manifiesto desde un análisis filosófico y más allá de lo empírico —que puede expresar esta explicación filosófica de fondo y da lugar a ella—, consiste en lo que McLuhan vislumbró: las redes sociales, al modificar nuestra percepción, aumentando la 'conexión' global, intentan dar lo que no tienen aparentando tenerlo a través del mensaje que dan actuando como el medio que son.

Las redes sociales, como medio, aunque aparentan extender nuestro marco de amistades y disminuir la soledad, realizando una apertura global e instantánea a la vida íntima de tantos seres personales —con los perfiles de usuario, el contenido publicado, las redes de contactos y la interactividad cuasi ilimitada en Internet— y ampliando nuestros sentidos más allá de los límites espaciales —con todo tipo de herramientas audiovisuales y de difusión, que permiten *ver* y *oír* lo que *no está presente* en cuanto tal para ser visto y oído—, en realidad son como un velo a la individualidad propia de la persona humana: su entendimiento y virtud expresados mediante la palabra directa y voluntaria suya, con signos sensibles que, aunque sean obstáculo de la aprehensión directa del intelecto del amigo, son propios de su ser mismo, en cuanto son *sus* signos corporales, y no los provistos por una plataforma digital *online*. Es el cuerpo efectivamente percibido el que provee de la experiencia del otro, pues el cuerpo *es* la persona en su materialidad y en él se expresa la interioridad inmaterial del intelecto y la voluntad. Aunque la red social puede ser buena como primera aproximación al otro, que se expresa remota y voluntariamente en ella, el problema es pensar que con eso basta. Pero justamente la red social en sí aparenta dar lo que no tiene, pues ella no es una expresión del otro en cuanto otro, como lo es el cuerpo, sino que es un artefacto distinto de la persona que nos aproxima a ella aparentando mostrárnosla en sí misma; pero no es su cuerpo el presente, sino una imagen de él.

La verdadera conexión amistosa requiere del conocimiento intelectual y sensible, espiritual y corporal, de la persona humana, lo que se ve directamente obstaculizado, ya no como el cuerpo obstaculiza el

pensamiento —el primero, que se sigue del ser del alma intelectual, manifiesta al segundo—, sino como una pintura obstaculiza el paisaje. Por muy hermoso que sea el óleo, este no es la pradera conocida en cuanto tal, sino solo su imagen. Por muy avanzada que esté la tecnología audiovisual y comunicacional, lo que aquella muestra, sí, aumenta el alcance de nuestros sentidos, pero oculta la verdadera y real presencia del amigo, en su ser corpóreo directamente conocido en la *convivencia* con él. En efecto, así como la pintura, obra externa, de alguien, expresa su individualidad, la red social puede hacerlo mediante avatares, fotografías, etc. Pero ello nunca será la real presencia de la persona, un sujeto corpóreo: el error yace en equiparar ambos planos, y la tecnología de la red social misma aparenta esta igualdad. Incluso, mientras más desarrollada esté la tecnología más aparentará mostrar la realidad del amigo, cuando por sí misma nunca expresará lo que la sustancia misma de este manifiesta: su racionalidad corporizada. Al contrario, mientras menos *real* parezca la tecnología, menos causará este espejismo.

De este modo, aunque es cierto que —sin hacer un análisis exhaustivo de su estructura intrínseca— las redes sociales puedan ser usadas para mantener contacto con amistades previas, desarrollarlas, así como para encontrar nuevas, el verdadero sustento del amor personal recíproco es la convivencia real, sustancial, con la persona que, recordemos, se dice tal por su subsistencia individual. Esta tecnología por sí sola, entonces, no provee de un sólido fundamento para la contemplación personal y, en muchas ocasiones, genera un obstáculo material adicional, por la aparente conectividad y transparencia que provee. Sin embargo, y especialmente ante jóvenes con psicopatologías previas, soledad y adicción, la ilusión del carácter social de las redes implica la falsa esperanza de una plenitud que solo puede dar la real contemplación del amigo, en mutua convivencia, conversación, conocimiento y amor, en cuerpo y espíritu.

Las redes sociales, por tanto, sin el sustento de la amistad, sí la impiden, en cuanto aparentan dar tal sustento. Cuando tal sustento existe, sea actualmente o como la finalidad misma de la conexión *online*, el

instrumento de la red social puede servir valiosamente al valor de la amistad[18]. Esta conclusión, no obstante, dada la metodología presentada al inicio, constituye solo una verdad general, aplicada a la mayoría de los casos en el tipo de 'amistades' fundadas en redes sociales, en un mundo contemporáneo inmerso en Internet. Lo que sí es seguro, por el análisis metafísico y ético de la persona y la amistad, es que solo en esta última hay felicidad, a la cual las redes sociales deben ordenarse para un uso correcto de ellas y para mitigar o eliminar los perjuicios que esta tecnología podría conllevar. La amistad es la verdadera impulsora del bien que la red social conlleva, en cuanto por aquella se realiza la felicidad humana; felicidad cuya esencial perfección provee del sentido último del que toda tecnología carece por sí.

Conclusión: la amistad como verdadera red social

Dados los fundamentos metafísicos, antropológicos y éticos de la amistad, podemos afirmar, con Ponzoni (en prensa), no solo que "… es como si no fuéramos totalmente racionales en nuestras relaciones con la tecnología…" (p. 3), sino que con seguridad no lo somos, porque "… desarrollamos y adoptamos con entusiasmo tecnologías que contribuyen a hacer de nuestras vidas más privadas de sentido y del verdadero gozo que solo una profunda *conexión humana* puede aportar" (p. 3). Ya lo decía Turkle (2011): "En el extremo, estamos tan envueltos en nuestras conexiones que nos descuidamos unos a otros" (p. 294). *En el extremo*, aquel en que las amistades tienen por único fundamento la red social y no apuntan a la verdadera convivencia humana. Y aunque la conclusión no sea absoluta, por la variabilidad de la vida humana, las redes sociales, sin el sustento de la amistad, aparecen como un espejismo de sociabilidad, cuando es en la persona del amigo —real, individual y corporal—, donde la felicidad humana en la amistad se realiza, y sin la cual cualquier uso *social* de la tecnología solo oculta la verdadera perfección de la vida

[18] Dejando de lado posibles problemas que van más allá de la conectividad directa con otros, como podría ser la adicción a videos cortos.

humana, más importante que la posesión de la ciencia: la conversación con *mi* amigo (Amado, 2014, p. 58). En el fondo, la verdadera *red social* es la amistad, a partir de la cual la tecnología en análisis adquiere su bondad y contribuye al bien humano.

En línea con la tesis defendida, surgen dos áreas de investigación: primero, un análisis del objeto moral a la hora de elegir usar redes sociales, con su intención y circunstancias (*S. Th.*, I-II, q. 18), para lo cual es necesario una explicación profunda de la estructura tecnológica intrínseca a las redes y una conexión con el conocimiento por connaturalidad (Buzeta, 2023, 2025) que podría verse modificado por la acción de esta tecnología; segundo, dado que hay un tipo de amistad según el bien buscado en ella (*S. Th.*, II-II, q. 26, a. 2c), el bien común político da lugar a la amistad política, fundamental para la existencia de la sociedad y su máxima perfección (Letelier, 2012, p. 169; *Prooemium*, *In Ethic.*; *Prooemium*, *In Politic.*; *S. Th.*, I-II, q. 90, a. 2; ad 2), lo cual da paso a un análisis de los efectos de las redes sociales en la amistad requerida y perfeccionada en la comunidad política.

REFERENCIAS

Activa. (2022). *Estudio Global. Privacidad de la información digital.* AIM Chile. https://aimchile.cl/wp-content/uploads/2022/05/211856_Estudio-Global-WIN-Tecnologia-V2-0516_compressed.pdf

Amado, A. (2014). Persona y verdad: para una metafísica de la amistad. En M. E. Gómez de Pedro, M. P. Cereceda Martínez & I. Serrano del Pozo (eds.), *La persona: divina, angélica, humana. Actas del Primer Congreso Internacional de Filosofía Tomista* (pp. 47–58. RIL editores.

Aristóteles. (1985). *Ética nicomáquea. Ética eudemia.* (J. Pallí Bonet, Trans.) Gredos.

Aristóteles. (1995). *Tratados de Lógica (Órganon). Vol II. Sobre la interpretación, Analíticos primeros, Analíticos segundos.* (M. Candel Sanmartín, Trans.) Gredos.

Aristóteles. (2015). *Ética nicomaquea.* (E. Sinnott, Trans.) Ediciones Colihue.

Aristóteles. (2021). *Metafísica.* (T. Calvo Martínez, Trans.) Gredos.

Buzeta Undurraga, S. (2023). *Sobre el conocimiento por connaturalidad* (Segunda ed.). Ediciones Saint Michael Archangel International University.

Buzeta Undurraga, S. (2025). La singular condición de existencia del conocimiento por connaturalidad en la amistad desde el pensamiento de Tomás de Aquino. *Scripta Medievalia*, *18*(1): 165–194. https://doi.org/10.48162/rev.35.049

Canals, Francisco (1989, Abril). *La persona, sujeto y término de amor y amistad.* Enciclopedia Mercabá: https://www.mercaba.org/Filosofia/Antropologia/la_persona_sujeto_termino.htm

de Aquino, T. (1952). *On the Power of God.* (English Dominican Fathers, Trans.). The Aquinas Institute. https://aquinas.cc/la/en/~QDePot

de Aquino, Tomás (1967). *Suma contra los gentiles. Volumen II. Libros 3.° y 4.°. Dios, fin último y gobernador supremo. Misterios divinos y postrimerías* (Segunda ed.). (L. Robles & A. Robles, Trans.) BAC.

de Aquino, Tomás (2010). *Comentario a la Ética a Nicómaco de Aristóteles* (Tercera ed.). (A. Mallea, Trans.) EUNSA.

de Aquino, Tomás (2015). *Cuestiones disputadas sobre el mal.* (E. Téllez, Trans) EUNSA.

de Aquino, Tomás (2016). *Cuestiones disputadas sobre la verdad.* Tomo I. (J. García López, Á. L. González, C. Llano, M. J. Soto-Bruna, J. F. Sellés, E. Téllez, S. Gelonch, & S. Argüello, Trans.) EUNSA.

de Aquino, Tomás (2019). *Opera omnia.* Recognovit et instruxit Enrique Alarcón automato electronico Pompaelone ad Universitatis Studiorum Navarrensis aedes a MM A.D. Corpus Thomisticum. https://www.corpusthomisticum.org/iopera.html

de Aquino, Tomás (2020). *Summa Theologiae.* (Trad. Fr. L. Shapcote OP) The Aquinas Institute. https://aquinas.cc/la/en/~ST.I

de Aquino, Tomás, & de Alvernia, Pedro (2001). *Comentario a la Política de Aristóteles.* (Trad. A. Mallea) EUNSA.

Echavarría, M. F. (2013). El modo de subsistir personal como reflexión sustancial según Tomás de Aquino. *Espíritu: Cuadernos del Instituto Filosófico de Balmesiana*, *62*(146), 277–310. https://revistaespiritu.istomas.org/modo-subsistir-personal-reflexion-sustancial-segun-tomas-aquino/

Echavarría, M. F. (2019). La corporalidad humana según Tomás de Aquino. *Revista Española de Teología*, *79*(3), 345–366. https://repositorio.sandamaso.es/handle/123456789/2333

Finnis, J. (2011a). *Natural law and natural rights* (Second ed.). Oxford University Press.

Finnis, J. (2011b). *Intention and identity. Collected essays: Volume II.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199580064.001.0001

Finnis, J. (2019). *Tomás de Aquino. Teoría moral, política y jurídica.* (F. Morales, Trads.) IES.

Fuentes Mendiola, A (Ed.) (2008). *Los evangelios.* (A. Fuentes Mendiela, Trans.) Rialp.

Giussani, L. (2008). *El sentido religioso* (Décima ed.). (J. Oriol, Trans.) Ediciones Encuentro.

Haidt, J. (2024). *La generación ansiosa.* (V. Puertollano López, Trans.) Paidós.

Heidegger, M. (2021). *La pregunta por la técnica.* (J. A. Escudero, Trans.) Herder.

Hobolt, S., Lawall, K., & Tilley, J. (2024). The polarizing effect of partisan echo chambers. *American Political Science Review*, *118*(3), 1464–1479. https://doi.org/10.1017/S0003055423001211

Hui, Y. (2024). *La pregunta por la técnica en China.* (M. Gonnet, Trans.) Caja Negra Editora.

Ivie, E., Pettitt, A., Moses, L., & Allen N. (2020). A meta-analysis of the association between adolescent social media use and depressive symptoms. *Journal of Affective Disorders*, *275*, 165–174. https://doi.org/10.1016/j.jad.2020.06.014

Kelly, Y., Zilanawala, A., Booker, C., & Sacker, A. (2018). Social media use and adolescent mental health: Findings from the UK millennium cohort study. *EClinicalMedicine*, *6*, 59–68. https://doi.org/10.1016/j.eclinm.2018.12.005

Kubin, E., & Sikorski, C. (2021). The role of (social) media in political polarization: a systematic review. *Annals of the International Communication Association*, *45*(3), 188–206. https://doi.org/10.1080/23808985.2021.1976070

Kuss, D., y Griffiths, M. (2017). Social Networking Sites and Addiction: Ten Lessons Learned. *International Journal of Environmental Research and Public Health*, *14*(3), 311. https://doi.org/10.3390/ijerph14030311

Letelier, G. (2012). La amistad como principio de la polis. *Derecho y Humanidades,* *1*(19), 153–177. https://derechoyhumanidades.uchile.cl/index.php/RDH/article/view/25723

Letelier, G. (2024). *Conversatio nostra in caelis*. La conversación amistosa como perfección del entendimiento según Tomás de Aquino. *Revista chilena de estudios medievales*(26), 1–30. http://dx.doi.org/10.4067/S0719-689X2024000200001

Lorenz-Spreen, P., Oswald., L., Lewandowsky, S., & Hertwig, R. (2022). A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behaviour*, 7, 74–101. https://doi.org/10.1038/s41562-022-01460-1

Martínez, E. (2006). Vida personal y comunicación interpersonal. *Revista Española de Filosofía Medieval*, 13, 81–88. https://doi.org/10.21071/refime.v13i.6274

McLuhan, M. (1969). *The Playboy Interview: Marshall McLuhan*. Playboy Magazine. https://web.cs.ucdavis.edu/~rogaway/classes/188/spring07/mcluhan.pdf

Millán-Puelles, A. (2013). *Obras completas. Volumen II. Fundamentos de filosofía (1955)*. Rialp.

Oberst, U., Wegmann, E., Stodt, B., Brand, M., & Chamarro, A. (2016). Negative consequences from heavy social networking in adolescents: The mediating role of fear of missing out. *Journal of Adolescence*, 55, 51-60. https://doi.org/10.1016/j.adolescence.2016.12.008

Odgers, C., & Jensen, M. (2020). Annual research review: Adolescent mental health in the digital age: Facts, fears, and future directions. *Journal of Child Psychology and Psychiatry*, *61*(3), 336–348. https://doi.org/10.1111/jcpp.13190

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin Books.

Platón (2023). *Diálogos. Volumen III. Fedón, Banquete, Fedro*. (C. García Gual, M. Martínez Hernández, E. Lledó Íñigo, L. A. de Cuenca y Prado, & J. L. Navarro González, Trans.). Gredos.

Ponzoni, Federico (en prensa). Cloaks and iPhones. *Artículo en espera de publicación.*

de Reuver, M., Sørensen, C., & Basole, R. (2018). The digital platform: a research agenda. *Journal of Information Technology*, *33*(2), 124–135. https://doi.org/10.1057/s41265-016-0033-3

Terren, L., & Borge, R. (2021). Echo chambers on social media: A systematic review of the literature. *Review of Communication Research*, 9, 99-118. https://doi.org/10.12840/ISSN.2255-4165.028

Tollefsen, C. (2018). Aquinas's Four Orders, Normativity, and Human Nature. *The Journal of Value Inquiry*, *52*(3), 243–256. https://doi.org/10.1007/s10790-018-9657-6

Treviño Benavides, T. B., Alcorta Castro, A. T., Garza Marichalar, S. A., Peña Cisneros, M., & Baker Suárez, E. C. (2023). *Social Media Addiction in Generation Z Consumers*. Springer. https://doi.org/10.1007/978-3-031-33452-8

Valkenburg, P., Meier, A., & Beyens, I. (2022). Social media use and its impact on adolescent mental health: An umbrella review of the evidence. *Current Opinion in Psychology*, 44, pp. 58–68. https://doi.org/10.1016/j.copsyc.2021.08.017

Turkle. S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.

Widow, F. (2020). Santo Tomás frente a la *ley de Hume*. En C. Casanova (ed.), *El obrar sigue al ser: metafísica de la persona, la naturaleza y la acción* (pp. 339–355). RIL editores.

Widow, J. (2014). La personalidad como perfección del ente. En M. E. Gómez de Pedro, M. P. Cereceda Martínez & I. Serrano del Pozo (eds.), *La persona: divina, angélica, humana. Actas del Primer Congreso Internacional de Filosofía Tomista* (pp. 59–70). RIL editores.

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 79-111
ISSN 3087-260X
Santiago de Chile

UN NATURALISMO MINIMALISTA[1]

Constantino Contreras Hernández[2]
Universidad Autónoma de México

Resumen

En este artículo propongo una caracterización minimalista del naturalismo como la tesis según la cual: 1) hemos de tener un mínimo, mas no total, compromiso con la ontología de nuestras ciencias, y 2) hemos de tener un compromiso *meta-metodológico* con el uso de métodos que permitan justificaciones y afirmaciones que puedan ser *intersubjetivamente evaluadas*, tales como argumentos formales e informales, o el uso de los métodos mismos de las ciencias. De esta forma, primero habré de criticar las definiciones clásicas del naturalismo entendido como una forma de fisicalismo y una restricción en la postulación de entidades y explicaciones sobrenaturales en las

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Contreras, C. (2026) Un naturalismo minimalista. *Revista de Filosofía Homónima, 1*(1), pp. 79-111. https://doi.org/xxxxxxx[Por asignar]

[2] Constantino Contreras es Estudiante de licenciatura en matemáticas, Universidad Autónoma de México (UNAM). Email: constantino.contreras@ciencias.unam.mx. ORCID: 0009-0000-2863-3465

ciencias. Después defenderé la prioridad de esta meta-metodología no sólo en las ciencias y la filosofía, sino en toda empresa intelectual de investigación. Concluyo que si la filosofía pretende ser una empresa intelectual de investigación, entonces debe orientarse según esta meta-metodología; es decir, debe ser naturalista.

Palabras Clave: *Naturalismo, naturalismo metodológico, fisicalismo, metodología científica, metodología filosófica, metafilosofía, intersubjetividad, compromiso ontológico*

A Minimalist Naturalism

Abstract

In this paper, I propose a minimalist characterization of naturalism as the theses that: 1) we must have a minimal, but not total, commitment to the ontology by our sciences, and 2) we must have a *meta-methodological* commitment to the use of methods that allow for justifications and claims that can be *intersubjectively evaluated*, such as formal and informal arguments, or the use of the methods from the sciences themselves. Thus, I will first criticize the classical definitions of naturalism understood as a form of physicalism and as a restriction on the positing of supernatural entities and explanations in the sciences. Then I'll defend the priority of this meta-methodology not only in the sciences and philosophy, but in every intellectual enterprise of inquiry. I conclude that if philosophy aims to be an intellectual enterprise of inquiry, then it must be oriented according to this meta-methodology; that is, it must be naturalist.

Key Words: *Naturalism, methodologic naturalism, physicalism, scientific methodology, philosophical methodology, metaphilosophy, intersubjectivity, ontological commitment*

1. Introducción

Aunque, según Jaegwon Kim, el naturalismo es *incuestionablemente* la ideología de la filosofía analítica actual (2003, p. 84), la tarea de definir y demarcar puntualmente al naturalismo ha resultado históricamente titánica; tanto así que recientemente se ha argumentado que esta ambigüedad es suficiente para abandonar al naturalismo como una postura (o un término en su totalidad) filosóficamente relevante (Raleigh, 2024).

Mi propósito en este artículo será combatir esta conclusión y ofrecer una caracterización mínima y laxa del naturalismo que considero razonable y defendible bajo pocos supuestos y asunciones. Por tanto, este *naturalismo minimalista* no será una postura ni una tesis novedosa, y, como mostraré adelante, se encuentra presente implícitamente en varias formulaciones del naturalismo a lo largo de la literatura. Así, mi ambición será argumentar que esta caracterización presenta los compromisos y metodologías mínimas que debería presentar toda filosofía, y empresa de investigación en general, que pretenda ser epistémicamente valiosa. Es decir, toda aquella filosofía que *nos gustaría hacer*.

En una primera instancia parecería que para poder defender este punto de forma rigurosa hace falta una demarcación bien definida de lo que constituye la filosofía y una delimitación clara de sus objetivos. Creo que esta tarea no es sólo excesivamente complicada sino fútil. Mi interés no será en precisar las líneas que demarcan la filosofía de la no-filosofía, sino las líneas que demarcan la filosofía naturalista de la no-naturalista, y por qué la primera constituye una empresa de investigación más confiable, rigurosa, prometedora y mejor sustentada que la segunda.

Para los propósitos de este artículo habré de entender el *naturalismo* según la definición estándar ofrecida por Papineau (2023): como la postura compuesta, por una parte, de un naturalismo ontológico (entendido, en líneas generales, como alguna variación del fisicalismo) y, por otra, de un naturalismo metodológico orientado al *quehacer filosófico*

(entendido, según la clásica interpretación de Quine (1960, 1969), como un continuismo de la filosofía con las ciencias).

Además, consideraré la interpretación del naturalismo metodológico orientado al *quehacer científico*, entendido bien sea como una prohibición, o rechazo pragmático, a la postulación de entidades y explicaciones *sobrenaturales* (Boudry et. al., 2010; Ruse, 1982), o bien sea como una prohibición al uso de justificaciones *sobrenaturales* en nuestras investigaciones científicas y filosóficas (Smith, 2017). Para diferenciarlas, habré de referirme al primero como naturalismo metodológico filosófico y al segundo como naturalismo metodológico científico.

Así, habré de ofrecer una caracterización minimalista de esta definición del naturalismo, junto con una breve defensa tanto de su formulación como de su adopción. Mi ruta será la siguiente: primero haré eco de las críticas notables a las interpretaciones clásicas del naturalismo ontológico y del naturalismo metodológico científico para argumentar en favor de una reformulación de ambas. Bajo esta reformulación, ambas tesis afirmarían, respectivamente, que debemos tener un compromiso con la ontología de nuestras mejores teorías científicas, y que debemos tener un compromiso con el uso de aquellas metodologías en la investigación que permitan la realización de afirmaciones y prácticas justificatorias intersubjetivamente evaluables. En el segundo caso, defenderé que esta reformulación del naturalismo metodológico científico puede, y debería, expandirse a todas nuestras empresas de investigación, proponiendo que el uso de estas metodologías puede entenderse como un criterio de demarcación para las ciencias. En particular, defenderé que la filosofía debe seguir también estas metodologías, implicando el continuismo entre la filosofía y las ciencias defendido por el naturalismo metodológico filosófico.

2. Un Naturalismo Minimalista

2.1. Reinterpretando el naturalismo ontológico

¿Por qué hacemos filosofía? Quizá la mejor forma de abordar tal pregunta es formularla de forma más general: ¿por qué hacemos investigaciones de cualquier índole?

Pareciera que la fuerza motriz de cualquiera de nuestras investigaciones es una curiosidad particular por dar una interpretación y explicación de los fenómenos tanto cotidianos como extraordinarios que rodean nuestra experiencia en el mundo. Estos fenómenos pueden tratar de objetos y relaciones lo suficientemente dispares como para motivarnos, al menos históricamente, a separar nuestras investigaciones en áreas y disciplinas locales con sus debidos compromisos y metodologías. Aunque esto, por supuesto, no evita que intersecccionen y se informen las unas de las otras para expandir sus métodos y capacidad explicativa.

No será mi intención dar una exhaustiva genealogía de la empresa intelectual de la humanidad. Para mis propósitos únicamente será relevante notar que tal fuerza motriz, y sus tribulaciones a lo largo de la historia, han parecido converger en una variedad de disciplinas que consideramos muy bien asentadas para el objetivo de ciertas investigaciones: las ciencias.

Parece ser consenso universal que, al menos restringidas a sus particulares dominios de investigación, las ciencias son nuestra mejor fuente de explicaciones e interpretaciones de los fenómenos presentes en sus dominios. No sólo las investigaciones en tales dominios tienden a orientarse según los resultados y métodos de la, o las, ciencias particulares encargadas de su estudio, sino que además tenemos un muy buen motivo para creer que las ciencias se aproximan a describir *verdaderamente* lo que ocurre detrás de tales fenómenos de interés: su éxito predictivo. Este es el famoso *argumento de no milagros* ofrecido por Putnam (1975). Según él, la tesis de que nuestras mejores teorías científicas son

aproximadamente ciertas, en el sentido de que "las entidades por ellas postuladas, o, en todo caso, entidades muy similares a las postuladas, habitan el mundo" (Psillos, 2005, p. 385),[3] es la única tesis bajo la cual el éxito predictivo de tales teorías no es milagroso. En palabras de Putnam: "si tales objetos [postulados por nuestras teorías] no existen realmente, entonces es un *milagro* que una teoría [...] prediga fenómenos exitosamente" (Putnam, 1978, p. 18).

Bajo esta primera consideración parece razonable pensar que si queremos buscar explicaciones sobre los fenómenos de los dominios tratados por las ciencias entonces tales ciencias son nuestra mejor apuesta para hallar tales explicaciones, pues la predicción de nuevos fenómenos en estos dominios parece inexplicable y arbitrariamente fortuita si nuestras explicaciones no son, en gran medida, *verdaderas*. Y si hemos de comprometernos con tales explicaciones, parece razonable también comprometernos con los objetos, relaciones y propiedades de los que tales explicaciones tratan. De lo contrario, ¿estas explicaciones son explicaciones de qué exactamente?

Bajo esta breve defensa podemos empezar a bosquejar una forma de naturalismo ontológico *laxa* y *mínima*. Mi principal propósito será alejarme del fisicalismo clásicamente concebido con el cual se tiende a relacionar al naturalismo ontológico. Con *clásicamente concebido* me refiero a la interpretación del fisicalismo como un compromiso ontológico únicamente con las entidades de la física, o un compromiso ontológico únicamente con entidades físicas y aquellas entidades que dependen, o están metafísicamente constituidas, por entidades físicas.[4] En forma de slogan: todo *es*, o *depende*, de lo *físico*.

[3] La traducción de esta cita, así como el resto de citas textuales de este artículo, es mía, por lo que omito mencionarlo en adelante.

[4] Esta dependencia o constitución metafísica puede entenderse de varias formas, sea bajo la noción de *superveniencia* (Davidson, 2001; Horgan, 1982; Pettit, 1993), la noción de

La principal preocupación con el fisicalismo así concebido es la falta de claridad con respecto al significado de *físico*. Si entendido en referencia a nuestras teorías físicas, este fisicalismo corre el riesgo del dilema de Hempel (1969): si adoptamos la ontología de nuestras teorías físicas actuales, el fisicalismo es falso, pues estas teorías no son *completas* (i.e. no incluyen a todas las entidades que consideramos existentes, así como otras que seguramente existen y desconocemos todavía), y si por otro lado adoptamos la ontología de una futura teoría física ideal, entonces el fisicalismo es simplemente trivial, pues, ¿quién puede predecir lo que la física del futuro pueda contener? Si entendida la fisicalidad en términos negativos, i.e. en oposición a entidades o propiedades mentales (Spurrett & Papineau, 1999; Wilson, 2006), el uso de esta *vía negativa* como defensa del fisicalismo en el *argumento inductivo* de Montero & Papineau (2005) (según el cual todos los efectos físicos son debidos a causas físicas, por tanto todo aquello con efectos físicos debe ser físico en sí mismo) tiene claros contraejemplos: políticas sociales, valores normativos o teorías científicas (Bishop, 2010). Y, más preocupante, es posible concebir ciertas propiedades tanto físicas como mentales y otras ni físicas ni mentales (Stoljar 2024, Secc. 4.5; también véase las críticas de Gillet & Witmer, 2001).

Otra solución propuesta al dilema de Hempel ha sido reinterpretar el fisicalismo no como una tesis evaluable sino como una *actitud*. En particular, como un juramento: "por la presente juro ir en mi ontología a todas partes y sólo a aquellas donde la física me lleve" (Ney, 2008, p. 9). Sin embargo, como nota Peels (2017, pp. 11-12), toda actitud *racional* puede ser reformulada como una tesis; como la tesis de que *deberíamos* tener tal actitud o que es *permisible* tener tal actitud. En este sentido, parece que la solución de Ney sigue siendo víctima todavía del dilema de Hempel (Spiegel, 2023; Stoljar 2024, Secc. 4.6).

grounding (Schaffer, 2009; Rosen, 2010) o la noción de realizabilidad de propiedades (Melnyk, 2003).

Otras críticas, como Dove (2018), se han centrado en que la interpretación actitudinal no logra distanciarse lo suficiente del dualismo, una de las características más clásicas del fisicalismo. Si una teoría física futura llegara a admitir propiedades mentales irreducibles, entonces todo fisicalista actitudinal sería también dualista. Así Ney, intentando recuperar esta oposición entre fisicalismo y dualismo, argumenta que, aunque en tal caso el fisicalista actitudinal y el dualista compartirían ontología, aun así el fisicalismo sería *falso* por haber llegado a tal ontología mucho más tarde que el dualista, así manteniendo su oposición. Pero esto parece una crítica débil al fisicalista y, además, el dualista pudo *sólo haber tenido suerte* (Dove, 2018, p. 517). La propia caracterización de Dove como un programa de investigación el cual "busca ofrecer una explicación positiva y empíricamente informada de cómo lo físico da origen o explica lo mental" (Dove, 2018, p. 519) está también diseñada de tal forma que, el día en que las propiedades mentales irreducibles sean parte de nuestras teorías científicas, entonces tal programa de investigación habría fallado por definición.

Esta rivalidad con el dualismo es, para mí, una segunda motivación para considerar abandonar el fisicalismo como constituyente del naturalismo. Y esto no es porque, secretamente, adhiera a un dualismo de sustancias o de propiedades. Mi propósito no es abandonar el fisicalismo, pero sí despejar al naturalismo de comprometerse con tesis restrictivas acerca de *lo que hay.* Así, considero que si realmente llegase el día en que nuestras mejores teorías científicas apoyaran la existencia de sustancias o propiedades mentales irreducibles, como por ejemplo argumenta Cucu (2023), el fisicalismo será, siguiendo su tradición, efectivamente falso. Pero el naturalismo no será falso, en tanto ser naturalista implica orientar, ante todo, las investigaciones propias según los resultados y métodos de nuestras mejores investigaciones. Es en este sentido que considero pertinente alejar al naturalismo ontológico de un fisicalismo clásicamente concebido, y formularlo en su lugar como un compromiso ontológico

mínimo, mas no necesariamente total, con la ontología de nuestras ciencias actuales.

Si ciertamente es razonable considerar a las ciencias como nuestra mejor fuente de explicaciones para los fenómenos de sus respectivos dominios, argumenté entonces que es necesario el compromiso con los objetos, relaciones y propiedades tratados por tales explicaciones. En este sentido considero defendible un compromiso con la ontología de nuestras mejores teorías científicas. Por un lado, considero pertinente incluir la totalidad de la ontología necesaria para preservar el poder explicativo de nuestras ciencias. Por otro, siguiendo el primer cuerno del dilema de Hempel, considero pertinente también considerar la incompletitud de nuestras ciencias actuales. Es bajo esta perspectiva que concedo la adición de otras entidades, relaciones y propiedades en nuestra ontología, además de aquellas con las que uno *debe* tener un compromiso mínimo para heredar el poder explicativo y éxito predictivo de nuestras mejores teorías científicas.

Esta visión del naturalismo ontológico parece encajar mucho más con el *naturalismo epistemológico* definido por Oppy: "los naturalistas no aceptan ninguna entidad causal, propiedad causal o explicación causal que no sea aceptada por la ciencia bien establecida" (Oppy, 2020, p. 9). Pero es quizá este compromiso epistemológico, como argumenta Collins (2015), la única y mejor apuesta para el naturalista en lugar de, como vimos anteriormente, compromisos particulares con entidades *naturales* o *físicas*. Tanona (2010), por ejemplo, también sugiere que "*lo natural* está caracterizado por desiderata epistemológica de intersubjetividad y realizar predicciones" (p. 82) después de una similar revisión a los intentos anteriores en la literatura.

¿Pero cómo lidiar con otras entidades, relaciones y propiedades (ERyP de aquí en adelante) que podríamos considerar agregar a nuestra ontología y que no forman parte de la ontología de nuestras ciencias? Las ERyP macroscópicas que son parte de nuestras observaciones cotidianas,

por ejemplo, son una primera clase que nos gustaría incluir en nuestra ontología y que no figuran en nuestras teorías científicas (e.g. no hay ninguna ciencia con un compromiso ontológico explícito con las mesas o sillas). Esta primera clase parece fácil de acomodar según nuestro compromiso ontológico mínimo en tanto, siguiendo la fuerza motriz mencionada al inicio, nuestras ciencias buscan, con las explicaciones de sus ERyP especiales, dar explicaciones de fenómenos más generales considerados compuestos o constituidos por estas ERyP especiales (e.g. explicaciones de fenómenos microscópicos que por su parte explican fenómenos macroscópicos formados por ellos). Así que podemos incluir en nuestra ontología también aquellas ERyP que parecen *bien relacionadas* con las ERyP de nuestras ciencias.[5]

¿Pero qué ocurre con aquellas ERyP con las cuales no es tan clara la existencia de una relación con las ERyP de nuestras ciencias, o con aquellos fenómenos que consideramos no estar bien explicados por las explicaciones de otros fenómenos en nuestras ciencias? Aquí es donde nos interesa, como naturalistas, determinar bien nuestras metodologías de investigación.

2.2. Reinterpretando el naturalismo metodológico científico

Por un lado, el naturalismo metodológico científico se ha entendido clásicamente como una restricción con respecto a la investigación científica, bajo la cual "en la medida en que uno hace ciencia, uno evita toda referencia teológica o religiosa" (Ruse, 2005, p. 46) o toda referencia

[5] Por *bien relacionadas* entiendo que estas ERyP estén constituidas o dependan (véase nota anterior) por las ERyP de nuestras ciencias. Esta constitución o dependencia debería ser verificada según prácticas justificatorias intersubjetivamente evaluables, o, aun si no verificada directamente, debería ser adoptada por inferencia a la mejor explicación en tanto esta constitución y dependencia explica de mejor manera sus fenómenos. Si ninguna de ambas condiciones se cumple, considero que las ERyP especiales no están *bien relacionadas* con las ERyP de las ciencias.

a entidades *sobrenaturales* en general. Una primera preocupación para esta caracterización es la de definir puntualmente a qué nos referimos con *sobrenatural*. Hay, al menos, dos vertientes bien conocidas: por un lado en oposición al *mundo natural* y las *entidades naturales* (Oppy, 2020; Stroud, 1996), y por otro, en línea con la *vía negativa*, en oposición a mentes incorpóreas y agentes *inmateriales* (Carrier, 2007, 2025; Fales, 2013). Ambas, como vimos, parecen retornar a la caracterización clásica del fisicalismo y cargan, de igual manera, los mismos problemas que el fisicalismo.

Fuera de esta preocupación inicial, quizá la mayor motivación para buscar una reformulación del naturalismo metodológico científico es que ninguna de sus dos defensas principales parecen ser del todo correctas. Por un lado, las defensas *intrínsecas* afirman que *por definición* la ciencia debe dejar de lado toda apelación a entes y explicaciones sobrenaturales (Haught, 2004; Miller, 2009; Ruse, 1982), o que debe dejarlas de lado en tanto las tesis sobrenaturales son inverificables e infalsables (Pennock, 2001; Scott, 1995). La primera defensa es problemática pues ignora "el hecho de que científicos reputados y escépticos hayan investigado fenómenos paranormales" (Boudry et al., 2010, p. 230). Esta defensa intrínseca implicaría que tales investigaciones son, por definición, fútiles y fuera del ámbito científico. Por otra parte, la segunda defensa intrínseca es falsa: podemos considerar el contraejemplo de Sober (2007) de la hipótesis del *diseño inteligente púrpura*, según la cual un ser omnipotente tendría como prioridad principal colorear absolutamente todo el universo de púrpura, una hipótesis claramente verificable y falsable.[6]

En respuesta a estos problemas, se ha propuesto una defensa *pragmática* según la cual la ciencia *sí* que puede tratar y juzgar hipótesis y entidades sobrenaturales, sin embargo, la historia de tales intentos ha sido lo

[6] Remito al lector a Boudry et al. (2010) y Fishman (2009) para una robusta y exhaustiva crítica a las varias defensas intrínsecas que se han presentado al naturalismo metodológico científico.

suficientemente infructífera como para pragmáticamente descartarlas (Boudry et al., 2010; Coyne, 2009; Ruse, 2005). En este sentido, el naturalismo metodológico científico "es simplemente la mejor guía metodológica que surgió de la historia de la ciencia, en particular, del éxito consistente de las explicaciones naturalistas" (Boudry et al., 2010, pp. 229–230).

Sin embargo, no parece que esta defensa pragmática sea del todo correcta tampoco, en tanto "la afirmación de que la ciencia, a lo largo de los siglos, ha acabado adoptando el naturalismo metodológico como reacción al fracaso de las teorías sobrenaturales es una afirmación que prácticamente no tiene ningún respaldo histórico y mucha evidencia en contra" (Smith, 2017, p. 326). Más bien, una forma particular de naturalismo metodológico científico ha caracterizado y distinguido a la ciencia desde la época medieval, no entendida como una prohibición a entidades sobrenaturales, sino como un rechazo al uso de *métodos sobrenaturales* (tales como los milagros o la fe) y una preferencia al uso de métodos empíricos y racionales. Los científicos, o filósofos naturales, de la Edad Media "tomaron como misión principal la explicación de la estructura y el funcionamiento del mundo en términos puramente racionales y seculares" (Grant, 2010, p. 11).

En este sentido, quizá la mejor forma de entender el naturalismo metodológico científico es no como una limitación con respecto a nuestra ontología, sino como una limitación, como indica su nombre, a los *métodos* aceptados para conducir nuestras investigaciones y *justificar* nuestras afirmaciones.

¿Y qué limitación podría ser esta? En tanto nos interesa entender los fenómenos que nos rodean, nos interesa poder *comprender* claramente las explicaciones que nuestras investigaciones ofrecen de tales fenómenos. No sólo eso, nos interesa, a la luz de nuevos fenómenos e investigaciones, o de ampliaciones de dominio, poder *comparar* si nuestras explicaciones se mantienen vigentes, si sufren cambios mínimos o mayores, o si deben

ser abandonadas ante circunstancias extraordinarias y reelaborarse en su totalidad. Con esto en mente, aunque "hay en la actualidad gran consenso [...] en que no hay (un único) método científico, y que la práctica científica consiste en comunidades de personas tomando decisiones por razones tanto empíricas como sociales, usando diversos métodos" (Oreskes & Macedo, 2019, p. 55), parece aun así que la pluralidad de métodos usados por la ciencia comparte algo en común: hacer públicas y accesibles sus afirmaciones y justificaciones para poder ser intersubjetivamente evaluadas y revisadas.

Así, el naturalismo metodológico científico defendería este estándar de evaluabilidad intersubjetiva en nuestras investigaciones. Este estándar no es sólo característico de los métodos de las ciencias, sino que es, en gran medida, el principal fundamento detrás de su asentamiento como nuestras fuentes primarias de conocimiento, descripción, explicación y predicción de fenómenos y ERyP.

Turunen et al. (2023), haciendo eco de esta observación, consideran que este estándar, el cual llaman el *meta-método de evaluabilidad epistémica*, sería suficiente como un criterio de demarcación para las ciencias. En tanto es un estándar razonable al que aspirar para satisfacer la fuerza motriz de nuestras investigaciones, entonces el hecho de que las ciencias cumplan con este meta-método es un motivo suficiente para ser *cientistas epistemológicos*, es decir, para pensar que las ciencias nos ofrecen la única, o la mejor, fuente de conocimiento, descripción, explicación y predicción de fenómenos y ERyP.[7]

Entonces, aunque podamos encontrarnos bajo la presencia de ERyP lo suficientemente distantes de las ERyP de nuestras conocidas teorías científicas como para dudar de su reducción y explicación mediante ellas, aun así podemos guiar nuestras investigaciones con respecto a estas ERyP

[7] Estrictamente, el cientismo epistemológico se entiende como, al menos, una familia de cuatro posturas que surgen de la división de cientismo *fuerte* y *débil* y de cientismo *amplio* y *estrecho,* véase Hietanen et al. (2020).

excepcionales ocupando la principal meta-metodología que todas nuestras mejores investigaciones realizan: el uso de afirmaciones y prácticas justificatorias intersubjetivamente evaluables.

De momento, he defendido una reformulación del naturalismo ontológico y el naturalismo metodológico científico, y he argumentado en favor de la adopción de ambos. ¿Pero dónde queda exactamente la filosofía en todo esto?

Es aquí, finalmente, donde como naturalistas nos interesa la metodología de nuestras propias labores filosóficas. Aunque, según nuestras consideraciones previas, quizá ya hemos bosquejado una solución a este problema.

2.3. ¿Por qué deberíamos ser naturalistas?

En esta sección argumentaré que si la filosofía ha de ser una empresa intelectual que realice investigaciones cualesquiera, entonces es razonable orientar sus esfuerzos según el estándar de la meta-metodología que caracteriza nuestras mejores investigaciones: realizar afirmaciones y prácticas justificatorias intersubjetivamente evaluables.

Esta será mi formulación del naturalismo metodológico filosófico, el cual no es nada distante de su clásica interpretación como un continuismo entre la filosofía y la ciencia. Así, mi defensa en favor del naturalismo metodológico filosófico será simplemente la defensa de este naturalismo minimalista en su totalidad: toda empresa intelectual que pretenda realizar investigaciones exitosas epistémicamente, deberá adherirse al meta-método de la evaluabilidad epistémica.

Pero también creo que podemos ir más allá y llevar la tesis del continuismo entre filosofía y ciencia a su expresión más radical: el continuismo entre ambas es trivial, pues la filosofía *puede* ser en sí misma una ciencia.

Defenderé, primero, esta tesis radical en función de la tesis de Turunen et al. (2023) de proponer el meta-método de evaluabilidad epistémica como un criterio de demarcación para las ciencias, tesis secundaria que habré de defender usando dos vías.

Mi primera vía de defensa es que subsume varios de los mayores intentos de demarcación de las ciencias. Por ejemplo, los intentos de demarcar a las ciencias a través de la metodología, sea por un único criterio (Popper, 1959) o por múltiples (Bunge, 1982; Fasce, 2017; Langmuir, 1989; Mahner, 2013; Thagard, 1988), se subsumen en tanto, como mencioné anteriormente, parece que todas las metodologías usadas por las ciencias parecen seguir el estándar de presentar afirmaciones y prácticas justificatorias intersubjetivamente evaluables.

Por otro lado, esta evaluabilidad epistémica permite varios procesos sociales y estructuras institucionales (como consenso, autorreflexividad crítica, revisión por pares, diversidad de perspectivas, mitigación de sesgos, interrogación transformativa, etc.) que autores como Ladyman et al. (2007), Longino (1990) y Oreskes & Macedo (2019) proponen como demarcación de las ciencias. También garantiza, priorizando la visibilidad, la accesibilidad pública y la valoración intersubjetiva, las cuatro normas que Merton (1973) clásicamente usa como criterio de demarcación: el *universalismo* (las afirmaciones deben ser sujetas a criterios preestablecidos e impersonales), el *comunalismo* (los resultados de las ciencias, en tanto productos de colaboración, no pertenecen a particulares), la *imparcialidad* (una mitigación institucional de las idiosincrasias e ideologías particulares de los científicos que puedan afectar la investigación) y el *escepticismo organizado* (un escrutinio imparcial de las creencias sostenidas por otras instituciones).[8]

[8] Aunque en la práctica no se vean reflejadas muchas, o ninguna, de las cuatro normas Mertonianas (por ejemplo, en investigación comercial las empresas pueden llegar a fabricar o falsificar datos, manipular conceptos científicos así como diseños de investigación para forzar ciertos resultados, o financiar selectivamente ciertas

Mi segunda vía de defensa es el hecho de que este criterio de demarcación ofrece, a mi parecer, la más parsimoniosa y simple visión de la labor científica. Si la ciencia se define como toda aquella empresa de investigación que realice afirmaciones y prácticas justificatorias evaluables intersubjetivamente, entonces por ciencia hemos de entender, simplemente, todas nuestras *mejores investigaciones disponibles*. En tanto las prácticas justificatorias son evaluables, son constantemente refinadas según nuestros varios subsecuentes tribunales de experiencia, y su accesibilidad pública e intersubjetiva permite el consenso entre ellas para orientar de mejor forma la investigación. Refinadas y mejoradas nuestras prácticas justificatorias, las afirmaciones realizadas por tales investigaciones se asientan con mayor vigor en tanto nuestras justificaciones de ellas se vuelven más robustas.

Y siendo este meta-método característico de la labor científica desde la ciencia aristotélica, la ciencia de la Edad Media, y nuestra ciencia moderna, esta demarcación también conecta nuestras empresas intelectuales contemporáneas con las empresas intelectuales de nuestros antecesores, con todos aquellos hombros de gigantes, y errores de antaño, sobre los cuales se erigen nuestros más brillantes resultados.

Creo que el propio Quine sintetiza bien mi consideración en ambas vías: "a cada hombre le es dada una herencia científica, junto a un bombardeo continuo de estímulos sensoriales; y las consideraciones que lo guían a la hora de deformar su herencia científica para que se ajuste a sus continuos estímulos sensoriales son, si racionales, pragmáticas" (Quine, 1951, p.

investigaciones externas (Doucet & Sismondo, 2008; Elliot, 2016)), aun así los científicos mantienen una gran afinidad por ellas como normas a las que el resto de científicos deberían adherirse. Tan así, que son apoyadas por más del 75% de científicos según la encuesta de Anderson et al. (2010), y más del 90% apoyan en particular la norma del comunalismo en la encuesta de Macfarlane & Cheng (2008). Sobra decir que las normas Mertonianas, aunque un ideal científico, son un ideal deseable por la vasta mayoría de científicos, y las cuatro, en especial el comunalismo, son garantizadas y priorizadas por el meta-método.

43). Esta larga herencia científica se ve, por una parte, unificada bajo el meta-método que ha perseguido desde sus raíces, meta-método el cual permite que este "bombardeo continuo de estímulos" pueda incorporarse continuamente tanto a nuestras prácticas justificatorias como a nuestras afirmaciones, verificándolas, revisándolas, enmendándolas, y, cuando necesario, abandonándolas.

Creo que no es controversial considerar que este meta-método es una condición necesaria, o al menos un ideal deseable, de nuestras investigaciones científicas. ¿Pero será un criterio suficiente? Por supuesto, existe el riesgo de contar dentro demasiadas disciplinas como científicas usando este criterio. En tanto considero, junto con Turunen et al. (2023), que los argumentos formales e informales cuentan como prácticas justificatorias intersubjetivamente evaluables, ¿podría la teología, usando argumentos formales (como el famoso argumento ontológico modal de Gödel), considerarse una ciencia? ¿Qué tal investigaciones cotidianas sobre cuestiones banales pero bien argumentadas?[9] Una teoría que cumpla el meta-método, pero sea *falsa*, también sería científica.

En algunos de estos casos, aunque chocantes según nuestra concepción clásica de *ciencia*, podríamos ceder y morder la bala. Si por ciencia hemos de entender nuestras mejores investigaciones sobre un cierto dominio, y si efectivamente la teología o tales investigaciones banales fueran nuestras mejores investigaciones en sus respectivos dominios, ¿por qué no considerarlas también? Una investigación acerca de términos y conceptos absolutamente banales puede verse informada de varias otras disciplinas afines a su dominio. Una investigación informada por la sociología, la historia o la lingüística puede hacer un recorrido de la historia del término, el cómo, cuándo y por qué de sus cambios, el porqué de su banalidad, etc. Aun si no una investigación tradicionalmente científica,

[9] Agradezco a un revisor anónimo por presionar estos puntos.

parece ser nuestra mejor investigación del tema, aunque sea un tema, aparentemente, irrelevante. A la vez, se considera este meta-método como adoptado por una empresa intelectual de investigación, por lo que afirmaciones aisladas o no sistematizadas, aun con una justificación intersubjetivamente evaluable, no contarían propiamente como ciencia.

Con respecto al segundo y último punto, como afirman Bright & Hessen (2023), la visión de la ciencia como un término que siempre connota *éxito epistémico* debería reconsiderarse seriamente. Las teorías científicas refutadas del pasado, como la mecánica newtoniana o el modelo de Bohr, serían descartadas. Además, ignora el hecho de que el error y el fallo científico es parte crucial de la labor científica. Los resultados nulos, los cuales de hecho deberían ocupar la mayoría de resultados publicados en varias disciplinas (Ioannidis, 2006), son un gran ejemplo. El punto es: la *mala* ciencia, la ciencia no exitosa, o la ciencia banal e irrelevante sigue siendo ciencia.

Sin embargo, aunque en algunos de estos casos ceder únicamente implicaría reconsiderar nuestras concepciones tradicionales de la ciencia, en otros casos podemos correr el riesgo de incluir teorías o disciplinas dañinas. Al final de cuentas, la demarcación entre ciencia y pseudociencia es una cuestión con varias implicaciones serias en políticas climáticas, salud pública, educación, o en testimonios legales (Hansson, 2025). Creo que, en estos casos, vale más la pena hacer una demarcación de la pseudociencia propiamente. Algunos casos notables pueden servirnos como guía: la homeopatía, la teoría del diseño inteligente y el negacionismo del cambio climático antropogénico. Las tres poseen un alto nivel de sistematización: tienen instituciones, libros de texto, conferencias, revistas revisadas por pares, y ofrecen hipótesis claras, apoyadas por argumentos, evidencia y referencia a otros trabajos previos en su cuerpo de literatura (Oreskes, 2019). Como empresas intelectuales de investigación, son casi indistinguibles de las ciencias naturales, y utilizan sus mismos métodos. Pareciera que, siguiendo nuestro criterio de

demarcación, serían catalogadas inmediatamente como ciencias. Pero aquí podemos hacer varios matices importantes.

A priori, ninguna de estas empresas tiene nada malo. Si efectivamente sus hipótesis llegaran a demostrarse ciertas (si efectivamente las diluciones sucesivas fueran tratamientos exitosos, o si efectivamente nuestras evidencias apuntaran a un *Dios de los vacíos*, o si efectivamente el cambio climático no tuviera causas antropogénicas) nuestras ciencias deberían ser las primeras interesadas. ¿Entonces cuál es el problema? Como reconoce Oreskes (2019, p. 898), tanto la teoría del diseño inteligente como la homeopatía son actividades que buscan conocimiento de forma honesta, postulando hipótesis interesantes para la ciencia, y usando varios de sus métodos institucionales para autentificarse. El problema radica en que consistentemente nuestras evidencias e investigaciones no han apoyado a ninguna.

En el caso de la homeopatía, para la mayoría de enfermedades la evidencia es insuficiente para juzgar su efectividad, y en otras no resulta más efectiva que un placebo (National Health and Medical Research Council, 2014). Lo mismo puede decirse de los negacionistas del cambio climático antropogénico: sus principales afirmaciones no son apoyadas por la evidencia de nuestras mejores investigaciones (United Nations Environmental Program, 2024), y es un consenso muy bien establecido en la literatura que el cambio climático está causado por el humano (Lynas et al., 2021). En el caso de la teoría del diseño inteligente, se ha argumentado que descansa en las premisas irrevocables de la existencia de Dios y su rol en el diseño del universo, lo que la vuelve no falsable y no revisable ante nueva evidencia (Albert, 1986, p. 26; Oreskes, 2019, p. 899). Esto haría que su hipótesis principal no sea evaluable, por lo que no satisface del todo el meta-método. Además, fenómenos como la convergencia evolutiva parecen ofrecer evidencia bastante convincente contra las principales tesis del diseño inteligente (Greener, 2007, p. 1108).

El hecho de que las tres sistematicen sus hipótesis y ocupen el meta-método para justificarlas precisamente las acerca bastante a la labor científica. Muchos artículos publicados en la revista *Homeopathy,* editada por Elsevier, ocupan los mismos métodos que aquellos publicados en cualquier revista de medicina tradicional. Como investigaciones aisladas, estas no son problemáticas, y podrían perfectamente ser parte de nuestro cuerpo de investigaciones médicas. O considérese el conocido artículo de Vardiman (2013): "Numerical Simulations of Winter Storms, Tropical Cyclones, and Nor'easters During the Ice Age Using the NCAR URF Model with a Warm Ocean". Este artículo, con un título indistinguible de un artículo cualquiera sobre ciencias climáticas, argumenta que "el aumento en la temperatura de la superficie del mar después del diluvio universal del Génesis fue causa más probable de la Edad de Hielo", usando el conocido modelo Weather Research and Forecasting (WRF) del Centro Nacional para la Investigación Atmosférica estadounidense (NCAR). Nuevamente, este artículo individualmente parece cumplir el meta-método, y, aunque entretenga una conclusión aparentemente ridícula, no parece tener motivos para no ser considerado seriamente también dentro del cuerpo de la literatura de las ciencias climáticas.

Lo que vuelve pseudocientíficas a cada una de las empresas mencionadas es el hecho de que sus principales hipótesis no son soportadas por nuestro grueso de evidencias. Ignorar la evidencia en contra de estas hipótesis, repetir ciertas afirmaciones ya refutadas u obsoletas, representar engañosamente la información científica, las teorías, los modelos, u oscurecer la evidencia para aparentar una falta de consenso donde no la hay son todas las faltas epistémicas que vuelven pseudocientíficas a estas disciplinas. Así, aunque prima facie parezca que el meta-método puede incluir ciertas empresas peligrosamente similares a la labor científica (lo que Oreskes (2019) llama *ciencia facsímil*), una investigación detallada nos lleva a observar que constituyen programas de investigación que sostienen afirmaciones refutadas e hipótesis no respaldadas por nuestras mejores investigaciones. Estos programas pueden ser descartados. Así,

podemos incluir esta heurística a la hora de proponer el meta-método de evaluabilidad epistémica como criterio de demarcación.

Para otros casos, tampoco parece complicado heredar las nociones de progreso teórico y empírico de Lakatos: podemos decir que un programa de investigación es *teóricamente progresivo* si cada nueva teoría propuesta por el programa produce nuevas predicciones y es *empíricamente progresivo* si algunas de estas nuevas predicciones son corroboradas empíricamente. Una cierta empresa de investigación es entonces *progresiva*, en general, si es tanto progresiva teóricamente como empíricamente, y es *degenerativa* si no es progresiva (1970, pp. 33–34). Por supuesto, este criterio de progresividad únicamente aplicaría en empresas de investigación donde la formación de predicciones es posible, verificable, y forma parte integral de la ciencia en particular. Para otros casos, nuestras consideraciones previas acerca de la obstinación en hipótesis no apoyadas por la evidencia, y el resto de malas praxis mencionadas, son una heurística más útil.

Hacer una defensa más detallada del meta-método como criterio de demarcación nos llevaría muy lejos del alcance de este artículo. Si las defensas aquí presentadas son mínimamente convincentes, entonces la tesis de que la filosofía que ocupa este meta-método puede considerarse como una ciencia es al menos plausible. Si concedida, entonces la postura del naturalismo metodológico de una continuidad entre la ciencia y la filosofía se vuelve trivial: la filosofía que sigue este meta-método puede considerarse como una ciencia en sí misma.

Pero aún si esta conclusión radical no me fuera concedida, creo que de todas formas puede concederse fácilmente la tesis original del naturalismo metodológico filosófico: "la filosofía debería […] 'estar informada' por los métodos y las afirmaciones de la ciencia […] cuando una afirmación filosófica *bien-justificada* y una afirmación científica *bien-justificada* son inconsistentes […] la afirmación científica gana sobre la filosófica" (Flanagan, 2009, p. 430).

Esta versión más laxa será la defendida, principalmente, por el naturalismo metodológico filosófico aquí caracterizado. En este sentido, argumentaré que adoptar este meta-método es un criterio *necesario* para toda empresa de investigación que pretenda tener éxito epistémico.

Por una parte, como mencioné anteriormente, la gran mayoría (si no es que todos) los métodos usados por las ciencias parecen compartir el principio de hacer accesibles sus resultados y prácticas justificatorias para ser evaluadas, principalmente por otros pares. Adoptar el meta-método deja a nuestra disposición todos estos métodos evaluables epistémicamente para su uso en el estudio de dominios afines, donde se ha demostrado su eficacia. También, como también mencioné antes, permiten varias de las estructuras institucionales que caracterizan el *rigor* y *objetividad* de las ciencias, entre ellas: la revisión por pares, la autorreflexividad crítica, la interrogación transformativa, la mitigación de sesgos, la replicación de estudios, la falsabilidad, la capacidad de prueba, la revisión ante adición de datos, el análisis crítico de las metodologías usadas, la realización de predicciones, etc. Además, garantizan la priorización de las normas Mertonianas, las cuales son ideales deseables para la gran mayoría de científicos (Anderson et al., 2010; Macfarlane & Cheng, 2008, véase nota 6).

El meta-método, en tanto antepone la constante evaluabilidad intersubjetiva, permite también la existencia de consenso entre expertos del área. Siendo la ciencia una "empresa colectiva de descripciones, predicciones y explicaciones basadas en datos [...] el consenso científico entre expertos tiene máxima autoridad" (Oppy, 2020, p. 12).

Creo que estas consideraciones constituyen una decente muestra de los beneficios epistémicos de la adopción del meta-método de evaluabilidad epistémica. Teniéndolos en cuenta, creo queda claro el por qué nuestras investigaciones deberían priorizar, como nuestras mejores ciencias han priorizado ya, este meta-método. Si tengo razón en este punto, la filosofía

por su parte debería adherirse también a él. ¿Pero de qué forma puede la filosofía seguir este meta-método?

Por un lado, el uso de argumentos formales e informales es una práctica justificatoria que cumple con el criterio, siempre y cuando las premisas y reglas de inferencia usadas también tengan una justificación previa que pueda evaluarse. Naturalmente, en la construcción de estos argumentos, la filosofía puede necesitar del auxilio de otras ciencias para la formación de premisas e hipótesis, usando de estas sus resultados y prácticas justificatorias. El argumento y sus premisas se presentan de forma evaluable y revisable, permitiendo así los beneficios epistémicos que conlleva esta evaluabilidad.

Por otro lado, es importante notar que, precisamente por su robustez explicativa y predictiva, muchas ciencias se auxilian las unas a las otras, ofreciendo sus ERyP y métodos particulares. Esta interdisciplinariedad es otra característica familiar de las ciencias, y es muy alentada por el meta-método. Si una, o varias,[10] empresas de investigación han estudiado ya, usando el meta-método, ciertos objetos o dominios afines a los que pretendemos estudiar con nuestra empresa de investigación, entonces es más que natural considerar la adopción de sus ERyP, sus resultados y sus

[10] Sobre un mismo dominio pueden haber varias empresas de investigación compitiendo. En ocasiones, ciertas disciplinas pueden triunfar sobre otras no necesariamente por sus méritos epistémicos sino por cuestiones contingentes como factores históricos o ciertos prejuicios. El meta-método, enfatizando la importancia de la evaluabilidad y revisabilidad epistémica, puede ofrecernos un marco para intentar despejar estas contingencias. Como mencioné anteriormente con los casos de la homeopatía, el diseño inteligente y el negacionismo del cambio climático, el meta-método puede considerar seriamente también algunas de estas investigaciones, siempre y cuando, como empresas o programas de investigación, no incurran en las faltas epistémicas ya mencionadas. El meta-método permite, y prioriza, una pluralidad de visiones y marcos para realizar investigaciones, siempre y cuando sean evaluables y revisables intersubjetivamente. Así, podemos descartar ciertas empresas por motivos epistémicos, y delimitar mejor qué disciplinas e investigaciones considerar a la hora de trabajar en un cierto dominio.

prácticas justificatorias para poder heredar la gama de explicaciones exitosas epistémicamente sobre el área de interés. De esta forma, la adherencia de la filosofía, o una empresa de investigación cualquiera, a la meta-metodología científica implica de por sí un compromiso ontológico mínimo con las ERyP de las ciencias, y con los métodos usados por estas en dominios similares a los estudiados por ciertas disciplinas filosóficas. Así, toda filosofía que ocupe los resultados y métodos de la ciencia en dominios afines, i.e. toda filosofía informada científicamente, o continuista con la ciencia, satisface también el meta-método.

Esta observación también nos da una segunda línea argumentativa a favor la adopción de esta meta-metodología: ofrece, como corolario, las reformulaciones del naturalismo ontológico y el naturalismo metodológico científico ya mencionadas y defendidas anteriormente.

Esa será, entonces, mi caracterización minimalista del naturalismo:

Naturalismo Minimalista: el compromiso de realizar toda empresa intelectual de investigación según el meta-método de evaluabilidad epistémica, implicando un compromiso mínimo con la ontología de las ciencias para preservar su eficacia explicativa y predictiva.

3. Conclusión

A lo largo de este artículo he defendido una nueva caracterización *minimalista* del naturalismo, criticando en primera instancia las definiciones clásicas del naturalismo ontológico y el naturalismo metodológico científico orientadas como restricciones e imposiciones *sobre lo que hay*. Concluí que ambos componentes son mejor caracterizados como compromisos, por una parte, con la ontología de nuestras mejores teorías científicas, y, por otra, con el uso de métodos que permitan la evaluabilidad epistémica de sus afirmaciones y prácticas justificatorias.

Defendí también una reinterpretación del naturalismo metodológico filosófico como la tesis según la cual la filosofía, y toda empresa de investigación, debería adherirse a esta meta-metodología. Mostré su potencial como un criterio de demarcación para las ciencias, dando al continuismo entre la filosofía y la ciencia su expresión más radical. Independientemente de su adopción como criterio de demarcación, argumenté que esta meta-metodología ofrece varios beneficios epistémicos para nuestras investigaciones, garantizando también muchos ideales científicos bien aceptados entre la comunidad.

Finalmente, ofrecí mi definición del naturalismo minimalista como la conjunción de estos tres compromisos: un compromiso con la ontología de las ciencias, un compromiso con el meta-método de evaluabilidad epistémica en las ciencias, y un compromiso con el mismo meta-método en la filosofía y en cualquier otro programa de investigación.

Como mencioné al inicio, mi propósito aquí no es demarcar a la filosofía. La filosofía es una disciplina demasiado vasta y amplia. No dudo que haya filosofía despreocupada por ofrecer conocimiento. Quizá podría objetarse que la filosofía no es un medio sino un fin en sí mismo, a pesar de que pueda ofrecernos ocasionalmente resultados útiles e instrumentales. En ese caso, diría que la ciencia también puede afirmarse como un fin en sí mismo, a pesar de sus aplicaciones. Creo que pocos científicos abandonarían el impulso de seguir investigando aún si estas investigaciones no ofrecieran nada más que satisfacer nuestra curiosidad.

En este sentido, tampoco dudo en absoluto que haya filosofía preocupada por ofrecer conocimiento. No dudo que exista filosofía que pretenda ser una empresa intelectual de investigación exitosa en clarificar, interpretar, o explicar de forma aproximadamente correcta sus fenómenos de interés. Tampoco creo estar demasiado errado al creer que, de hecho, la mayoría de la filosofía pretende cumplir esta ambición.

De ser así, si la filosofía tiene el objetivo de satisfacer nuestra incansable curiosidad por el mundo, y de ser exitosa epistémicamente al hacerlo,

entonces afirmo que *debería* ser naturalista. Porque ser naturalista implica solamente comprometernos con todo aquello que nos dicen nuestras mejores investigaciones, con realizar afirmaciones que puedan ser evaluadas y revisadas intersubjetivamente, y con el uso de prácticas justificatorias que permitan un verdadero progreso en la investigación.

¿No es aquella la filosofía que nos gustaría hacer?

REFERENCIAS

Albert, L. H. (1986). "Scientific" Creationism as a Pseudoscience. *Creation/Evolution Journal*, *6*(2), 25–34. https://ncse.ngo/files/pub/CEJ/pdfs/CEJ_18.pdf

Anderson, M. S., Ronning, E. A., De Vries, R. & Martinson, B. C. (2010). Extending the Mertonian Norms: Scientists' Subscription to Norms of Research. *The Journal of Higher Education*, *81*(3), 366–393. https://dx.doi.org/10.1353/jhe.0.0095

Bishop, R. (2010). The Via Negativa: Not the Way to Physicalism. *Mind and Matter*, *8*(2), 203–214.

Boudry, M., Blancke, S. & Braeckman, J. (2010). How not to attack Intelligent Design Creationism: Philosophical misconceptions about methodological naturalism. *Foundations of Science*, *15*(3), 227–244. https://doi.org/10.1007/s10699-010-9178-7

Bright, L. K., & Heesen, R. (2023). To Be Scientific Is To Be Communist. *Social Epistemology*, *37*(3), 249–258. https://doi.org/10.1080/02691728.2022.2156308

Bunge, M. (1982). Demarcating Science from Pseudoscience. *Fundamenta Scientiae*, *3*(3-4), 369–388.

Carrier, R. (2007, January 18). *Defining the Supernatural.* Richard Carrier Blogs: https://richardcarrier.blogspot.com/2007/01/defining-supernatural.html

Carrier, R. (2025, February 7). *Defining Naturalism: The Definitive Account*. Richard Carrier Blogs: https://www.richardcarrier.info/archives/33004

Collins, J. (2015) Naturalism without metaphysics. In E. Fischer & J. Collins (eds.), *Experimental Philosophy, Rationalism, and Naturalism* (pp. 85–109). Routledge.

Coyne, J. A. (2009, February 4). *Seeing and Believing.* The New Republic: https://newrepublic.com/article/63388/seeing-and-believing

Cucu, A. C. (2023). Turning the Tables: How Neuroscience Supports Interactive Dualism. *Mind and Matter*, *21*(2), 219–239. https://doi.org/10.53765/MM2023.219

Davidson, D. (2001). The Material Mind. In *Essays on Actions and Events* (pp. 245–260). Oxford University Press. https://doi.org/10.1093/0199246270.003.0013

Doucet, M., & Sismondo, S. (2008). Evaluating Solutions to Sponsorship Bias. *Journal of Medical Ethics*, *34*(8), 627–630. https://doi.org/10.1136/jme.2007.022467

Dove, G. (2018). Redefining Physicalism. *Topoi*, *37*(3), 513–522. https://doi.org/10.1007/s11245-016-9405-0

Elliott, K. C. (2016). Environment. In A. J. Angulo (ed.), *Miseducation: A History of Ignorance-Making in America and Abroad* (pp. 96–119). Johns Hopkins University Press. https://doi.org/10.1353/book.72117

Fales, E. (2013). Is a Science of the Supernatural Possible? In M. Pigliucci & M. Boudry (eds.), *Philosophy of Pseudoscience: Reconsidering the Demarcation Problem* (pp. 247–262). University of Chicago Press. https://doi.org/10.7208/9780226051826-014

Fasce, A. (2017). What Do We Mean When We Speak of Pseudoscience? The Development of a Demarcation Criterion Based on the Analysis of Twenty-one Previous Attempts. *Disputatio. Philosophical Research Bulletin*, *6*(7), 459–488. https://doi.org/10.5281/zenodo.1433737

Fishman, Y. I. (2009). Can science test supernatural worldviews? *Science & Education 18*, 813–837. https://doi.org/10.1007/s11191-007-9108-4

Flanagan, O. (2009). Varieties of naturalism. In P. Clayton & Z. Simpson (eds.), *The Oxford Handbook of Religion and Science* (pp. 430–452). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199543656.003.0027

Gillett C. & Witmer G. (2001): A 'physical' need: Physicalism and the via negativa. *Analysis*, *61*, 302–308. http://dx.doi.org/10.1093/analys/61.4.302

Grant, E. (2010). *The nature of natural philosophy in the late middle ages*. The Catholic University of America Press. https://doi.org/10.2307/j.ctt284vbb

Greener M. (2007). Taking on creationism. Which arguments and evidence counter pseudoscience? *EMBO Reports*, *8*(12), 1107–1109. https://doi.org/10.1038/sj.embor.7401131

Hansson, S. O. (2025, July 28). Science and Pseudoscience. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2025/entries/pseudo-science/

Haught, J. F. (2004). Darwin, design, and divine providence. In M. Ruse & W. A. Dembski (eds.), *Debating Design: From Darwin to DNA* (pp. 229–245). Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511804823.013

Hempel, C. (1969). Reduction: ontological and linguistic facets. In W. Morgenbesser (ed.), *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel* (179–199). St. Martin's Press.

Hietanen, J., Turunen, P., Hirvonen, I., Karisto, J., Pättiniemi, I. & Saarinen, H. (2020). *How Not to Criticise Scientism. Metaphilosophy*, *51*(4), 522–547. https://doi.org/10.1111/meta.12443

Horgan, T. (1982). Supervenience and Microphysics. *Pacific Philosophical Quarterly*, *63*, 29–43. http://dx.doi.org/10.1111/j.1468-0114.1982.tb00084.x

Ioannidis, J. P. (2006). Journals Should Publish All "Null" Results and Should Sparingly Publish "Positive" Results. *Cancer Epidemiology Biomarkers & Prevention*, *15*(1), 186. https://doi.org/10.1158/1055-9965.EPI-05-0921

Kim, J. (2003). The American Origins of Philosophical Naturalism. Journal of Philosophical *Research 28*(Suppl), 83–98. https://doi.org/10.5840/jpr200328supplement28

Ladyman, J., Ross, D., Spurrett, D., & Collier, J, (2007). *Every Thing Must Go: Metaphysics Naturalized.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199276196.001.0001

Lakatos, I. (1978). Falsification and the methodology of scientific research programmes. In J. Worrall & G. Currie (eds.), *The Methodology of Scientific Research Programmes: Philosophical Papers Volume I* (pp. 8–101). Cambridge University Press. https://doi.org/10.1017/CBO9780511621123.003

Langmuir, I. (1989). Pathological Science. *Physics Today*, *42*(10), 36–48. https://doi.org/10.1063/1.881205

Longino, H. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry.* Princeton University Press. http://dx.doi.org/10.1515/9780691209753

Lynas, M., Houlton, B. Z., & Perry, S. (2021). Greater than 99% consensus on human caused climate change in the peer-reviewed

scientific literature. *Environmental Research Letters*, *16*(11), 114005. https://doi.org/10.1088/1748-9326/ac2966

Macfarlane, B., & Cheng, M. (2008). Communism, Universalism and Disinterestedness: Re-Examining Contemporary Support Among Academics for Merton's Scientific Norms. *Journal of Academic Ethics*, *6*(1), 67–78. https://doi.org/10.1007/s10805-008-9055-y

Mahner, M. (2013). Science and Pseudoscience How to Demarcate after the (Alleged) Demise. In M. Pigliucci & M. Boudry (eds.), *Philosophy of Pseudoscience: Reconsidering the Demarcation* Problem (29–44). University of Chicago Press. https://doi.org/10.7208/chicago/9780226051826.003.0003

Melnyk, A. (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism.* Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511498817

Merton, R. K. (1973). The Normative Structure of Science. In N. W. Storer (ed.), *The Sociology of Science: Theoretical and Empirical Investigations*? (pp. 267–278). University of Chicago Press.

Miller, K. B. (2009). The misguided attack on methodological naturalism. In J. S. Schneiderman & W. D. Allmon (eds.), *For the Rock Record: Geologists on Intelligent Design* (pp. 117–140). University of California Press. http://dx.doi.org/10.1525/california/9780520257580.003.0008

Montero, B. & Papineau, D. (2015). Naturalism and Physicalism. En K. J. Clark (ed.), *The Blackwell Companion to Naturalism* (pp. 182–195). Wiley-Blackwell. https://doi.org/10.1002/9781118657775.ch13

National Health and Medical Research Council. (2014, April). Evidence On The Effectiveness Of Homeopathy For Treating Health Conditions. *National Health and Medical Research Council.* https://www.hri-research.org/wp-content/uploads/2014/07/nhmrcdrafthomeopathyinformationpaper140408.pdf

Ney, A. (2008). Physicalism as an attitude. *Philosophical Studies*, *138*(1), 1–15. https://doi.org/10.1007/s11098-006-0006-4

Oppy, G. (2020). Naturalism. *Think*, *19*(56), 7–20. https://doi.org/10.1017/s1477175620000172

Oreskes, N., & Macedo, S. (2019). *Why trust science?* Princeton University Press. https://doi.org/10.1515/9780691189932

Papineau, D. (2020, March 31). Naturalism. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2023/entries/naturalism/

Peels, R. (2017). Ten reasons to embrace scientism. *Studies in History and Philosophy of Science Part A*, *63*, 11–21. https://doi.org/10.1016/j.shpsa.2017.04.001

Pennock, R. (2001). Naturalism, evidence, and creationism. En R. Pennock (ed.), *Intelligent design creationism and its critics* (pp. 77–98). MIT Press.

Pettit, P. (1993). A Definition of Physicalism. *Analysis*, *53*(4), 213–223. http://dx.doi.org/10.1093/analys/53.4.213

Popper, Karl. (1959). *The logic of scientific discovery.* New York: Routledge.

Psillos, S. (2005). Scientific realism and metaphysics. *Ratio*, *18*(4), 385–404. http://dx.doi.org/10.1111/j.1467-9329.2005.00301.x

Putnam, H. (1975). What is mathematical truth? *Historia Mathematica*, *2*(4), 529–533. https://doi.org/10.1016/0315-0860(75)90116-0

Putnam, Hilary. (1978). *Meaning and the Moral Sciences.* Routledge and Kegan Paul.

Quine, W. V. O. (1951). Two Dogmas of Empiricism. *Philosophical Review*, *60*(1), 20–43. https://doi.org/10.2307/2181906

Quine, Willard van Orman. (1960). *Word & Object*. MIT Press.

Quine, W. V. O. (1969). Epistemology Naturalized. In W. V. O. Quine (ed.), *Ontological Relativity and Other Essays* (pp. 69–90). Columbia University Press. http://dx.doi.org/10.7312/quin92204-004

Raleigh, T. (2024). The Emptiness of Naturalism. *Philosophy*, *99*(4), 597–623. https://doi.org/10.1017/s0031819124000172

Rosen, G. (2010). Metaphysical Dependence: Grounding and Reduction. En B. Hale & A. Hoffman (eds.), *Modality: Metaphysics, Logic, and Epistemology* (pp. 109–136). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199565818.003.0007

Ruse, Michael. (1982). *Darwinism defended: a guide to the evolution controversies.* Addison-Wesley.

Ruse, M. (2005). Methodological Naturalism Under Attack. *South African Journal of Philosophy*, *24*(1), 44–60. https://doi.org/10.4314/sajpem.v24i1.31413

Schaffer, J. (2009). On What Grounds What. En R. Wasserman, D. Manley & D. Chalmers (eds.), *Metametaphysics: New essays on the foundations of ontology* (pp. 347–383). Oxford University Press. http://dx.doi.org/10.1093/oso/9780199546046.003.0012

Scott, E. (1995). Creationism, ideology, and science. *Annals of the New York Academy of Sciences*, *775*(1), 505–522. http://dx.doi.org/10.1111/j.1749-6632.1996.tb23167.x

Smith, T. (2017). Methodological naturalism and its misconceptions. *International Journal for Philosophy of Religion*, *82*(3), 321–336. https://doi.org/10.1007/s11153-017-9616-3

Sober, E. (2007). What is wrong with intelligent design? *The Quarterly Review of Biology*, *82*(1), 3–8. http://dx.doi.org/10.1086/511656

Spiegel, T. J. (2023). Why Naturalism cannot (Merely) be an Attitude. *Topoi*, *42*(3), 745–752. https://doi.org/10.1007/s11245-022-09846-6

Spurrett, D. & Papineau, D. (1999). A note on the completeness of "physics". *Analysis*, *59*(1), 25–29. http://dx.doi.org/10.1093/analys/59.1.25

Stoljar, D. (2024, May 25). Physicalism. *The Stanford Encyclopedia of Philosophy* https://plato.stanford.edu/archives/spr2024/entries/physicalism/

Stroud, B. (1996). The Charm of Naturalism. *Proceedings and Addresses of the American Philosophical Association*, *70*(2), 43–55. https://doi.org/10.2307/3131038

Tanona, S. (2010). The pursuit of the natural. *Philosophical Studies*, *148*(1), 79–87. https://doi.org/10.1007/s11098-010-9497-0

Thagard, Paul. (1988). *Computational Philosophy of Science.* MIT Press. http://dx.doi.org/10.7551/mitpress/1968.001.0001

Turunen, P., Hirvonen, I. & Pättiniemi, I. (2023). Epistemological scientism and the scientific meta-method. *European Journal for Philosophy of Science*, *13*(2), 26. https://doi.org/10.1007/s13194-023-00526-9

United Nations Environmental Program. (2024, June 4). Debunking eight common myths about climate change. *United Nations Environmental Program.* https://www.unep.org/news-and-stories/story/debunking-eight-common-myths-about-climate-change

Vardiman, L. (2013, September 12). Numerical Simulations of Winter Storms, Tropical Cyclones, and Nor'easters During the Ice Age Using the NCAR WRF Model With a Warm Ocean. In M. Horstemeyer (ed.), *Proceedings of the Seventh International Conference on Creationism.* https://www.icr.org/i/pdf/technical/Numerical-Simulations-Storms-Cyclones-Noreasters-During-Ice-Age.pdf

Wilson, J. (2006). On characterizing the physical. *Philosophical Studies*, *131*(1), 61–99. https://doi.org/10.1007/s11098-006-5984-8

HOMÓNIMA
Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 112-127
ISSN 3087-260X
Santiago de Chile

ONTOLOGÍA DEL COLOR Y TESIS DE LA CONTINUIDAD:

Relacionalismo adverbial como fusión de estilos de razonamiento[1]

Rodolfo Aldea[2]
Universidad Alberto Hurtado

Resumen

Este artículo propone una lectura metafilosófica del debate ontológico sobre el color, centrada en los "estilos de razonamiento" (Hacking, 1992) que subyacen al conflicto histórico entre Helmholtz y Hering. Frente al estancamiento entre objetivismo, eliminativismo y relacionalismo, se argumenta que el problema no radica tanto en la localización ontológica del color cuanto en una tensión metodológica entre un estilo fisicalista —que prioriza la explicación causal externa—

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Aldea, R. (2026). Ontología del color y tesis de la continuidad: relacionalismo adverbial como fusión de estilos de razonamiento. *Revista de Filosofía Homónima, 1*(1), pp. 112-127. https://doi.org/xxxxxxx[Por asignar]

[2] Rodolfo Aldea es Doctor en Filosofía de la Universidad Alberto Hurtado (UAH). Miembro de: La Asociación Latinoamericana de Filosofía Analítica (ALFAn); Red Latinoamericana en Filosofía y Psiquiatría (RLFP); La Sociedad Chilena de Filosofía de las Ciencias (SOCHIFIC); Grupo de Investigación Santiago Mind and Cognition (SM&C). Email: roaldea@uahurtado.cl. ORCID: 0009-0002-7506-8504

y uno fenomenológico —que parte de la estructura interna de la experiencia vivida. Inspirado en la tesis de la continuidad de Quine (1957), se sostiene que esta tensión puede resolverse al concebir la ciencia perceptiva no como sustituta del sentido común, sino como su extensión crítica y sistemática. En este marco, el relacionalismo adverbial emerge como la articulación ontológica más coherente de dicha continuidad, redefiniendo el color como un modo interactivo de aparición perceptual. Esta estrategia no solo disuelve el problema ontológico tradicional, sino que invita a una metafilosofía naturalizada que integra estilos opuestos en una metodología unificada, con implicaciones para dominios afines como la ontología de la emoción.

Palabras clave: *metafilosofía; estilos de razonamiento; ontología del color; conflicto Helmholtz-Hering; tesis de continuidad; relacionalismo adverbial.*

Ontology of color and the continuity tesis

Abstract

This article proposes a metaphilosophical reading of the ontological debate on color, centered on the "styles of reasoning" (Hacking, 1992) underlying the historical conflict between Helmholtz and Hering. Faced with the stalemate between objectivism, eliminativism, and relationalism, it is argued that the problem lies less in the ontological location of color than in a methodological tension between a physicalist style—prioritizing external causal explanation—and a phenomenological one—starting from the internal structure of lived experience. Inspired by Quine's "thesis of continuity" (1957), it is maintained that this tension can be resolved by conceiving perceptual science not as a substitute for common sense, but as its critical and systematic extension. Within this framework, adverbial relationalism

emerges as the most coherent ontological articulation of such continuity, redefining color as an interactive mode of perceptual appearance. This strategy not only dissolves the traditional ontological problem but also invites a naturalized metaphilosophy that integrates opposing styles into a unified methodology, with implications for related domains such as the ontology of emotion.

Keywords: *metaphilosophy; styles of reasoning; color ontology; Helmholtz-Hering conflict; continuity thesis; adverbial relationalism.*

1. Introducción

El debate ontológico sobre el color se ha caracterizado por una persistente oscilación entre posiciones que intentan localizar esta cualidad sensible en el mundo objetivo, en la mente subjetiva o en alguna relación intermedia. Sin embargo, tras décadas de refinamientos teóricos —desde el objetivismo fisicalista de Byrne y Hilbert (2003, 2021) hasta las defensas primitivistas de Allen (2016) o Gert (2017)— el campo parece atrapado en un punto muerto, donde cada teoría enfrenta objeciones que la remiten invariablemente a sus rivales. Sostengo que este impasse no es meramente accidental, sino el síntoma de una tensión metafilosófica más profunda: un choque entre estilos de razonamiento opuestos que estructuran implícitamente el modo en que abordamos la percepción cromática. Siguiendo a Hacking (1992, pp. 1-4), estos estilos representan marcos conceptuales históricos que determinan qué cuenta como evidencia válida y cómo se construyen las explicaciones, manifestándose en el legado del conflicto entre Helmholtz y Hering.

Helmholtz representaba un estilo fisicalista y causal-externo que priorizaba la explicación de la percepción cromática como un proceso inferencial inconsciente a partir de estímulos físicos externos, tratando el ojo como un instrumento pasivo que reconstruye el mundo objetivo mediante mecanismos causales reductivos. Por el contrario, Hering encarnaba un estilo fenomenológico y estructural-interno que tomaba la

experiencia vivida como dato primordial, centrándose en la organización oposicional interna de los fenómenos cromáticos (como los pares rojo-verde y azul-amarillo) y considerando que dicha estructura revela propiedades fundamentales de la visión sin necesidad de subordinarla inmediatamente a causas físicas externas.

Esta diferencia no era solo técnica o empírica, sino que reflejaba estilos de razonamiento distintos en el sentido de Hacking (1992): mientras el helmholtziano considera evidencia legítima principalmente los datos físicos medibles y las cadenas causales externas, subordinando la apariencia fenomenológica a procesos reductivos impersonales, el heringiano valida como punto de partida epistemológico la descripción rigurosa y sistemática de la estructura interna de la experiencia, sin requerir su traducción inmediata a términos físicos. Tal conflicto de estilos persiste en las teorías contemporáneas del color, como ha analizado recientemente Gutiérrez (2021, pp. 40-50).

Ahora bien, sostengo que esta tensión no es irresoluble. La hipótesis central de este trabajo es que puede superarse mediante una metodología inspirada en la *tesis de la continuidad* de Quine (1957, p. 5), la cual postula que la ciencia no reemplaza al sentido común, sino que lo extiende de manera crítica y sistemática. Bajo esta perspectiva metafilosófica, el enfoque fenomenológico de Hering no se opone al fisicalismo helmholtziano, sino que lo complementa como una forma legítima de refinamiento perceptivo, tal como sugieren análisis recientes sobre estilos de razonamiento en la visión del color (Pelucchi et al., 2024). De este modo, el relacionalismo adverbial —particularmente en la versión de Chirimuuta (2015, p. 139, 2020)— se presenta no como una teoría ontológica adicional, sino como la articulación natural de esta continuidad metodológica, al redefinir el color como un modo de aparición interactiva en lugar de una propiedad estática.

Para desarrollar esta propuesta, el artículo se organiza de la siguiente manera. La Sección 2 examinará en detalle el choque de estilos de razonamiento, ilustrando cómo las teorías contemporáneas del color

heredan los legados de Helmholtz y Hering. La Sección 3 introducirá la tesis de la continuidad quineana como herramienta metafilosófica para mediar esta tensión, argumentando que el método de Hering ejemplifica una extensión sistemática del sentido común. La Sección 4 defenderá el relacionalismo adverbial como la culminación ontológica de este enfoque continuista. Finalmente, la Sección 5 ofrecerá conclusiones sobre las implicancias de esta metodología para la filosofía de la percepción en general.

2. El Conflicto de las Imágenes y los Estilos de Razonamiento en la Percepción

El estancamiento en la filosofía del color puede interpretarse como un caso paradigmático de la tensión que Wilfrid Sellars (1963, pp. 5-10) describió entre la *imagen manifiesta* —el marco conceptual cotidiano poblado por cualidades sensibles como los colores— y la *imagen científica* —la teorización de un mundo de procesos causales impersonales. Sin embargo, la naturaleza exacta de esta imagen manifiesta, especialmente respecto al color, dista de ser unívoca; estudios en filosofía experimental han comenzado a cuestionar si las creencias intuitivas sobre el color son tan objetivistas como se presupone tradicionalmente (Roberts & Schmidtke, 2019; Adams & Hansen, 2020). En contraste, la imagen científica reduce el color a fenómenos físicos como reflectancias espectrales o longitudes de onda, priorizando explicaciones causales que a menudo marginan la experiencia vivida.

Siguiendo a Hacking (1992, p. 19) esta brecha entre imágenes no es solo descriptiva, sino que refleja un choque fundamental entre estilos de razonamiento, esto es, marcos históricos que definen qué se considera evidencia legítima y cómo se construyen las hipótesis. En el contexto de la percepción del color, este choque se remonta al siglo XIX con las figuras de Hermann von Helmholtz y Ewald Hering. Helmholtz, representante de un estilo fisicalista, abordaba la visión como un proceso inferencial donde el ojo actúa como un instrumento pasivo que recibe

estímulos luminosos, reconstruyendo el mundo externo a través de mecanismos causales inconscientes. Su enfoque, heredero de la tradición newtoniana, prioriza la explicación reductiva: el color se reduce a propiedades físicas como longitudes de onda o reflectancias, integrando la percepción en un marco científico impersonal (Byrne & Hilbert, 2021; Sharp, 2023).

En oposición, Hering defendía un estilo fenomenológico que parte de la experiencia vivida como dato primordial, enfatizando la estructura interna de la percepción cromática —como los pares oponentes rojo-verde y azul-amarillo— sin subordinarla inmediatamente a causas físicas externas. Este enfoque no rechaza la ciencia, sino que la reorienta hacia un análisis descriptivo riguroso de la apariencia, cuestionando la primacía de las explicaciones causales sobre la fenomenología (Conway et al., 2023). Como ha argumentado Gutiérrez (2021, p. 52), esta controversia no fue meramente técnica, sino un conflicto profundo de estilos de razonamiento que persiste en las teorías contemporáneas del color. Así, el objetivismo fisicalista (e.g., Jackson, 2019; Byrne & Hilbert, 2021) hereda el legado helmholtziano al privilegiar como evidencia legítima las propiedades físicas medibles (como reflectancias espectrales) y buscar una explicación reductiva que subsume la apariencia cromática bajo cadenas causales externas impersonales. Por su parte, versiones primitivistas o relacionalistas (e.g., Gert, 2017; Chirimuuta, 2020) resuenan con el estilo heringiano al tomar la estructura interna de la experiencia vivida como dato epistemológico primordial, defendiendo su irreductibilidad y cuestionando la necesidad de reducirla inmediatamente a términos físicos.

La persistencia de este conflicto en la ontología perceptual contemporánea no solo ilustra la durabilidad de estos estilos de razonamiento, sino que también revela una dimensión metafilosófica crucial, a saber: *la forma en que nuestros métodos filosóficos predeterminan las posibles soluciones*. Por ejemplo, defensas recientes del objetivismo fisicalista, como las de Jackson (2019) o Watkins y Shech (2025), continúan privilegiando un estilo helmholtziano al buscar reducir el color

a propiedades conmensurables como reflectancias espectrales, asumiendo que la imagen científica debe subsumir por completo la manifiesta.

En cambio, enfoques como el de Chirimuuta (2024, p. 98), con su giro neopragmatista, incorporan implícitamente elementos heringianos al enfatizar la irreductibilidad de la experiencia y cuestionar metafilosóficamente el estatuto mismo del debate ontológico tradicional. En particular, estos enfoques sugieren que dicho debate —con su oscilación aparentemente ineludible entre objetivismo, eliminativismo y relacionalismo— podría ser en gran medida un artefacto de la adopción exclusiva de estilos de razonamiento reductivos (principalmente helmholtzianos), que presuponen que la única forma legítima de explicación científica consiste en reducir la apariencia fenomenológica a propiedades físicas externas medibles, marginando así dimensiones estructurales y contextuales de la percepción que no encajan en ese marco.

Al considerar estos estilos reductivos como inadecuados —no por falsos, sino por parciales e insuficientemente inclusivos— tales enfoques reconfiguran el problema ontológico, i.e., deja de ser un dilema sustantivo sobre la *localización* del color para convertirse en una tensión metodológica superable mediante la integración de estilos opuestos. Esta dinámica sugiere que resolver el *impasse* requiere una metafilosofía explícita que examine y reconcilie estos estilos de razonamiento, en lugar de perpetuar su oposición.

3. La Tesis de la Continuidad Quineana como Puente Metafilosófico

La clave para superar el conflicto entre estilos de razonamiento radica en adoptar una perspectiva metafilosófica inspirada en la *tesis de la continuidad* de W. V. Quine (1957, p. 5), quien argumentaba que la ciencia no es una ruptura radical con el sentido común, sino su refinamiento sistemático y crítico. En el contexto de la ontología perceptual, esta tesis implica que el estilo fenomenológico de Hering no

contradice el fisicalismo helmholtziano, sino que lo extiende al incorporar un análisis riguroso de la experiencia como base para hipótesis científicas más inclusivas. Lacey (2021), en su exploración neoquineana de la metaontología, refuerza esta idea al proponer que compromisos ontológicos como los del color deben evaluarse en términos de continuidad pragmática, evitando dualismos —y consecuentes conflictos— innecesarios entre lo *manifiesto* y lo *científico*.

Esta reinterpretación quineana no solo ofrece un puente conceptual entre estilos opuestos, sino que también invita a una metafilosofía naturalizada, en la línea de Maddy (2007, pp. 86-91), donde la filosofía examina sus propios métodos como *extensiones del razonamiento ordinario*. Aplicado al color, el estilo de Hering —con su énfasis en la fenomenología descriptiva— se revela como una encarnación de esta continuidad quineana, i.e., lejos de constituir un rechazo subjetivista o una alternativa ontológica irreconciliable con el fisicalismo helmholtziano, representa un refinamiento sistemático del sentido común perceptivo que integra la experiencia vivida como evidencia legítima dentro del marco científico más amplio.

Esta lectura continuista se justifica porque el enfoque de Hering no introduce compromisos ontológicos dualistas ni niega la relevancia de los mecanismos causales externos, sino que los complementa al sistematizar la estructura fenomenológica (e.g., pares oponentes rojo-verde, azul-amarillo) como dato empírico que cualquier teoría científica completa debiera acomodar. Precisamente esta integración se evidencia en modelos matemáticos recientes de percepción cromática (Pelucchi et al., 2024), que incorporan explícitamente las estructuras oponentes heringianas dentro de marcos computacionales y empíricos más amplios, sin reducirlas eliminativamente ni tratarlas como incompatibles con explicaciones físicas causales, sino como refinamientos que enriquecen la predicción y comprensión científica del color. Así, la tesis de continuidad disuelve la aparente oposición al tratar el conflicto Helmholtz-Hering no como un dilema ontológico insoluble, sino como etapas complementarias en una metodología perceptual unificada.

De este modo, la ontología del color deja de ser un problema de localización estática —objetiva o subjetiva— para convertirse en una cuestión de integración pragmática. Siguiendo a Carranante (2020), quien aplica perspectivas quineanas a la ingeniería conceptual en percepción, podemos argumentar que compromisos metafilosóficos como estos permiten redefinir categorías perceptuales sin caer en reduccionismos, fomentando una continuidad que valora tanto la evidencia causal (helmholtziana) como la estructural (heringiana). Esta aproximación no solo resuelve tensiones internas en la filosofía del color, sino que también abre vías para extender el análisis a otros dominios ontológicos, como la emoción, donde patrones similares de estilos opuestos persisten.

En última instancia, esta metodología continuista, al disolver dualismos metafilosóficos entre estilos de razonamiento, permite una reevaluación de la ontología perceptual que prioriza la funcionalidad pragmática sobre la localización absoluta. Inspirados en enfoques neoquineanos como los de Lacey (2021), podemos concebir la percepción del color no como un dominio fragmentado por conflictos históricos, sino como un *continuum* donde el análisis heringiano enriquece el helmholtziano, generando hipótesis más robustas que integran evidencia fenomenológica con mecanismos causales. Esta perspectiva no solo resuelve tensiones internas en la filosofía del color, sino que también establece un marco metafilosófico extensible, donde la continuidad quineana actúa como herramienta crítica para examinar cómo nuestros métodos filosóficos evolucionan a partir del sentido común hacia explicaciones sistemáticas.

4. El Relacionalismo Adverbial como Culminación Ontológica de la Continuidad

Bajo el lente de esta metodología continuista, el relacionalismo adverbial se erige no como una posición ontológica entre muchas, sino como la expresión natural de una metafilosofía que reconcilia estilos opuestos. Siguiendo a Chirimuuta (2015, p. 187, 2020), esta variante del relacionalismo redefine el color no como una propiedad inherente a los

objetos o mentes, sino como un *modo de aparecer* en la interacción perceptual entre observador, objeto y entorno. Esta concepción adverbial —donde el color se describe en términos relacionales como *apareciendo-rojo-para-un-sujeto-en-ciertas-condiciones*— evita los reduccionismos del estilo helmholtziano al incorporar la irreductibilidad fenomenológica heringiana, alineándose con la tesis quineana al tratar la experiencia cromática como una extensión pragmática del sentido común perceptivo. Esta redefinición adverbial no solo resuelve tensiones ontológicas tradicionales —como las variaciones perceptuales sin color objetivo, tal como explora Brown (2022)— sino que también encarna una metafilosofía de la continuidad al fusionar metodológicamente estilos de razonamiento aparentemente irreconciliables. Ahora bien, lejos de constituir un mero reemplazo del estilo helmholtziano por el heringiano, esta fusión conserva elementos centrales del primero —la relevancia explicativa de los mecanismos causales externos (reflectancias espectrales, condiciones de iluminación, procesos neurofisiológicos) como factores que co-determinan la aparición cromática— mientras integra los del segundo —la irreductibilidad estructural de la experiencia fenomenológica (organización oposicional, constancias perceptivas) como dato que modula y enriquece dicha explicación causal—. En el relacionalismo adverbial, el color se concibe como un modo interactivo de aparición (*apareciendo-rojo-de-tal-manera-en-tales-condiciones*), donde los factores físicos helmholtzianos se mantienen como condiciones necesarias, pero se subordinan funcionalmente a una descripción relacional que incorpora la perspectiva heringiana sin reducirla eliminativamente. De este modo, Chirimuuta (2015, 2020, 2024) logra una síntesis continuista que no privilegia un estilo sobre el otro, sino que los unifica en una metodología híbrida más inclusiva.

Dicho de otro modo, en el marco helmholtziano, el color se reduce a causas físicas externas, pero el enfoque adverbial integra esta causalidad con la fenomenología heringiana, tratando la experiencia cromática como un proceso relacional que emerge de interacciones contextuales. Chirimuuta (2024), en su defensa neopragmatista, refuerza esta

integración al argumentar que el color no requiere una "localización" metafísica estática, sino una descripción funcional que extiende el sentido común perceptivo hacia explicaciones científicas, alineándose con la tesis quineana de refinamiento sistemático sin ruptura.

Desde esta perspectiva metafilosófica, el relacionalismo adverbial disuelve el problema ontológico al cuestionar los presupuestos mismos de los estilos opuestos, en la medida en que, en lugar de oponer lo causal a lo estructural, propone una metodología híbrida donde la percepción del color se entiende como un *continuum pragmático* en el que los factores biológicos (e.g., adaptación cromática y sensibilidad individual de los conos), ambientales (e.g., condiciones de iluminación y reflectancias espectrales variables) y culturales (e.g., categorización lingüística y convencional del color, como los estudios sobre variabilidad intercultural en la percepción de azules y verdes) co-determinan el modo interactivo de aparición cromática. Esta aproximación encuentra eco en análisis recientes que aplican marcos continuistas a la percepción, como los de Decock (2018), quien aboga por una metafísica cognitiva quineana que prioriza la integración sobre la reducción. Así, el color deja de ser un enigma ontológico para convertirse en un caso ejemplar de cómo la metafilosofía puede reconciliar legados históricos en una ontología perceptiva más inclusiva y funcional.

Esta culminación ontológica del relacionalismo adverbial, vista a través de una lente metafilosófica, no solo ofrece una resolución al debate sobre el color, sino que también ilustra cómo la metodología continuista puede extenderse a otros desafíos perceptivos. Por ejemplo, en el ámbito de la ontología de la emoción —donde se reproducen tensiones similares entre modelos dimensionales reductivos (Russell, 1980) y enfoques fenomenológicos descriptivos— la aplicación de estilos híbridos inspirados en Quine permite disolver dualismos análogos, priorizando interacciones relacionales sobre propiedades monádicas. Como hemos explorado en trabajos previos (Autor/Autora [información eliminada] y Autor/Autora [información eliminada], en preparación), esta extensión revela una "brecha representacional" en la fenomenología emocional que

puede cerrarse mediante una integración continuista, similar a cómo el adverbialismo reconcilia Helmholtz y Hering en el color.

5. Conclusiones

El debate ontológico sobre el color, lejos de ser un impasse irresoluble, se revela como el síntoma de una tensión metafilosófica profunda entre estilos de razonamiento opuestos, heredados del conflicto Helmholtz-Hering. Hemos argumentado que esta tensión puede diagnosticarse como un choque entre enfoques fisicalistas causales y fenomenológicos estructurales, y que encuentra una vía de superación en la tesis de la continuidad de Quine (1957, p. 5), la cual concibe la ciencia perceptiva como una extensión crítica del sentido común. En este marco, el método de Hering ejemplifica dicha continuidad al guiar hipótesis científicas desde la experiencia vivida, mientras que el relacionalismo adverbial completa esta integración al redefinir el color como un modo relacional de aparición, disolviendo así el problema tradicional de localización.

Las implicancias de este giro metafilosófico trascienden el caso específico del color, ofreciendo un modelo para abordar tensiones similares en otros dominios de la percepción y la ontología. Por ejemplo, en la ontología de la emoción, donde modelos evolutivos o dimensionales (Plutchik, 1984; Russell, 1980) reproducen un estilo helmholtziano reductivo, mientras que tradiciones fenomenológicas buscan estructuras internas heringianas, la metodología continuista quineana permite una integración pragmática que cierra brechas representacionales (Autor/Autora [información eliminada] y Autor/Autora [información eliminada], en preparación). Esta extensión no solo enriquece la filosofía de la percepción al promover estilos híbridos, sino que también subraya cómo una metafilosofía atenta a los métodos históricos puede transformar debates estancados en oportunidades para un refinamiento sistemático del sentido común.

En resumen, al adoptar la tesis de la continuidad como puente metafilosófico, el relacionalismo adverbial no se presenta como una mera alternativa ontológica, sino como la culminación de una metodología que

reconcilia estilos opuestos en una ontología perceptiva funcional y relacional. La pregunta deja de centrarse en "¿dónde reside el color?" para enfocarse en "¿qué rol cumple en nuestra interacción perceptiva y práctica?", abriendo un horizonte donde la filosofía de la percepción se alinea con un pragmatismo naturalizado. Esta perspectiva, enriquecida por análisis recientes como los de van der Merwe (2025) sobre influencias quineanas en el pragmatismo, invita a futuras investigaciones que apliquen esta continuidad a fenómenos perceptivos más amplios, fortaleciendo así el diálogo entre la imagen manifiesta y la científica.

REFERENCIAS

Adams, Z., & Hansen, N. (2020). The myth of the common sense conception of color. In T. Marques & Å. Wikforss (Eds.), *Shifting Concepts: The Philosophy and Psychology of Conceptual Variability* (pp. 106–127). Oxford University Press.

Allen, K. (2016). *A Naïve Realist Theory of Colour*. Oxford University Press.

Brown, D. H. (2022). Colour variation without objective colour. *Philosophy and the Mind Sciences*, 3, 1–31. https://doi.org/10.33735/phimisci.2022.9183

Byrne, A., & Hilbert, D. R. (2003). Color realism and color science. *Behavioral and Brain Sciences*, 26(1), 3–21. https://doi.org/10.1017/S0140525X03000013

Byrne, A., & Hilbert, D. R. (2021). Objectivist reductionism. In D. H. Brown & F. Macpherson (Eds.), *The Routledge Handbook of Philosophy of Colour*. Routledge.

Cappelen, H. (2018). *Fixing Language: An Essay on Conceptual Engineering*. Oxford University Press.

Carranante, G. (2020). The boundaries of perception: A conceptual engineering perspective. *Thèse de doctorat*. Université Paris Sciences et Lettres. https://theses.hal.science/tel-03662720v1/document

Chirimuuta, M. (2015). *Outside Color: Perceptual Science and the Puzzle of Color in Philosophy*. MIT Press.

Chirimuuta, M. (2020). Naturalism and the philosophy of colour ontology and perception. *Philosophy Compass*, 15(2), e12649. https://doi.org/10.1111/phc3.12649

Chirimuuta, M. (2023). Neopragmatism in the Philosophy of Perception? The Case of Primitive Color. In J. Gert (Ed.), *Neopragmatism: Interventions in First-Order Philosophy*. Oxford University Press. https://doi.org/10.1093/oso/9780192894809.003.0004

Conway, B. R., Malik-Moraleda, S., & Gibson, E. (2023). Color appearance and the end of Hering's Opponent-Colors Theory. *Trends in Cognitive Sciences*, 27(9), 791–804. https://doi.org/10.1016/j.tics.2023.06.003

Davies, W. (2020). Colour relations in form. *Philosophy and Phenomenological Research*, 102(3), 574–594. https://doi.org/10.1111/phpr.12679

Decock, L. (2018). Cognitive Metaphysics. *Frontiers in Psychology*, 9, Article 1700. https://doi.org/10.3389/fpsyg.2018.01700

Gert, J. (2017). *Primitive Colors: A Case Study in Neo-Pragmatist Metaphysics and Philosophy of Perception*. Oxford University Press.

Gutiérrez, J. (2021). Hermann von Helmholtz, Ewald Hering and color vision: A controversy over styles of reasoning? *Manuscrito: Revista Internacional de Filosofia*, 44(1), 37–97. https://doi.org/10.1590/0100-6045.2021.V44N1.JG

Hacking, I. (1992). 'Style' for historians and philosophers. *Studies in History and Philosophy of Science Part A*, 23(1), 1–20. https://doi.org/10.1016/0039-3681(92)90024-Z

Jackson, F. (2019). How to be an objectivist about colour. *Phenomenology and the Cognitive Sciences*, 18, 819–831. https://doi.org/10.1007/s11097-017-9551-x

Jackson, F. (2020). Learning from what color experiences are good for. *Harvard Review of Philosophy*, 27, 49–58. https://doi.org/10.5840/harvardreview202072529

Lacey, M. V. (2021). Neo-Quinean and neo-Aristotelian metaontology. *Doctoral Thesis*. Lingnan University. https://commons.ln.edu.hk/cgi/viewcontent.cgi?article=1139&context=otd

Maddy, P. (2007). Second Philosophy: A Naturalistic Method. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199273669.001.0001

Pelucchi, F., Berthier, M., & Provenzi, E. (2024). On the Philosophical Standpoint of a Recent Mathematical Color Perception Model. *Foundations of Science*. https://doi.org/10.1007/s10699-024-09951-2

Plutchik, R. (1984). Emotions: A general psychoevolutionary theory. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 197–219). Lawrence Erlbaum Associates.

Quine, W. V. (1957). The Scope and Language of Science. *The British Journal for the Philosophy of Science*, 8(29), 1–17. http://www.jstor.org/stable/685377

Roberts, P., & Schmidtke, K. A. (2019). Folk Core Beliefs about Color. *Review of Philosophy and Psychology*, 10, 849–869. https://doi.org/10.1007/s13164-019-00437-w

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. https://doi.org/10.1037/h0077714

Sellars, W. (1963). Philosophy and the scientific image of man. In Science, perception and reality (pp. 1–40). Routledge & Kegan Paul.

Sharp, W. A. (2025). Spectral reflectances and commensurateness. *Erkenntnis*, 90, 1399-1414. https://doi.org/10.1007/s10670-023-00762-8

van der Merwe, R. (2025). Whewell's fundamental antithesis: A lineage of influence. *South African Journal of Philosophy*, *44*(1), 55–73. https://doi.org/10.1080/02580136.2025.2462434

Watkins, M., & Shech, E. (2025). *The Metaphysics of Color*. Cambridge University Press. https://doi.org/10.1017/9781009324212

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 128-161
ISSN 3087-260X
Santiago de Chile

EL FILÓSOFO DE LA CIENCIA COMO PENSADOR PÚBLICO[1]

Estándares Profesionales Para La Práctica Filosófica

Gabriel Donoso[2]

Resumen

Sostengo que la filosofía de la ciencia recupera su legitimidad pública cuando vincula su progreso interno con formas efectivas de impacto. Para ello, reconstruyo primero la genealogía de su repliegue institucional en el siglo XX y su posterior re-emergencia como una práctica socialmente comprometida. Luego, integro evidencia reciente sobre las vías reales de influencia académica para mostrar que el impacto filosófico es fundamentalmente relacional antes que bibliométrico. Con ese trasfondo, argumento que la métrica interna

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Donoso, G. (2026). El filósofo de la ciencia como pensador público. Estándares profesionales para la práctica filosófica. *Revista de Filosofía Homónima, 1*(1), pp. 128-161. https://doi.org/xxxxxxx[por asignar]

[2] Gabriel Donoso es Profesor de Estado en Filosofía y Magíster en Filosofía de las Ciencias de la Universidad de Santiago de Chile, Santiago, Chile. Email: gabriel.donoso.u@usach.cl. ORCID: https://orcid.org/0000-0002-4842-1387.

más fértil para valorar el progreso filosófico es el aumento de la comprensión estructural: la capacidad de articular mapas de dependencia entre conceptos, fenómenos y prácticas que faciliten coordinación epistémica y decisiones confiables. Finalmente, propongo una ética profesional para el filósofo de la ciencia con cuatro principios operativos —integridad epistémica, colaboración horizontal, responsabilidad social y transparencia institucional—, acompañados de indicadores y métricas evaluables más allá de la citación. El resultado es un marco normativo que ofrece lineamientos institucionales para medir y exigir impacto.

Palabras clave: *filosofía de la ciencia; impacto social; comprensión estructural; confianza epistémica; interdisciplinariedad; ética profesional.*

The philosopher of science as a public thinker: professional standards for philosophical practice

Abstract

I maintain that philosophy of science regains its public legitimacy when it connects its internal progress with effective forms of impact. To that end, I first reconstruct the genealogy of its institutional retreat in the twentieth century and its subsequent re-emergence as a socially engaged practice. I then integrate recent evidence on the actual pathways of academic influence to show that philosophical impact is primarily relational rather than bibliometric. Against that background, I argue that the most fertile internal metric for assessing philosophical progress is an increase in structural understanding: the capacity to articulate maps of dependence among concepts, phenomena, and practices that enable epistemic co-ordination and support reliable decision-making. Finally, I propose a professional ethics for the contemporary philosopher of science built around four operational principles—epistemic integrity, horizontal collaboration, social

responsibility, and institutional transparency—each accompanied by indicators and evaluative metrics beyond citation counts. The result is a normative framework that offers institutional guidance for measuring and demanding impact.

Keywords: *philosophy of science; social impact; structural understanding; epistemic trust; interdisciplinarity; professional ethics.*

1. Introducción

Durante el siglo XX, la profesionalización reorientó la filosofía de la ciencia hacia la especialización técnica, apartándola de las preocupaciones públicas que habían motivado su institucionalización moderna. Si en los años treinta el manifiesto de la Círculo de Viena proyectaba la concepción científica del mundo como una fuerza de transformación social (Asociación Ernst Mach, 2002, p. 124), las décadas posteriores presenciaron un progresivo aislamiento institucional de la disciplina.

Suele subestimarse o ignorarse la dimensión política del Círculo de Viena. Su propuesta no fue meramente un programa teórico intraepistemológico, sino también una actitud práctica y pública mediante la cual el grupo se constituyó como un actor político reconocible, en interlocución con movimientos liberales y socialdemócratas, especialmente durante su "fase pública" entre 1928 y 1936 (Romizi, 2012, pp. 211–216). Ese compromiso, orientado a la transformación social mediante la educación científica, la laicidad y la planificación racional, se plasmó en iniciativas como la Sociedad Ernst Mach y las conferencias populares, con el propósito de intervenir en el

espacio político y oponerse a fuerzas reaccionarias como el clericalismo y el nacionalismo.[3]

A mediados del siglo XX, la filosofía de la ciencia reorientó su autocomprensión. La *Philosophy of Science Association* (PSA, en adelante) modificó en 1959 su misión original —que incluía explícitamente promover prácticas "en beneficio de los científicos y filósofos en particular, y de las personas de buena voluntad en general"— por una puramente intra-académica, centrada en la discusión libre "desde diversos puntos de vista en el campo de la filosofía de la ciencia" (Cartieri & Potochnik, 2014, pp. 902–903). Desde entonces, la disciplina osciló entre la búsqueda del rigor analítico y la pérdida de impacto social.

En las últimas décadas, la disciplina se ha replanteado, en registros diversos, la relación entre ética profesional, prácticas laborales e impacto disciplinar.[4] Entre ellas, la *filosofía de la ciencia socialmente comprometida* (*socially engaged philosophy of science*) ha revitalizado la idea de que la reflexión filosófica puede tener efectos medibles tanto en la práctica científica como en el espacio cívico (Plaisance & Elliott, 2021). Paralelamente, estudios empíricos sobre las vías de influencia de la filosofía han apuntado a que el impacto efectivo se produce primordialmente mediante la interacción personal entre filósofos y científicos, y no solo por la circulación bibliométrica (Plaisance et al., 2021). Esto sugiere que el problema no es una falta de pertinencia, sino la ausencia de una ética profesional capaz de alinear las prácticas internas con las condiciones reales de producción de impacto.

[3] Tal compromiso no ha sido plenamente tematizado en la literatura, en parte por la lectura que redujo al Círculo a neutralidad valorativa y cientificismo apolítico (Richardson 2009a; 2009b; Sorell 1991).

[4] Un ejemplo pionero fue la filosofía feminista, especialmente en su vertiente analítica orientada a la epistemología y filosofía de la ciencia. Desde 1980, esta se ha ocupado de estudiar cómo las normas, estructuras sociales y valores —incluidos los relacionados con el género y las prácticas institucionales— configuran el contenido del conocimiento, así como las prácticas y jerarquías sociales (Anderson, 2024; Crasnow & Intemann, 2020; Grasswick, 2011).

Sostendré que maximizar el impacto filosófico exige concebir la labor del filósofo de la ciencia como una práctica socialmente involucrada. A partir de la literatura reciente sobre el compromiso social de la filosofía de la ciencia y las teorías contemporáneas del progreso filosófico, argumentaré que el núcleo de dicha ética reside en entender el progreso como el aumento de la comprensión estructural y en comprometerse con su traducción a contextos sociales específicos.

Organizo la propuesta en cuatro secciones. Primero, reconstruyo la genealogía de la preocupación por la incidencia de la filosofía (§ 2). Segundo, examino modelos de influencia académica y las tensiones entre prácticas de compromiso y estructuras de incentivo (§ 3). Tercero, reformulo el progreso filosófico como criterio de valor (§ 4). Finalmente, desarrollo los principios normativos de una ética profesional del filósofo de la ciencia (§ 5). En conjunto, sostengo que una filosofía con pretensión de impacto es aquella que hermana la precisión analítica con la responsabilidad social.

2. De la retirada al re-compromiso: genealogía del impacto

2.1. La desvinculación institucional de la filosofía de la ciencia

El retraimiento de la filosofía de la ciencia respecto de la esfera pública no fue accidental, sino el efecto de una redefinición institucional de su identidad. Ello se advierte en los estatutos de la PSA: en 1948 su misión incluía promover la utilidad práctica del pensamiento filosófico y científico "en beneficio de los científicos y filósofos en particular, y de las personas de buena voluntad en general". En 1959, esa orientación fue reemplazada por un objetivo más acotado: "fomentar la libre discusión de los diversos puntos de vista en filosofía de la ciencia". Para 1961, las referencias al beneficio social habían desaparecido por completo (Cartieri & Potochnik, 2014, pp. 902–903).

Esta modificación simbólica condensa un giro más amplio: la filosofía de la ciencia dejó de concebirse como una actividad mediadora entre la práctica científica y los problemas humanos, para entenderse

ahora como un campo especializado en el análisis lógico-conceptual. Siguiendo a Howard (2003, p. 26), es 1959 el punto de inflexión que marca la transición desde una filosofía con vocación pública hacia una comunidad académica crecientemente autorreferente. Tal retraimiento se reforzó, además, con la salida de la PSA de las reuniones anuales de la *American Association for the Advancement of Science* (AAAS) a inicios de los sesenta, foro donde filósofos y científicos compartían paneles sobre problemas de interés público. Desde entonces, la filosofía de la ciencia se institucionalizó como una disciplina universitaria autónoma, con sus propias revistas, congresos y criterios de publicación.[5] Esto supuso un avance profesional, pero acentuó su aislamiento social y político.

La figura del filósofo como mediador externo del conocimiento científico fue impugnada por los propios filósofos, en la medida en que le atribuye una autoridad metalingüística y metodológica para fijar exógenamente el significado y uso de los conceptos (Bricmont, 2015; Feyerabend, 1963; Laudan, 1983). Tal pretensión puede interferir con la dinámica autocorrectiva de las ciencias, cuyo progreso depende en parte de revisar marcos, sustituir teorías y ensayar alternativas desde dentro. Cuando el mediador opera como custodio de un vocabulario estándar o de una "gran imagen" metodológica, tiende a favorecer una uniformidad que protege a la teoría dominante frente a la crítica. En cambio, la confrontación entre programas rivales y la proliferación de alternativas funcionan como recursos internos contra cierres prematuros.

Así, la noción de impacto filosófico se redujo al ámbito interno: la circulación de artículos entre pares y la acumulación de citaciones desplazaron el diálogo con el público y con comunidades científicas. El resultado fue un modelo endogámico de productividad, donde el éxito se mide por visibilidad académica más que por contribución social o epistémica. Esta retirada no obedeció a desinterés moral, sino a incentivos

[5] La primera reunión bienal autónoma de la PSA se realizó del 11 al 13 de octubre de 1968 en Pittsburgh (Douglas, 2010). Desde entonces, la asociación mantiene congresos bienales propios.

institucionales: el temor a la politización durante la Guerra Fría, la expansión del paradigma analítico anglófono y la creciente tecnificación del discurso filosófico (Reisch, 2009, pp. 14–17, 21–23, 455). He ahí la paradoja: una disciplina que esclarece la racionalidad científica sin participar plenamente en las prácticas públicas donde esa racionalidad adquiere sentido.

2.2. Confianza y *expertise* como puente

Un posicionamiento como el anterior arrastra tesis problemáticas. En particular, sugiere que la mediación filosófica equivale a una imposición externa e irreflexiva de significados y normas sobre comunidades científicas determinadas. La literatura contemporánea ofrece, sin embargo, un cuadro distinto: en numerosos casos, los filósofos no operan como árbitros que "dictan" desde fuera, sino como interlocutores técnicamente responsables que, mediante herramientas filosóficas, abordan problemas propiamente científicos y formulan propuestas que los científicos pueden evaluar, discutir o rechazar dentro de los canales ordinarios de la práctica científica (Pradeu et al., 2024). Cuando un debate se enreda por ambigüedades semánticas o por usos cruzados de vocabularios, una distinción filosófica empíricamente neutral puede revelar que parte del desacuerdo era un pseudodebate y reorientar la investigación hacia cuestiones empíricas mejor planteadas y públicamente escrutables.[6]

He aquí la clave. El compromiso público aparece cuando la filosofía de la ciencia recupera su papel mediador en la relación entre conocimiento experto y ciudadanía. Desde inicios de los 2000, la epistemología contemporánea se ha hecho eco de este giro. La confianza

[6] Hay casos donde marcos conceptuales propuestos por filósofos fueron adoptados por científicos y usados para producir resultados y clarificar controversias. Véase la distinción entre replicadores e interactores para ordenar los debates sobre unidades de selección (Pradeu et al., 2024, pp. 386-388). Esto es posible cuando la alfabetización científica del filósofo es acusada. De lo contrario, la pertinencia de su intervención resulta nula (Ladyman et al., 2007, p. 10; Maudlin, 2007, p. 1).

epistémica es condición de posibilidad de la práctica científica: sin deferencia racional hacia comunidades especializadas no hay acumulación de conocimiento, transferencia tecnológica ni orientación informada de la acción colectiva (Whyte & Crease, 2010, p. 413). El aporte filosófico, entonces, no se agota en la crítica abstracta de conceptos, sino que consiste en articular normativamente los vínculos de credibilidad que sostienen la cooperación entre expertos y no expertos. Esto desplaza la pregunta guía. La pregunta guía deja de ser únicamente "¿qué es una explicación científica?" para ser, además: "¿cuándo es racional y moralmente justificable deferir a *X* en *Y* bajo condiciones *Z*?". La respuesta exige criterios explícitos y auditables.

La filosofía de la ciencia aporta aquí dos movimientos complementarios. Primero, clarifica reglas de deferencia: cuándo la ciudadanía o los decisores particulares deben confiar, desconfiar o diferir prudentemente; cuándo elevar el umbral de evidencia (e.g., en contextos de alto riesgo) y cuándo basta el mejor juicio experto disponible (e.g., bajo severas restricciones temporales). Segundo, diseña y evalúa mecanismos de confianza: las prácticas institucionales que materializan esos criterios (e.g., declaraciones de conflicto de interés, datos abiertos, informes explicativos para audiencias no técnicas, protocolos de revisión cruzada entre disciplinas). Así, la confianza deja de concebirse como un estado psicológico interno —inaccesible y volátil— para convertirse en un bien epistémico institucionalizable y evaluable.

La consecuencia metodológica es decisiva: la credibilidad se vuelve una magnitud normativa —no es un dato sociológico bruto— y requiere justificación pública mediante pautas claras. Por ello, la filosofía de la ciencia puede identificar y hacer operativos esos criterios, de modo que la confianza no sea un acto opaco, sino un compromiso justificable y auditado entre expertos y no expertos. Desarrollaré esto en § 5. Por ahora, basta con adelantar que la mediación ejercitada por el filósofo de la ciencia no se reduce a la difusión unidireccional de conceptos, sino a la articulación reflexiva de los vínculos epistémicos que sostienen la cooperación científica y democrática.

Que lo anterior no se lea como un clamor moralizante. Se trata es de institucionalizar la "confianza bien puesta": especificar las condiciones bajo las cuales deferir es racional, disentir es responsable y explicar es obligatorio. Con ello, la filosofía de la ciencia logra impactar sin renunciar al rigor analítico, fortaleciendo las bases normativas de la credibilidad experta y aportando una infraestructura epistémica para la toma de decisiones en sociedades complejas. Si la confianza epistémica se reconfigura como un bien normativo institucionalizable, entonces la pregunta por la *expertise* deja de ser un problema puramente metateórico y pasa a describir un programa de intervención, empujando a la disciplina fuera de su autorreferencia y hacia los espacios donde esos estándares se negocian, se implementan y se auditan. Pero, ¿qué forma adopta ese cometido?

2.3. Re-emergencia de la filosofía de la ciencia socialmente comprometida

Uno de los avatares del retorno de la filosofía de la ciencia a la esfera pública cristalizó en el movimiento conocido como *filosofía de la ciencia socialmente comprometida* (SEPOS, abreviado en inglés). Esta corriente surge como respuesta a la percepción de que el retiro institucional del siglo XX dejó a la disciplina sin voz en los debates donde se deciden las condiciones de producción y legitimación del conocimiento científico.

SEPOS se nutre de dos líneas convergentes. Primero, la creciente autoconciencia de los científicos sobre las dimensiones sociales y normativas de su trabajo, institucionalizadas a través de distintos programas de gobernanza científica.[7] Segundo, la reflexión filosófica

[7] En Europa, el marco de *Responsible Research and Innovation* (RRI) del programa *Horizon 2020* articula la responsabilidad social de la investigación (Owen et al., 2012; Stilgoe et al., 2013). En EE.UU., la *National Science Foundation* introdujo el criterio de *Broader Impacts* en la dimensión pública de la evaluación de proyectos (Holbrook, 2012; Watts et al., 2015; Drummond Otten, 2022). En Reino Unido, el *Research Excellence Framework* exige estudios de casos de impacto que demuestren la relevancia social de la investigación (Research England et al., 2019), mientras que el

sobre el contextualismo en epistemología, que reconoce que los valores epistémicos se construyen socialmente y requieren instituciones que los sostengan (McKenna, 2015; Venturinha, 2018). En el cruce de ambos itinerarios, la filosofía de la ciencia encuentra suelo para recuperar su función mediadora: explicitar los supuestos normativos y conceptuales de la práctica científica, evaluar su coherencia con los valores democráticos y contribuir al diseño de políticas del conocimiento.

Cartieri & Potochnik (2014) identifican este renacer en proyectos y programas que, a diferencia de la filosofía analítica clásica, nacen directamente del contacto con comunidades científicas concretas y con respaldo institucional. Un caso destacado es el *Toolbox Project*, que emplea instrumentos filosóficos para facilitar la comunicación interdisciplinaria. En lugar de solo aplicar teorías filosóficas preexistentes, este crea un entorno de coproducción reflexiva donde los equipos examinan sus propias categorías de explicación, evidencia y valor. Es una práctica *in situ* que hace visible la función epistémica de la reflexión filosófica en contextos reales de investigación. En § 3 volveré sobre este caso.

El redescubrimiento de esta vocación pública se extiende más allá de la ciencia misma. La *filosofía de la ciencia ampliamente involucrada* (Plaisance & Elliott 2021) amplía el espectro del compromiso hacia el ámbito civil.[8] La idea central es que la filosofía de la ciencia puede desempeñar un papel deliberativo y normativo en la configuración de sociedades epistémicamente justas, donde la distribución de confianza,

programa *ELSI* del *Human Genome Project* aborda la reflexión ética y social dentro de la práctica científica (Hanna, 1995; Dolan et al., 2022).

[8] No debe confundirse la SEPOS con la *broadly engaged philosophy of science*. Esta última no designa una escuela: es un término paraguas para referirse a múltiples formas de vinculación entre la filosofía de la ciencia y el mundo extra-académico. Su objetivo es mapear y fomentar estrategias de participación de diverso tipo, no *necesariamente* centradas en problemas de bienestar social. La SEPOS, por su parte, es una corriente específica, orientada a cuestiones sociales y de política pública (e.g., salud, educación), trabajando en alianza con los actores sociales para aclarar aspectos descriptivos y normativos de la ciencia en democracia.

autoridad y credibilidad responda a razones públicas. Plaisance et al. (2021, p. 4877) han dado, además, evidencias premilitares de que las formas más efectivas de influencia filosófica ocurren en escenarios de interacción directa, especialmente en colaboraciones prolongadas y talleres interdisciplinarios. Estos generan comprensión mutua y fortalecen la confianza epistémica; siendo este último un recurso escaso en sociedades polarizadas y altamente tecnificadas como las nuestras.

Esta agenda reposiciona la autoimagen de la filosofía de la ciencia, devolviéndola a una visión próxima a lo que el Círculo de Viena defendía. No basta con analizar los fundamentos del conocimiento. Es preciso participar en su construcción, articulando puentes entre el análisis riguroso de los productos de las ciencias y la responsabilidad social que estos comportan. El desafío contemporáneo consiste en entender ese rigor como responsabilidad epistémica. El compromiso social no amenaza la cientificidad o el profesionalismo del filósofo: los renueva al reubicar el ideal de la claridad analítica en el terreno de la comunicación intersubjetiva, donde la precisión conceptual es condición de posibilidad del dialogo.

Cierro con una razón adicional para este re-compromiso: la creciente asunción de responsabilidades epistémicas colectivas. En las últimas décadas se ha incrementado la complejidad técnica de los sistemas de investigación y se han intensificado sus consecuencias sociales. Nunca hemos podido intervenir tanto en nuestro entorno. Frente a los riesgos de la especialización excesiva y la desconfianza social hacia la ciencia, la filosofía del siglo XXI parece demandar prácticas de integración interdisciplinaria. La ética del impacto comienza, así, con la tarea de reconectar la reflexión filosófica con los sistemas de decisión social. Ya no basta con influir mediante modelos idealizados de racionalidad, pues la credibilidad pública depende, importantemente, de intervenciones pertinentes y concretas en la orientación del conocimiento. He mencionado que las vías más eficaces de influencia son las interacciones cara a cara, los proyectos colaborativos y la formación compartida. Corresponde ahora precisar tales estrategias de impacto.

3. Modelos de influencia y sus tensiones

3.1. De la difusión al intercambio: las vías reales de impacto

La noción tradicional de impacto filosófico se apoya en un modelo lineal de comunicación científica: las ideas influirían en la medida en que se difunden y son apropiadas por otros agentes epistémicos. La implementación de este modelo tiende a presuponer que la publicación y citación bastan como vehículos de transferencia conceptual. Sin embargo, esta metáfora del "viaje de las ideas" es teóricamente pobre y empíricamente cuestionable (Plaisance et al., 2021, pp. 4869, 4876–4877).[9] A partir de entrevistas semiestructuradas con 35 filósofos de la ciencia implicados en colaboración interdisciplinaria, la evidencia sugiere que el impacto más relevante no ocurre cuando la filosofía circula como producto textual, sino cuando opera como práctica situada en contextos de investigación. Allí, la intervención se realiza mediante acciones epistémicas: explicitar supuestos ontológicos, coordinar estándares de evidencia, mediar conflictos de valores, y diseñar criterios operativos de decisión. La influencia no es un efecto distal del discurso, sino un resultado de la interacción comunicativa en la que el filósofo comparte agencia en la construcción del conocimiento.

Esto exige desplazar el análisis desde la propiedad de los *outputs* hacia la estructura de las relaciones epistémicas. Los autores identifican que la filosofía incide con mayor fuerza cuando:

a) participa en equipos interdisciplinarios orientados a la resolución de problemas;

[9] Entiéndase esta afirmación con cautela. No es extraño que, si un filósofo de la ciencia busca incidir significativamente en su área, deba publicar en revistas posicionadas por cuartiles del factor de impacto (e.g., el indicador *JIF Quartile* en WoS) para alcanzar la discusión de frontera; resultando menester que publique en revistas Q1 y Q2 de historia y filosofía de las ciencias. Desatender la relevancia de la revista pone al académico en una posición donde su trabajo no solo corre el riesgo de no ser leído, sino que además pierde acceso a instancias de alto valor, como la retroalimentación de pares selectos que las revistas de mayor prestigio pueden ofrecer.

b) contribuye a la coproducción normativa, asesorando aspectos metodológicos y de diseño experimental; y
c) interviene en instancias de formación mutua, especialmente mediante programas interdisciplinares de posgrado que integran competencias filosóficas en la educación científica.

Empero, estas formas de impacto se encuentran sistemáticamente infravaloradas por los sistemas de evaluación académica, pues no se traducen directamente en méritos cuantificables (e.g., artículos en revistas indexadas, citaciones, invitaciones a congresos y productos con autoría claramente asignable), sino en logros colaborativos distribuidos.[10] Se produce así una asimetría estructural: las prácticas filosóficas más efectivas suelen ser las menos visibles en el régimen de recompensas disciplinar, lo que instala una ética profesional implícita que desincentiva la participación pública y premia el rendimiento escritural por sobre la incidencia colaborativa.

La situación se agrava bajo la mercantilización de la publicación científica. La concentración oligopólica en la industria, donde cinco editoriales —Reed-Elsevier, Wiley-Blackwell, Springer, Taylor & Francis y Sage Publications— controlan más del 50 % de la producción científica, consolida un régimen en el que la progresión académica depende de acceder a esos circuitos comerciales (Larivière et al., 2015, pp. 3–7). Estas empresas operan con márgenes de beneficio superiores al 30 %, comparables o mayores a los del sector farmacéutico o bancario, monetizando trabajo académico no remunerado y vinculando el valor del conocimiento a su rentabilidad editorial (Larivière et al., 2015, pp. 10–11). Además, los precios se negocian según la capacidad de pago

[10] Para muestra, el *Reglamento del Académico* de la Universidad de Antofagasta incorpora como indicadores de investigación la publicación en revistas indexadas y la acreditación de citas (Universidad de Antofagasta, 2017, pp. 10, 12). En el contexto anglosajón, la política de *Appointment, Promotion and Tenure* del New Mexico Institute of Mining and Technology explicita el uso de métricas como número de citas e invitaciones como evidencia de impacto académico, vinculadas a decisiones de continuidad profesional (New Mexico Institute of Mining and Technology, 2020, p. 4).

institucional, no según el valor epistémico o social del contenido. Lo que se publica y se lee está condicionado por barreras económicas, no por su pertinencia pública (Bergstrom et al., 2014, pp. 9425–9427).

Bajo este modelo, la filosofía que interviene en contextos sociales queda estructuralmente invisibilizada al no producir bienes evaluables en el mercado editorial que organiza el reconocimiento académico. El resultado es un modelo autorreferencial de excelencia, donde la influencia social aparece como efecto colateral y no como un propósito perseguido. De aquí la necesidad de replantear los mecanismos institucionales de reconocimiento si se aspira a una filosofía de la ciencia que intervenga con responsabilidad en la orientación de la investigación. De otro modo, esta permanecerá invisible aunque sea epistémicamente indispensable.

Antes de pasar a la sección siguiente, conviene ponderar una objeción al artículo académico como género para producir y comunicar filosofía. Reconocer que el papel de los *papers* en el impacto es matizable no equivale a atacar el medio como tal. Distinto es el diagnóstico de José Santos Herceg, quien, desde una perspectiva foucaultiana, entiende la hegemonía del *paper* como un dispositivo de control y normalización. Este impondría un modelo textual de origen en las ciencias naturales —transitorio, especializado, modesto en pretensiones teóricas y orientado a la eficiencia comunicativa— que subordina formas históricas de expresión humanística, como el diálogo, la carta, el poema o los aforismos. Operaría como un filtro de acceso a la comunidad epistémica, marginando autores no acreditados, enfoques no proposicionales y tradiciones que articulan conocimiento mediante dispositivos retóricos, narrativos o estéticos, particularmente el pensamiento latinoamericano de vocación crítica (Santos Herceg, 2012, pp. 208–213). De ahí su denuncia de la "tiranía del paper", que devaluaría la producción local y sometería la filosofía a una lógica cienciométrica centrada en la cuantificación, la mercantilización y la competitividad, en detrimento de la densidad conceptual, el alcance histórico y la creatividad (Santos Herceg, 2020, pp. 19–20).

Salvando las preocupaciones por los incentivos mercantiles y por la centralidad de los méritos cuantificables —cuestiones que ameritan discusión—, esta denuncia incurre en una caricatura normativamente ingenua y empíricamente objetable. Frente a la hipérbole del dispositivo de domesticación discursiva, el artículo académico funciona como una infraestructura epistémica que posibilita el intercambio rápido de argumentos, el escrutinio público y la trazabilidad del impacto intra y extrafilosófico; siendo complementario —no sustitutivo— de las interacciones cara a cara (Plaisance et al., 2021, p. 4877). Rasgos como la modestia teórica, la renuncia al discurso con pretensión de absolutez, la especialización temática y la eficiencia comunicativa son virtudes epistémicas basales para una intervención filosófica estructurada que mejore la comunicación interdisciplinaria. El *paper* obliga a formular tesis como hipótesis revisables en un lenguaje estandarizado y accesible para filósofos y no filósofos, mitigando cuellos de botella epistémicos (O'Rourke & Crowley, 2013, pp. 1939–1944).

Para lo que nos ocupa, una *filosofía de la ciencia ampliamente involucrada* y *socialmente comprometida* exige instituciones de confianza entre expertos y ciudadanía. Aquí, el *paper* cumple una función doble, endosando fiabilidad epistémica y responsabilidad pública. Contra la lectura del artículo como artefacto uniformador, este opera antes como una condición habilitante que como un obstáculo para ejercitar una filosofía con impacto comprobable, evaluable y exigible. Con todo, la objeción de Santos Herceg introduce una pregunta pertinente sobre las estructuras institucionales que legitimamos; pero, tomada en sus propios términos, abre dudas retroactivas:

- ¿qué forma de filosofía se avala al relativizar los estándares elementales de escritura, rigor y pertinencia académica?;
- ¿qué se incentiva cuando se premia la ocurrencia vistosa y el gesto retórico por sobre un lenguaje compartido que facilite la colaboración intra e interdisciplinaria?; y

- ¿qué entendemos por filosofía cuando se diluye la diferencia entre la investigación filosófica articulada en continuidad crítica con nuestra mejor ciencia y una filosofía de salón anclada en doxografía o en la repetición irreflexiva de tópicos importados?

Precisamente porque estas dudas remiten a los criterios y prácticas mediante los cuales una comunidad filosófica produce y distribuye legitimidad epistémica, conviene reformular el problema no solo en términos de *quién* tiene conocimiento, sino de *cómo* coordinamos estándares, mediamos conceptos y generamos confianza entre pares y con otras disciplinas.

3.2. Filosofía como intervención: el modelo del *Toolbox Project*

Resulta menester complementar el paradigma de la autoridad epistémica —que concibe al filósofo como custodio de estándares racionales— con el de la interacción epistémica. El valor del impacto filosófico ha de ligarse a la calidad de nuestros procesos de mediación, coordinación conceptual y confianza. Esta orientación no solo delimita en un sentido no trivial qué cuenta como mérito disciplinar, sino que vuelve visible una dimensión relacional del rigor filosófico, de la cual dependen tanto el progreso interno de la filosofía como su relevancia en la gobernanza epistémica de las sociedades democráticas.

Superar esta situación paradójica, en donde la filosofía de la ciencia mantiene un alto potencial de incidencia pública a la vez que resulta ser una de las disciplinas más aisladas del debate social (Conix et al., 2022, pp. 11–12), exige reconfigurar los criterios de excelencia académica. Una ética del impacto coherente debe reconocer como mérito disciplinar las prácticas colaborativas que generan confianza epistémica. Para ello, es necesario institucionalizar la cooperación como parte constitutiva del oficio del filósofo de la ciencia: ampliar las métricas de evaluación para incluir contribuciones cualitativas a proyectos interdisciplinarios, acciones formativas y aportes al diseño normativo de

la ciencia; y crear espacios estables de interacción entre la filosofía y la sociedad civil. En ese contexto, resulta de interés el caso del *Toolbox Project*.

El *Toolbox Project*, iniciativa desarrollada por Michael O'Rourke & Stephen J. Crowley, constituye un caso de estudio para comprender cómo la filosofía puede interactuar epistémicamente dentro de la investigación interdisciplinaria (Eigenbrode et al., 2007; Hubbs et al., 2022; O'Rourke & Crowley, 2013). Lejos de concebirse como una consultoría ocasional o una instancia de transferencia unidireccional de conceptos, el proyecto diseña talleres estructurados donde los participantes —entre los que se cuentan biólogos, ingenieros, científicos sociales y filósofos— emplean instrumental filosófico explícito (e.g., matrices de preguntas, taxonomías de supuestos, mapeos argumentativos) para identificar y evaluar las presuposiciones que orientan su práctica científica. Este proceso no presupone consenso: parte de la premisa de que las diferencias conceptuales poseen valor epistémico cuando son identificadas y articuladas públicamente.

El impacto epistémico del *Toolbox Project* consiste en optimizar la integración del conocimiento mediante el alineamiento conceptual del equipo. La filosofía opera como un dispositivo metacognitivo, explicitando criterios metodológicos, clarificando el estatuto de las evidencias relevantes para cada disciplina y detectando incompatibilidades ontológicas que podrían entorpecer la cooperación. En lugar de resolver desacuerdos mediante jerarquías de autoridad o imposiciones metodológicas, el proyecto habilita procesos deliberativos que amplían la comprensión del problema investigado y de las contribuciones diferenciales de cada experto. La filosofía no produce aquí resultados empíricos directos, pero sí condiciones epistémicas para que esos resultados sean alcanzables. Junto a este componente epistémico, el proyecto incorpora un impacto ético. Reconoce que la colaboración interdisciplinaria es también una práctica de distribución de responsabilidades. Al garantizar que todas las voces disciplinarias dispongan de recursos para explicitar sus perspectivas, el proyecto

fortalece la justicia epistémica interna del grupo. Además, al hacer visibles los valores implicados en la toma de decisiones, fomenta la corresponsabilidad y evita que las tensiones normativas se oculten bajo una apariencia de neutralidad.

El *Toolbox Project* ha mostrado impactos sociales indirectos verificables al fortalecer la coordinación interdisciplinaria en ámbitos con externalidades públicas significativas. Estudios empíricos muestran que sus talleres producen mejoras cuantificables en la comprensión epistémica mutua —donde el 84,9 % de participantes reporta mayor claridad respecto del enfoque de sus colegas y un intervalo entre 29-42 % modifica sustancialmente sus respuestas pre y pos-intervención, según el área temática (Schnapp et al., 2012, p. 471)—, lo que constituye un mecanismo proximal reconocido en la literatura de ciencia en equipo (*team science*) como predictor de efectividad en contextos de bienestar social (O'Rourke & Crowley, 2013, p. 1943). Asimismo, su aplicación en gobernanza ambiental ha facilitado el diálogo entre actores académicos y comunitarios para la toma de decisiones locales sobre, por ejemplo, resiliencia climática (Cwik et al., 2022). Este reconocimiento se ve reforzado por distinciones institucionales como el *APA/PDC Prize for Excellence and Innovation in Philosophy Programs* y el apoyo que les ha cedido la *National Science Foundation*.

Lo anterior comporta una primera constatación de que la intervención filosófica no solo optimiza los procesos de coordinación cognitiva en entornos científicos, sino que también contribuye de manera estructural a la formación de capacidades sociales relevantes para la toma de decisiones en políticas públicas de alto impacto, especialmente en dominios donde la integración interdisciplinaria es condición habilitante para la acción colectiva. Con todo, conviene ser prudentes. La literatura evaluativa se apoya, hasta ahora, en indicadores proximales de impacto —como la mejora en la comunicación conceptual, la explicitación de valores y la alineación metodológica— que, aunque plausibles como antecedentes del impacto social, no constituyen evidencia de resultados finales. Una agenda rigurosa debería incorporar métricas longitudinales,

estudios de trazabilidad institucional de decisiones derivadas y análisis comparados que permitan estimar en qué medida estas intervenciones se traducen en cambios normativos y prácticos sostenidos, tales como el número de políticas implementadas, mejoras particulares en gobernanza científica o beneficios poblacionales medibles en el mediano y largo plazo.

4. Progreso filosófico y estándares de valor

El renovado interés por el impacto social de la filosofía exige revisar cómo se evalúa su propio progreso interno. Se ha señalado que las métricas tradicionales —resolver definitivamente problemas filosóficos o alcanzar consenso disciplinario— son insuficientes y empíricamente insostenibles para captar el progreso en filosofía (Norton, 2023, pp. 11–12). La propuesta reciente de Dellsén et al. (2022, 2024) constituye una alternativa robusta y que me gustaría rescatar para la ocasión: el progreso filosófico consiste en aumentar la comprensión estructural, i.e., mejorar la representación de relaciones de dependencia entre conceptos, fenómenos y prácticas.

Este estándar rechaza la obsesión con las respuestas finales y desplaza la atención hacia los procesos de articulación conceptual. Por articulación conceptual entenderé prácticas que aumentan el entendimiento al construir representaciones sistemáticas y factuales de cómo un fenómeno se correlaciona con otros. Sistemáticas, porque sustituyen "*bits*" sueltos por una estructura amplia que muestra "cómo las cosas encajan". Factuales, porque ese mapa de dependencias debe ser más exacto y más comprehensivo que otros previos para contar como progreso (Dellsén et al., 2024, pp. 686-687). Esto excluye logros triviales que no informan sobre dependencias existentes y explica por qué ciertas distinciones o modelos pueden ser progresivos sin requerir justificación o conocimiento en sentido fuerte, siempre que mejoren el mapa de dependencias (Dellsén et al., 2022, pp. 829-830; 2024, pp. 675-677).

Los beneficios se aprecian cuando se mezclan inadvertidamente dos consideraciones bajo un mismo término. Considérese el concepto de responsabilidad. A veces hablamos de ser responsables por las consecuencias de nuestros actos. En otras, hablamos de ser moralmente responsables incluso cuando esas consecuencias no dependen totalmente de nosotros (e.g., una omisión, una intención fallida o un accidente). Bajo una misma palabra, es fácil confundir qué consideraciones pertenecen a un tipo de responsabilidad y cuáles al otro. La articulación conceptual separa ambas nociones y permite ver qué depende de qué: cuándo importa la intención, cuándo importa el resultado y cómo cambian las conclusiones éticas según el caso. El resultado es un mejor entendimiento del fenómeno, porque distinguimos piezas antes aglutinadas.

Lo mismo ocurre si el término forma parte del lenguaje científico. Piénsese en la noción de inteligencia. Una definición usual la caracterizaría como "la capacidad resolver problemas". Hoy sabemos que las capacidades cognitivas dependen también de factores sociales y tecnológicos, como el acceso a herramientas, el lenguaje, la educación y la cooperación. Una articulación conceptual más avanzada integraría esos elementos en un solo marco, mostrando que cierta habilidad individual depende de un entorno colaborativo o de un artefacto cultural. Gracias a ese mapa vemos cómo encajan las cosas y qué deberíamos modificar para optimizar la inteligencia humana.

Concebir el progreso como comprensión tiene una consecuencia decisiva: el trabajo filosófico aplicado no es secundario, sino un modo privilegiado de comprobar y ampliar esa comprensión. Cuando los filósofos colaboran con científicos o con instituciones, ponen a prueba su aparato conceptual y lo refinan mediante la interacción. Así, el impacto pasa a ser un indicador epistémico del progreso logrado.[11] Aquí se

[11] Podemos apreciar lo fructífero de este enfoque en las discusiones sobre metafísica naturalizada y metafísica analítica. En ese debate se discute cómo la fricción entre proyectos metafísicos anclados en supuestos tradicionales y los resultados de nuestra mejor ciencia puede enriquecer la discusión y promover progreso disciplinar, especialmente al sentar bases para nuevas metafísicas surgidas de un diálogo continuista y sistemático con las ciencias (Jaksland,

unifican dos dimensiones: el valor epistémico de la filosofía —la clarificación conceptual y la coordinación teórica— y su valor social —la responsabilización por nuestras prácticas científicas y deliberativas. Lo distintivo es que ambas dependen de una misma capacidad: producir comprensión compartida. La Tabla 1 organiza esta convergencia mediante un esquema bidimensional que conecta criterios internos de evaluación del progreso con criterios externos de impacto social.

Función epistémica	**Dimensión interna (progreso filosófico)**	**Dimensión externa (impacto social)**
Organización conceptual	Claridad y coherencia de las redes conceptuales.	Claridad y accesibilidad de los marcos explicativos.
Integración y cooperación teórica	Capacidad de integrar perspectivas teóricas divergentes.	Capacidad de coordinar actores y saberes distintos.
Gestión de tensiones	Detección y resolución de tensiones conceptuales.	Detección y resolución de conflictos de valor o de confianza.
Comprensión orientada a la acción	Producción de comprensión compartida.	Generación de decisiones informadas y confiables.

Tabla 1. Funciones epistémicas y criterios de evaluación del progreso e impacto filosófico.

Por organización conceptual entiendo la tarea de hacer explícitas las dependencias que estructuran un dominio, de modo que los términos

2023; Ladyman et al., 2007; Le Bihan & Barton, 2021; Morganti & Tahko, 2017; Silva Silva, 2025). Este número incluye, además, una traducción de Rasmus Jaksland dedicada a esta temática.

tengan roles estables y comparables. Con ello se mitigan desacuerdos y se explicitan inferencias legítimas. En la dimensión interna, ello exige redes conceptuales claras y coherentes; en la externa, marcos explicativos accesibles sin sacrificar precisión. La integración y cooperación teórica, por su lado, consiste en articular perspectivas divergentes sin forzar el consenso, coordinando vocabularios y estándares de evidencia para que actores y saberes distintos contribuyan a ocuparse del problema relevante al caso.

La gestión de tensiones apunta a identificar incompatibilidades conceptuales y a resolverlas mediante la distinción, revisión o reasignación de cargas de justificación. En su proyección pública, la misma se aplica a conflictos de valor o de confianza, volviendo explícitos los puntos de desacuerdo y ayudando a establecer cuándo deferir, disentir o explicar es obligatorio. La comprensión orientada a la acción marca el umbral en que la comprensión compartida se traduce en reglas de decisión: fija criterios para actuar bajo incertidumbre y habilita decisiones informadas y confiables.

Quienes defienden una filosofía exclusivamente teórica suelen asociar el rigor con la clausura y la autosuficiencia disciplinar. Sin embargo, bajo esta concepción, el rigor —entendido como precisión en la mediación— es el que hace explícitos valores, supuestos y estructuras conceptuales que permiten decisiones más confiables en entornos complejos. Siendo así, el progreso filosófico se verifica cuando, al clarificar nuestra comprensión del mundo, clarificamos también cómo debemos actuar en él.

5. Ética profesional del trabajo filosófico

Si el impacto filosófico es primariamente relacional (§ 3) y el progreso filosófico puede entenderse como el aumento de comprensión estructural (§ 4), entonces se requiere de una ética profesional que traduzca esas tesis en normas de oficio. Con ello en mente, propongo

cuatro principios: integridad epistémica, colaboración horizontal, responsabilidad social y transparencia institucional (ver Tabla 2).

Principio ético	Competencia filosófica asociada	Indicador operativo
Integridad epistémica	Rigor conceptual	Clarificación de supuestos en investigación aplicada
Colaboración horizontal	Mediación conceptual	Integración funcional de perspectivas diversas
Responsabilidad social	Deliberación pública	Aporte a decisiones y políticas basadas en evidencia
Transparencia institucional	Comunicación y trazabilidad	Visibilización de procesos y criterios

Tabla 2. Principios de ética profesional para el trabajo filosófico.

Estos principios han de entenderse como requisitos operativos complementarios para intervenir en entornos donde se negocia la credibilidad experta y se toman decisiones bajo incertidumbre. A continuación, desarrollo qué implican en la práctica cotidiana, cómo se implementarían institucionalmente y qué tensiones estructurales enfrentan.

5.1. Integridad epistémica

Llamo integridad epistémica a la obligación de preservar, en contextos colaborativos y aplicados, el control de calidad epistémico de nuestras intervenciones —claridad, trazabilidad, disciplina conceptual y

ajuste a la evidencia—, resistiendo presiones que incentiven la simplificación oportunista o las conclusiones no respaldadas.

Considérese un equipo interdisciplinar de salud pública evalúa un piloto de triaje algorítmico que asigna horas médicas. El desacuerdo interno no es empírico, sino conceptual: unos hablan de "precisión", otros de "equidad", otros de "calidad". Una intervención íntegra consistiría en:

(i) exigir definiciones operativas mínimas (¿equidad como igualdad de error por grupo, como igualdad de acceso, como no discriminación legal?);
(ii) distinguir niveles normativos (criterios de evaluación del modelo vs. criterios de justicia del procedimiento);
(iii) mapear dependencias ("si priorizamos la sensibilidad, aumentamos los falsos positivos; si imponemos paridad de error, cambia el umbral; si cambiamos el umbral, cambia la carga asistencial");
(iv) proponer un protocolo de decisión que separe qué balance de pérdidas y ganancias resulten aceptable y quién está autorizado a aceptarlo.

El producto es un marco de decisión y un glosario controlado que vuelve auditables las elecciones del equipo. La contribución puede evaluarse con indicadores simples: trazabilidad —que el informe final explicite supuestos y balances—, estabilidad —que el vocabulario se mantenga consistente—, y utilidad —que el equipo reporte la reducción de malentendidos metodológicos. Un obstáculo por superar es que el reconocimiento institucional no debe entenderse como un gesto caritativo. Debe haber coautoría en informes metodológicos, un reconocimiento del rol como coinvestigador en el proyecto, y una valoración explícita en los reportes de impacto.

5.2. Colaboración horizontal

Llamo colaboración horizontal al principio según el cual se interviene como par epistémico —no como mero árbitro o asesor externo—, co-diseñando marcos, protocolos y criterios con los demás expertos, bajo reglas de deliberación y reconocimiento.

Pensemos en un equipo interdisciplinario de resiliencia climática integrado por ecólogos, ingenieros, científicos sociales y autoridades locales. El conflicto es la divergencia de estándares de evidencia: el ingeniero exige modelos predictivos, el ecólogo trabaja con incertidumbre irreductible, la autoridad local exige umbrales claros para actuar. Aquí, el papel del filósofo consistiría en asistir como mediador horizontal, algo materializable en un taller estructurado con tres salidas:

(i) una matriz de supuestos (qué cuenta como "daño", qué horizonte temporal importa, qué distribución de riesgos es aceptable);

(ii) un mapa de dependencias entre decisiones y evidencias (qué datos cambian qué decisión);

(iii) un protocolo de desacuerdo: cuando no hay consenso, se explicita el punto de fractura y se asigna la carga de justificación.

Institucionalmente, esto exige roles estables de "facilitación epistémica" en proyectos, con tiempo financiado y criterios de evaluación. El obstáculo más común aquí es cultural: se suele interpretar como prescindibles las habilidades de mediación "sin un papel tangible". Por eso insisto en la presencia de productos verificables como matrices o protocolos.

5.3. Responsabilidad social

Llamo responsabilidad social a la obligación del filósofo de la ciencia de producir intervenciones normativamente trazables y públicamente justificables que mejoren la calidad de las decisiones tecnocientíficas con efectos colectivos, sin sustituir la deliberación democrática.

Es habitual que una agencia, al evaluar la regulación de un contaminante nuevo, cuente con evidencia incompleta. Allí la pregunta clave no es "¿qué es evidencia?", sino "¿cuándo es racional elevar el umbral de precaución?". El papel del filósofo involucrado consistiría en articular, mancomunadamente con los otros expertos, un marco público:

(i) distinguir la incertidumbre epistémica de la incertidumbre moral;
(ii) explicitar criterios de severidad y reversibilidad del daño; y
(iii) proponer una regla de decisión con mecanismos de actualización condicional (regulación provisional, monitoreo reforzado, revisión a 12 meses).

La contribución se mide por su impacto deliberativo, aportando en la claridad de la justificación de las decisiones de cara al público, asegurando la consistencia de los criterios, y estableciendo medidas de inspección. El obstáculo estructural es la tentación tecnocrática. El filósofo de la ciencia ha de entrar como parte de una infraestructura normativa que haga la decisión más transparente, no como un sustituto de la decisión.

5.4. Transparencia institucional

Llamo transparencia institucional al principio según el cual los estándares, procesos y criterios de evaluación y decisión que estructuran la práctica filosófica con impacto deben quedar explícitos y ser auditables, de modo que la credibilidad se apoye en la trazabilidad y la rendición de cuentas, no en autoridad informal.

Usemos un ejemplo local. Un departamento de Filosofía adopta una pauta interna de promoción donde, además de publicaciones, se exigen al menos dos "evidencias de contribución":

(i) participación formal en proyectos interdisciplinarios con productos documentados;
(ii) incidencia en guías o protocolos públicos;

(iii) conformación de equipos (e.g., talleres metodológicos, grupos de estudio);
(iv) productos de vinculación con el medio (e.g., informes para audiencias no técnicas).

La fortaleza es obvia: alinea los incentivos institucionales con prácticas reales de impacto. La limitación también es evidente: puede burocratizarse. Por eso, la transparencia debe incluir mecanismos para prevenir la manipulación de métricas: revisión externa interdisciplinaria y criterios *ex ante* (qué cuenta como evidencia, qué estándar de trazabilidad se exige).

5.5. Obstáculos institucionales y tensiones reales

Este marco sólo es aplicable si se reconocen las fricciones que lo vuelven costoso. Identifico al menos cuatro obstáculos estructurales:

(i) *desalineación de incentivos*: los sistemas de contratación y promoción siguen premiando aportes bibliométricos rápidos y penalizan el trabajo relacional lento;
(ii) *ambigüedad de rol*: el filósofo puede ser incorporado como "sello de legitimidad", pero sin agencia real, como un consultor periférico sin incidencia en decisiones relevantes;
(iii) *asimetrías jerárquicas y de crédito*: en equipos dominados por una disciplina, la mediación tiende a la subordinación, y las contribuciones filosóficas son invisibilizadas;
(iv) *restricciones de trazabilidad*: la confidencialidad, la propiedad intelectual y los "tiempos políticos" limitan qué puede documentarse y cómo se comunica públicamente.

Estas tensiones no refutan la ética profesional que propongo, pero exigen condiciones mínimas para su implementación: roles formales con tiempo financiado, acuerdos *ex ante* de gobernanza y reconocimiento, criterios de evaluación que admitan evidencia documental auditables, y

revisión externa interdisciplinaria. Sin estas condiciones institucionales, la ética del impacto se reduce a una declaración de buenas intenciones.

6. Conclusión

El recorrido realizado en este trabajo demanda comprometerse con la siguiente afirmación: la filosofía de la ciencia sólo está a la altura de su objeto cuando asume plenamente la responsabilidad epistémica que le corresponde. Buscar obtener y ofrecer comprensión estructural de los contenidos y problemas que nos entrega nuestra mejor ciencia no es un ejercicio ornamental. Supone, más bien, comprometerse con las condiciones bajo las cuales las sociedades toman decisiones informadas y confiables. Allí se juega la dignidad de la práctica filosófica, en su capacidad de convertir nuestras distinciones y modelos en una orientación efectiva para quienes deben actuar bajo cotas de incertidumbre. Si la filosofía rehúye este deber, abdica de su razón de ser y se refugia en un formalismo autosatisfecho o, peor aún, en una retórica de la crítica que opera como una exención de responsabilidad.

El filósofo de la ciencia no se puede permitir sostener que rendir cuentas por el impacto social de su trabajo sea una concesión al pragmatismo. Es, por el contrario, una exigencia de honestidad intelectual para toda disciplina que aspira a establecer cómo debe ejercitarse la razón humana en contextos colectivos. Asumir una ética profesional sensible al impacto social constituye una condición relevante —aunque no exclusiva— de la legitimidad de nuestra autoridad epistémica ante la ciudadanía. El llamado, entonces, es a optar deliberadamente por una filosofía que no tema ser evaluada en su contribución efectiva a la comprensión colectiva. Rehuir esta exigencia, amparándonos en la supuesta inmunidad de la teoría frente a los criterios de la práctica, incrementa el riesgo de debilitar la vocación que dio origen a la disciplina: esclarecer y orientar el uso público del conocimiento.

Esto no niega el valor intrínseco de la investigación teórica; subraya, más bien, que en tiempos donde la confianza epistémica es un

bien frágil, la ausencia de mecanismos de transparencia, trazabilidad y rendición de cuentas puede volver opaca nuestra contribución y erosionar su recepción pública. En tal escenario, lo que está en juego no es sólo el prestigio de la filosofía, sino su contribución a la racionalidad democrática.

REFERENCIAS

Anderson, E. (2024, August 1). *Feminist Epistemology and Philosophy of Science.* Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/feminism-epistemology/

Asociación Ernst Mach. (2002). La concepción científica del mundo: el Círculo de Viena. *Redes 9*(18), 105-149. https://doi.org/10.48160/18517072re18.579

Bergstrom, T. C., Courant, P. N., McAfee, R. P. & Williams, M.A. (2014). Evaluating big deal journal bundles, *Proceedings of the National Academy of Sciences of the United States of America*, *111*(26) 9425–9430, https://doi.org/10.1073/pnas.1403006111

Bricmont, J. (2015). Why there is no scientific method. And why it is not a problem. *Metode Science Studies Journal*(5), 183–187. https://doi.org/10.7203/metode.0.4040

Cartieri, F., & Potochnik, A. (2014). Toward Philosophy of Science's Social Engagement. *Erkenntnis*, *79* (Suppl 5), 901–916. https://doi.org/10.1007/s10670-013-9535-3

Conix, S., Lemeire, O., & Chi, P.-S. (2022). The public relevance of philosophy. *Synthese*, *200*(15). https://doi.org/10.1007/s11229-022-03546-9

Crasnow, S., & Intemann, K. (Eds.). (2020). *The Routledge Handbook of Feminist Philosophy of Science.* Routledge. https://doi.org/10.4324/9780429507731

Cwik, B., Gonnerman, C., O'Rourke, M., Robinson, B., & Schoonmaker, D. (2022). Building community capacity with philosophy: Toolbox dialogue and climate resilience. *Ecology and Society*, *27*(2), 21. https://doi.org/10.5751/ES-13282-270221

Dellsén, F., Lawler, I., Norton, J., & Firing, T. (2022). Thinking about Progress: From Science to Philosophy. *Noûs*, *56*(4), 814–840. https://doi.org/10.1111/nous.12383

Dellsén, F., Lawler, I., Norton, J., & Firing, T. (2024). What is philosophical progress? *Philosophy and Phenomenological Research*, 109(2), 663–693. https://doi.org/10.1111/phpr.13067

Dolan, D. D., Lee, S. S.-J., & Cho, M. K. (2022). Three decades of ethical, legal, and social implications research: Looking back to chart a path forward. *Cell Genomics*, *2*(7), 100150. https://doi.org/10.1016/j.xgen.2022.100150

Douglas, H. (2010). *A History of the PSA Before 1970.* Philosophy of Science Association. https://philsci.org/history_of_the_association.php

Drummond Otten, C. (2022). Assessing broader impacts of funded research: The US National Science Foundation v. Lamar Smith. *Science and Public Policy*, *49*(2), 313–323. https://doi.org/10.1093/scipol/scab082

Eigenbrode, S. D., O'Rourke, M., Wulfhorst, J. D., Althoff, D. M., Goldberg, C. S., Merrill, K., Morse, W., Nielsen-Pincus, M., Stephens, J., Winowiecki, L., & Bosque-Pérez, N. A. (2007). Employing Philosophical Dialogue in Collaborative Science. *BioScience*, *57*(1), 55–64. https://doi.org/10.1641/B570109

Feyerabend, P. (1963). How to Be a Good Empiricist: A Plea for Tolerance in Matters Epistemological. In B. Baumrin (Ed.), *Philosophy of Science: The Delaware Seminar, Vol. II* (pp. 3–39). Interscience Publishers.

Grasswick, H. E. (Ed.). (2011). *Feminist Epistemology and Philosophy of Science. Power in Knowledge.* Springer. https://doi.org/10.1007/978-1-4020-6835-5

Hanna, K. E. (1995). *The Ethical, Legal, and Social Implications Program of the National Center for Human Genome Research: A Missed Opportunity?* In R. E. Bulger, E. M. Bobby, & H. V. Fineberg (Eds.), *Society's Choices: Social and Ethical Decision Making in Biomedicine* (pp. 432–457). National Academies Press. https://doi.org/10.17226/4771

Holbrook, J. B., & Frodeman, R. (2012). *Science: For Science's or Society's Sake? Owning the National Science Foundation's Broader Impacts Criterion*. Center for American Progress. https://digital.library.unt.edu/ark:/67531/metadc86186/

Howard, D. (2003). Two Left Turns Make a Right: On the Curious Political Career of North American Philosophy of Science at Midcentury. In A. Richardson & G. Hardcastle (Eds.), *Logical empiricism in North America* (pp. 25–93). University of Minnesota Press. https://www.jstor.org/stable/10.5749/j.cttttvrh.5

Hubbs, G., O'Rourke, M., & Orzack, S. H. (Eds.). (2020). *The Toolbox Dialogue Initiative: The Power of Cross-Disciplinary Practice*. CRC Press. https://doi.org/10.1201/9780429440014

Jaksland, R. (2023). Naturalized metaphysics or displacing metaphysicians to save metaphysics. *Synthese*, *201*(199). https://doi.org/10.1007/s11229-023-04207-1

Ladyman, J., Ross, D., Spurrett, D., & Collier, J. (2007). *Every Thing Must Go: Metaphysics Naturalized.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199276196.001.0001

Larivière, V., Haustein, S., & Mongeon, P. (2015). The Oligopoly of Academic Publishers in the Digital Era. *PLOS ONE*, *10*(6), e0127502. https://doi.org/10.1371/journal.pone.0127502

Laudan, L. (1983). The Demise of the Demarcation Problem. In R. S. Cohen, & L. Laudan (Eds.), *Physics, Philosophy and Psychoanalysis. Essays in Honor of Adolf Grünbaum* (pp. 111–127). Springer. https://doi.org/10.1007/978-94-009-7055-7_6

Le Bihan, B., & Barton, A. (2021). Analytic Metaphysics Versus Naturalized Metaphysics: The Relevance of Applied Ontology.

Erkenntnis, *86*, 21–37. https://doi.org/10.1007/s10670-018-0091-8

Maudlin, T. (2007). *The Metaphysics Within Physics.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199218219.001.0001

McKenna, R. (2015). Contextualism in epistemology. *Analysis*, *75*(3), 489–503. https://doi.org/10.1093/analys/anv029

Morganti, M., & Tahko, T. E. (2017). Moderately naturalistic metaphysics. *Synthese*, *194*, 2557–2580. https://doi.org/10.1007/s11229-016-1068-2

New Mexico Institute of Mining and Technology. (2020, March 13). *Policy and Procedure for Appointment, Promotion, and Tenure.* New Mexico Tech: https://www.nmt.edu/academicaffairs/NMT-Appointment-Promotion-Tenure-Policy.pdf

Norton, J. (2023). Why philosophy needs a concept of progress. *Metaphilosophy*, *54*(1), 3–16. https://doi.org/10.1111/meta.12603

O'Rourke, M., & Crowley, S. J. (2013). Philosophical intervention and cross-disciplinary science: The story of the Toolbox Project. *Synthese, 190*(11), 1937–1954. https://doi.org/10.1007/s11229-012-0175-y

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy*, *39*(6), 751–760. https://doi.org/10.1093/scipol/scs093

Plaisance, K. S., Michaud, J., & McLevey, J. (2021). Pathways of influence: Understanding the impact of philosophy of science in scientific domains. *Synthese*, *199*, 4865–4896. https://doi.org/10.1007/s11229-020-03007-1

Plaisance, K. S., & Elliott, K. C. (2021). *A Framework for Analyzing Broadly Engaged Philosophy of Science. Philosophy of Science*, *88*(4), 594–615. https://doi.org/10.1086/713891

Pradeu, T., Lemoine, M., Khelfaoui, M., & Gingras, Y. (2024). Philosophy in Science: Can Philosophers of Science Permeate

through Science and Produce Scientific Knowledge? *The British Journal for the Philosophy of Science, 75*(2), 375–416. https://doi.org/10.1086/715518

Reisch, G. A. (2009). *Cómo la Guerra Fría transformó la filosofía de la ciencia: Hacia las heladas laderas de la lógica* (D. Blanco, Trans.). Universidad Nacional de Quilmes.

Research England; Scottish Funding Council; Higher Education Funding Council for Wales; & Department for the Economy, Northern Ireland. (2019). *Guidance on submissions (REF 2019/01).* REF 2021. https://2021.ref.ac.uk/publications-and-reports/guidance-on-submissions-201901/index.html

Richardson, S. S. (2009a). The Left Vienna Circle. Pt. 1: Carnap, Neurath, and the Left Vienna Circle Thesis. *Studies in History and Philosophy of Science*, *40*(1), 14–24. https://doi.org/10.1016/j.shpsa.2008.12.002

Richardson, S. S. (2009b). The Left Vienna Circle. Pt. 2: The Left Vienna Circle, disciplinary history, and feminist philosophy of science. *Studies in History and Philosophy of Science*, *40*(2), 167–174. https://doi.org/10.1016/j.shpsa.2009.03.010

Romizi, D. (2012). The Vienna Circle's 'scientific world-conception': Philosophy of Science in the Political Arena. *HOPOS*, *2*(2), 205–242. https://doi.org/10.1086/666659

Santos Herceg, J. (2012). Tiranía del paper. Imposición institucional de un tipo discursivo. *Revista Chilena de Literatura*, *82*, 197–217. http://dx.doi.org/10.4067/S0718-22952012000200011

Santos Herceg, J. (2020). *La tiranía del paper: De la mercantilización a la normalización de las textualidades*. Ediciones Universidad Austral de Chile.

Schnapp, L. M., Rotschy, L., Hall, T. E., Crowley, S. J., & O'Rourke, M. (2012). How to talk to strangers: Facilitating knowledge sharing within translational health teams with the Toolbox dialogue method. *Translational Behavioral Medicine, 2*(4), 469–479. https://doi.org/10.1007/s13142-012-0171-2

Silva-Silva, J. (2025). Addressing McKenzie's Challenge Regarding Progress in Naturalized Metaphysics. *Culturas Científicas*, *6*(1), 48–62. https://doi.org/10.35588/cc.v6i177424

Sorell, T. (1991). *Scientism*. Routledge.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, *42*(9), 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

Universidad de Antofagasta. (2017, Diciembre 4). *Reglamento del Académico.* Universidad de Antofagasta: https://www.uantof.cl/wp-content/uploads/2021/11/Reglamento-del-Academico-.pdf

Venturinha, N. (2018). *Description of Situations: An Essay in Contextualist Epistemology*. Springer. https://doi.org/10.1007/978-3-030-00154-4

Watts, S. M., George, M. D., & Levey, D. J. (2015). Achieving Broader Impacts in the National Science Foundation, Division of Environmental Biology. *BioScience*, *65*(4), 397–407. https://doi.org/10.1093/biosci/biv006

Whyte, K. P., & Crease, R. P. (2010). Trust, expertise, and the philosophy of science. *Synthese*, *177*, 411–425. https://doi.org/10.1007/s11229-010-9786-3

HOMÓNIMA
Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 162-177
ISSN 3087-260X
Santiago de Chile

¿QUÉ ES LEER Y ESCRIBIR FILOSOFÍA?[1]

Oscar Morales[2]
Universidad de Chile

Resumen

Este breve ensayo tiene el propósito de explorar los diversos significados que adquieren la lectura y la escritura filosófica, no desde su papel pedagógico en la formación del estudiante, sino desde la comprensión de su naturaleza crítica y dialógica. Se examina su función tanto en el ámbito interno del autor, como en la responsabilidad pública que este asume al entregar sus palabras al escrutinio colectivo, como si fueran hijos abandonados expuestos a la interpretación ajena. Lejos de pretender rebatir a otros autores o probar una hipótesis, el texto se desarrolla a partir de una escritura libre, que busca caracterizar

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Morales, O. (2026) ¿Qué es leer y escribir filosofía? *Revista de Filosofía Homónima*, *1*(1), pp. 162-177. https://doi.org/xxxxxxx[Por asignar]

[2] Oscar Morales es Licenciado en Filosofía, Universidad de Chile y estudiante de Magíster de la misma casa de estudios. Email: moralesbravo58@outlook.com. ORCID: https://orcid.org/0009-0002-3996-0028

el quehacer filosófico y reflexionar sobre su sentido como ejercicio vivo de pensamiento, comunicación y autoconocimiento.

Palabras clave: *lectura, escritura, metafilosofía, educación filosófica, pensamiento crítico.*

What does it mean to read and write philosophy?

Abstract

This brief essay aims to explore the various meanings assumed by philosophical reading and writing, not from the perspective of their pedagogical role in the student's formation, but rather from an understanding of their critical and dialogical nature. It examines their function both in the author's inner domain and in the public responsibility undertaken when offering one's words to collective scrutiny, as if they were abandoned children left to the interpretation of others. Far from seeking to refute other authors or to prove a hypothesis, the text unfolds through free writing that endeavors to characterize the philosophical enterprise and to reflect on its meaning as a living exercise of thought, communication, and self-knowledge.

Keywords: *Reading, writing, metaphilosophy, philosophical education, critical thinking*

Introducción

Para muchos estudiantes, la aproximación a la lectura y a la escritura de textos filosóficos constituye un desafío considerable. Leer textos filosóficos implica un ejercicio distinto y más exigente que la lectura de otros textos argumentativos o literarios. La filosofía exige no solo entender argumentos complejos, sino también formular juicios propios con precisión y responsabilidad; escribir filosóficamente conlleva someter esas opiniones al escrutinio público de quien las lea. Platón, a través de la voz de Sócrates en el Fedro, ofrece una imagen elocuente de esta condición: el texto escrito queda fijado y silencioso, incapaz de defenderse o aclararse, como hijos que uno ha engendrado sin cuidado y que quedan abandonados a su suerte (Platón, *Fedro* 275c–e). Esa metáfora subraya la vulnerabilidad de lo escrito y la exigencia que supone dejar bien plasmado aquello que se piensa.

A partir de esa dificultad inicial se hace evidente que la escritura no es un mero trámite técnico, sino un instrumento formativo y epistemológico central. A diferencia del dogmatismo que Sócrates denuncia, la práctica de redactar obliga a revisitar y poner a prueba nuestros juicios pasados y presentes; nos obliga a explicitar supuestos implícitos, ordenar razones y anticipar objeciones. En consecuencia, leer y escribir filosofía no tiene una importancia meramente pedagógica, sino que constituyen procedimientos mediante los cuales se configura nuestra identidad intelectual, pues en la página se reconstruyen y articulan saberes, experiencias y preocupaciones personales.

Desde esta función constitutiva conviene atender también a la dimensión íntima del texto, ya que la escritura articula la propia subjetividad del autor. En la lectura nuestra intimidad encuentra un aliciente que, aunque breve, permite reconocernos en algo que creemos que tiene valor y merece ser escrito. La subjetividad aparece en el texto como un espejo donde afloran convicciones, dudas y vacíos; la escritura toma esas impresiones y las transforma en proposiciones articuladas y comunicables. El autor no escribe solo para sí mismo ni se entrega a un

solipsismo contemplativo, sino que proyecta su voluntad en el lenguaje con la esperanza de que otro reconozca en sus palabras una verdad o quede persuadido por ellas. Si nadie más existiera en el mundo, escribir o tal vez incluso el usar un lenguaje perdería gran parte de su razón de existir. Por eso la obra es al mismo tiempo expresión íntima y oferta pública, y exige la precisión y la responsabilidad necesarias para someter las propias ideas al examen intersubjetivo.

La responsabilidad pública que conlleva publicar una obra filosófica remite a una exigencia metodológica esencial presente en cualquier estudio. Antes de emprender una crítica conviene reconstruir fielmente la doctrina ajena. Este ejercicio debe propiciar un verdadero encuentro de ideas, evitando trivializar las posiciones y procurando una exposición que sea fiel tanto a la intención del autor como al contexto histórico y político en que se inscriben. Aun en la crítica debemos ser capaces de familiarizarnos con el cometido del autor e intentar presentar cada doctrina en su mejor versión antes de someterla al cuestionamiento, pues este es un requisito de rigor. Para ello resulta indispensable haber leído y aprendido al autor, valorar su empeño por formular una idea bien cuidada y ser consciente de las objeciones previsibles, tal como señala Rawls al abogar por reconstruir las teorías de los pensadores de la manera más fuerte posible (Rawls, 2007, p. xiii).

Por otra parte, ni el estudiante ni el profesor se aproximan a la filosofía con la expectativa de acumular en la memoria un repertorio enciclopédico de ideas, pues su auténtica asimilación exige ejercitar el talento de la razón. La máxima "se hace filosofía filosofando" reafirma que se trata ante todo de una actividad. En este horizonte, la escritura se presenta como una práctica filosófica en sí misma. Al exteriorizar la subjetividad con sus convicciones, dudas e inquietudes, el autor inaugura un intercambio dialéctico en el que convergen lectores, tradiciones y el propio lenguaje. Dicho intercambio constituye, en último término, la esencia misma del filosofar.

El ejercicio de la escritura filosófica no se circunscribe a un reducido círculo de especialistas encerrados en torres de marfil, sino que constituye una empresa pública que exige claridad, modestia y rigor intelectual. En palabras de Ortega y Gasset, la claridad debe entenderse como una cortesía hacia el lector: "He de hacer el más leal esfuerzo para que a todos ustedes, aun sin previo adiestramiento, resulte claro cuánto diga. Siempre he creído que la claridad es la cortesía del filósofo" (Ortega y Gasset, 1958, p. 7). Frente al hermetismo que pretende imponer un lenguaje opaco como signo de erudición, el esfuerzo por expresar el pensamiento con transparencia no lo empobrece, sino que lo somete a escrutinio crítico, al contraste de perspectivas y, en última instancia, a su transformación colectiva.

La simplicidad del lenguaje, sin embargo, no equivale a superficialidad. Nietzsche observa que la verdadera hondura intelectual se manifiesta en la claridad, mientras que la oscuridad deliberada suele ser un artificio para aparentar profundidad: "Quien se sabe profundo, se esfuerza por alcanzar la claridad; quien pretende parecer profundo a la multitud, se esfuerza por alcanzar la oscuridad" (Nietzsche, 2009, p. 461). Pero el filósofo alemán también advierte que ideas complejas requieren un tiempo de maduración antes de volverse plenamente legibles (Nietzsche, 2019, p. 36). La incomprensión inicial no siempre revela falta de gusto, sino, en ocasiones, la necesidad de elaborar conceptualmente el material hasta hacerlo comunicable. Esta tensión entre claridad y complejidad no debe resolverse mediante una simplificación empobrecedora. La escritura filosófica no es solo una cuestión de estilo; las ideas complejas, como las tratadas en filosofía, pueden ciertamente expresarse en términos del lenguaje ordinario cuando ello sea posible, pero en muchas ocasiones resulta necesario romper con las convenciones lingüísticas hegemónicas para dar cuenta de fenómenos conceptuales que el lenguaje cotidiano no logra capturar adecuadamente. De aquí se desprende una doble exigencia para quien escribe filosofía: por un lado, buscar la mayor claridad posible; por otro, preservar la complejidad

inherente a ciertos conceptos, guiando al lector en su acceso mediante ejemplos precisos, descomposiciones conceptuales y metáforas cuidadosamente dosificadas.

La cuestión de la claridad constituye un paradigma que todo estudiante debe interiorizar; para aspirar a escribir con sofisticación, primero debe aprender a ordenar y reconstruir ideas que no le son propias. Desde los ejercicios más modestos, como el protocolo, hasta el comentario y el análisis crítico, cada estadio constituye un paso necesario para dominar este quehacer. La escritura filosófica clara refleja una mente igualmente clara, que comprende el lenguaje no solo como instrumento retórico, sino también como expresión de una disciplina intelectual basada en el orden conceptual y la fidelidad argumentativa (Martinich, 2016, p. 17). Desde esta perspectiva, la modestia del filósofo se revela como activa y constructiva: no consiste en exhibir padecimientos personales ni en regodearse en la dificultad, sino en decidir deliberadamente cómo presentar las razones y en asumir la responsabilidad pública que implican sus afirmaciones.

Ahora bien, aunque la claridad constituye un criterio de deseabilidad con el que se rige buena parte de la disciplina filosófica, el verdadero desafío aparece en el ámbito de su transmisión, ya que la formación académica, tanto universitaria como secundaria, continúa mostrando dificultades para convertir esa exigencia en una práctica constante y sostenida. La elocuencia, ciertamente, es un talento excepcional y escasamente innato; pero la filosofía no demanda poetas, sino autores capaces de responder a las estrictas exigencias formales del discurso argumentativo. No es raro encontrar trabajos que, pese a dar testimonio de un esfuerzo intelectual genuino, se aproximan más a la divulgación que a la investigación, pues en la propia literatura proliferan propuestas expositivas carentes de intención crítica o de vocación por contribuir al proceso de producción del conocimiento. Esa confusión de géneros termina por empobrecer la disciplina, ya que, aunque la difusión

constituye una labor valiosa, no puede sustituir la articulación rigurosa de una tesis, la reconstrucción crítica de los argumentos ni la atención minuciosa a las condiciones de validez que demanda toda proposición filosófica. Formar una opinión informada, desarrollar una crítica lúcida, asumir una postura filosófica o contribuir con solvencia a una disputa pública son estadios superiores del pensamiento que exigen reflexión pausada y sin atajos. La filosofía se vitaliza cuando quienes la practicamos formulamos preguntas originales y pertinentes. Los diálogos expositivos cumplen una función pedagógica insustituible, aunque requieren ser complementados con ejercicios que desarrollen la capacidad de pensar con rigor, analizar supuestos y someter los argumentos a examen y refutación.

No menos decisivo resulta atender a las condiciones concretas de producción del pensamiento. Un elemento que suele pasarse por alto es el valor de la lectura cuidadosa y la revisión sistemática de los borradores; ese trabajo silencioso es a menudo la diferencia entre una idea inacabada y un argumento verdaderamente sólido. La filosofía no es el oficio de un ermitaño; exige conversación y comunitariedad. Como señala Lipman, la comunidad de investigación se caracteriza por compartir procesos y compromisos orientados a alcanzar objetivos comunes (Lipman, 2003, pp. 83–84). Esta dimensión comunitaria del filosofar implica una exploración que crece mediante el diálogo entre distintas individualidades, donde la exigencia de entregar algo novedoso y directo hace que el argumento de uno se convierta en el motor de superación de otro. Compartir ideas con colegas, someterlas a la crítica incluso de quienes nos son cercanos, y aceptar la corrección honesta de amigos que no rehúyen la franqueza, son prácticas esenciales para detectar fallos, matizar supuestos y afinar los argumentos. Solo en la interacción sostenida, como en seminarios, coloquios y procesos de evaluación, la argumentación se enraíza en el diálogo con los pares y la disciplina demuestra su vigor intelectual y su pertinencia social. Por eso conviene agradecer públicamente a quienes comentan nuestros textos con rigor y

generosidad, y reconocer la labor de los evaluadores anónimos y editores, pues en su función de curadores de la calidad académica sostienen los estándares que permiten que la conversación filosófica prospere.

Adentrarse en la lectura filosófica implica internarse en un vasto horizonte de escenarios y disputas. La pluralidad de estilos que atraviesa la historia del pensamiento, desde la dialéctica platónica y el método cartesiano hasta los aforismos de Nietzsche y la austeridad analítica de Wittgenstein, pasando por la oscuridad evocadora de Heidegger y la prosa comprometida del existencialismo sartreano, conforma tradiciones casi autónomas que erigen verdaderas arquitecturas conceptuales. Términos como voluntad de poder, dialéctica, *cogito ergo sum*, fenomenología, ideología o contrato social no constituyen un mero repertorio técnico; son nodos semánticos que han ido modelando maneras distintas de plantear preguntas, de construir problemas y de formular respuestas. Leer filosofía exige por tanto no solo atención al contenido, sino una particular sensibilidad hermenéutica para reconocer géneros, contextos históricos y presupuestos disciplinarios, y con ello poder reconstruir las alternativas conceptuales desde las que cada autor piensa y habla.

Acceder a este universo temático e histórico exige adentrarse en el canon filosófico occidental, un acervo que se despliega a lo largo de cerca de veintiséis siglos y que hunde sus raíces en el *logos* griego. Ello no equivale a acumular información de interés erudito ni a trazar una mera cronología de autores y doctrinas; más bien, supone abrirse a un depósito vivo de ideas y problemas fundamentales sobre los que descansa la arquitectura intelectual de la civilización. Como advierte Giannini, la filosofía "constituye un depósito de ideas, en cierto sentido platónicas, sobre las que está cimentado el proceso y el orden de todo el mundo civilizado. Filosofar, dentro de tal presupuesto, es el acto de volverse continuamente hacia aquellas ideas, y hacia los textos en que han sido fijadas de una vez para siempre" (Giannini, 1978, p. 26). De ahí que el

estudio de la historia de la filosofía signifique, ante todo, incorporarse a una conversación siempre inconclusa, donde ninguna cuestión queda definitivamente clausurada. Pensar en el presente exige situarse en la trama de debates que generaciones anteriores abrieron y transmitieron; de este modo, la tarea filosófica consiste en prolongar, matizar e incluso transformar una discusión que nunca se extingue, sino que se renueva incesantemente.

Esta visión, sin embargo, forma parte de una hegemonía intelectual que merece ser cuestionada. Existen autores que van más allá de las categorías tradicionales de la filosofía occidental para desarrollar un pensamiento propio que testimonia una suerte de emancipación intelectual. Pensadores latinoamericanos como Dussel, Zea, Roig y de Sousa Santos, así como tradiciones orientales representadas por Confucio, Buda y Lao-Tsé —más antiguas incluso que la propia filosofía griega— son muestra de que el canon occidental no agota el universo de la reflexión filosófica. Reconocer esta pluralidad de voces y tradiciones no implica abandonar el estudio riguroso del canon occidental, sino más bien situarlo en un horizonte más amplio y admitir que las arquitecturas conceptuales desde las cuales se filosofa son múltiples y que cada tradición ofrece recursos singulares para abordar problemas fundamentales. La filosofía contemporánea se enriquece cuando dialoga con estas tradiciones diversas y cuando se abre a la posibilidad de que existen otros modos de plantear y resolver las cuestiones que nos conciernen como seres humanos.

Esta discusión siempre abierta pone de relieve la modestia inherente al filosofar, que consiste en reconocer los límites cognoscibles que condicionan muchos de los problemas abordados por la disciplina. Tales problemas, en gran medida, desbordan el horizonte de las respuestas definitivas o plenamente satisfactorias. A diferencia de las ciencias empíricas, dotadas de métodos experimentales rigurosos para contrastar hipótesis, la filosofía carece de procedimientos equivalentes

que permitan verificar de manera directa sus supuestos y proposiciones. Ello no significa, sin embargo, que deba descartarse toda especulación abstracta o metafísica como un incesante "derroche de tinta", tal como sostuvo con escepticismo David Hume al examinar críticamente la tradición racionalista. El gesto humeano, pionero en esta actitud de sospecha, anticipa lo que dos siglos más tarde encontrarían su expresión más sistemática en la filosofía analítica temprana de Wittgenstein y en el positivismo lógico, corrientes que llevaron este rechazo al extremo al considerar carentes de sentido todos los enunciados que no fueran verificables empírica o lógicamente.

En este contexto la noción socrática de *aporía* resulta especialmente pertinente. El reconocimiento explícito de la propia ignorancia y la perplejidad ante problemas complejos constituyen, en la *praxis* filosófica, un procedimiento esencial para poner en juicio todas nuestras certezas y abrirse a una investigación genuina. La célebre perplejidad de Sócrates frente al oráculo de Delfos, que lo proclamó "el más sabio entre los atenienses" pese a su conciencia de no poseer saber, ejemplifica esta actitud epistemológica (Platón, *Apología* 21a–23c). De este modo la filosofía se presenta como una invitación ineludible a implicarse profundamente con los problemas, no con la expectativa de hallar soluciones fáciles o definitivas, sino con la disposición a sostener un ejercicio constante de cuestionamiento, análisis y reconstrucción crítica de los conceptos y creencias que orientan nuestra comprensión del mundo y de nosotros mismos. Comprender y aceptar que no existe una doctrina final forma parte de la madurez intelectual que la disciplina procura cultivar.

La perplejidad del filósofo no busca lo insólito, sino que se detiene en lo familiar, aquello que por cotidiano suele escapar al examen. Las preguntas filosóficas, a menudo semejantes a las de un niño, surgen de advertir que lo ordinario encierra problemas conceptuales que tendemos a dar por resueltos sin haberlos pensado con detención (García, 1999, p.

21). Desde esta perspectiva cobra pleno sentido la noción wittgensteiniana de la filosofía como terapia. La imagen de la mosca en la botella, que por más que intenta salir queda atrapada por un cristal que le impide ver la salida, ilustra un tipo de labor intelectual cuya finalidad no es erigir teorías grandilocuentes sino deshacer las confusiones que nos atrapan. Para Wittgenstein la filosofía consiste en desenredar los enredos lingüísticos y conceptuales que originan nuestros problemas, y así aclarar el terreno en que se plantean las cuestiones (Wittgenstein, 2017, p. 203).

Sin embargo, pese al aprecio que he concedido a la filosofía a lo largo de este ensayo, la disciplina afronta la amenaza persistente de ser marginada en las instituciones educativas. Desde el *Conflicto de las Facultades* (1798/1963) de Kant, la filosofía ha sido con frecuencia relegada a la condición de "saber de segunda categoría", de modo que muchos planes de estudio la consideran prescindible. No obstante, quienes nos dedicamos a las humanidades hemos defendido su enseñanza con firmeza, poniendo de relieve su contribución irreemplazable al cultivo del pensamiento reflexivo y crítico. Como recuerda con perspicacia Carlos Peña en sus frecuentes intervenciones públicas, la filosofía no es una entretención para tener una buena sobremesa.[3]

El valor de la filosofía no reside en una utilidad meramente instrumental o productiva, sino en la capacidad de enriquecer el capital humano y de formar a ciudadanos aptos para la deliberación pública. No se trata únicamente de una labor teórica destinada a desentrañar el sentido de nuestros conceptos fundamentales; su horizonte trasciende el aula. La filosofía no concluye cuando el profesor cierra su libro; perdura en la formación de sujetos dotados de herramientas intelectuales que les permiten intervenir con juicio y responsabilidad en el debate democrático. De igual forma, la filosofía no es un lujo reservado a quienes no padecen necesidades materiales ni es un privilegio de puras élites

[3] Véase: https://www.cnnchile.com/lodijeronencnn/carlos-pena-becas-chile-humanidades_20250601/

intelectuales, como a veces se sugiere al evocar ciertos rasgos de la antigua Grecia. Por el contrario, es necesario explorar con rigor su función social y reconocer que la reflexión crítica constituye un derecho y una necesidad para todos los miembros de una comunidad política.

Hoy, el asedio contra la filosofía y las humanidades ha adquirido nuevos matices. La hegemonía del neoliberalismo y la preponderancia de los saberes técnicos han impulsado una visión utilitarista de la educación que valora únicamente los resultados cuantificables. A esa lógica se suma la arrogancia de quienes, ajenos a la disciplina, la desestiman por considerarla inútil, así como la indiferencia de quienes no comprenden el papel público de las humanidades.[4] Al mismo tiempo, la crisis de la universidad contemporánea, subsumida a la lógica de producción, ha convertido a muchas instituciones en fábricas de *papers* que privilegian la productividad bibliométrica antes que la calidad de trabajos o tratados con ideas bien maduradas. Esta presión por publicar de manera constante y acelerada conspira contra la reflexión pausada y profunda que caracteriza el mejor pensamiento filosófico. Sin embargo, esta crítica falla en su medición de la utilidad. Como observa Carlos Peña: "con la filosofía nada puede hacerse; aunque acto seguido hay que preguntar si ella podría hacer algo por nosotros, y la respuesta me parece rotundamente positiva [...] el lugar de la filosofía en la universidad debe estar asegurado, porque si la universidad moderna debe cultivar los saberes técnicos, el único modo de hacerlo sin dejarse vencer por el ánimo supersticioso que suele acompañarlos, es contar a la mano con el antídoto de la filosofía" (Peña, 2018, pp. 186–187).

[4] Me refiero, en concreto, al tipo de declaraciones que menosprecian el valor de las humanidades, como la formulada por Sebastián Edwards en el programa Mirada Económica de CNN Chile, cuando propuso: "Yo cerraría las Becas Chile en humanidades por 10 años. Solo las otorgaría a estudiantes de ingeniería aplicada" (sic.) Véase: https://www.cnnchile.com/programas-completos/edwards-asegura-que-no-hay-que-darle-facilidades-a-carreras-que-no-tienen-futuro-cerraria-las-becas-chile-en-humanidades_20240612/

El planteamiento de Peña subraya que la formación profesional no puede reducirse a saberes técnicos; la educación debe formar sujetos críticos capaces de analizar y evaluar las ideas que orientan la vida social. En este sentido las humanidades y, en particular, la filosofía, resultan insustituibles, pues aportan las herramientas para clarificar supuestos, distinguir argumentos sólidos de meras elocuciones y proteger la esfera pública frente a prejuicios. Esta defensa encuentra eco en David Hume. En plena Ilustración escocesa se sostenía que "una ventaja considerable que se deriva de la filosofía consiste en el antídoto supremo que ofrece contra la superstición y la falsa religión" (Hume, 1980, p. 97). Para Hume, la filosofía no es un entretenimiento ni la sierva de la religión o de otras ciencias; antes bien, funciona como una clarificación racional que inmuniza frente a la ceguera ideológica y el fanatismo. Al mismo tiempo, prepara a los individuos para el ejercicio del juicio crítico, y contribuye a preservar la esfera pública de dogmas y prejuicios, fortaleciendo así la vocación emancipadora de la educación. En este sentido, la filosofía no debe estar desconectada de la realidad, sino que debe constituir un medio privilegiado para comprenderla críticamente y transformarla. Explorar las potencialidades de la filosofía implica reconocer su capacidad para iluminar las condiciones materiales, políticas y simbólicas que estructuran nuestra existencia colectiva, y para ofrecer herramientas conceptuales que permitan imaginar y construir formas de vida más justas y dignas.

A modo de cierre, conviene subrayar que interrogar qué son la lectura y la escritura filosóficas equivale a cuestionar la propia naturaleza de la filosofía. Toda filosofía, de manera explícita o implícita, comporta una dimensión metafilosófica en tanto presupone ciertos objetos, métodos y fines. De ahí la dificultad de ofrecer una definición unívoca de la disciplina, pues en ella conviven tradiciones y procedimientos diversos —el análisis del lenguaje, la fenomenología, la historia de las doctrinas, la ética o el existencialismo—, cada uno con criterios propios de legitimidad y alcance. Sin embargo, esa pluralidad constituye también la riqueza de la filosofía: nos obliga a buscar una comprensión de ella misma, a revisar

constantemente sus supuestos y a reformular sus marcos conceptuales. En consecuencia, enseñar y practicar la lectura y la escritura filosóficas implica cultivar una disposición crítica que, en último término, consiste en filosofar sobre la filosofía.

Volviendo a la metáfora socrática del *Fedro* con la que iniciamos este recorrido, resulta preciso matizar aquella imagen de los hijos abandonados a su suerte. Si bien es cierto que el texto escrito queda fijado y silencioso, incapaz de defenderse ante cada lector, ello no significa que debamos resignarnos a una suerte de orfandad irresponsable. Por el contrario, precisamente porque sabemos que nuestras palabras quedarán expuestas al escrutinio público, debemos prepararlas con toda la cautela y exigencia posibles. La lectura y la escritura filosóficas constituyen el medio mediante el cual equipamos a esos "hijos" para que puedan defenderse por sí mismos: les damos argumentos sólidos, anticipamos objeciones, clarificamos conceptos y tejemos conexiones lógicas que resistan el examen crítico. No se trata de hijos totalmente abandonados, sino de textos cuidadosamente formados con la esperanza de que alcancen vida propia en el diálogo intelectual.

Más aún, esos textos que entregamos al mundo no se alejan definitivamente de nosotros como hijos que emigran sin retorno. Volveremos a revisitarlos en más de una oportunidad, los releeremos con ojos distintos, veremos cómo han sido recibidos, interpretados y cuestionados por otros. En ese proceso de retorno descubriremos cómo nuestras propias ideas nos han formado, cómo el acto de articularlas nos ha transformado, y cómo participan en la formación de nuevas generaciones de lectores y pensadores. La filosofía es así un saber productivo, pero entiende los términos de la producción desde una lógica distinta a la instrumental o mercantil: no produce mercancías desechables ni resultados cuantificables, sino ideas vivas que germinan en el tiempo, que se transforman en el diálogo y que contribuyen a la construcción colectiva del sentido.

REFERENCIAS

García, A. (1999). Hacia una caracterización de la filosofía analítica. *Cuaderno Gris*, *4*, 13–32. https://files01.core.ac.uk/download/pdf/71859262.pdf

Giannini, H. (1978). Experiencia y filosofía. *Revista de Filosofía*, *16*(1-2), 25–32. https://doi.org/10.5354/0718-4360.1978.44438

Hume, D. (1980). *Dialogues concerning natural religion, and the posthumous essays Of the immortality of the soul and Of suicide* (R. H. Popkin, Ed.). Hackett.

Kant, I. (1963). *El conflicto de las facultades* (E. Tabernig, Trad.). Losada (Obra original publicada en 1798).

Lipman, M. (2003). *Thinking in education* (Second ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511840272

Martinich, A. (2016). *Philosophical writing: An introduction* (Fourth ed.). Wiley Blackwell.

Nietzsche, F. (2009). La ciencia jovial. En *El nacimiento de la tragedia; El caminante y su sombra; La ciencia jovial* (G. Cano, Trad., pp. 158–305). Gredos.

Nietzsche, F. (2019). *La genealogía de la moral: Un escrito polémico* (A. Sánchez Pascual, Trad.). Alianza Editorial.

Ortega y Gasset, J. (1958). *¿Qué es filosofía?* Revista de Occidente.

Peña, C. (2018). *¿Por qué importa la filosofía?* Taurus.

Platón. (1985). Apología. En *Diálogos I* (E. Calonge Ruiz, C. Lledó Íñigo y C. García Gual Trads., pp. 137–186). Gredos.

Platón. (2019). Fedro. En *Diálogos: Fedón, Fedro, Banquete* (L. Gil y M. Araujo, Trads., pp. 165-266). Penguin Clásicos.

Rawls, J. (2007). *Lectures on the history of political philosophy*. Cambridge University Press.

Wittgenstein, L. (2017). *Tractatus logico-philosophicus; Investigaciones filosóficas* (I. Reguera, Trad.). Gredos.

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 178-213
ISSN 3087-260X
Santiago de Chile

PROFUNDIDAD, ESTÉTICA Y LA ACTUAL FILOSOFÍA DE LAS MATEMÁTICAS: un caso de Estudio del Teorema de Picard-Lindelöf[1]

Facundo Camargo Platas[2]
Unai León Rico[3]

Resumen

El objetivo de este manuscrito es ofrecer al lector un acercamiento general al concepto de profundidad en matemáticas, ausente en la literatura filosófica en lengua hispana, y proponer posteriormente una formulación original de la profundidad, fundamentada en la idea de franjas de verdad. El trabajo se divide en tres partes. Después de la

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Camargo, F., & León, U. (2026). Profundidad, estética y la actual filosofía de las matemáticas: un caso de estudio del teorema de Picard-Lindelöf. *Revista de Filosofía Homónima, 1*(1), pp. 178-213. https://doi.org/xxxxxxx[Por asignar]

[2] Facundo Camargo Platas es estudiante de Ingeniería Físico-Matemática y Licenciatura en Matemática de la Universidad de la República (Uruguay). Email: facundo.camargo.05@gmail.com. ORCID: https://orcid.org/0009-0000-0241-8628
[3] Unai León Rico es estudiante de Licenciatura en Filosofía de la Universidad Nacional de Educación a Distancia (España). Email: unaileonrico@gmail.com. ORCID: https://orcid.org/0009-0009-0605-4476

introducción general al tópico del artículo, se presenta una aproximación a la noción de profundidad desde el estado del arte matemático, mostrando su existencia, relevancia y revisando las problemáticas e intentos de solución propuestos por los autores que han abordado esta cuestión. Posteriormente, se desarrolla un estudio de caso en el que se ilustran las características de la profundidad a través del teorema de Picard–Lindelöf. Por último, se examina de qué modo este concepto y la investigación en torno a él dan lugar a nuevos indicios sobre la profundidad, en particular su comprensión como relación interdisciplinaria, los cuales pueden resultar útiles, interesantes y provechosos para la filosofía contemporánea de las matemáticas.

Palabras clave: *profundidad, estética, filosofía de las matemáticas, teorema de Picard-Lindelöf, teorema del punto fijo, metateoría matemática, práctica matemática, teoría de ecuaciones diferenciales ordinarias, axiomas.*

Depth, Aesthetics, and the actual Philosophy of Mathematics: a case study of the Picard–Lindelöf Theorem

Abstract

The aim of this manuscript is to offer the reader a general introduction to the concept of depth in mathematics, which has been largely absent from the philosophical literature in the Spanish language, and subsequently to propose an original formulation of depth grounded in the idea of bands of truth. The work is divided into three parts. After a general introduction to the topic of the article, an approach to the notion of depth is presented from the perspective of the mathematical state of the art, showing its existence and relevance, and reviewing the main problems and attempted solutions proposed by authors who have

addressed this issue. Subsequently, a case study is developed in which the characteristics of depth are illustrated through the Picard–Lindelöf theorem. Finally, the article examines how this concept and the research surrounding it give rise to new indications concerning depth, in particular its understanding as an interdisciplinary relation, which may prove useful, interesting, and fruitful for contemporary philosophy of mathematics.

Keywords: *depth, aesthetics, philosophy of mathematics, Picard–Lindelöf theorem, fixed-point theorem, mathematical metatheory, mathematical practice, theory of ordinary differential equations, axioms.*

1. Introducción

Los filósofos de las matemáticas han tendido a entender esta disciplina como un todo relativamente homogéneo y, en consecuencia, han intentado abordar problemas de carácter extremadamente general.[4] Aunque esta perspectiva es legítima y necesaria, resulta insuficiente para comprender la riqueza de la práctica matemática concreta y el modo en que nuestras creencias matemáticas adquieren justificación. En las últimas décadas, varios autores (De Toffoli, 2021; Franklin, 2014; Lakatos, 2015) han impulsado una corriente que aborda la matemática desde su operatividad, su historia, su desarrollo y el trabajo individual del matemático. Nuestro objetivo no es defender esa corriente —tarea para la cual existe ya una literatura amplia—, sino ejercitarla mediante el análisis de un concepto estético que, desde hace ya una década, ha

[4] Podrían ponerse como ejemplos de este tipo de problemáticas el debate sobre los fundamentos de las matemáticas, las preguntas acerca de la posible reducción de la disciplina (i.e., a la lógica, a la teoría de conjuntos, a la teoría de categorías o a la teoría homotópica de tipos), así como la cuestión de la validez universal de ciertos teoremas lógicos (e.g., el principio del tercio excluso) y las epistemologías que buscan justificación para toda la disciplina matemática (e.g., el holismo empirista de Quine). De Toffoli (2022) y Lakatos (1976) hacen referencia a este fenómeno.

comenzado a estudiarse seriamente y cuya incorporación puede resultar especialmente fructífera: la *profundidad*.

Hasta hace apenas unos años, el concepto de profundidad había recibido muy poca atención. Los matemáticos carecían de una definición formal y los filósofos no habían ofrecido un análisis sistemático del fenómeno. Sin embargo, forma parte de la práctica matemática desde hace largo tiempo, al menos desde la obra de Gauss (Gray, 2015, p. 177). Este término se emplea de manera general para referirse al contenido evaluativo considerado relevante en un resultado matemático (Arana, 2015, p. 163); siendo muy amplio el abanico de entidades a las que puede atribuirse dicha predicación, como teoremas, demostraciones, conjeturas, resultados, métodos, conexiones e intuiciones.

El debate acerca de cuál es realmente el portador de profundidad está lejos de resolverse. Algunos consideran que los portadores son únicamente los teoremas (Cheng, 2022, p. 174), mientras que otros extienden esta noción a las demostraciones (Lange, 2015, p. 198) o incluso a otros tipos de resultados matemáticos. Esta cuestión será retomada más adelante.

2. Desarrollo de la problemática

2.1. Principales interrogantes sobre la profundidad

Aunque pródigamente empleada, la noción de profundidad es extremadamente vaga y problemática. Los autores que se han embarcado en su estudio comparten un conjunto de temas y preguntas fundamentales que toda tesis sobre la profundidad debe abordar (Ernst et al., 2015, p. 155):

1) ¿Se puede llegar a un consenso sobre cuáles son los portadores de profundidad y en qué casos se manifiesta esta propiedad?

2) ¿Existen regularidades entre los distintos ejemplos de profundidad y superficialidad dados por los profesionales?

3) ¿Puede la profundidad reducirse a otros valores como prolificidad, importancia, belleza, elegancia, simplicidad, complejidad, fundamentalidad o poder explicativo?

4) ¿Es la profundidad una propiedad independiente de la mente humana (es decir, objetiva) o está ligada a nuestros intereses, fines, creencias y percepciones? ¿La profundidad matemática está más o menos vinculada a estos rasgos subjetivos que la profundidad en el resto de ciencias naturales?

La mayor parte del debate existente gira, de un modo u otro, en torno a estas problemáticas. El propio concepto de profundidad resulta difícil de ubicar, pues suele situarse en el terreno de la estética, aunque con facilidad se extiende hacia otras áreas como la epistemología, la metodología o incluso la ontología. Estas preguntas son amplias y exigen una perspectiva interdisciplinaria, lo que enriquece enormemente esta empresa.

2.2. Las cuatro concepciones de profundidad

Existen diversas concepciones sobre el significado de la profundidad, o más precisamente, múltiples intuiciones al respecto. Como señalan varios autores, no se dispone de una definición formal del concepto (Arana, 2015, p. 163), e incluso se ha llegado a cuestionar que una definición de ese tipo sea posible (Stillwell, 2015, p. 230). A partir del análisis de la literatura matemática, Arana identifica cuatro concepciones generales de

la profundidad (2015, pp. 167-174).[5] Conviene detenerse y escrutar los elementos generales de dicha taxonomía.

2.2.1. Tesis Genética

Esta visión, en términos propios de la filosofía analítica, se aproximaría a una postura antirrealista. Fundamenta la profundidad en criterios sociológicos, reduciéndose a factores como el talento, el impacto, la relevancia o el renombre de su autor (de ahí el nombre de esta tesis). Responde así al problema de la objetividad y consenso con un claro talante subjetivista: si la profundidad de un teorema, demostración o conjetura depende únicamente de las creencias de los profesionales en un momento histórico determinado, entonces la profundidad es también relativa al tiempo.

Esta tesis resulta problemática, pues implicaría que la profundidad depende del talento del matemático, lo que imposibilita demarcar cuáles de sus resultados son profundos y cuáles no, o bien conduce a aceptar que todos lo son. Además, esta postura no ofrece criterios claros para identificar en qué consiste dicho talento ni métodos que permitan discernir cuándo y por qué un matemático debe considerarse talentoso.

2.2.2 Tesis consecuencialista

Podría argumentarse que los criterios sociológicos son relativamente independientes del matemático que los origina, liberando así a la tesis anterior de parte de sus dificultades. Sin embargo, persiste el problema: ¿cómo explicar que ciertos teoremas se consideren profundos a lo largo del tiempo? ¿Por qué los profesionales juzgan algunos resultados como profundos y otros no?

[5] Su desarrollo proviene del análisis de los comentarios de matemáticos profesionales acerca de la profundidad, diferenciando en cada caso los distintos usos y matices del término. Se trata, por tanto, de un trabajo de análisis lógico del lenguaje empleado por los propios matemáticos, y no de una construcción impuesta externamente por los filósofos.

De no poder responder estas cuestiones, se caería en un cierto escepticismo respecto de la pregunta 2 planteada al inicio. Pero podría pensarse lo contrario, tal vez no sean los criterios sociológicos los que generan la profundidad, sino que es el hecho de que un teorema sea profundo lo que lo lleva a satisfacer dichos criterios.

La tesis consecuencialista adopta precisamente esta inversión. Un teorema o demostración posee la propiedad de profundidad debido a sus consecuencias: su fecundidad, su interés y su capacidad de generar nuevo conocimiento matemático. Se elimina así el intermediario del autor y los criterios sociológicos se aplican directamente sobre el portador de profundidad.

Esta posición, sin embargo, continúa siendo ambigua respecto de las preguntas 1, 2, 3 y 4, pues solo puede defender una noción objetiva de profundidad si asume que esos intereses y valores sociológicos son indicadores de algo que los trasciende. Alternativamente, y siguiendo a Putnam (2002, pp. 5-64), podría sostenerse que los valores son objetivos y que no existe una distinción sólida entre hecho y valor.

Si no es así, el resultado es una visión deflacionada y antirrealista de la profundidad, similar a la de la tesis anterior. Esta postura puede observarse en autores como Maddy:

> Dudo que un intento de dar una caracterización general de lo que 'la profundidad matemática' es realmente resulte productivo; me parece que la expresión se entiende mejor como un término paraguas para los diversos tipos de virtudes especiales que percibimos claramente en nuestros ejemplos ilustrativos de formación de conceptos y elección de axiomas. (Maddy, 1980, p. 81)[6]

2.2.3 Tesis evidencialista

[6] Todos los textos citados presentes en el manuscrito son traducciones de nuestra autoría, a excepción del correspondiente a Manin (1981, 173), cuya traducción del ruso corre a cargo de E. M. Kotenko.

La tesis evidencialista busca una mayor objetividad en su intento de responder a las preguntas 3 y 4. Centra su atención en las demostraciones y en sus propiedades. Algunos autores (Lange, 2015, p. 198) consideran que las demostraciones son portadoras de profundidad, incluso con independencia del enunciado del teorema. Es decir, se propone que puede haber teoremas con enunciados superficiales y demostraciones profundas, o viceversa.

Se podría entender la tesis de Lange (2015) como una versión de la tesis evidencialista, aunque no todo evidencialista debe adoptar esa medida. También puede sostenerse que los teoremas son los portadores de profundidad y que las demostraciones, al ser profundas o no, confieren esta propiedad al teorema.

Las estrategias evidencialistas son diversas. Algunas concepciones (Arana 2015, p. 169) asocian la profundidad con demostraciones complejas, elegantes, difíciles, extensas o impuras. Entre estas características, la impureza ocupa un lugar especialmente relevante. La pureza, a su vez, constituye otro concepto estético problemático, aunque puede describirse de manera sencilla: una demostración es pura cuando no recurre a términos o conceptos ajenos al área en la que se formula el resultado.

Esta concepción presenta serios problemas: existen demostraciones largas pero sencillas, y otras breves pero extremadamente difíciles; también las hay difíciles y puras, o fáciles e impuras. También, y para responder a las preguntas 1, 2, 3 y 4, el teórico evidencialista tendría que establecer jerarquías de valores entre las propiedades de las demostraciones, una tarea particularmente compleja. Además, surgen nuevas interrogantes, como la relativa a si las demostraciones asistidas por ordenador pueden considerarse profundas, del mismo modo que el teorema de los cuatro colores.

Otra alternativa más viable es sostener que la pluralidad de demostraciones impuras contribuye a la profundidad de un teorema, ya que esta acumulación de impurezas y de relaciones entre áreas genera una red de conocimiento entre distintos dominios matemáticos.[7] Esta tesis adopta un enfoque más epistémico y da cuenta de ciertos comportamientos de la práctica matemática, como la pregunta retórica de Madrid Casado (2018b, p. 2), que aquí parafraseamos: ¿por qué los matemáticos ofrecen tantas demostraciones de un mismo teorema? ¿Acaso lo hacen por pasatiempo? Si una sola demostración fuera suficiente para garantizar su validez universal, este fenómeno sería inexplicable. Pero bajo la concepción evidencialista, cobra sentido: se enlazan o agrupan diversos conceptos (Madrid Casado, 2018b, p. 53). Aquí parece reintroducirse el concepto de impureza, pues precisamente la existencia de una pluralidad de demostraciones de distintas áreas constituye una colección de demostraciones impuras. Es precisamente esta impureza la que incrementa la profundidad del teorema. Madrid Casado describe este fenómeno mediante la noción de *franja de verdad*, un concepto un tanto obscuro que alude a la idea de que el "área" o radio de acción de un teorema se amplía a medida que se multiplican sus demostraciones desde dominios diversos. Esta expansión genera nudos en la red del conocimiento matemático, en los que confluyen distintas teorías, métodos y estructuras conceptuales. En la siguiente sección profundizaremos en esta noción y examinaremos hasta qué punto puede extrapolarse a casos en los que tales nudos no solo se producen entre

[7] Este es un caso especialmente ilustrativo. Aunque Madrid Casado (2018a, pp. 109-110), no ha tratado explícitamente el concepto de profundidad, es un filósofo de las matemáticas centrado en las prácticas y metodologías de la disciplina, y ha llegado a conclusiones afines a las que aquí se sostienen. En sus propios términos, la verdad material de un teorema puede "engrosarse" mediante sucesivas demostraciones que enlazan distintas áreas de las matemáticas. Esto sugiere que nociones de este tipo ya estaban presentes, aunque de manera implícita, en autores que adoptan esta metodología de estudio.

áreas matemáticas, sino también entre teorías pertenecientes a disciplinas distintas.

Pese a sus virtudes, la tesis evidencialista presenta varios problemas. Si aceptamos que una pluralidad de demostraciones impuras convierte al teorema en un nodo dentro de la red de conocimiento matemático, surge entonces la cuestión de determinar cuándo la impureza es realmente relevante. Arana, por ejemplo, muestra que es posible construir demostraciones artificialmente impuras cuya impureza no añade nada a la prueba. También permanece abierta la dificultad de precisar los criterios que permiten distinguir cuándo una demostración debe considerarse impura. Si conseguimos sortear estos problemas, o simplemente decidimos quitar el factor de la impureza de las demostraciones, nos queda el problema de la individuación de las demostraciones: no siempre es trivial determinar cuándo dos demostraciones aparentemente distintas son, en realidad, la misma. Aun así, en ciertos casos (e.g., el teorema de Pitágoras o los teoremas del punto fijo), la multiplicidad de vías demostrativas es claramente observable.

2.2.4. Tesis cosmológica

La tesis cosmológica es quizá la más interesante, aunque también la más confusa. Sostiene que la profundidad surge de la aparición de cierta estructura u orden inesperado revelado por el teorema. A pesar de su formulación ambigua, esta postura busca explícitamente una concepción objetiva y realista de la profundidad, en contraste con los enfoques subjetivos, prácticos o sociológicos.

Arana ofrece una formalización particularmente sugerente:

> Supongamos que existe una función S de los enunciados a R que mide el orden expresado por un enunciado, y una relación A << B que se cumple cuando A es mucho menor que B. Entonces podemos decir que un teorema es (S, <<)-cosmológicamente profundo si S(premisas) << S(conclusión). Podríamos interpretar

> S como una medida del grado de orden de las entidades especificadas por un enunciado, o del estado de cosas especificado por un enunciado. (Arana, 2015, p. 173)

Esta formulación permite entender la idea de orden o estructura implícita en la tesis cosmológica, aunque asimismo, Arana reconoce sus limitaciones: ni siquiera el teorema que toma como caso de estudio cumple de forma clara este criterio, y muchos teoremas intuitivamente profundos tampoco lo harían.

Además, Arana no aborda un conjunto de consideraciones que resultan pertinentes. En la lógica, si una inferencia tal que $P \vDash Q$ es válida, ello implica que todo modelo en el que P sea verdadero será también un modelo de Q, de tal manera que $P \wedge \neg Q$ no puede darse. La deducción preserva la verdad entre enunciados. En el plano epistémico, las conclusiones están aseguradas en la medida en que lo están las premisas, junto con la validez de la inferencia. En casos como el de los axiomas, que fueron defendidos como autoevidentes hasta hace apenas unos siglos, la confianza en las premisas se asume como máxima y, en consecuencia, como infalible. La información contenida en Q se encuentra ya implícita en P, de modo que la inferencia deductiva no introduce información nueva ni amplía nuestro conocimiento.

Pero si tomamos en serio la idea de que las premisas "afirman menos" que las conclusiones, surgen entonces varias preguntas: ¿derivan los teoremas profundos de inferencias ampliativas? ¿Serían, en ese caso, teoremas sintéticos? ¿Existen inferencias deductivas ampliativas? A partir de aquí se abren cuestiones adicionales, como la naturaleza misma de la distinción entre deducción e inducción: ¿puede una reducirse a la otra, ya sea entendiendo la deducción como un caso particular de inducción, la inducción como una forma encubierta de deducción, o negando directamente que exista una distinción sustantiva entre ambas? Estas cuestiones, aunque fascinantes, dificultan aún más la defensa de la

tesis cosmológica, pues exigirían un fundamento mucho más sofisticado de la noción de deducción.

Pese a la aparente contraposición entre estas posturas, no parece necesario adoptar una posición rígida. Como sugiere la tesis consecuencialista, la fecundidad, relevancia e impacto de un resultado podrían indicar la existencia de algo más allá de los factores sociológicos, lo que permitiría combinar esta perspectiva con la evidencialista o la cosmológica. De hecho, algunos autores reconocen que emplean simultáneamente múltiples nociones de profundidad en sus análisis, del mismo modo que Lange (2015) distingue entre portadores de profundidad independientes, como los teoremas y las demostraciones. Además, esto podría esclarecer el significado de orden o estructura dentro de la visión cosmológica, pues en ciertos momentos diversos autores parecen referirse a que dicho orden y dicha estructura están dados por relaciones entre proposiciones.

Tal es el caso de la última estrategia evidencialista propuesta: la capacidad de ciertas demostraciones para conjugar distintas ramas de la matemática da lugar a una red de conocimiento que podría ser la responsable de este tipo de orden o de estructura. Este pluralismo podría ofrecer una respuesta a las dificultades particulares de cada tesis, al proporcionar una perspectiva más ajustada sobre un fenómeno complejo.

Sin embargo, también presenta problemas. La combinación de la tesis consecuencialista con la evidencialista exige una concepción objetiva de la profundidad, lo que implica asumir un cierto realismo de valores. Esta exigencia entra en tensión con la noción de impureza defendida por la evidencialista, ya que la práctica matemática ha tendido a valorar más las demostraciones puras que las impuras. En consecuencia, algunas estrategias pluralistas corren el riesgo de generar contradicciones entre los valores que postulan y los que realmente operan en la práctica.

2.3 Sobre los portadores de profundidad

Hemos venido mencionando reiteradamente la cuestión de los portadores de profundidad desde el inicio del artículo. Parece, pues, un buen momento para detenernos en ella y examinarla con mayor detalle.[8]

No parece haber controversia en que los teoremas lo son, todos los autores que han abordado esta problemática coinciden en ello. Pero ¿hay más? Algunos autores (Cheng, 2022, p. 174) limitan la profundidad exclusivamente a los teoremas; otros (Lange, 2015, pp. 198-204), sostienen que también las demostraciones pueden ser portadoras de profundidad, incluso con independencia de la profundidad del teorema que establecen. Algunos hablan de una enorme variedad de portadores de profundidad: intuiciones, conjeturas, métodos, áreas de la matemática, conceptos o definiciones, e incluso ejemplos (Ernst et al., 2015, p. 156).

Ante esta diversidad, cabe preguntarse —siguiendo la máxima de Peirce— qué tanto se modifica nuestra comprensión del fenómeno al adoptar una noción más o menos inclusiva de portadores de profundidad. Consideramos que este problema puede abordarse desde varios ejes. En primer lugar, desde la intuición: ¿qué queremos decir, intuitivamente, con "profundidad"? Un autor podría aceptar distintos portadores de profundidad además de los teoremas si entiende que ello se ajusta mejor a nuestra intuición del término. Esto suele darse, especialmente, en enfoques antirrealistas o subjetivistas, donde la profundidad se interpreta como una forma de expresar éxito, fecundidad o relevancia. Un segundo eje es el problema de la definición. La noción de profundidad ya es, de por sí, vaga y ambigua, y agregar nuevos portadores de profundidad puede volver aún más compleja su delimitación.

Las cuatro tesis discutidas anteriormente, con sus respectivas variantes, fueron concebidas para dar cuenta principalmente de teoremas profundos. En el caso de las demostraciones, estas pueden integrarse con relativa facilidad en las tesis genética, consecuencialista y, con mayor

[8] La profundidad no debe entenderse como un valor booleano. No se es profundo o no. La profundidad es una cuestión de grado y, en tanto tal, un teorema, una demostración, un axioma o una definición pueden ser más o menos profundos en relación con otros.

desarrollo, en la evidencialista. Sin embargo, en la cosmológica surgen dificultades, ya que esta se formula entre teoremas y los enunciados de los cuales se derivan, y la demostración aparece solo como medio o enlace entre ambos. Por ello, aceptar las demostraciones como portadoras de profundidad no resulta especialmente atractivo para un teórico que adopte una versión rigurosa de la tesis cosmológica. La situación no mejora al seguir ampliando la lista de portadores potenciales (conceptos, intuiciones, ejemplos o métodos), pues con cada adición se diluye más el alcance y la precisión de la noción misma de profundidad. La tesis evidencialista deja igualmente de ser factible y, sin otro apoyo, las tesis restantes, genética y consecuencialista, aunque posibles, no logran ofrecer una explicación objetiva de este fenómeno. Parecería, entonces, que todo intento de dar cuenta de la profundidad de manera objetiva y realista —no reducible a otros términos estéticos— requiere ser selectivo con respecto a los portadores de profundidad que se admitan.

2.3.1 Los Axiomas

Existe un tipo de portador de profundidad que apenas ha sido considerado en la literatura (Ernst et al., 2015, p. 156) y que consideramos que merece atención: los axiomas. En este apartado analizaremos su papel como posibles portadores de profundidad.

En primer lugar, cabe preguntarse si un axioma puede ser considerado profundo en alguno de los sentidos previamente mencionados. Esto depende de la tesis epistemológica que adoptemos frente a ellos. Muchos anti-realistas, como los nominalistas (Field, 1988), convencionalistas (Ayer, 1946, pp. 71-86; Carnap, 1937, p. 51) o incluso ciertos realistas aprioristas (Fine, 2005), sostienen que nuestro conocimiento matemático comienza con la elección —en el caso de los antirrealistas— o descubrimiento —en el caso de los realistas— de unos axiomas.

Todo el conocimiento matemático se deduce analíticamente a partir de ellos, y la verdad de los axiomas, junto con la preservación de la verdad en la deducción, constituye la única justificación necesaria para la disciplina (Field, 1988, pp. 57-59; Hahn, 1959, pp. 158-159). Bajo esta concepción, no parece haber un interés significativo ni una motivación clara para calificar a un axioma como profundo, pues o bien se los considera autoevidentes, o bien se los entiende como elecciones convencionales que no implican ningún descubrimiento genuino.

Sin embargo, existen otras perspectivas no aprioristas, de corte cuasi empirista, que defienden el carácter científico no excepcional de las matemáticas y conciben su metodología como mucho más próxima a la de las ciencias naturales. Autores como Lakatos (2015), precursor de esta tradición, se oponen a las concepciones anteriores y, además de presentar virtudes propias, ofrecen un marco especialmente fértil para comprender por qué un axioma podría ser considerado portador de profundidad.

Bajo ciertas concepciones variadas se sostiene, en contra de lo propuesto por las tesis anteriores, que el conocimiento matemático no comienza con los axiomas.[9] Estos autores ponen en duda la distinción entre ciencias formales y naturales, y tratan de mostrar que el conocimiento matemático no posee una ontogénesis muy diferente a la de otras ciencias naturales. Algunos de ellos argumentan que, al igual que en las ciencias naturales, el conocimiento matemático se apoya en hechos o verdades intermedias (Irvine, 1989; Russell, 1907/1973; Shaffer 2017). Este tipo de hechos matemáticos constituye una forma de conocimiento elemental, accesible mediante la manipulación de objetos y la observación de procesos físicos, como ocurre con enunciados del tipo "2 + 2 = 4", o con el conocimiento de las ternas pitagóricas previo tanto a la

[9] Véase, a modo de referencia, una amplia nómina de filósofos de las matemáticas que se comprometen con esta tesis: De Toffoli (2021), Franklin (2014), Gödel (1944), Irvine (1989), Kim (1982), Kitcher (1983), Lakatos (1976), Maddy (1980), Madrid Casado (2018a), Peirce (1992), Putnam (1975) y Russell (1907/ 1973).

demostración del teorema de Pitágoras como a la axiomatización de cualquier sistema formal (Ratner, 2009). Se trata de un conocimiento que puede obtenerse a partir de la experiencia sensorial. Esto podría parecer, a primera vista, una idea extraña, pero con una mirada atenta a la historia de la matemática cobra sentido y permite explicar numerosos fenómenos. Hemos tenido acceso a las tablas pitagóricas y a demostraciones del teorema de Pitágoras antes de contar con axiomas de la geometría; hemos hablado de conjuntos y trabajado con conceptos afines siglos antes del surgimiento de la teoría de conjuntos de Zermelo-Fraenkel; hemos practicado la aritmética básica antes de los axiomas de Peano.

Son múltiples los autores que defienden la capacidad de percibir perceptualmente ciertos hechos matemáticos. Algunos abordan esta cuestión desde la fisiología de nuestros mecanismos perceptivos en relación con los hechos matemáticos (Fogelin, 1995, pp. 162-163; Maddy, 1980). Otros tratan el tema en términos de formas específicas de adquisición del conocimiento matemático (Brown, 2008; De Toffoli, 2021), sosteniendo que este puede darse a través de diagramas, a los cuales otorgan un papel mucho más relevante del que tradicionalmente se les ha reconocido. Asimismo, en la ontogénesis del conocimiento matemático en los niños (Bijeljac-Babic et al., 1993; Dienes, 1960) puede observarse cómo estas verdades intermedias son las primeras en manifestarse durante el aprendizaje, y cómo gran parte del conocimiento matemático parece originarse en medios sensorialmente estimulantes.

Los axiomas de las matemáticas se entienden de una forma similar a las leyes de nuestras ciencias naturales (Irvine, 1989; Russell, 1907/1973). Se infieren a partir de conocimiento previo, especialmente de las llamadas verdades intermedias mencionadas anteriormente. Del mismo modo en que las leyes de nuestras teorías científicas "salvan el fenómeno" a partir de la experiencia sensorial previamente obtenida, el rol epistémico de los axiomas surge de la necesidad de justificar deductivamente un

conocimiento intermedio anterior a esos mismos axiomas. Como señala Gödel:

> Me parece que la suposición de tales objetos es tan legítima como la suposición de cuerpos físicos, y que hay tantas razones para creer en su existencia como en la de estos últimos. Son, en el mismo sentido, necesarios para un sistema satisfactorio de las matemáticas como los cuerpos físicos lo son para una teoría satisfactoria de nuestras percepciones sensibles. (Gödel, 1944, p. 128)

Paradójicamente, bajo esta consideración del rol de los axiomas en la ontogénesis matemática, los principios de las matemáticas suelen ser, en muchas investigaciones, lo que obtenemos al final de las mismas. Así, los axiomas se revelan como una forma de conocimiento adquirido, algo que descubrimos y que emerge de nuestras prácticas y de nuestra interacción con los objetos matemáticos. En consecuencia, difícilmente puedan considerarse triviales, arbitrarios o simplemente evidentes: los axiomas son, en sí mismos, portadores de profundidad.

Ahora bien, cabe preguntarse: ¿son los axiomas profundos del mismo modo que los teoremas o las demostraciones? Y si no lo son, ¿en qué sentido lo son? Si compartieran la misma forma de profundidad, ¿qué tesis sobre la profundidad sostendría su admisibilidad?

Una vía clara es la tesis consecuencialista: los axiomas son profundos en la medida en que sus consecuencias para la disciplina lo son. En matemáticas es fácil ver cómo la adopción o la omisión de un axioma produce diferencias sustantivas en una teoría. Por ejemplo, la teoría de conjuntos de Zermelo-Fraenkel sin teorema de elección, la teoría de conjuntos intuicionista o las teorías de conjuntos no bien fundadas muestran marcos teóricos con consecuencias muy distintas al modificar o prescindir de ciertos axiomas. El descubrimiento de un axioma, o de alguna de sus propiedades, puede tener una gran relevancia

para la disciplina. Un ejemplo es la constatación de que el quinto postulado es independiente de los demás postulados de la geometría euclidiana, lo cual permitió el desarrollo de una amplia variedad de geometrías no euclidianas.

Podríamos intentar aplicar una concepción evidencialista, pero surge inmediatamente un problema: esta postura depende de las demostraciones, y los axiomas, en principio, no se demuestran. Al menos, no se demuestran en un sentido formal o estrictamente deductivo. Sin embargo, si seguimos la línea de pensamiento de los autores mencionados, pueden explorarse otras vías para fundamentar su papel. Manin, por ejemplo, señala:

> Las representaciones intuitivas acerca de los conjuntos, a los cuales recurrimos en la exposición anterior construyendo el universo V, son materia primaria. El lenguaje L_1 Set ha sido ideado para escribir textos formales equivalentes a nuestros razonamientos no formales respecto de V. Los axiomas de L_1 Set (incluso los axiomas lógicos) se han obtenido como resultado del análisis de las demostraciones no formales. (Manin, 1981, p. 173)

Aunque resulta problemático construir un criterio evidencialista o cosmológico a partir de razonamientos no formales —ya que el proceso de individuación de demostraciones formales es de por sí complejo, y más aún lo sería con demostraciones informales—, sigue siendo posible considerar el modo en que estos razonamientos generan una red de conocimiento matemático, especialmente cuando logran articular saberes procedentes de distintas áreas.

Algunos podrían sostener que el método de adquisición de axiomas descrito puede entenderse en términos abductivos (Irvine, 1989, p. 322). En este sentido, cabe preguntarse si las demostraciones abductivas creativas podrían constituir una vía para dotar de profundidad a un axioma. En la tesis cosmológica ya nos hemos encontrado con formas

de inferencia singulares, deductivas y ampliativas, o al menos aparentemente deductivas.

Aún existen otras estrategias para dar cuenta de la profundidad de los axiomas. Aunque no podamos demostrarlos formalmente, sí podemos demostrar propiedades acerca de ellos. Por ejemplo, sabemos que el axioma de elección es independiente de la teoría de conjuntos de Zermelo-Fraenkel, y que por ello podemos trabajar tanto con él como sin él, obteniendo teorías distintas. De manera análoga, el descubrimiento de la independencia del quinto postulado de Euclides con respecto a los demás postulados de la geometría, por autores como Gauss, Lobachevsky y Bolyai, dio lugar a las geometrías no euclidianas.

Podemos sostener, entonces, que los axiomas con más propiedades demostradas poseen una profundidad superior. Estas características admiten al menos dos lecturas. Por un lado, una interpretación consecuencialista, según la cual tales propiedades generan repercusiones a nivel disciplinar, impulsando nuevos programas de investigación o transformando los ya existentes. Por otro lado, una interpretación evidencialista, dado que la capacidad de vincular un axioma con distintas teorías amplía su red de relaciones y su alcance conceptual.

Por último, las consecuencias de aceptar o rechazar un axioma, así como la manera en que estas decisiones influyen en teorías particulares, pueden considerarse también un indicador de profundidad en sentido cosmológico. En casos como el descubrimiento de la posibilidad de construir geometrías consistentes negando el quinto postulado, puede hablarse de un hallazgo acerca del orden, la estructura y las propiedades de la geometría, tanto euclidiana como no euclidiana. En este contexto, podría reinterpretarse la relación (S, <<) aplicándola al axioma en lugar de a las premisas, y al sistema formal resultante en lugar de a la conclusión.

3. Nuestra noción de profundidad y el teorema de Picard-Lindelöf

Habiendo introducido y examinado el concepto de profundidad, resulta oportuno intentar dar cuenta de un caso concreto de profundidad asociado a un resultado matemático específico con conexiones muy fuertes con la física.

En la sección anterior se mencionó la tesis de la *franja de verdad* propuesta por Madrid Casado; en lo que sigue intentaremos refinarla y mostrar que puede entenderse como una aproximación a nuestra noción de profundidad.

En esta tesis se distingue el valor verificativo funcional de una proposición —es decir, su verdad formal dentro de un sistema determinado— de lo que podríamos denominar su rango de acción o contexto determinante. El rango de acción puede caracterizarse como la propiedad de una determinada verdad de ser efectivamente aplicable y operativizable, conservando un rol inferencial o estructural en una pluralidad de dominios o marcos lingüísticos. Cuanto mayor sea el número de dominios en los que dicha verdad desempeña ese rol, mayor será su rango de acción.

Conviene no confundir esta noción con los casos en los que una misma proposición resulta verdadera en una teoría y falsa en otra, como ocurre con el teorema de Pitágoras en geometrías no euclidianas, o con la identidad "1+1=1" en ciertas álgebras de Boole frente a su invalidez en la aritmética clásica. En estos ejemplos se produce un cambio en el valor verificativo funcional debido a la modificación del marco lingüístico o teórico. La noción que aquí nos interesa, en cambio, no remite a variaciones en el valor de verdad, sino a la transferencia conceptual de un resultado entre distintos dominios.

Un teorema puede conservar su estatuto formal y, aun así, carecer de sentido operativo fuera del marco para el cual fue formulado. Así, preguntar por la validez de la Segunda Ley de Newton en la aritmética carece de sentido, no porque dicha ley sea falsa, sino porque el dominio de cuantificación de dicha ley incluye magnitudes físicas y no entidades aritméticas. En contraste, ciertos resultados matemáticos, como el teorema de Pitágoras, poseen una estructura lo suficientemente abstracta como para ser reinterpretados y utilizados en marcos distintos del original, tal como ocurre en el álgebra o en el análisis funcional, donde reaparecen como propiedades de normas o productos internos. En estos casos, el enunciado amplía su rango de acción sin que ello implique una modificación de su valor verificativo funcional, por motivos que se explicarán más adelante.

Esta idea es expresada con claridad por Madrid Casado cuando afirma:

> "Cuando los teoremas se demuestran mediante cursos de operaciones heterogéneos se convierten en auténticos nudos de la red matemática, en el sentido de que su franja de verdad aumenta (no porque su valor de verdad formal cambie, sino porque su verdad en sentido material engrosa). Así, el Teorema del Punto Fijo de Brouwer fue originariamente demostrado usando técnicas de topología algebraica, pero su conversión en piedra angular de la matemática moderna se fraguó gracias a las demostraciones usando teoría de retractos y análisis, que permitieron su empleo en otros dominios." (Madrid Casado, 2018a, pp. 109-110)

Madrid Casado apela a que las demostraciones de un mismo teorema realizadas desde distintas áreas de las matemáticas producen un engrosamiento de su franja de verdad, en la medida en que permiten concatenar, engranar o articular fenómenos inicialmente dispares bajo un mismo enunciado. Este proceso da lugar a un aumento del dominio de

aplicación de la verdad del teorema, dando lugar al nudo en esta red de conocimiento.[10]

Sostenemos que la noción de franja de verdad, entendida de este modo, constituye una de las formas más adecuadas de captar la profundidad matemática, al menos en el caso de los teoremas. Pues esta tesis permite articular en una única noción los enfoques evidencialista, cosmológico y consecuencialista. Por un lado, la franja de verdad ofrece una interpretación intuitiva del orden y la estructura que generan ciertos resultados matemáticos, fundamentada en los métodos de demostración que los sustentan; en particular, en aquellos cursos de operaciones heterogéneos que, en nuestros términos, corresponden a demostraciones impuras. Por otro lado, esta concepción permite explicar por qué ciertos teoremas poseen una repercusión tan notable dentro de la práctica matemática, algo que en la tesis consecuencialista aparecía, en gran medida, como un hecho convencional o sociológicamente determinado.

Desde esta perspectiva, se puede dar cuenta objetivamente al por qué son precisamente ciertos resultados los que alcanzan un estatus privilegiado en la disciplina, pudiendo apelar a algo mas que un mero interés sociológico contingente ni llegar a una metafísica demasiado inflada y abstracta. Los teoremas con mayor impacto histórico tienden a ser aquellos que admiten un gran número de demostraciones impuras y heterogéneas, y que, por ello, funcionan como auténticos nodos dentro de la red del conocimiento matemático, otra ventaja de esta tesis es que puede ser apropiada desde las propias matemáticas y su conocimiento existente, pues el fundamento de esta son las demostraciones de los matematicos. Ejemplos de este fenómeno son el teorema de Pitágoras, los teoremas de punto fijo, los métodos de diagonalización, la desigualdad de

[10] Madrid Casado no define explícitamente la naturaleza de la red de conocimiento a la que alude, aunque sí recurre a su noción de *contextos determinantes* para dar cuenta de este fenómeno. Nuestra interpretación se inspira en la noción quineana de *red de creencias* (Quine y Ullian, 1970).

Cauchy–Schwarz o el teorema fundamental del álgebra, todos los cuales cuentan con decenas o incluso centenares de demostraciones y ocupan un lugar central en la práctica matemática.[11]

A partir de aquí surge una cuestión adicional: ¿existen casos de resultados que no solo generen nodos de conocimiento entre distintos dominios de una misma disciplina, sino también entre disciplinas diferentes? Para Madrid Casado, la respuesta es afirmativa de manera casi trivial, puesto que considera a las distintas ramas de las matemáticas como disciplinas particulares (Madrid Casado, 2018, p. 111). Sin embargo, nuestra propuesta va más allá de este punto. Cabe preguntarse si existen resultados profundos, en el sentido del engrosamiento de la franja de verdad, que articulen dominios pertenecientes a ciencias distintas, como la matemática y la física.

Sostenemos que la respuesta es afirmativa y que es posible ofrecer un ejemplo claro de este fenómeno: el teorema de Picard–Lindelöf, cuyo análisis permitirá mostrar cómo un mismo resultado puede operar como nodo no solo entre áreas matemáticas, sino también entre marcos conceptuales matemáticos y físicos.

De forma intuitiva, este teorema garantiza que, dada una ecuación diferencial que no varía de manera "excesivamente abrupta", existe una única solución determinada por una condición inicial. Por ejemplo, si a una partícula en el espacio, sometida a determinadas fuerzas, se le aplica la Segunda Ley de Newton, se obtiene una ecuación diferencial que describe su movimiento. Si dicha ecuación no varía de manera tan

[11] Un caso de especial impacto es el del teorema de Pitágoras, para el cual se han registrado aproximadamente 371 demostraciones (Ratner 2009, p. 230). Loomis (1940) agrupa estas demostraciones en cuatro categorías principales: algebraicas (109), geométricas (255), cuaterniónicas (4) y dinámicas, basadas en consideraciones de masa y velocidad (2). También se han encontrado pruebas que hacen uso de cuadrados mágicos (5).

abrupta, entonces el movimiento de la partícula queda totalmente determinado por la posición inicial. En este sentido, el determinismo implícito en la Segunda Ley de Newton es formulado rigurosamente por el teorema de Picard–Lindelöf.

Formalmente, si consideramos una ecuación diferencial sujeta a cierto valor inicial (lo que se denomina "problema de Cauchy"), obtenemos el siguiente sistema:

$$\begin{cases} \dot{x} = f(t,x) \\ x(t_0) = x_0 \end{cases}$$

Este problema se puede traducir en términos de una ecuación integral del siguiente modo:

$$x(t) = x_0 + \int_{t_0}^{t} f\big(s, x(s)\big)\, ds.$$

Corroboremos la equivalencia. Si evaluamos en $t = t_0$, obtenemos

$$x(t_0) = x_0 + \int_{t_0}^{t_0} f\big(s, x(s)\big)\, ds,$$

observamos que el segundo término se anula y resulta que $x(t_0) = x_0$, por lo que la equivalencia se cumple para la condición inicial. Ahora derivemos ambos miembros con respecto a t:

$$\dot{x}(t) = \frac{d}{dt}\left(x_0 + \int_{t_0}^{t} f\big(s, x(s)\big)\, ds\right)$$

el primer término es constante y al derivarlo se anula, entonces observamos que

$$\dot{x}(t) = \frac{d}{dt}\left(\int_{t_0}^{t} f\big(s, x(s)\big)\, ds\right) = f\big(t, x(t)\big),$$

lo cual demuestra que ambas formulaciones son equivalentes.

Definimos entonces el operador $\Gamma: C(I) \to C(I)$ donde $C(I)$ denota el espacio de funciones continuas definidas sobre un intervalo compacto I, provisto de la norma supremo, dado por

$$\Gamma(\varphi(t)) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds.$$

Observemos que hallar soluciones al problema de Cauchy es equivalente a hallar puntos fijos del operador Γ, al que le llamamos operador de Picard. En efecto, una función $\varphi \in C$ es solución del problema si y solo si $\Gamma(\varphi) = \varphi$, pues en ese caso $\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds$ que, como vimos, es equivalente al problema diferencial inicial.

Esta reformulación del problema no parece, a primera vista, revelar una estructura subyacente; parece más bien una traducción arbitraria. Sin embargo, consideremos un ejemplo concreto, en particular el siguiente problema de Cauchy:

$$\begin{cases} \dot{x} = ax \\ x(0) = 1 \end{cases}$$

definamos ahora la siguiente sucesión de funciones continuas: $\Gamma^n(\varphi) = \Gamma\left(\Gamma^{n-1}(\varphi)\right)$ con $\Gamma^0(\varphi) = x(0) = 1$. Desarrollando sucesivamente:

$$\Gamma^0(x) = 1$$

$$\Gamma(x) = 1 + \int_0^t a\, ds = 1 + at$$

$$\Gamma^2(x) = \Gamma\left(\Gamma(x)\right) = 1 + \int_0^t a(1 + at)\, ds = 1 + at + \frac{(at)^2}{2}$$

$$\Gamma^3(x) = \Gamma\left(\Gamma^2(x)\right) = 1 + \int_0^t a\left(1 + at + \frac{(at)^2}{2}\right) ds$$
$$= 1 + at + \frac{(at)^2}{2} + \frac{(at)^3}{3!}$$
$$\vdots$$
$$\Gamma^n(x) = 1 + at + \frac{(at)^2}{2} + \frac{(at)^3}{3!} + \cdots + \frac{(at)^n}{n!} = \sum_{k=0}^{n} \frac{(at)^k}{k!}.$$

Notemos que si tendemos n a infinito, obtenemos $\sum_{k=0}^{+\infty} \frac{(at)^k}{k!} = e^{at}$, ya que este es precisamente el desarrollo de Taylor de e^{at}. Se trata de un caso particular en el que la convergencia puede describirse explícitamente.

Verifiquemos ahora si e^{at} es punto fijo del operador Γ, es decir, si satisface el problema de Cauchy:

$$\Gamma(e^{at}) = 1 + \int_0^t ae^{as}\, ds = 1 + e^{at} - 1 \Rightarrow \Gamma(e^{at}) = e^{at}.$$

Acabamos de mostrar que, si este no es un caso particular y se pueden fijar hipótesis adecuadas para garantizar, la sucesión Γ^n converge a una determinada función; en particular, la solución al problema de Cauchy.

Existe un resultado matemático perteneciente a la rama del análisis que no suele ser asociado a la teoría de ecuaciones diferenciales ordinarias: el teorema del punto fijo de Banach. El enunciado de este teorema es el siguiente:

Sean (X, d) un espacio métrico completo con distancia d y $F: X \to X$ una contracción. Entonces existe un único punto fijo de F que llamamos p_∞. Más aún, p_∞ es un atractor de F, es decir, $lim\ F^n(p) = p_\infty$ para todo $p \in X$.

Donde una contracción sobre el espacio (X, d) se define como una función $F: X \to X$ donde existe una constante $k \in [0,1)$ tal que $d\big(F(x), F(y)\big) \leq kd(x, y)$.

Y un espacio métrico completo (X, d) es un conjunto de puntos X con una noción de distancia d que no presenta "agujeros". Más formalmente, se trata de un espacio en el que toda sucesión de Cauchy converge a un elemento del propio espacio.

Es decir, si el operador $\Gamma: C(I) \to C(I)$ es una contracción sobre el espacio métrico completo $C(I)$, entonces existe y es única la solución al problema de Cauchy, y esta se obtiene aplicando sucesivamente Γ sobre sí mismo. Esto resulta verdaderamente revelador, pues hemos identificado un comportamiento dinámico global altamente estructurado que emerge a partir de condiciones con un menor contenido estructural: la completitud del espacio de funciones continuas y la condición contractiva del operador de Picard. Conviene precisar, sin embargo, que el carácter "mínimo" de estas condiciones no debe entenderse en términos absolutos, sino de manera relativa al comportamiento concreto de las soluciones de la ecuación diferencial. Aunque abstractas y rigurosamente estructuradas, dichas hipótesis no constituyen información dinámica sobre las soluciones, sino que operan como condiciones generales.

A este fenómeno se le puede atribuir la noción de profundidad cosmológica propuesta por Arana: un orden no trivial que se manifiesta a partir de premisas con escasa estructura en relación a las conclusiones.

Con el fin de culminar el análisis cosmológico del teorema de Picard-Lindelöf, se presentará su enunciado formal, en el formato que habitualmente presenta en la literatura matemática:

Sea $f: \Omega \subset R \times R^n \to R^n$ con Ω abierto, continua y localmente lipschitziana en la variable espacial. Dado $(x_0, t_0) \in \Omega$, existe un intervalo $I = [t_0 - \alpha, t_0 + \alpha]$ con $\alpha > 0$ tal que el problema de Cauchy

$$\begin{cases} \dot{x} = f(t, x) \\ x(t_0) = x_0 \end{cases}$$

tiene solución única en I.[12]

La condición de Lipschitz es la que asegura que el operador de Picard $\Gamma: C(I) \to C(I)$ es una contracción sobre el espacio de funciones continuas definido en el intervalo I.

La tesis del teorema de Picard–Lindelöf reduce el conjunto de posibles soluciones a un único elemento, a partir de hipótesis con un contenido estructural mínimo sobre la función f. Esto le confiere una profundidad cosmológica en el sentido de Arana. Sin embargo, la profundidad cosmológica no es la única que puede atribuirse a este teorema: su propia demostración es también merecedora de un análisis evidencialista.

Urquhart (2015, pp. 238-240) propone una analogía sumamente acertada entre las estrategias de demostración y las decisiones en una partida de ajedrez. Una jugada profunda, en ese contexto, se caracteriza por su escasa obviedad y gran efectividad. Extrapolado al caso del teorema de Picard–Lindelöf, esta "jugada profunda" corresponde a la reformulación del problema de Cauchy como la búsqueda de puntos fijos del operador de Picard. Dicha estrategia permitió vincular dos áreas distintas de las matemáticas: la teoría de ecuaciones diferenciales ordinarias y el análisis matemático, haciendo posible utilizar un resultado auxiliar propio de esta última para completar la demostración del teorema. Estos "puentes" entre teorías revelan propiedades estructurales ocultas, aumentando la profundidad de las demostraciones matemáticas y, en consecuencia, de los propios teoremas (a menos que uno opte por separar la profundidad de una prueba de la profundidad de un teorema).

[12] Una demostración completa de este teorema puede encontrarse en Teschl (2012, pp. 33-38).

En este sentido, la demostración del teorema de Picard–Lindelöf también puede considerarse impura, entendiendo la impureza en un sentido técnico, es decir, como el uso de conceptos y resultados provenientes de otra área. Sin embargo, a diferencia de la mayoría de los ejemplos presentes en la literatura filosófica actual sobre la profundidad matemática, este teorema posee una parte significativa de su origen y consecuencias dentro del marco de la mecánica newtoniana, en particular en relación con la Segunda Ley de Newton. En este contexto, su resultado actúa como una condición para que el determinismo en la mecánica clásica se cumpla, en la medida en que el teorema garantiza que, bajo ciertas condiciones de regularidad (es decir, en ausencia de cambios excesivamente abruptos), una ecuación diferencial admite una única solución, quedando completamente determinada por el estado inicial del sistema.

El descubrimiento de un teorema matemático nos proporciona aquí conocimiento físico acerca del mundo, y no de un tipo trivial. En términos de la noción de profundidad como franja de verdad, introducida al inicio del capítulo, este hecho puede interpretarse como un engrosamiento del rango de acción del teorema. En la literatura filosófica contemporánea se discute la posibilidad de que ciertos hechos matemáticos expliquen fenómenos físicos (Lange, 2013; Lyon, 2011). Casos de este tipo no solo abren la posibilidad de hablar de profundidad asociada también a tesis fundamentales de otras ciencias naturales, sino que permiten, además, introducir una nueva forma de interpretar la explicación matemática de hechos físicos: una explicación más indirecta, pero no por ello menos relevante desde el punto de vista filosófico.[13]

[13] El problema de los puentes de Königsberg pudo resolverse mediante técnicas propias de la teoría de grafos; en ese caso, asumiendo la tesis de la explicación matemática, el conocimiento matemático permitió dar cuenta de un problema del mundo físico de manera directa, sin requerir conocimientos específicos de una teoría física. En cambio, para sostener rigurosamente el determinismo en la mecánica newtoniana fue necesario formular la Segunda Ley de Newton en términos de ecuaciones diferenciales y aplicar el

Ahora bien, si el engrosamiento de la franja de verdad y la constitución de nodos interdisciplinarios dependen, en buena medida, de las estrategias demostrativas empleadas —y en particular de la impureza de dichas demostraciones—, entonces resulta natural volver a interrogarse por el estatuto de las demostraciones mismas como posibles portadoras de profundidad.

Tal y como se dijo anteriormente en la tesis de Lange, las demostraciones también pueden ser portadoras de profundidad: una demostración es más profunda que otra cuando logra responder las preguntas del tipo "¿por qué?" que la otra deja sin resolver (Lange, 2015, pp. 197-204). Entonces surge la siguiente cuestión: ¿una demostración del teorema de Picard–Lindelöf que incorpore la demostración del teorema del punto fijo de Banach es más profunda que una que simplemente lo utiliza sin demostrarlo? Y, en caso afirmativo, ¿hasta qué punto puede seguir considerándose una demostración del teorema de Picard–Lindelöf si, al mismo tiempo, demuestra otros teoremas? Más generalmente, ¿cuál es el límite de resultados previos que pueden incluirse o demostrarse dentro de la prueba de un teorema sin que esta deje de ser considerada como la prueba de dicho teorema? Esta es una derivación del problema de la individuación planteado en la primera sección, y constituye un obstáculo para determinar qué teoremas o demostraciones poseen más profundidad evidencial que otros.[14] Esta es una cuestión que sigue vigente.

teorema de Picard–Lindelöf. Se trata, entonces, de un proceso que exige un paso adicional: la relación entre conocimiento propio de la mecánica newtoniana y conocimiento perteneciente al dominio matemático, lo que da lugar a una forma de explicación más indirecta.

[14] El problema se agrava al considerar otros posibles portadores de profundidad, como los diagramas o los métodos: ¿cómo podrían individuarse? Incluso las intuiciones o ejemplos pondrían en riesgo la tesis evidencialista (Ernst et al., 2015, p. 156).

4. Conclusiones

Mediante la recopilación del trabajo existente en torno al concepto de profundidad, hemos mostrado que ciertos valores provenientes de la estética matemática, pese a su escaso desarrollo en la literatura actual, plantean problemáticas significativas para la filosofía de la práctica del matemático y para la historia de la disciplina. Asimismo, estos valores abren líneas de investigación con conexiones metodológicas, metateóricas, ontológicas, lógicas y, especialmente, epistemológicas.

El estudio de caso del teorema de Picard–Lindelöf ha permitido ilustrar de manera concreta cómo emerge la profundidad en un resultado matemático específico. Este teorema resultó especialmente adecuado para mostrar que la profundidad puede entenderse como un fenómeno vinculado a la capacidad de ciertos resultados para funcionar como nodos dentro de redes de conocimiento cada vez más amplias. En particular, se hizo explícito su papel en la fundamentación rigurosa del determinismo implícito en la Segunda Ley de Newton, lo que pone de manifiesto su proyección más allá del dominio estrictamente matemático. La reformulación de la noción de franja de verdad ha permitido dar cuenta de este engrosamiento del rango de acción del teorema, ofreciendo una vía para integrar, dentro de una misma caracterización, las intuiciones evidencialistas, cosmológicas y consecuencialistas.

Esto invita a reconsiderar el lugar de la estética en la matemática y el papel que desempeñan los valores no epistémicos en su práctica. Incluso cabe preguntarse si la profundidad debe ser considerada un valor epistémico o si, como sugieren algunos autores (Pournari, 2008; Rooney, 1992), la distinción entre valores epistémicos y no epistémicos puede sostenerse.

Agradecimientos

Queremos expresar nuestro más sincero agradecimiento a Gabriel Donoso por su tiempo, apoyo y valiosas contribuciones a lo largo del desarrollo de este proyecto, sin las cuales este artículo no habría alcanzado su forma actual. también le damos gracias asimismo a Oceana por sus críticas constructivas, y a Constantino Contreras por sus pertinentes y enriquecedoras sugerencias bibliográficas. Por último, también agradecemos a los revisores por sus observaciones y comentarios, que contribuyeron de manera significativa a mejorar la calidad del manuscrito.

REFERENCIAS

Arana, A. (2015). On the Depth of Szemerédi's Theorem. *Philosophia Mathematica, 23*(2), 163–176. https://doi.org/10.1093/philmat/nku036

Ayer, A. J. (1946). *Language, Truth and Logic* (Second ed.). Victor Gollancz Ltd.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology, 29*(4), 711–721. https://doi.org/10.1037/0012-1649.29.4.711

Brown, J. R. (2008). *Philosophy of Mathematics: A Contemporary Introduction to the World of Proofs and Pictures* (Second ed.). Routledge. https://doi.org/10.4324/9780203932964

Carnap, R. (1937). *The Logical Syntax of Language.* Kegan Paul, Trench, Trubner & Co., Ltd.

Cheng, Y. (2022). On the Depth of Gödel's Incompleteness Theorems. *Philosophia Mathematica, 30*(2), 173–199. https://doi.org/10.1093/philmat/nkab034

De Toffoli, S. (2021). Groundwork for a Fallibilist Account of Mathematics. *The Philosophical Quarterly, 71*(4), pqaa076. https://doi.org/10.1093/pq/pqaa076

Dienes, Z. P. (1960). *Building Up Mathematics.* Hutchinson Educational.

Ernst, M., Heis, J., Maddy, P., McNulty, M. B., & Weatherall, J. O. (2015). Foreword to Special Issue on Mathematical Depth. *Philosophia Mathematica, 23*(2), 155–162. https://doi.org/10.1093/philmat/nkv003

Field, H. (1988). Realism, Mathematics and Modality. *Philosophical Topics, 16*(1), 57–107. https://doi.org/10.5840/philtopics19881613

Fine, K. (2005). Our Knowledge of Mathematical Objects. In T. S. Gendler, & J. Hawthorne (Eds.), *Oxford Studies In Epistemology* (pp. 89–109). Oxford University Press. https://doi.org/10.1093/oso/9780199285891.003.0004

Fogelin, R. J. (1995). *Wittgenstein* (Second ed.). Routledge. https://doi.org/10.4324/9780203219065

Franklin, J. (2014). *An Aristotelian Realist Philosophy of Mathematics: Mathematics as the Science of Quantity and Structure.* Palgrave Macmillan. https://doi.org/10.1057/9781137400734

Gödel, K. (1944). Russell's Mathematical Logic. In S. Feferman (Ed.), *Collected Works. Volume II. Publications 1938-1974* (pp. 119–141). Oxford University Press.

Gray, J. (2015). Depth — A Gaussian Tradition in Mathematics. *Philosophia Mathematica, 23*(2), 177–195. https://doi.org/10.1093/philmat/nku035

Hahn, H. (1959). Logic, Mathematics, and Knowledge of Nature. In A. J. Ayer (Ed.), *Logical positivism* (pp. 147–161). The Free Press.

Irvine, A. D. (1989). Epistemic logicism & Russell's regressive method. *Philosophical Studies, 55*(3), 303–327. https://doi.org/10.1007/BF00355328

Kim, J. (1982). Perceiving Numbers and Numerical Relations. *Noûs, 16*(1), 93–94. https://doi.org/10.2307/2215418

Kitcher, P. (1983). *The Nature of Mathematical Knowledge.* Oxford University Press.

Lakatos, I. (1976). A Renaissance of Empiricism in the Recent Philosophy of Mathematics. *The British Journal for the Philosophy of Science, 27*(3), 201–223. https://doi.org/10.1093/bjps/27.3.201

Lakatos, I. (2015). *Proofs and Refutations. The Logic of Mathematical Discovery.* (J. Worrall, & E. Zahar, Eds.) Cambridge University Press. https://doi.org/10.1017/CBO9781316286425

Lange, M. (2013). *W*hat Makes a Scientific Explanation Distinctively Mathematical*?. The British Journal for the Philosophy of Science, 64*(3), 485–511. https://doi.org/10.1093/bjps/axs012

Lange, M. (2015). Depth and Explanation in Mathematics. *Philosophia Mathematica, 23*(2), 196–214. https://doi.org/10.1093/philmat/nku022

Loomis, E. S. (1940). *The Pythagorean Proposition. Its Demonstrations Analyzed and Classified and Bibliography of Sources For Data of The Four Kinds of "Proofs"* (Second ed.). Edwards Brothers.

Lyon, A. (2012). Mathematical Explanations Of Empirical Facts, And Mathematical Realism. *Australasian Journal of Philosophy, 90*(3), 559–578. https://doi.org/10.1080/00048402.2011.596216

Maddy, P. (1980). Perception and Mathematical Intuition. *The Philosophical Review, 89*(2), 163–196. https://doi.org/10.2307/2184647

Madrid Casado, C. M. (2018a). *Filosofía de la Cosmología: Hombres, teoremas y leyes naturales.* Pentalfa Ediciones.

Madrid Casado, C. M. (2018b). Respuesta a un crítico de la teoría del cierre. *El Catoblepas: Revista crítica del presente,* (183), 2. https://www.nodulo.org/ec/2018/n183p02.htm

Manin, Y. I. (1981). *Lo demostrable e indemostrable.* (E. M. Kotenko, Trans.) Editorial MIR.

Peirce, C. S. (1992). *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898.* (K. L. Ketner, Ed.) Harvard University Press.

Pournari, M. (2008). The Distinction Between Epistemic and Non-Epistemic Values in the Natural Sciences. *Science & Education, 17*(6), 669–676. https://doi.org/10.1007/s11191-007-9101-y

Putnam, H. (1975). What is mathematical truth? *Historia Mathematica, 2*(4), 529–533. https://doi.org/10.1016/0315-0860(75)90116-0

Putnam, H. (2002). *The Collapse of the Fact/Value Dichotomy and Other Essays.* Harvard University Press. https://doi.org/10.2307/j.ctv1pdrpz4

Quine, W. V. O., & Ullian, J. S. (1970). *The web of belief.* Random House.

Ratner, B. (2009). Pythagoras: Everyone knows his famous theorem, but not who discovered it 1000 years before him. *Journal of Targeting, Measurement and Analysis for Marketing, 17*, 229–242. https://doi.org/10.1057/jt.2009.16

Rooney, P. (1992). On Values in Science: Is the Epistemic/Non-Epistemic Distinction Useful? *PSA: Proceedings of the Biennial*

Meeting of the Philosophy of Science Association, 1992(1), 13–22. https://doi.org/10.1086/psaprocbienmeetp.1992.1.192740

Russell, B. (1973). The Regressive Method of Discovering the Premises of Mathematics. In D. Lackey (Ed.), *Essays in Analysis* (pp. 272–283). Allen and Unwin. (Obra original publicada en 1907).

Shaffer, M. J. (2017). Knowledge of Abstract Objects in Physics and Mathematics. *Acta Analytica, 32*(4), 397–409. https://doi.org/10.1007/s12136-017-0318-3

Stillwell, J. (2015). What Does 'Depth' Mean in Mathematics? *Philosophia Mathematica, 23*(2), 215–232. https://doi.org/10.1093/philmat/nku033

Teschl, G. (2012). *Ordinary Differential Equations and Dynamical Systems.* American Mathematical Society. https://doi.org/10.1090/gsm/140

Urquhart, A. (2015). Mathematical Depth. *Philosophia Mathematica, 23*(2), 233–241. https://doi.org/10.1093/philmat/nkv004

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 214-236
ISSN 3087-260X
Santiago de Chile

LA SUPERACIÓN DE LOS DESACUERDOS PROFUNDOS:

Revisión de bisagras y racionalidad dinámica[1]

Leandro Lema[2]

Universidad de Buenos Aires
IES N°1 «Dra. Alicia Moreau de Justo»

Resumen

Los desacuerdos profundos plantean un desafío central para la epistemología del desacuerdo, al involucrar conflictos persistentes y sistemáticos entre pares epistémicos que no pueden resolverse mediante argumentación ni apelando a nueva evidencia. Este trabajo busca explorar algunos aspectos de la estrategia de revisión de bisagras desarrollada en Lema (2023) y compararla con otras respuestas propuestas

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima.* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Leandro, L. (2026) La superación de los desacuerdos profundos: Revisión de bisagras y racionalidad dinámica. *Revista de Filosofía Homónima, 1*(1), pp. 214-236. https://doi.org/xxxxxxx[Por asignar]

[2] Leandro Lema es Profesor en Filosofía por el IES N°1 «Dra. Alicia Moreau de Justo» y estudiante de grado de la Lic. en Filosofía en la UBA. Email: leandrolema18@gmail.com. ORCID: orcid.org/0009-0002-4934-777X.

para afrontar este tipo de desacuerdos. Sostendremos que la revisión de bisagras constituye una estrategia racional apta para disolver y superar ciertos desacuerdos profundos sin por ello renunciar a las pretensiones teóricas propias de la investigación filosófica y científica. Esta propuesta permite, además, ofrecer un modelo para reconstruir la superación de determinados desacuerdos en la historia de la filosofía y de la ciencia, y contribuye a delinear una concepción dinámica de la racionalidad epistémica.

Palabras clave: *Desacuerdos profundos, Pares Epistémicos, Epistemología del desacuerdo*

Overcoming Deep Disagreements: Hinge Revision and Dynamic Rationality

Abstract

Deep disagreements pose a central challenge for the epistemology of disagreement, as they involve persistent and systematic conflicts between epistemic peers that cannot be resolved through argumentation or by appealing to new evidence. This paper aims to explore several aspects of the hinge revision strategy developed in Lema (2023) and to compare it with other responses that have been proposed to address this kind of disagreement. It is argued that hinge revision constitutes a rational strategy capable of dissolving and overcoming certain deep disagreements without thereby abandoning the theoretical ambitions of philosophical and scientific inquiry. Moreover, this proposal provides a model for reconstructing the overcoming of some past

disagreements in the history of philosophy and science, and helps to articulate a more dynamic conception of epistemic rationality.

Keywords: *Deep disagreements, Epistemic Peer, Epistemology of Disagreement.*

1. Introducción

Pocos negarían que la filosofía es la disciplina que contiene el mayor y más variado número de desacuerdos. A lo largo de su historia existen disputas que se despliegan en múltiples dimensiones y donde los problemas parecen renacer cíclicamente articulados de diversas maneras. Un gran número de estos debates han persistido por años y, independientemente de la controversia existente sobre si son perennes o no, se vuelve innegable su insensibilidad frente a nueva evidencia o nuevos caminos argumentativos. Por el contrario, con el pasar de los años pareciera aumentarse la distancia entre las posiciones rivales y, con ello, la sospecha de que se tratan de "pseudoproblemas". Por el carácter persistente y sistemático de las diferencias entre los litigantes en disputa, esta ha sido entendida como la sintomatología inevitable de un mal uso del lenguaje o la encrucijada donde nos acorrala un fundamento filosófico errado. En la historia de la filosofía fue común diagnosticar que determinadas ideas, principios o valores eran la génesis de grandes querellas evitables, por lo que, al mismo tiempo, se pretendió encontrar en estos cimientos filosóficos el talón de Aquiles para desanudar estos conflictos filosóficos aparentemente indisipables.

En la epistemología del desacuerdo han recibido un tratamiento especial aquellos desacuerdos en los cuales las diferencias entre los litigantes hacen que, en principio, parezcan ser irresolubles de manera

racional. A partir del texto de Fogelin (2019), a estos desacuerdos se los ha llamado *profundos* y es tema de debate tanto la naturaleza de los mismos como las consecuencias filosóficas que de ellos se desprenden en torno a la racionalidad humana. Uno de los tópicos más controversiales fue si efectivamente había alguna forma de disolverlos (aún si no fuera argumentalmente) o si debíamos admitirlos como un rasgo inherente de la finitud de la razón humana.

Para que tenga sentido mantener alguna actitud doxástica con respecto de P, es necesario asumir ciertos presupuestos. Únicamente podemos tener un desacuerdo profundo acerca de P si compartimos el conjunto de presupuestos que hacen posible adoptar una actitud doxástica sobre P. En Lema (2023) planteamos que una posible estrategia para superar desacuerdos profundos teóricos es llevar adelante una *revisión de bisagras,* es decir, reformar aquellos presupuestos que son condición de posibilidad de este tipo de desacuerdos (y de su profundidad) y, por lo tanto, disolverlos.

En este escrito sostendremos que el núcleo de presupuestos mínimos (sin los cuales no se puede desenvolver la discusión racional sobre el objeto del desacuerdo profundo) es el que determina y circunscribe el área de problemas a los cuales nos dirigimos al adoptarlo y que, en última instancia, es ese núcleo el que tenemos que analizar cuando queremos explicar la 'profundidad' del desacuerdo mismo. Argumentaremos en favor de la aplicabilidad de la *revisión de bisagras* en casos de desacuerdos filosóficos y científicos como una manera para superarlos. Por otra parte, analizaremos la manera en la que esta estrategia puede modificar la manera en la que entendemos la racionalidad en los contextos de desacuerdos.

Para ello, en la segunda sección haremos una breve presentación de qué es un desacuerdo profundo y qué características tiene. Su primera subsección estará dedicada a presentar sus principales características a partir del texto de Ranalli (2021) y mostraremos su aplicabilidad a casos de

desacuerdos en filosofía y en ciencia. En la segunda subsección indagaremos en cómo entender la profundidad del desacuerdo y qué la origina. La tercera sección explicaremos la *revisión de bisagras* que presentamos en Lema (2023) como una posible solución a ciertos desacuerdos profundos y defenderemos su consistencia comparándola con otras estrategias posibles planteadas para afrontar este tipo de desacuerdos.

2. En los límites de la argumentación

2.1. Paridad e Irresolubilidad

En primer lugar, debemos preguntarnos si existen desacuerdos filosóficos y científicos[3] que puedan considerarse profundos. La noción de desacuerdo profundo, aunque su caracterización sigue siendo objeto de debate (véase Ranalli & Lagewaard, 2022a, 2022b), fue introducida por primera vez por Robert Fogelin en un artículo de 1985 titulado "La lógica de los desacuerdos profundos" (2019). Aunque inicialmente se dirigía a un público interesado en los problemas de la lógica informal, este trabajo terminó introduciendo nociones y problemas que más tarde fueron recogidos y debatidos en un marco más amplio de discusión que luego se configuró como epistemología del desacuerdo. En ese artículo, Fogelin caracterizó un tipo de desacuerdos en el cual las partes que litigan comparten un trasfondo común que resulta insuficiente para resolver el conflicto mediante medios argumentativos. Inspirado en cómo Wittgenstein entiende el intercambio de razones en *Sobre la certeza* (2009), Fogelin sostiene que, para que la argumentación sea efectiva, debe existir una base común entre ambos litigantes. De esto resulta que, cuando el

[3] Aquí vamos a unificar distintos desacuerdos teóricos acerca de qué actitud doxástica debemos adoptar frente a una proposición P, tanto si son científicos o filosóficos. Desde ya que son muchas las diferencias entre ambos casos, más todo lo sustancial de nuestra propuesta para modelar los desacuerdos profundos puede ser aplicado indistintamente.

trasfondo compartido se encuentra socavado y los argumentos presentados se apoyen en un suelo no compartido, serán vacíos para el oponente.

Recientemente Ranalli (2021) ofreció una caracterización de los desacuerdos profundos que, además de su simplicidad, presenta la virtud de ser neutral frente a las posibles controversias acerca de su naturaleza. Para este autor, una buena teoría de los desacuerdos profundos debe dar cuenta de la articulación de sus principales características, a saber: ser genuino, persistente, sistemático y que sus partes sean sensibles a dar y recibir razones. Veamos qué desacuerdos filosóficos y científicos cumplen con estas cuatro características que ahora pasaremos a explicar.

En primer lugar, existen varios debates entre filósofos (e incluso, en algunos casos, entre científicos) donde la irresolubilidad se debe a una diferencia en lo que cada parte entiende que está discutiendo, al significado que atribuyen a los términos en cuestión. En estos casos, pareciera que si los interlocutores se pusieran de acuerdo en su lenguaje y determinaran con precisión a qué se refieren con sus palabras o qué quieren afirmar no habría tal confusión. Estos casos no pueden ser considerados como desacuerdos genuinos, sino meras diferencias verbales cuya disolución elimina la apariencia misma de desacuerdo. Sin comprometernos con el proyecto analítico de creer que toda disputa filosófica es resuelta por medio de un análisis adecuado del lenguaje, podemos comprender que numerosos desacuerdos filosóficos parecen deberse a confusiones en la forma en la que las partes usan términos como “es”, “existe”, “rojo” o “infinito”. A nosotros nos interesará el caso en el cual los filósofos y científicos tengan un uso del lenguaje que les permita entenderse lo suficiente como para tener un desacuerdo genuino y que ni aun así logren estar de acuerdo.

En segundo término, nos interesarán los casos donde el desacuerdo es persistente. Esto no sólo alude a que se perpetúe indefinidamente, sino que su resolubilidad pareciera ser independiente de la nueva evidencia o de los nuevos argumentos que entren en juego. Si la categoría de desacuerdo profundo depende de la persistencia temporal del debate y sólo se puede

aplicar retrospectivamente cuando parece haber evidencia suficiente de su prolongación, entonces estaríamos frente a un conjunto de desacuerdos profundos ante los cuales existe la amenaza de que sean resueltos en el mañana. ¿Qué relevancia puede tener el contar con una categoría de desacuerdos análoga a las esmeraldas verzules de Goodman[4]? La irresolubilidad a través del tiempo debe ser considerada como uno de los síntomas de la naturaleza del desacuerdo profundo, no como una de sus características esenciales.

El suelo de creencias comunes, como dijimos en un principio, no es suficiente para que sobre él se apoyen argumentos que definan la cuestión en uno u otro sentido. Sin embargo, la evidencia tampoco podrá servir para resolver la cuestión, no sólo por limitaciones contingentes (falta de equipamiento tecnológico o de presupuesto), sino porque en principio parece imposible conseguir evidencia relevante. Toda evidencia será interpretada de manera distinta por cada uno de los litigantes, no hay una base observacional neutral que sirva para zanjar la discusión. Un ejemplo de esto es el caso del *Homo Floresiensis* presentado en De Cruz & De Smedt (2013), en el cual no hubo un consenso sobre si los fósiles hallados eran de una nueva especie de homínido o, por el contrario, solo eran el resultado de una enfermedad que causaba problemas de crecimiento. Aquí los científicos ni siquiera podían ponerse de acuerdo en cómo interpretar la evidencia proporcionada por estos restos fósiles.

[4] Nota del Editor: para quienes no conozcan las "esmeraldas verzules", son parte de una paradoja planteada por Goodman, N. (1979), que versa sobre la inducción. "Verzul" es un adjetivo que se predica correctamente de algo (por ejemplo una esmeralda) cuando ese objeto al ser observado, hasta cierto momento dado en el tiempo, es detectado como verde, y desde ese momento en adelante al ser observado es detectado como azul. Lo paradójico es que basándonos en la información de observaciones realizadas antes de dicho momento podríamos inducir adecuadamente tanto que se trata de un objeto verde como de un objeto "verzul". Para este artículo, la característica importante de las esmeraldas verzules es solo lo primero: este cambio drástico que tienen en su forma de ser, de un instante a otro.

En tercer lugar, el hecho de que todo argumento o evidencia remite a elementos no compartidos está íntimamente conectado con el hecho de que el desacuerdo es sistemático. La incompatibilidad de las actitudes doxásticas de los agentes con respecto a P en realidad está instanciando dos sistemas de creencias distintos, donde es imposible aislar una afirmación para evaluarla de manera independiente. Poner en duda la actitud doxástica que un agente mantiene con respecto a P es hacerlo de todo su sistema articulado sobre cómo entender un fragmento del mundo.

Por último, en el caso de los desacuerdos filosóficos y científicos idealmente las personas están abiertas a dar y recibir razones, no se limitan a cerrarse dogmáticamente sobre su posición o a creer que la otra persona no es un digno rival argumentativo. Ha sido de mucha relevancia el auge a mitades de la década de los 2000 de trabajos que específicamente discutían las implicancias de estudiar casos ideales de desacuerdos entre "pares epistémicos" (véase Christensen, 2007)[5]. En un desacuerdo se considera que los actores son pares epistémicos si comparten la misma evidencia relevante para la cuestión sobre la que desacuerdan, tienen la misma capacidad para procesar esa evidencia y carecen de vicios epistémicos que puedan obstruir la racionalidad frente a la evidencia y la cuestión a discutir. Al entender los debates filosóficos y científicos como desacuerdos entre pares epistémicos no buscamos elevar un peldaño en la idealización de las discusiones reales, sino que perseguimos despejar el análisis de explicaciones que transformen la profundidad del desacuerdo en un síntoma de una brecha cognitiva o evidencial. La idea es que los integrantes de este tipo de desacuerdos estarán abiertos a la argumentación porque perciben a su interlocutor como un par epistémico.

[5] No obstante, el tratamiento de la noción de par epistémico fue introducida anteriormente en Gutting (1982).

2.2. Origen de la profundidad

La cuestión que surge aquí es que, si los desacuerdos que caracterizamos no son resultado de una confusión, de una disparidad evidencial/cognitiva o de no mantener una actitud argumental, ¿qué es lo que origina este tipo de desacuerdos? Aquí nos vamos a suscribir al "enfoque wittgensteiniano", según el cual, lo que origina los desacuerdos profundos es una divergencia de hinge propositions (proposiciones bisagra) que sirven de base a todo el resto de las creencias6. Siguiendo algunas ideas wittgensteinianas expuestas en *Sobre la Certeza* (2009)7, antes de hacer cualquier afirmación, negación o duda debemos contar con una base doxástica mínima que esté fuera de toda duda. Aquello que tomamos por cierto configura el marco que nos permite adoptar una actitud doxástica frente a un determinado conjunto de proposiciones. Eso no implica que sea suficiente contar con dichos presupuestos para desenvolver de forma fructífera la argumentación como medio de resolución. Estas certezas poseen una incuestionabilidad que no se deben a ninguna característica intrínseca de ellas (como ser tautológica, autoevidente o cualquier otra característica), sino por la función que cumplen en un determinado sistema de creencias, al servir de fundamento alrededor del cual se teje toda la red de creencias. La incuestionabilidad solo es relativa al juego de lenguaje que habilita, no por ser un dogma de nuestras prácticas. Como explica Wittgenstein (2009, §§114, 166 y 253), es un presupuesto necesario para llevar adelante toda investigación o prácticas discursivas.

Para Fogelin (2019) en los desacuerdos profundos la argumentación es ineficaz debido al choque de *proposiciones bisagra* distintas; lo que es tomado como una certeza difiere para cada una de las partes que están en

6 No vamos a defender aquí la superioridad de este enfoque frente a otras reconstrucciones alternativas. Una reconstrucción integral del debate puede verse en Ranalli (2020) & Ranalli & Lagewaard (2022a, 2022b).

7 Aquí voy a tomar la caracterización que hace Nudler (2002) de las ideas wittgensteinianas de *Sobre la Certeza*.

disputa. Esto esparce el desacuerdo a través del resto del sistema de creencias que articula la evidencia y funciona como suelo para los argumentos, haciendo que el desacuerdo persista a lo largo de tiempo. Esto hizo que la mayor parte de la literatura sobre desacuerdos profundos se enfoque en las características de lo que diferencia cada posición, sin conectar aquellas creencias compartidas con la persistencia y la sistematicidad del desacuerdo. Sin embargo, existe una profunda conexión entre el desacuerdo profundo y las creencias compartidas que lo enmarcan.

El *common ground* doxástico que es condición de posibilidad para discutir sobre una proposición mantiene una relación directa con esta última. No es suficiente asumir un grupo de compromisos[8] como hablante de un idioma o como sujeto de una comunidad para adoptar una actitud doxástica sobre una proposición, sino que debo tener algunas presuposiciones que hagan inteligible plantearse el contenido de esta proposición en particular. Para que no sea absurdo discutir si la entidad que título "estrella matutina" y la que nombro como "vespertina" son o no la misma, debo tomar por cierto que ambos términos son referenciales. Un ejemplo concreto es ofrecido en Lema & Madroñal (2023), donde, para el caso del debate entre realistas y antirrealistas científicos, se determinan un núcleo de creencias que es necesario compartir para participar de la discusión, ya que son suficientes para poder plantear el problema del estatus de las entidades inobservables postuladas por nuestras mejores teorías. Se vuelve absurdo debatir esto último si no aceptamos que existe un mundo externo, que los términos científicos parecen referir a entidades externas o que podemos ajustar nuestro conocimiento a través de nuestra interacción con el mundo.

La anterioridad lógica y metodológica del núcleo de creencias que

[8] En sentido estricto, hablar de *proposiciones bisagra* y de *compromisos bisagra* puede tener algunas diferencias conceptuales en la forma en la que entendemos su naturaleza y el lugar que desempeña en nuestro sistema epistémico. No obstante, pasaremos por alto esta distinción para una mayor fluidez.

compartimos circunscribe qué debates podemos o no tener y la naturaleza de los mismos. Entre este conjunto de presupuestos habrá algunos que nos guíen en que actitudes doxásticas adoptar frente a determinadas circunstancias. Regulan lo que, a grandes rasgos, podemos entender cómo nuestras *políticas epistémicas*[9]. Esto resulta relevante cuando estas políticas epistémicas acordadas no son suficientes para determinar qué actitud doxástica aplicar en un caso particular. Sin embargo, no es su formulación lingüística la que contiene una ambigüedad, sino que hay locaciones de nuestro espacio epistémico donde ellas funcionan de una forma precisa y otros donde no están condicionados plenamente por el marco. Debido a esto, los argumentos esgrimidos necesitan apoyarse en creencias no compartidas que obedecen a visiones incompatibles del problema, no siendo convincentes para el interlocutor. Sin embargo, desde ya que sería vacuo intentar establecer una ecuación precisa entre los presupuestos que nos determinan qué actitud epistémica *debemos* adoptar y aquellos presupuestos más laxos que solo nos *guían* a adoptarlas[10].

Cuando la persistencia y la sistematicidad del desacuerdo emerge, aparecen de manera patente aquellos lugares de nuestro espacio lógico donde nuestro suelo de discusión genera indeterminación. La idea es que los requisitos doxásticos que posibilitan tomar posiciones distintas sobre una misma proposición son igual de "exigentes" que aquellas condiciones necesarias para que haya acuerdo. El desacuerdo no es más simple de explicar que el acuerdo, también requiere una serie de condiciones mínimas

9 Sigo aquí el término de Borge & Lo Guercio (2019), aunque sin usarlo en su sentido técnico para analizar la función que tienen dentro del debate.

10 Serebrinsky (2025) interpreta que en Lema & Madroñal (2023) defendemos que el debate entre realistas y antirrealistas científicos se debe a una divergencia de significados. Sin embargo, como sostenemos aquí, hay un desacuerdo genuino donde se comparten los presupuestos suficientes para hablar de lo mismo y que haya una comprensión mutua mínima, pero se produce una divergencia en la aplicación de las reglas que guían nuestras prácticas epistémicas debido a la adopción de bisagras adicionales y de otras configuraciones de nuestro marco de certezas.

sin las cuales se diluye.

Por otra parte, la imposibilidad para llevar adelante la argumentación de manera exitosa no impugna la razonabilidad de los agentes. Considerar que podemos ser llevados a admitir actitudes doxásticas incompatibles frente a una misma evidencia sin que haya un error racional de por medio nos pone en discusión, en primer lugar, con quienes sostienen la famosa *tesis de la unicidad*, según la cual, frente a un mismo conjunto de evidencias, solo hay una actitud doxástica adecuada frente a una proposición determinada. Es decir, según ella, si en el encuentro con un amigo ambos somos testigos de que un coche se está incendiando, debería haber una (y solamente una) actitud doxástica legítima y racional frente a ese conjunto de evidencias con respecto a la proposición "el coche se incendió". Sin embargo, el punto es que solo desde un sistema epistémico podemos hablar de razonabilidad epistémica en la asunción de cierta actitud doxástica y es eso, precisamente, en lo que difieren los interlocutores de un desacuerdo profundo. El problema es que esas proposiciones están por fuera de la evaluación del marco porque configuran el marco, están más allá de la evidencia porque son el soporte desde el cual un hecho aumenta la probabilidad de que una creencia sea verdadera. Aún si se pudiera predicar la verdad o la falsedad acerca de las proposiciones bisagra tal como sugiere Feldman (2005), no nos interesaría dicha cuestión, dado que siempre partimos de un sistema epistémico dado y la pregunta es qué sucede cuando nuestro interlocutor parte de un sistema lo suficientemente similar para que acordemos en que una proposición merece que adoptemos una actitud doxástica hacia ella pero diferimos irremediablemente en cómo hacerlo.

3. La revisión de bisagras y la racionalidad dinámica para superar los desacuerdos

Entonces, si en un determinado marco es posible adoptar dos actitudes doxásticas opuestas frente a una misma proposición y no hay una forma

consensuada de determinar qué posición es racional adoptar, ¿debemos conformarnos con un pesimismo epistémico?

Para Castro (2022) la única estrategia que es factible llevar adelante para superar los desacuerdos profundos teóricos (es decir, no aquellos acerca de lo que hay que hacer, sino acerca de un hecho) es la suspensión del juicio. El reconocimiento de que la evidencia no será suficiente para reconocer la disputa y los límites de la argumentación lleva a resignar toda pretensión de verdad o correctitud. No obstante, esto padece el defecto de no ser una salida convincente en contextos científicos, e incluso podría defenderse que tampoco lo sería en discusiones filosóficas. Difícilmente podemos renunciar a todo problema teórico solo por el hecho de que exista una diferencia persistente y sistemática. La esterilidad teórica de esta forma de superar los desacuerdos profundos es patente si, por ejemplo, lo llevamos al caso del homo floresiensis, donde esa estrategia hubiera sido heurísticamente vacua para la ciencia.

Por otra parte, tampoco es factible ver al *relativismo* como la única estrategia meta-filosófica posible frente al desacuerdo profundo teórico cuando la argumentación falla, como es sugerido por Serebrinsky (2025). Para la autora, como toda justificación epistémica es circular e incluso la justificación pragmática desarrollada en Lynch (2011) sería circular, la única estrategia legítima es que las partes mantengan su posición reconociendo la legitimidad y racionalidad de la posición del adversario. Más allá de las desventajas teóricas y pragmáticas que son propias del relativismo por sí mismo, no parece ser útil para modelar cómo se comporta la racionalidad en las discusiones de la historia de la ciencia y la filosofía. Claro que podemos concebir adoptar una posición relativista sobre un desacuerdo profundo contemporáneo; no obstante, cuando miramos el decurso de las discusiones pasadas, parecen abandonarse las discusiones o perder significatividad.

Si estuviéramos encerrados en marcos epistémicos de manera permanente, habría problemas perennes cuya persistencia se mantendría a

lo largo de los años. En ese contexto tendría sentido una visión pesimista que para estos casos solo conciba como una salida posible el suspender el juicio o adoptar una posición relativista acerca de la proposición en disputa. No obstante, si bien algunos filósofos han apostado por estrategias de este estilo, no es la norma dentro de la indagación científica.

Que las proposiciones bisagra sean asumidas como irrevisables para un determinado campo de indagación no constituye que no podamos revisarlas por fuera de él. La imposibilidad se limita a rechazar las proposiciones bisagra al mismo tiempo que se llevan adelante prácticas discursivas apoyadas en ellas, es relativa a dichas prácticas. Corrientemente hacemos ajustes en el marco en el cual planteamos los problemas y sostenemos nuestros argumentos.

De acuerdo a lo que dijimos en la sección anterior, si todo desacuerdo (incluso los profundos) tiene como condición de posibilidad un conjunto de creencias y sin ellas se disolvería, la modificación del *common ground* del debate puede causar la disolución de este. Esto es lo que en Lema (2023) titulamos revisión de bisagras, la cual consiste en superar los desacuerdos cambiando las reglas del juego que los posibilitan.

En la historia de la ciencia son notorios los casos donde los científicos, a pesar de estar inmersos en un paradigma con los mismos presupuestos, herramientas, valores y objetivos, presentan desacuerdos que cumplen con ser genuinos, sistemáticos, persistentes y donde las partes parecen estar abiertas al intercambio argumentativo. También en estos casos corroboramos (incluso de forma más patente) que, al abandonar el paradigma, los científicos posteriores miran con desconcierto la problematicidad de la discusión pasada.

Reconstruyamos un caso de la historia de la ciencia para entender la *revisión de bisagras.* Durante fines del siglo XVIII y comienzos del XIX, había una parte de la comunidad científica que estudiaba los fenómenos

térmicos y compartía una serie de presupuestos que tomaba como ciertos en sus investigaciones. En primer lugar, conciben que (C1) el calor es una sustancia material, denominada calórico; en segundo lugar, creen que (C2) el calórico se conserva y puede transferirse de un cuerpo a otro; y, finalmente, presuponen que (C3) la tarea de la física térmica consiste en describir los fenómenos caloríficos en términos de la distribución y el flujo de dicha sustancia. Bajo estos compromisos, se vuelve pertinente la cuestión acerca de si ciertos procesos físicos implican o no la liberación de calórico, constituyendo un problema genuino de investigación. En particular, estaba en tela de juicio la verdad o no de la proposición P: "Durante la fricción, un cuerpo libera una cantidad de calórico previamente contenida en él". Dentro de este marco común, los científicos adoptaron posiciones divergentes acerca de ella.

Por un lado, existían autores como Antoine-Laurent Lavoisier o Pierre-Simon Laplace, los cuales sostenían que la fricción provocaba la liberación de calórico almacenado en los cuerpos, interpretación coherente con la concepción del calor como una sustancia sutil y conservada. Por otro lado, otros investigadores, sin cuestionar la existencia del calórico ni el marco teórico general con los compromisos C, ponían en duda que la fricción pudiera explicarse en términos de una simple descarga de una cantidad previamente contenida. A su juicio, los resultados experimentales no permitían establecer con claridad ni el modo en que el calórico se encontraba alojado en los cuerpos ni las razones por las cuales su producción parecía no exhibir un límite definido.

Ambas posiciones se concebían a sí mismas como aplicaciones legítimas de los principios C1–C3 y podían acomodar la evidencia empírica mediante el recurso a hipótesis auxiliares y partiendo de presupuestos metodológicos distintos, tales como diferentes supuestos acerca de la capacidad de los materiales para contener calórico o sobre su transmisión. En este contexto, los argumentos presentados por una de las partes podían ser sistemáticamente neutralizados por la otra, sin que se dispusiera de un criterio empírico independiente que permitiera zanjar

de manera definitiva la cuestión acerca de la verdad o falsedad de P.

En este contexto, cuando Benjamin Thompson abandona los compromisos C1-C3, rompe con el marco desde el cual surgía el desacuerdo. Rumford sospecha que el calor no podía ser una entidad almacenada en los cuerpos, sino que debe vincularse de alguna manera al movimiento. Con la ruptura del suelo doxástico compartido, la proposición P pierde su estatuto de afirmación significativa: la controversia acerca de si un cuerpo libera o no calórico durante la fricción deja de ser inteligible, puesto que ya no se reconoce la existencia de aquello cuya liberación estaba en disputa. El nuevo marco teórico reconfigura así el conjunto de problemas y soluciones legítima, conllevando una pérdida en la inteligibilidad de las preguntas que estructuraban el debate previo. Aquello que antes se entendía como una cuestión acerca de la verdad o falsedad de P pasa a ser para el nuevo marco una pregunta mal planteada o un pseudoproblema.

Aquello sobre lo que desacordamos es relativo al tipo de proposiciones bisagras que adoptemos. Sin embargo, como señalamos en Lema (2023), esto no es una forma de resolver racionalmente un desacuerdo, sino de superarlo. No disolvemos el desacuerdo llegando a un consenso entre las partes a través de argumentos que sean convincentes, ya que carecemos de un suelo común para hacerlo, sino que decidimos abandonar los cimientos sobre los cuales se desarrolla el debate.

Es evidente que no es factible abandonar toda base epistémica que enmarque una discrepancia persistente y sistemática. Para cualquiera sería absurdo que, impulsado por la discordancia acerca de cómo fundamentar las matemáticas, alguien sugiriera abandonar la investigación en dicha disciplina. No obstante, revisar las proposiciones bisagra sobre las cuales nos estamos apoyando bajo la sospecha de que ellas nos pueden estar conduciendo a callejones sin salida no puede ser tachada de actitud "irracional". Por el contrario, es síntoma de una mente lo

suficientemente crítica para cambiar elementos centrales del sistema de creencias y mudar de marco. Quienes abandonan marcos epistémicos de discusión parecen estar empujados por buenas razones, incluso cuando sea imposible ofrecer un argumento definitivo en favor de ello (por carecer de un fundamento último que permita apoyar una estrategia inferencial de este tipo)[11]. La revisión de bisagras se encuentra implícita en varios episodios de la historia de la filosofía y la ciencia como una manera de reconstruirlos[12].

Esto nos brinda una imagen dinámica de la argumentación, donde no tenemos agentes epistémicos encerrados en eternos desacuerdos profundos, sino que hay sujetos abiertos a, en determinadas circunstancias, cambiar la indagación teórica allí donde parecieran fallar dichos compromisos bisagra o donde otros logran convencernos. La argumentación dentro de marcos epistémicos compartidos es una dimensión fundamental de la racionalidad, pero también lo es el cambiar las reglas del juego desde donde se mueve nuestra racionalidad epistémica, aquella que es fundacionista e inferencial.

Esta estrategia permite que la suspensión del juicio y el relativismo no sean las únicas salidas frente a desacuerdos profundos teóricos. Disuelve el desacuerdo para abrir caminos a otras sendas teóricas. En él no estamos admitiendo la imposibilidad de llegar a un resultado (tal como sucede en la

[11] Excede a este trabajo abordar qué criterios guían la racionalidad al poner en cuestión las proposiciones bisagra que configuran nuestro marco de argumentación, pero baste señalar que seguramente la revisabilidad requiera criterios pragmáticos que determinen qué circunstancias habilitan ponerlas en cuestión. En este sentido, el contextualismo wittgensteiniano de Michael Williams (2007) hace una lectura de *Sobre la Certeza* que hace consistente la posibilidad de revisar las proposiciones bisagra como una estrategia para disolver los desacuerdos profundos.

[12] Un ejemplo que puede ilustrar esto es la defensa que hace Richard Rorty del abandono del programa representacionalista en epistemología en su famoso libro *La Filosofía el Espejo de la Naturaleza* (1989). Allí destaca que, si bien no hay argumentos conclusivos que pueda brindar en apoyo de su posición pragmatista, puede defenderse las virtudes de su adopción.

suspensión del juicio) ni tampoco nos mantenemos en nuestra posición reconociendo que nuestro adversario puede estar justificado desde su posición (como en el caso del relativismo). Podemos reconocer, en cambio, que el marco del que partimos nuestro adversario y nosotros es infértil, que es inconducente y que es necesario explorar otras sendas teóricas. Ni hay una ausencia de error ni es imposible determinarlo, el fallo está en el meta-nivel de la configuración del marco de discusión. Piénsese en todas las discusiones teológicas o metafísicas que, si no hubiéramos abandonado el marco de discusión, seguirían generando polémica hasta hoy sin poder ser superadas.

Modelar esos cambios bajo el modelo de la *revisión de bisagras* puede darnos herramientas adicionales para analizarlos, desglosar sus detalles y poder comprender de qué forma opera la racionalidad en esos episodios de la historia de la ciencia y la filosofía[13].

---

[13] Un ejemplo filosófico puede ser ofrecido para mostrar la fertilidad de nuestro modelo. Recientemente, con la introducción de la noción de *stance epistémica* por parte de van Fraassen (2002) y el posterior desarrollo por Chakravartty (2004, 2017), algunos autores se han inclinado por una lectura meta-filosófica más relativista del debate entre realistas y antirrealistas científicos (puede encontrarse un ejemplo de ello en Madroñal, 2023 o en Serebrinsky, 2025). En vez de caracterizar cada posición como un conjunto de proposiciones que conforman una doctrina, se ha optado por entender al debate como el choque de distintos valores que guían nuestras prácticas epistémicas. Estas *stances*, en sentido estricto, carecen de valor de verdad y solo puede evaluarse su racionalidad.
Si lo que hemos defendido hasta aquí es correcto, nuestro modelo acerca de los cambios de marco puede dar
cuenta de por qué para ciertos filósofos no tiene sentido la discusión entre realistas y antirrealistas científicos, incluso cuando se presenten como *stances.* Ambas *stances* en disputa son parte de una *meta-stance* a la cual podemos o no adherir. Una estrategia para superar el mencionado desacuerdo entre realistas y antirrealistas será realizar modificaciones en esta *meta-stance* que hagan que carezca de sentido tomar posición acerca del estatus de las entidades inobservables o del conocimiento acerca de ellas. Posiciones pragmatistas como las de Frápolli (2014; 2020) logran disolver el debate, haciendo que pierdan significatividad las actitudes doxásticas acerca de las proposiciones que se ponen en cuestión en este debate. Queda para otros trabajos analizar de qué forma podemos evaluar la aplicación de la revisión de bisagras, dado que existen casos exitosos (como el

4. Conclusión

A lo largo de este trabajo hemos sostenido que los llamados desacuerdos profundos, tal como los caracterizó Robert Fogelin (2019), no nos encierran en una racionalidad epistémica asfixiante, sino que hay caminos posibles para salir de aquellos casos donde falle sistemáticamente la argumentación. Frente a posiciones pesimistas o que creen que solo es posible adoptar un relativismo o suspender el juicio acerca de la cuestión discutida como una estrategia para este tipo de desacuerdos, defendimos que pueden haber otros caminos igualmente válidos. Los desacuerdos entre pares epistémicos acerca de cuestiones teóricas que son persistentes y sistemáticos han cautivado nuestra atención allí donde no pueden atribuirse a fallas cognitivas ni a déficits de evidencia, sino a divergencias más elementales en los compromisos que estructuran el pensamiento de los interlocutores. Estas divergencias, originadas en las llamadas proposiciones bisagra, delimitan aquel espacio lógico donde asumimos actitudes doxásticas acerca de una proposición y permiten explicar qué sucede cuando la argumentación resulta impotente para alcanzar consensos y servir de puente para una resolución racional de la disputa.

Defendimos que la revisión de bisagras que presentamos en Lema (2023) constituye una posible estrategia para superar algunos desacuerdos profundos teóricos en filosofía y en ciencia que, a diferencia del relativismo o la suspensión del juicio no es teóricamente infértil. Ella no consiste, desde ya, en una resolución argumentativa, sino una reconfiguración del marco epistémico que sostiene el conflicto y toda argumentación en él. Revisar bisagras equivale a someter a crítica los

abandono de paradigmas pasados que conllevo disolver los desacuerdos que existían dentro de ellos) y casos donde sería absurdo llevarla a cabo.

fundamentos mismos de nuestra práctica racional, un gesto que no implica irracionalidad, sino una forma más profunda de autocrítica teórica. Que el relativismo y la suspensión del juicio no sean la única estrategia para afrontar un desacuerdo profundo teórico, amplía el repertorio de conceptos posibles para hacer reconstrucciones de la dinámica de los desacuerdos profundos en ciencia y en filosofía.

Si bien esta estrategia no puede aplicarse universalmente, ya que no todo desacuerdo permite o exige una revisión de sus compromisos básicos, su reconocimiento amplía el horizonte de lo que puede considerarse una respuesta racional al disenso. En lugar de aceptar la irresolubilidad como un límite insalvable de la razón, la revisión de bisagras invita a pensar la filosofía y la ciencia como prácticas en permanente reconstrucción, capaces de transformar su propio suelo de inteligibilidad. Hemos defendido, por otra parte, que la revisión de bisagra es aplicable a varios casos de desacuerdos profundos dentro de la ciencia y la filosofía, no siendo un recurso excepcional en su historia. Contar con este concepto no solo nos ayuda a entender los desacuerdos profundos, sino que permite construir nuevos modelos acerca de cómo fueron abandonadas viejas disputas en el pasado.

REFERENCIAS

Borge, B., & Lo Guercio, M. (2019). Tipos de desacuerdo científico y metacientífico. *Ciência e Conhecimento*, 13, 185–204.

Castro, D. (2021). Argumentation and persistent disagreement. *Informal Logic, 41*(2), 245–280.

Chakravartty, A. (2004). Stance relativism: empiricism versus metaphysics: The Empirical Stance Bas C. Van Fraassen; Yale University Pres, London & New Haven, 2002, pp. xix-282, ISBN 0300088744. *Studies in History and Philosophy of Science Part A, 35*(1), 173–184.

Chakravartty, A. (2017). *Scientific Ontology: Integrating naturalized metaphysics and voluntarist epistemology*. New York: Oxford University Press. https://doi.org/10.1093/oso/9780190651459.001.0001

Christensen, D. (2007). Epistemology of Disagreement: The Good News. *Philosophical Review, 116(2),* 187–217. https://www.jstor.org/stable/20446955

De Cruz, H., & De Smedt, J. (2013). The value of epistemic disagreement in scientific practice: The case of *Homo floresiensis*. *Studies in History and Philosophy of Science Part A, 44*(2), 169–177. https://doi.org/10.1016/j.shpsa.2013.02.002

Feldman, R. (2005). Deep disagreement, rational resolutions, and critical thinking. *Informal Logic, 25*(1), 12–23. https://doi.org/10.22329/il.v25i1.1041

Fogelin, R. (2019). La lógica de los desacuerdos profundos (2005 [1985]). (D. Mejía Saldarriaga, Trad.). *Revista Iberoamericana de Argumentación, 19*, 84–99. (Obra original publicada en 1985.) https://doi.org/10.15366/ria2019.19.004

Frápolli, M. (2014). No miracles: What does it mean that science seeks the truth? *Zagadnienia Nauznawstwa, 4*(202). 263–280. https://www.ugr.es/~frapolli/No%20miracles.pdf

Frápolli, M. (2020). Teorías verdaderas. En B. Borge & N. Gentile (Eds.), *La ciencia y el mundo inobservable: Discusiones contemporáneas en torno al realismo científico* (pp. 255–286). Buenos Aires: Eudeba.

Goodman, N. (1979). *Fact, Fiction and Forecast*. Harvard University Press.

Gutting, G. (1982). *Religious belief and religious skepticism*. Notre Dame: University of Notre Dame Press.

Lema, L. (2023). Revisión de bisagras: una vía racional para superar los desacuerdos profundos. En G. Arroyo et al. (Eds.), *VII Jornadas de Lógica y Argumentación: Desacuerdos Profundos* (pp. 48–60). Buenos Aires: Secretaría de Investigación, Universidad Nacional General Sarmiento. http://repositorio.ungs.edu.ar:8080/bitstream/handle/UNGS/1733/Actas_Arroyo.pdf?sequence=1&isAllowed=y

Lema, L., & Madroñal, I. F. (2023). Debate entre realismo y antirrealismo científicos: ¿en qué acordar para desacordar? *Revista Instante, 5*(3), 476–493. https://revista.uepb.edu.br/revistainstante/en/article/view/2747

Madroñal, I. (2023). Realismo y antirrealismo científico, stances en desacuerdo. *Revista Colombiana de Filosofía de las Ciencia, 23*(46), 11–40. https://doi.org/10.18270/rcfc.v23i46.3809

Nudler, O. (2002). Las controversias filosóficas: el caso de la filosofía de la ciencia del siglo XX. En P. Lorenzano & F. Tula Molina (Eds.), *Filosofía e historia de la ciencia en el cono sur* (pp. 43–48). Buenos Aires: Universidad Nacional de Quilmes. https://www.afhic.com/wp-content/uploads/2017/11/Filosofia-e-historia-de-la-ciencia-en-el-cono-sur-01.pdf

Ranalli, C. (2020). Deep disagreement and hinge epistemology. Synthese, 197, 4975–5007. https://doi.org/10.1007/s11229-018-01956-2

Ranalli, C. (2021). What is deep disagreement? *Topoi, 40*, 983–998. https://doi.org/10.1007/s11245-018-9600-2

Ranalli, C., & Lagewaard, T. (2022a). Deep Disagreement (Part 1):

Theories of Deep Disagreement. *Philosophy Compass, 17*(12), e1288G. https://doi.org/10.1111/phc3.12886.

Ranalli, C., & Lagewaard, T. (2022b). Deep Disagreement (Part 2): Epistemology of Deep Disagreement. *Philosophy Compass, 17*(12), e12887. https://doi.org/10.1111/phc3.12887.

Rorty, R. (1989). *La filosofía y el espejo de la naturaleza.* Madrid: Cátedra. (Segunda ed.). (J. Fernández Zulaica, Trad.)

Serebrinsky, D. (2025). Desacuerdos profundos: ¿vale la pena seguir argumentando? *Revista Kriterion, 66*(161), e-41328. https://doi.org/10.1590/0100-512X2025n16104ds

van Fraassen, B. (2002). *The Empirical Stance.* New Haven: Yale University Press. https://www.jstor.org/stable/j.ctt1nprp8.

Williams, M. (2007). Why (Wittgensteinian) Contextualism Is Not relativism. *Episteme: A Journal of Social Epistemology, 4*(1), 93–114. https://doi.org/10.3366/epi.2007.4.1.93.

Wittgenstein, L. (2009). *Tractatus Logico-Philosophicus. Investigaciones Filosóficas. Sobre la Certeza.* Madrid: Gredos. (J. Muñoz Veiga, I. Reguera Pérez, A. García Suárez, C. U. Moulines, J. L. Prades, & V. Raga, Trans.) Gredos.

HOMÓNIMA
Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 237-309
ISSN 3087-260X
Santiago de Chile

MODELOS EN CIENCIA[1]

Publicado por primera vez el lunes 27 de febrero de 2006; última revisión sustancial el miércoles 2 de abril de 2025.

Entrada Original de *The Stanford Encyclopedia of Philosophy*[2]
<https://plato.stanford.edu/archives/sum2025/entries/models-science/>

[1] NT: Traducción al español de la entrada "Models in Science" de Roman Frigg y Stephan Hartmann —quienes gentilmente autorizaron su realización y publicación— se basa en la versión archivada (Summer 2025 Edition) de The Stanford Encyclopedia of Philosophy: Frigg, Roman; Hartmann, Stephan, "Models in Science", The Stanford Encyclopedia of Philosophy (Summer 2025 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/sum2025/entries/models-science/. En consecuencia, esta traducción puede diferir de la versión vigente de la entrada, que podría haber sido actualizada. La versión actual puede consultarse en: https://plato.stanford.edu/entries/models-science/. Salvo por los ajustes estilísticos propios de la traducción, el trabajo presenta una serie de modificaciones, entre ellas las notas a pie de página que comienzan con "NT" (nota del traductor) y las notas introductorias sobre la entrada, la enciclopedia, los autores y el traductor. Asimismo, se realizaron modificaciones en el sistema de referencias y de citas al estilo APA 7.

[2] Enciclopedia de filosofía del *laboratorio de investigación metafísica* (*Metaphysical Research Lab*) del departamento de filosofía de la Universidad de Stanford (licencia exclusiva para su publicación en línea). Coeditores principales: Edward N. Zalta (del departamento de filosofía de la Universidad de Stanford) y Uri Nodelman (del departamento de filosofía de la Universidad de Stanford). Editores asociados: Colin Allen (del departamento de filosofía de la Universidad de California/Santa Bárbara), Hannah Kim (del departamento de filosofía de la Universidad de Arizona) y Paul Oppenheimer (del departamento de filosofía de la Universidad de Adelaide). Dirección: Stanford, CA 94305; ISSN: 1095-5054; Contacto: *editors@plato.stanford.edu*

Por Roman Frigg[3] & Stephan Hartmann[4]
Traducción por Nibaldo Lorca Améstica[5]

Los modelos tienen una importancia central en muchos contextos de la investigación científica. Un ejemplo de ello es la centralidad de los modelos en sus respectivos dominios, como los modelos inflacionarios en cosmología, los modelos de circulación general en climatología global, el modelo de doble hélice del ADN, los modelos evolutivos en biología, los modelos basados en agentes en las ciencias sociales y los modelos de equilibrio general de mercados (La sección Otros recursos de internet al final de esta entrada contiene enlaces a recursos online que discuten estos modelos). Los científicos invierten una cantidad significativa de tiempo en construir, probar, comparar y revisar los modelos; asimismo, las revistas dedican un espacio considerable a interpretar y discutir sus implicaciones.

[3] Profesor de filosofía y jefe del departamento de Filosofía, Lógica y Método Científico de la *London School of Economics and Political Science* (LSE), en Londres, UK. Contacto: r.p.frigg@lse.ac.uk; ORCID: https://orcid.org/0000-0003-0812-0907; sitio web: https://romanfrigg.org/

[4] Profesor y decano de la Facultad de Filosofía, Filosofía de las Ciencias y Estudios de la Religión en la *Ludwig-Maximilians-Universität* y codirector del *Munich Center for Mathematical Philosophy* (MCMP) en Munich, Alemania. Contacto: S.Hartmann@lmu.de; ORCID: https://orcid.org/0000-0001-8676-2177; sitio web: https://www.stephanhartmann.org/

[5] Profesor de Filosofía y Doctorando en Filosofía de la Pontificia Universidad Católica de Chile (PUC), en Santiago de Chile, Chile. Contacto: nplorca@uc.cl

Por consiguiente, los modelos han llamado la atención de los filósofos y actualmente hay una gran cantidad de literatura sobre las diversas características del modelamiento científico. Un resultado tangible del compromiso filosófico con los modelos es la proliferación de distintos tipos de modelos reconocidos en la literatura filosófica. Algunas de las nociones usadas para categorizar a los modelos son: *Modelos probatorios*, *modelos fenomenológicos, modelos computacionales, modelos de desarrollo, modelos explicativos, modelos empobrecidos, modelos de prueba, modelos idealizados, modelos teóricos, modelos a escala, modelos heurísticos, modelos de caricatura, modelos exploratorios, modelos didácticos, modelos de fantasía, modelos mínimos, modelos de juguete, modelos imaginarios, modelos matemáticos, modelos mecánicos, modelos sustitutos, modelos icónicos, modelos formales, modelos análogos y modelos instrumentales.* Mientras que, a primera vista, tal abundancia es abrumadora, ésta puede controlarse al reconocer que dichas nociones conciernen a distintos problemas que surgen en relación con los modelos. Los modelos plantean preguntas relativas a la semántica (¿Cómo, de ser posible, representan los modelos?), la ontología (¿Qué clase de cosas son los modelos?), la epistemología (¿Cómo aprendemos y explicamos mediante los modelos?) y, por supuesto, a otros dominios de la filosofía de las ciencias.

1. Semántica: Modelos y Representación

2. Ontología: ¿Qué son los modelos?
   - 2.1 Objetos físicos
   - 2.2 Objetos ficticios y objetos abstractos
   - 2.3 Estructuras teóricas de conjuntos
   - 2.4 Descripciones y Ecuaciones

3. Epistemología: las funciones cognitivas de los modelos
   - 3.1 Aprendiendo sobre los modelos
   - 3.2 Aprendiendo acerca de los sistemas objetivo
   - 3.3 Explicaciones con modelos
   - 3.4 Entendimiento con modelos

3.5 Otras funciones cognitivas

4. Modelos y teoría
4.1 Modelos como subsidiarios de la teoría
4.2 Los modelos como independientes de la teoría

5. Modelos y otros debates en filosofía de las ciencias
5.1 Modelos, realismo y las leyes de la naturaleza
5.2 Modelos y Reduccionismo

Bibliografía

Herramientas académicas

Otros recursos de internet

Entradas relacionadas

1. Semántica: Modelos y Representación

Muchos modelos científicos son representacionales: ellos representan una parte o aspecto determinado del mundo, que constituye el sistema objetivo del modelo. [Algunos] ejemplos estándar son el modelo de bolas de billar de un gas, el modelo atómico de Bohr, el modelo Lotka-Volterra de la interacción depredador-presa, el modelo Mundell-Fleming de una economía abierta y el modelo a escala de un puente. Elliott-Graves (2020) discute cómo se individúan los objetivos [del modelo].

Esto plantea la pregunta de qué significa que un modelo represente un sistema objetivo. Este problema es bastante complejo y se desglosa en varios subproblemas. Para un resumen detallado del problema de la representación, véase Nguyen y Frigg (2022) y la entrada de Representación Científica. Llegados a este punto, en lugar de abordar el problema de qué significa que un modelo represente, nos enfocaremos en distintos tipos de representaciones que desempeñan un papel importante en la práctica de la ciencia basada en modelos; a saber: modelos a escala,

modelos analógicos, modelos idealizados, modelos de juguete, modelos mínimos, modelos fenomenológicos, modelos exploratorios y modelos de datos. Estas categorías no son mutuamente excluyentes y un modelo determinado puede pertenecer a varias al mismo tiempo.

Modelos a escala. Algunos modelos son copias de un sistema objetivo en tamaño reducido o en tamaño agrandado (Black, 1962). Un ejemplo típico es el de un pequeño coche de madera que se coloca en un túnel de viento para estudiar las propiedades aerodinámicas de un coche real. La intuición nos dice que un modelo a escala es una réplica realista o una imagen especular veraz del objetivo; por ello, a veces, los modelos a escala se denominan "modelos verdaderos" (Achinstein, 1968, Ch. 7). Sin embargo, no hay tal cosa como un modelo a escala perfectamente fidedigno; la fidelidad siempre está restringida a ciertos aspectos. El modelo de madera a escala del coche ofrece un retrato fidedigno de su forma, pero no de su material. Incluso en los aspectos en los que el modelo es una representación fidedigna, la relación entre las propiedades del modelo y las del objetivo no suele ser directa. Cuando los ingenieros usan, digamos, un modelo a escala 1:100 de un barco para investigar la resistencia que un barco real experimenta al moverse por el agua, ellos no pueden simplemente medir la resistencia del modelo y multiplicarla por la escala. De hecho, la resistencia enfrentada por el modelo no se traduce en la resistencia enfrentada por el barco real de un modo directo (es decir, uno no puede simplemente escalar la resistencia del agua con la escala del modelo: el barco real no tiene, necesariamente, cien veces la resistencia al agua de su modelo 1:100). Las dos magnitudes se encuentran en una complicada relación no lineal la una con la otra, y la forma exacta de dicha relación es, con frecuencia, no trivial y emerge como resultado de un estudio exhaustivo de la situación (Sterrett, 2006, 2020; Pincock, 2022).

Modelos analógicos. [Algunos] ejemplos estándar de modelos analógicos incluyen el modelo de las bolas de billar de un gas, el modelo hidráulico de un sistema económico y el modelo de agujero mudo de un agujero negro. En el sentido más básico, dos cosas son análogas si presentan ciertas similitudes relevantes entre sí. En un texto clásico,

Hesse (1963) distingue distintos tipos de analogías según el tipo de relación de similitud en la que se encuentran los dos objetos. Un tipo simple de analogía es la que se basa en propiedades compartidas. Hay una analogía entre la Tierra y la Luna, basada en el hecho de que ambas son grandes cuerpos, sólidos, opacos y esféricos que reciben calor y luz del Sol, giran alrededor de sus ejes y gravitan hacia otros cuerpos. No obstante, la igualdad de propiedades no es una condición necesaria. Una analogía entre dos objetos también puede basarse en similitudes relevantes entre sus propiedades. En este sentido más libre, podemos decir que hay una analogía entre el sonido y la luz porque el eco [del sonido] es similar a la reflexión [de la luz], el volumen al brillo, el tono al color, la detectabilidad por el oído a la detectabilidad por el ojo, y así sucesivamente.

Las analogías también pueden basarse en la igualdad o semejanza de las relaciones entre las partes de dos sistemas, en lugar de en sus propiedades monádicas. Es en este sentido que la relación de un padre con su hijo se considera análoga a la relación de un estado con sus ciudadanos. Las analogías mencionadas hasta ahora son las que Hesse denomina "analogías materiales". Para obtener una noción más formal de la analogía, [debemos] abstraer las características concretas del sistema y centrarnos únicamente en su configuración formal. Entonces, lo que el modelo analógico comparte con su objetivo no es un conjunto de características, sino el mismo patrón de relaciones abstractas (*e.g.,* la misma estructura, entendida en sentido formal). Esta noción de analogía está estrechamente relacionada con lo que Hesse denomina "analogía formal". En este sentido, dos objetos son [formalmente análogos] si ambos son interpretaciones del mismo cálculo formal. Por ejemplo, hay una analogía formal entre un péndulo que se balancea y un circuito eléctrico oscilante, ya que ambos se describen por la misma ecuación matemática.

Otra distinción importante de Hesse es [la que se establece] entre las analogías positivas, negativas y neutras. La analogía positiva entre dos objetos consiste en las propiedades o relaciones que comparten (las

moléculas de gas y las bolas de billar tienen masa); la analogía negativa consiste en las propiedades que no comparten (las bolas de billar están coloreadas, las moléculas de gas no); la analogía neutra comprende las propiedades que no sabemos (aún) si pertenecen a la analogía negativa o positiva (¿las bolas de billar y las moléculas tienen la misma sección eficaz en los procesos de dispersión?). Las analogías neutras desempeñan un papel importante en la investigación científica porque fomentan las preguntas y sugieren nuevas hipótesis. Por esta razón, muchos autores han enfatizado el rol heurístico que las analogías desempeñan en la construcción de teorías y modelos, así como en el pensamiento creativo (Bailer-Jones y Bailer-Jones, 2002; Bailer-Jones, 2009, Ch. 3; Hesse, 1974; Holyoak y Thagard, 1995; Kroes, 1989; Psillos, 1995; y los ensayos recopilados en Helman, 1988). Véase también la entrada sobre analogía y razonamiento analógico.

También se ha discutido si el uso de modelos analógicos puede, en algunos casos, ser confirmatorio en un sentido bayesiano. Hesse (1974, pp. 208-219) argumenta que esto es posible si la analogía es material. [En cambio,] Bartha (2010, 2013, [2019]) discrepa y argumenta que los modelos analógicos no pueden ser confirmatorios en un sentido bayesiano, ya que la información encapsulada en un modelo analógico es parte de los conocimientos previos relevantes; por consiguiente, la probabilidad posterior de una hipótesis respecto de un sistema objetivo no puede cambiar como resultado de lo observado en la analogía. Por lo tanto, los modelos analógicos sólo pueden establecer la *plausibilidad* de una conclusión, en el sentido de justificar una asignación de probabilidad previa no despreciable (Bartha, 2010, §8.5).

Más recientemente, estas preguntas han sido discutidas en el contexto de los denominados experimentos análogos, que prometen promover el conocimiento sobre un sistema objetivo experimentalmente inaccesible (*e.g.*, un agujero negro) mediante la manipulación de otro sistema, el sistema fuente (*e.g.*, un condensado Bose-Einstein). Dardashti, Thébault y Winsberg (2017) y Dardashti, Hartmann *et al.* (2019) han argumentado que, dadas ciertas condiciones, una simulación análoga de

un sistema mediante otro sistema puede confirmar afirmaciones sobre el sistema objetivo (*e.g.*, que los agujeros negros emiten radiación de Hawking). Véase Crowther *et al.* (2021) para una discusión crítica, así como la entrada de Simulaciones computacionales en ciencia.

Modelos idealizados. Los modelos idealizados son aquellos que implican una deliberada simplificación o distorsión de algún [aspecto] complejo, con el objetivo de hacerlo más manejable o entendible. Planos sin fricción, masas puntuales, sistemas completamente aislados, agentes omniscientes totalmente racionales y mercados en completo equilibrio son ejemplos bien conocidos. Las idealizaciones son medios cruciales para que la ciencia enfrente sistemas muy difíciles de estudiar en toda su complejidad (Potochnik, 2017).

Los debates filosóficos acerca de la idealización se han centrado en dos tipos generales de idealización: las denominadas idealizaciones aristotélicas y galileanas. Las idealizaciones aristotélicas equivalen a "despojar", en nuestra imaginación, todas las propiedades de un objeto concreto que creemos que no son relevantes para el problema en cuestión. Hay un desacuerdo sobre cómo hacer esto. Jones (2005) y Godfrey-Smith (2009) ofrecen un análisis de la abstracción en términos de verdad: a pesar de que una abstracción permanece en silencio sobre ciertas características o aspectos del sistema, aun así, no dice nada falso y ofrece una descripción verdadera, aunque restringida. Esto permite que los científicos se enfoquen en un conjunto limitado de propiedades aisladas. Un ejemplo es un modelo mecánico clásico del sistema planetario que describe la posición de un objeto como función del tiempo e ignora todas las demás propiedades de los planetas. Cartwright (1989, Ch. 5), Musgrave (1981) – que usa el término "supuestos insignificantes" – y Mäki (1994) – que habla de "método de aislamiento" – permiten que las abstracciones presenten falsedades, por ejemplo, al descuidar un factor causalmente relevante.

[Por otro lado,] las idealizaciones galileanas son aquellas que implican distorsiones deliberadas: los físicos construyen modelos de masas puntuales que se mueven en planos sin fricción; los economistas

asumen que los agentes son omniscientes; los biólogos estudian poblaciones aisladas; y así sucesivamente. El uso de simplificaciones de este tipo, siempre que una situación resultaba muy difícil de abordar, era característico de la aproximación de Galileo a las ciencias. Por esta razón, es común referirse a las idealizaciones "distorsionantes" de esta clase como "idealizaciones galileanas" (McMullin, 1985). Un ejemplo de estas idealizaciones es un modelo de movimiento en una pista de hielo que asume que el hielo no tiene fricción, cuando, en realidad, tiene un [coeficiente] de fricción muy bajo, pero no [igual a] cero.

A veces, las idealizaciones galileanas se caracterizan como idealizaciones controladas; es decir, como aquellas que permiten la deidealización mediante la sucesiva eliminación de las suposiciones distorsionadas (McMullin, 1985; Weisberg, 2007). Interpretadas de este modo, las idealizaciones galileanas no abarcan todos [los casos de] idealizaciones distorsionantes. Batterman (2002, 2011) y Rice (2019, 2021) discuten idealizaciones distorsionantes que no pueden eliminarse en tanto no pueden ser removidas del modelo sin desmantelarlo por completo.

¿Qué nos dice un modelo —que involucra distorsiones— acerca de la realidad? Laymon (1991) formuló una teoría que entiende las idealizaciones como límites ideales: imagínese una serie de ajustes de situación real que se aproxima al límite postulado; entonces, se requiere que, mientras más se acercan las propiedades del sistema al límite ideal, más se acerca su comportamiento al comportamiento del sistema en el límite (monotonicidad). Si este es el caso, los científicos pueden estudiar el sistema en el límite y trasladar sus conclusiones de este a sistemas alejados del límite. Sin embargo, estas condiciones no siempre son necesarias. De hecho, puede ocurrir que el sistema limitante no se aproxime al sistema en el límite. Si esto ocurre, nos enfrentamos a un límite singular (Berry, 2002). En tales casos, el sistema en el límite puede exhibir ciertos comportamientos que difieren de los comportamientos de los sistemas alejados del límite. Los límites de esta clase se presentan en una variedad de contextos, más notablemente en la teoría de transiciones

de fase de la mecánica estadística. No obstante, no existe un acuerdo sobre [cuál es] la interpretación correcta de dichos límites. Batterman (2002, 2011) los considera indicativos de fenómenos emergentes, mientras que Butterfield (2011a, b) los ve como compatibles con la reducción (véanse también las entradas acerca de Relaciones interteóricas en física y Reducción científica).

Las idealizaciones galileanas y aristotélicas no son mutuamente excluyentes; muchos modelos exhiben ambos [tipos] en tanto tienen en consideración un conjunto reducido de propiedades que distorsionan. Considérese nuevamente el modelo mecánico clásico del sistema planetario: el modelo sólo contempla un conjunto reducido de propiedades y las distorsiona, por ejemplo, al describir los planetas como esferas ideales con una distribución de masa rotacional simétrica.

Un concepto estrechamente relacionado con la idealización es la aproximación. En un sentido amplio, *A* puede llamarse una aproximación de B si A es, de algún modo, próxima a *B*. No obstante, esto es demasiado amplio, pues permite calificar cualquier semejanza como una aproximación. Rueger y Sharp (1998) limitan las aproximaciones a la proximidad cuantitativa y Portides (2007) plantea la noción de aproximación como un concepto esencialmente matemático. Bajo esta noción, *A* es una aproximación de *B* si y sólo si *A* es próximo a *B* en un sentido matemático especificable, donde el sentido relevante de "próximo" está dado por el contexto. Un ejemplo es la aproximación de una curva a otra mediante la expansión de una función en series de potencias, manteniendo sólo los primeros dos o tres términos. En distintas situaciones, aproximamos una ecuación a otra al permitir que un parámetro de control tienda a cero (Redhead, 1980). Esto plantea la pregunta de cómo las aproximaciones difieren de las idealizaciones, que también pueden implicar una proximidad matemática. Norton (2012) establece la distinción entre ambas como referencial: una aproximación es una descripción inexacta del [sistema] objetivo, mientras que una idealización introduce un sistema secundario (real o ficticio) que ocupa el lugar del sistema objetivo (siendo distinto de este). Si decimos que el

periodo del péndulo en la pared es más o menos dos segundos, entonces es una aproximación; si razonamos acerca del péndulo real, asumiendo que la bola del péndulo es una masa puntual y que su cuerda no tiene masa (i.e., si asumimos que el péndulo es, digamos, un péndulo ideal), entonces usamos una idealización. Separar las idealizaciones y las aproximaciones de este modo no implica que no pueda haber relaciones interesantes entre ambas. Por ejemplo, una aproximación puede justificarse al señalar que es una expresión matemática de una idealización aceptable (*e.g.*, cuando ignoramos un término disipativo en una ecuación de movimiento porque suponemos, de modo idealizado, que el sistema no tiene fricción).

Los modelos idealizados también desafían la filosofía bayesiana de las ciencias. ¿Cómo puede un agente tener un grado positivo de creencia en un modelo fuertemente idealizado? Esta es una pregunta apremiante porque los modelos idealizados son ubicuos en ciencias y son ampliamente usados en estadística (bayesiana). Para una discusión sobre las distintas respuestas, véase Sprenger y Hartmann (2019, Ch. 12).

Modelos de juguete. Los modelos de juguete son representaciones extremadamente simplificadas y fuertemente distorsionadas de sus objetivos; comúnmente sólo representan un pequeño número de factores causales y explicativos (Hartmann, 1995; Reutlinger *et al.*, 2018; Nguyen, 2020). [Algunos] ejemplos típicos son el modelo de Lotka-Volterra en ecología de poblaciones (Weisberg, 2013) y el modelo de Schelling de segregación en ciencias sociales (Sugden, 2000). Usualmente, los modelos de juguete no se desempeñan muy bien en términos de predicción y de adecuación empírica, pues parecen servir a otros propósitos epistémicos (más acerca de estos [propósitos] en la Sección 3). Esto plantea la pregunta de si deben considerarse representaciones en primer lugar (Luczak, 2017).

Algunos modelos de juguetes se caracterizan como "caricaturas" (Gibbard y Varian, 1978; Batterman y Rice, 2014). Los modelos de caricatura aíslan un pequeño número de características sobresalientes de un sistema y las distorsionan [hasta convertirlas] en un caso extremo. Un

ejemplo clásico es el modelo de Akerlof (1970) del mercado de autos ("el mercado de limones"), que explica la diferencia de precio entre autos nuevos y usados únicamente en términos de información asimétrica; de este modo, se hace caso omiso de todos los otros factores que pueden influir en el precio de los autos (véase también Sugden, 2000). Sin embargo, es controvertido si dichos modelos altamente idealizados todavía pueden considerarse representaciones informativas de sus sistemas objetivos. Para una discusión sobre los modelos de caricatura, en particular en economía, véase Reiss (2006).

Modelos mínimos. Los modelos mínimos están estrechamente relacionados con los modelos de juguete, en tanto que estos también son altamente simplificados. Es tal su simplificación que algunos argumentan que no son representacionales, [pues] carecen de cualquier similitud, isomorfismo o relación de semejanza con el mundo (Batterman y Rice, 2014). Se ha argumentado que muchos modelos económicos pertenecen a esta clase (Grüne-Yanoff, 2009). Los modelos mínimos de economía tampoco están restringidos por las leyes naturales ni aíslan ningún factor real (*Ibid.*). A pesar de esto, el estudio de los modelos mínimos puede ayudarnos a aprender algo sobre el mundo. Sin embargo, es controvertido si los modelos mínimos pueden ayudar a los científicos a aprender algo sobre el mundo si no representan nada (Fumagalli, 2016). Los modelos mínimos, que supuestamente carecen de similitud o representación, también se usan en distintas partes de la física para explicar los comportamientos a escala macro de varios sistemas cuyos comportamientos a escala micro son extremadamente diversos (Batterman y Rice, 2014; Rice, 2018, 2019; Shech, 2018). [Algunos] ejemplos típicos incluyen las características de las transiciones de fase y del flujo de fluidos. Defensores de los modelos mínimos argumentan que aquello que proporciona una explicación del comportamiento a escala macro de un sistema, en estos casos, no es una característica que el sistema y el modelo tengan en común, sino el hecho de que el sistema y el modelo pertenecen a la misma *clase universal* (una clase de modelos que exhiben el mismo comportamiento límite, aunque muestren comportamientos

muy distintos en escalas finitas). No obstante, es controversial si las explicaciones de este tipo son posibles sin hacer referencia a al menos alguna característica en común (Lange, 2015; Reutlinger, 2017).

Modelos subrogantes. Los modelos subrogantes se utilizan en las ciencias para estudiar sistemas [que resultan] difíciles o imposibles de investigar directamente. Estos modelos actúan como sustitutos de sus sistemas objetivos, permitiendo que los investigadores infieran sobre el objetivo a partir del comportamiento del subrogante. Un ejemplo clásico es el uso de organismos modelo en biología, tales como el *Drosophila melanogaster* (mosca de la fruta) o el *Mus musculus* (ratón) que sirven como *intermediarios* (*proxies*) para estudiar procesos biológicos en humanos. La elección de un modelo subrogante implica cuidadosas consideraciones sobre su adecuación representacional y la extensión a la que los hallazgos pueden generalizarse al sistema objetivo (Bolker, 2009; Burian, 1993). Los filósofos han debatido [acerca de] los desafíos epistemológicos de usar modelos subrogantes, en particular cuando el sistema subrogante y el objetivo difieren significativamente, como en el caso de los organismos modelo que a menudo presentan mecanismos subyacentes o historias evolutivas distintas de sus objetivos (Dietrich *et al.*, 2020; Ankeny & Leonelli, 2020). Para una discusión sobre los organismos modelo, véase la entrada de Experimentos en biología.

Modelos fenomenológicos. Los modelos fenomenológicos han sido definidos de diferentes – aunque relacionadas – maneras. Una definición común los considera como modelos que sólo representan propiedades observables de sus objetivos y se abstienen de postular mecanismos ocultos y similares (Bokulich, 2011). Otra aproximación, propuesta por McMullin (1968), define los modelos fenomenológicos como independientes de las teorías. Esto, no obstante, parece ser [una afirmación] demasiado fuerte. Muchos modelos fenomenológicos, a pesar de no ser derivables de una teoría, incorporan principios y leyes asociados a las teorías. El modelo de la gota líquida del núcleo atómico, por ejemplo, retrata al núcleo como una gota líquida y lo describe como poseyendo muchas propiedades (como la tensión de superficie y la carga,

entre otras) que se derivan de distintas teorías (como la hidrodinámica y la electrodinámica, respectivamente). Ciertos aspectos de estas teorías – aunque usualmente no toda la teoría – se utilizan para determinar tanto las propiedades estáticas como las dinámicas del núcleo. Finalmente, resulta tentador identificar los modelos fenomenológicos con *los modelos de un fenómeno*. Aquí, "fenómeno" es un término paraguas que abarca todas las características relativamente estables y generales del mundo que resultan interesantes desde un punto de vista científico. El debilitamiento de un sonido como función de la distancia a la fuente, el decaimiento de las partículas alfa, el crecimiento de una población de conejos y la dependencia del precio de una casa en función de la tasa de referencia de la Reserva Federal son fenómenos en este sentido. Para una discusión más exhaustiva, véase Bailer-Jones (2009, Ch. 7), Bogen y Woodward (1988) y la entrada de Teoría y observación en ciencias.

Modelos exploratorios. Los modelos exploratorios son modelos que no son propuestos, en primer lugar, para aprender algo acerca de un sistema objetivo específico o de un fenómeno particular establecido experimentalmente. Los modelos exploratorios funcionan como puntos de partida para futuras exploraciones, en las que se modifica y se refina el modelo. Gelfert (2016) señala que los modelos exploratorios nos pueden ofrecer de *pruebas-de-principio* y sugerir explicaciones *cómo-posiblemente* (2016, Ch. 4). Gelfert menciona los primeros modelos en ecología teórica a modo de ejemplo, tales como el modelo de Lotka-Volterra de la interacción depredador-presa que imita el comportamiento cualitativo de aceleración y desaceleración del crecimiento poblacional en un bioma con recursos limitados (2016, p. 80). Tales modelos no nos proporcionan una caracterización precisa del comportamiento de una población real, sino que nos ofrecen un punto de partida para el desarrollo de modelos más realistas. Fisher (2009) considera estos modelos como herramientas para examinar las características de una teoría determinada. Massimi (2019) señala que los modelos exploratorios proporcionan conocimiento modal. Sjölin Wirlin y Grüne-Yanoff (2021)

presentan una inspección de los distintos tipos de conocimiento modal que se reporta que los modelos proporcionan.

Modelos de datos. Un modelo de datos es una versión corregida, rectificada, regimentada y muchas veces idealizada de los datos que obtenemos de la observación inmediata, los denominados datos brutos (Suppes, 1962), Característicamente, uno primero elimina los errores (*e.g.*, se remueven los puntos del registro que son culpa de una observación defectuosa) y entonces se presentan los datos de forma "limpia"; por ejemplo, al dibujar una curva lisa a través de un conjunto de puntos. Estos dos pasos suelen denominarse "reducción de datos" y "ajustes de curvas". Cuando investigamos, por ejemplo, la trayectoria de un planeta determinado, primero eliminamos los puntos erróneos de los registros observacionales y luego los puntos restantes se ajustan a una curva lisa. Los modelos de datos juegan un rol crucial en la confirmación de teorías, ya que estas se someten a prueba frente a los modelos de datos y no contra los datos brutos, comúnmente desordenados y complejos.

La construcción de un modelo de datos puede resultar extremadamente compleja. Se requieren técnicas estadísticas sofisticadas y se plantean serias preguntas tanto metodológicas como filosóficas. ¿Cómo decidimos qué puntos del registro deben eliminarse? Y, dado un conjunto limpio de datos, ¿a qué curva los ajustamos? La primera pregunta ha sido abordada principalmente en el contexto de la filosofía de la experimentación (véase, por ejemplo, Galison, 1997 y Staley, 2004). En el núcleo de la última pregunta yace el denominado problema del ajuste de curvas: los datos [por sí] mismos no indican ni la forma de la curva ajustada ni las técnicas estadísticas que los científicos deben usar para construirla. La elección y racionalización de las técnicas estadísticas son materia de la filosofía de la estadística, y referimos al lector a la entrada de Filosofía de la estadística, así como a Bandyopadhyay y Forster (2011) y Otsuka (2023) para una discusión sobre estos temas. Una discusión más exhaustiva sobre los modelos de datos puede encontrarse en Bailer-Jones (2009, Ch. 7), Brewer y Chinn (1994), Harris (2003), Hartmann (1995), Laymon (1982), Mayo (1996, 2018) y Suppes (2007).

La reunión, procesamiento, diseminación, análisis, interpretación y almacenamiento de datos plantean muchas preguntas importantes más allá del tema relativamente reducido de los modelos de datos. Leonelli (2016, 2019) investiga el estatus de los datos en la ciencia, argumentando que estos deben definirse no por su procedencia, sino por su función evidencial, y estudia cómo los datos viajan entre distintos contextos.

2. Ontología: ¿Qué son los modelos?

¿Qué son los modelos? Es decir, ¿con qué tipo de objeto tratan los científicos al trabajar con modelos? Algunos autores han manifestado escepticismo acerca de si esta pregunta tiene una respuesta significativa porque los modelos no pertenecen a una categoría ontológica distintiva y cualquier cosa puede ser un modelo (Callender y Cohen, 2006; Giere, 2010; Suárez, 2024; Swoyer, 1991; Teller, 2001). Contessa (2010) replica que esta conclusión no se sigue. Incluso si, desde un punto de vista ontológico, cualquier cosa puede ser un modelo y la clase de objetos a los que se refieren como modelos contiene una colección heterogénea de distintos objetos, no se sigue que sea imposible o inútil desarrollar una ontología de los modelos. Esto se debe a que, aunque no todos los modelos pertenezcan a una [misma] clase ontológica particular, uno puede, no obstante, preguntar a qué clases ontológicas pertenecen las cosas que, de facto, se usan como modelos. Puede haber una plétora de estas clases y cada una puede analizarse por sí misma. El tipo de objetos que los científicos utilizan como modelos tiene una importante repercusión en cómo estos desempeñan funciones relevantes, tales como la representación y la explicación; por ello, este problema no puede ser descartado como [algo] "meramente sociológico".

Los objetos que comúnmente sirven como modelos pertenecen, en efecto, a distintas clases ontológicas: objetos físicos, objetos ficticios, objetos abstractos, estructuras teóricas de conjuntos[6], descripciones,

[6] NT: La versión original es "*set-theoretical structure*", pero el adjetivo "*set-theoretical*" presenta ciertos problemas para traducirse al español. En la versión en español de *La*

ecuaciones o combinaciones de estas clases son frecuentemente referidos como modelos y algunos modelos podrían incluso pertenecer a otras clases. Aquellos interesados en la ontología pueden ver esto como un objetivo en sí mismo. Sin embargo, vale la pena señalar que la pregunta tiene reverberaciones más allá de la ontología e influye en cómo uno entiende la semántica y la epistemología de los modelos.

2.1 Objetos físicos

Algunos modelos son objetos físicos. Tales modelos suelen denominarse "modelos materiales". [Algunos] ejemplos estándar de modelos de este tipo son los modelos a escala de objetos como puentes o barcos (véase Sección 1), el modelo del ADN de Watson y Crick (Schaffner, 1969), el modelo hidráulico de una economía de Pillips y Newlyn (Morgan y Boumans, 2004), el modelo de la bahía de San Francisco del USACE (Weisberg, 2013), el modelo de plastilina de la mioglobina de Kendrew (Frigg y Nguyen, 2016) y los organismos modelo en ciencias de la vida (Ankeny y Leonelli, 2020; Leonelli, 2010; Levy y Currie, 2015; Sartori, *en prensa*). Todos estos son objetos materiales que sirven como modelos. Los modelos materiales no dan lugar a dificultades ontológicas más allá de los reconocidos problemas relacionados a los objetos con los que tratan los

introducción a la lógica simbólica de Patrick Suppes, traducida por Gabriel Aguirre (Suppes, P., 1966 [1957], *Introducción a la lógica simbólica*. Compañía Editorial Central S.A.). Aguirre señala que "*set-theoretical*" es un adjetivo que sólo puede traducirse perifrásticamente como "*relativo a la teoría de conjuntos*", por lo que no existe un equivalente directo en español (Suppes, 1970, p. 287, NT). Apoyo la crítica del traductor, pero no su propuesta de traducir "*set-theoretical*" como "*sinforemático*", apelando a la raíz griega clásica *συμφόρημα* (*symphorema*), que, según el traductor (ibid., p. 287, NT), se refiere a un compuesto o conjunto. Otras traducciones de este concepto en distintas obras en español no aportan mayor claridad. Las propuestas de traducción varían de "*teórico-conjuntista*" o "*teórico de conjuntos*" a simplemente "*teóricas*" o "*teoréticas*", omitiendo la parte de "*conjunto*". Para mantener cierto grado de parsimonia en español, mantengo la traducción "*set-theoretical*" como "*teórico de conjuntos*".

metafísicos, por ejemplo, respecto de la naturaleza de las propiedades, la identidad de los objetos, las partes y los todos, y así sucesivamente.

Sin embargo, muchos modelos no son modelos materiales. El modelo atómico de Bohr, un péndulo sin fricción o una población aislada, por ejemplo, están más bien en la mente de los científicos que en el laboratorio y no tienen que estar físicamente realizados ni experimentados para servir como modelo. Estos modelos "no físicos" plantean serias preguntas ontológicas, y el debate sobre cómo deben analizarse de manera más adecuada es controversial. En lo que resta de esta sección, reseñamos algunas sugerencias de la literatura reciente sobre modelos que pueden atraer nuestra atención.

2.2 Objetos ficticios y objetos abstractos

Aquello que ha llegado a ser conocido como la *concepción ficticia de los modelos* considera a los modelos como similares a objetos imaginarios en la ficción literaria – es decir, similares a personajes ficticios como Sherlock Holmes o a lugares ficticios como la Tierra Media (Godfrey-Smith, 2007). Así, cuando Bohr introdujo su modelo atómico, él introdujo un objeto ficticio de la misma clase que el objeto introducido por Conan Doyle cuando él inventó a Sherlock Holmes. Esta visión cuadra bien con la práctica científica, donde los científicos comúnmente hablan de los modelos como si estos fuesen objetos y a menudo se consideran a sí mismos [como si] describiesen átomos, poblaciones o economías *imaginarias*. Además, cuadra bien con las concepciones filosóficas que consideran la construcción y manipulación de los modelos como un aspecto esencial de la investigación científica (Morgan, 1999), incluso si los modelos no son objetos materiales, ya que estas prácticas parecen estar dirigidas a *algún* tipo de objeto.

¿Qué preguntas filosóficas resuelve esta movida? El discurso ficcional y las entidades ficticias confrontan preguntas filosóficas reconocidas, y bien uno puede argumentar que la simple comparación de los modelos con las ficciones equivale a explicar *obscurum per obscurius*

(para una discusión acerca de estas preguntas, véase la entrada Entidades ficticias). Un modo de contrarrestar esta objeción y motivar la concepción ficticia de los modelos es señalar el poder heurístico de esta concepción. En este sentido, Frigg (2010b) identifica cinco problemas específicos que una ontología de modelos tiene que abordar y entonces señala que estos problemas surgen de formas muy similares en la discusión sobre las ficciones ([estos] problemas son las condiciones de identidad, la atribución de propiedades, la semántica de enunciados comparativos, las condiciones de verdad y la epistemología de objetos imaginados). Entonces, la comparación de los modelos con las ficciones tiene un valor heurístico, ya que existe una rica literatura sobre las ficciones que ofrece una variedad de soluciones a estos problemas.

Sólo una pequeña porción de las opciones disponibles en la extensa literatura sobre las ficciones ha sido realmente explorada en el contexto de los modelos científicos. Contessa (2010) formula lo que denomina "aproximación dualista", según la cual un modelo es un objeto abstracto que suple a un posible objeto concreto. El modelo atómico de Rutherford, por ejemplo, es un objeto abstracto que actúa como suplente de un posible sistema que contiene un electrón orbitando alrededor del núcleo en una órbita bien definida. Barberousse y Ludwig (2009) y Frigg (2010b) toman una ruta distinta y desarrollan una caracterización de los modelos como ficciones, basada en la teoría de la pretensión de la ficción de Walton (1990). De acuerdo con esta caracterización, las oraciones en un pasaje de texto que introducen a un modelo deben considerarse como un *atrezo*[7] en un juego de simulación de creencias, y el modelo es el

[7] NT: Original *prop*, puede traducirse como *'pieza de utilería'* en español. De acuerdo con el diccionario de Oxford (*s. v.)*, un *prop* (abreviatura de *properties*) puede definirse como un objeto de utilería usado en teatro o cine (sus otras definiciones, como objeto de soporte físico o de soporte emocional, no me parecen las más adecuadas para abarcar el sustantivo *prop* en este contexto particular). Dicha definición teatral, como utilería o atrezo (palabra italiana para la utilería de teatro, aceptada por la RAE), se acerca más al sentido de *prop* en los juegos de simulación de creencia. Asimismo, este es un concepto técnico de la teoría de la pretensión de Walton (1990), en la que un *prop* es un objeto que funciona como

producto de un acto de pretensión. Esta es una posición antirrealista en tanto que toma el hablar de "objetos" modelo como figuras del discurso, porque, en último término, no hay objetos modelo — los modelos sólo viven en la imaginación de los científicos. Salis (2021) reformula esta propuesta para convertirla en lo que denomina "la nueva concepción ficticia de los modelos". La diferencia central radica en que lo que se considera un modelo son las descripciones del modelo y su contenido, en lugar de las imaginaciones que prescriben. Esta es una concepción realista de los modelos, ya que las descripciones existen.

La concepción ficticia no está exenta de críticas. Giere (2009), Magnani (2012), Pincock (2012), Portides (2014) y Teller (2009) rechazan la aproximación ficticia y argumentan, de distintas maneras, que los modelos no deberían considerarse ficciones. Weisberg (2013) argumenta por una posición intermedia que ve a las ficciones desempeñando un rol heurístico, pero niega que deban considerarse como parte de un modelo científico. El núcleo común de estas críticas [radica en] que la concepción ficticia malinterpreta la posición epistémica de los modelos. Llamar algo una ficción equivale a afirmar que es falso, y es injustificado llamar a todo un modelo una ficción – y, por consiguiente, afirmar que falla en capturar cómo es el mundo – sólo porque el modelo implica ciertas presuposiciones falsas o elementos ficticios. En otras palabras, una representación no cuenta automáticamente como ficción sólo porque presenta algunas imprecisiones. Defensores de la concepción ficticia coinciden con este punto, pero niegan que la noción de ficción deba analizarse en términos de falsedad. Aquello que hace que una obra sea una ficción no es su falsedad (o alguna proporción [entre] afirmaciones falsas y verdaderas): no todo lo que se menciona en una novela es falso (*La guerra y la paz* de Tolstoy contiene muchas afirmaciones verdaderas acerca de las guerras napoleónicas franco-rusa), ni todos los textos que

generador de verdades ficticias en la simulación de creencias (Walton, 1990, p. 37). Opto por traducir *'prop'* como *'atrezo'* para remarcar su sentido técnico en el contexto particular de la teoría de Walton.

contienen afirmaciones falsas califican como ficción (las noticias falsas son sólo eso, no son ficciones). La característica definitoria de una ficción es que se supone que el lector imagina los eventos y personajes descritos, no que sean falsos (Frigg, 2010a; Salkis, 2021).

Giere (1988) defendió la concepción de que los modelos "no físicos" son entidades abstractas. Sin embargo, hay poco consenso sobre la naturaleza de los objetos abstractos y Hale (1988, pp. 86-87) enumera no menos de veinte posibles caracterizaciones distintas (para una revisión de las opciones disponibles, véase la entrada de Objetos abstractos). En publicaciones recientes, Thomasson (2020) y Thomson-Jones (2020) desarrollan lo que denominan la "concepción artefactualista" de los modelos, basada en la teoría de los artefactos abstractos de Thomasson (1999). Esta concepción concuerda con la teoría de la pretensión en tanto que el contenido del texto que introduce a un personaje ficticio o a un modelo debe entenderse como ocurriendo dentro de la pretensión. Sin embargo, al mismo tiempo, insiste en que los autores, en la producción de tales descripciones, crean artefactos culturales abstractos que, entonces, existen de modo independiente tanto del autor como del lector. El artefactualismo concuerda con el platonismo en tanto que los objetos abstractos existen, pero insiste, en contra del platonismo, en que estos son traídos a la existencia mediante un acto creativo y no son eternos. Esto permite que los artefactualistas preserven las ventajas de la teoría de la pretensión mientras, al mismo tiempo, mantienen una concepción realista que [sostiene que] realmente existen los personajes ficticios y los modelos.

2.3 Estructuras teóricas de conjuntos

Un punto de vista influyente sostiene que los modelos son estructuras teóricas de conjuntos. Esta posición puede rastrearse hasta Suppes (1960) y, actualmente, aunque con algunas variaciones menores, es sostenida por la mayoría de los defensores de la llamada concepción semántica de las teorías; para una discusión acerca de esta concepción, véase la entrada de

La estructura de teorías científicas y Frigg (2023). Hay diferencias entre las [distintas] versiones de la concepción semántica, pero, con la excepción de Giere (1988), todas las versiones concuerdan en que los modelos son estructuras de algún tipo u otro (Da Costa y French, 2000).

Esta concepción de los modelos ha sido criticada por diversos motivos. Una crítica persistente es que muchos tipos de modelos – que desempeñan un rol importante en las ciencias – no son estructuras y no pueden ser acomodados dentro del marco de la concepción estructuralista de los modelos, el cual no puede dar cuenta ni de cómo estos modelos son construidos ni de cómo ellos funcionan en el contexto de la investigación (Cartwright, 1999; Downes, 1992; Morrison, 1999). [Algunos] ejemplos de tales modelos son los modelos interpretativos y los modelos mediadores, que se discuten más adelante en la Sección 4.2. Otra acometida mantenida en contra de la aproximación teórica de conjuntos es que las estructuras teóricas de conjuntos, por sí mismas, no pueden ser modelos representacionales – al menos no si esto requiere que ellas compartan alguna estructura con el objetivo – porque la atribución de una estructura al sistema objetivo – que forma parte del mundo físico – depende de una descripción sustantiva (no-estructural) del objetivo, lo que va más allá de aquello que la aproximación estructuralista puede costear (Nguyen y Frigg, 2021).

2.4 Descripciones y Ecuaciones

Una posición que ha sido honrada por el tiempo [sostiene] que un modelo es una descripción estilizada de un sistema objetivo. Se ha argumentado que esto es lo que los científicos muestran en artículos y libros de texto al presentar un modelo (Achinstein, 1968; Black, 1962). Esta concepción no ha sido objeto de críticas explícitas. Sin embargo, algunas de las críticas que han sido reunidas en contra de la denominada concepción sintáctica de las teorías amenazan, de igual modo, el entendimiento lingüístico de los modelos (para una discusión acerca de esta concepción, véase la entrada sobre La estructura de las teorías

científicas). En primer lugar, una crítica estándar [en contra] de la concepción sintáctica es que, al asociar una teoría con una formulación particular, la concepción malinterpreta su identidad, pues cualquier cambio en su formulación da lugar a una nueva teoría (Suppe, 2000). Una concepción que asocia a los modelos con las descripciones pareciera ser vulnerable a la misma crítica. En segundo lugar, los modelos tienen propiedades distintas de las descripciones: el modelo newtoniano del sistema solar consiste en esferas que orbitan, pero no tiene sentido afirmar esto acerca de su descripción. Por el contrario, las descripciones tienen propiedades que los modelos no tienen: una descripción puede ser escrita en inglés y consistir en 517 palabras, pero lo mismo no puede decirse de un modelo. Una forma de evadir estas dificultades es asociar el modelo al contenido de la descripción, en lugar de a la descripción misma. Para una discusión sobre una posición respecto de los modelos construidos a partir del contenido de una descripción, véase Salis (2021).

Una versión contemporánea del descriptivismo se encuentra en la denominada concepción de representación directa de Levy (2012, 2015) y de Toon (2012). Esta concepción, al igual que la concepción ficticia de los modelos (Sección 2.2), comparte la sujeción a la teoría de la pretensión de Walton, pero la utiliza de un modo distinto. La diferencia principal es que las concepciones previamente discutidas ven el modelamiento como la introducción de un vehículo de representación, el modelo, que es distinto del objetivo, y subrayan que el problema es aclarar qué clase de objeto es el modelo. En la concepción de representación directa no hay modelos distintos del objetivo; sólo hay descripciones-modelo y objetivos, sin modelos intermedios. El modelamiento, desde esta perspectiva, consiste en ofrecer una descripción imaginaria de objetos reales. Una descripción-modelo prescribe imaginaciones acerca del sistema real; el péndulo ideal, por ejemplo, prescribe a los usuarios del modelo que imaginen la cuerda real como perfectamente elástica y la bola como una masa puntual. Esta aproximación evita los problemas previamente mencionados porque las condiciones de identidad del modelo se derivan de las condiciones del juego de simulación de creencias (no de la sintaxis

de la descripción), y la atribución de propiedades ocurre en la pretensión. No obstante, hay interrogantes acerca de cómo esta concepción lidia con los modelos sin objetivo (como los modelos del éter o las poblaciones de cuatro sexos) y sobre cómo los modelos, entendidos de este modo, lidian con las idealizaciones. Para una discusión sobre estos puntos, véanse Frigg y Nguyen (2016), Poznic (2016) y Salis (2021).

Una aproximación bastante cercana considera los modelos como ecuaciones. Esta es una versión de la concepción según la cual los modelos son descripciones, ya que las ecuaciones son objetos sintácticos que describen una estructura matemática. Los problemas que esta aproximación enfrenta son similares a los que ya nos hemos encontrado: en primer lugar, uno puede describir la misma situación usando diferentes tipos de coordenadas y, como resultado, obtener diferentes ecuaciones sin que, de este modo, también se obtengan distintos modelos. En segundo lugar, el modelo y la ecuación tienen propiedades distintas. Un péndulo contiene una cuerda sin masa, pero la ecuación que describe su movimiento no; y una ecuación puede ser no homogénea, pero el sistema que describe no. Es una cuestión abierta si estos problemas pueden evadirse apelando a una caracterización de la pretensión.

3. Epistemología: las funciones cognitivas de los modelos

Una de las principales razones por las que los modelos desempeñan un papel tan importante en las ciencias es que realizan una serie de funciones cognitivas. Por ejemplo, los modelos son vehículos para aprender sobre el mundo. Partes significativas de la investigación científica se llevan a cabo en modelos en lugar de en la realidad misma porque, al estudiar un modelo, podemos descubrir características y determinar hechos acerca del sistema que el modelo reemplaza: los modelos habilitan un "razonamiento subrogante" (Swoyer, 1991). Por ejemplo, estudiamos la naturaleza del átomo de hidrógeno, las dinámicas de una población o el comportamiento de un polímero mediante el estudio de sus respectivos modelos. Esta función cognitiva de los modelos ha sido ampliamente

reconocida en la literatura y algunos [autores] incluso han sugerido que los modelos suscitan un nuevo estilo de razonamiento, "el razonamiento basado en modelos", de acuerdo con el cual "las inferencias se realizan mediante la creación de modelos y su manipulación, adaptación y evaluación" (Nersessian, 2010, p. 12; véase también Magnani, Nersessian y Thagard, 1999; Magnani y Nersessian, 2002; y Magnani y Casadio, 2016).

3.1 Aprendiendo sobre los modelos

El aprendizaje sobre un modelo ocurre en dos instancias: en la construcción del modelo y en su manipulación (Morgan, 1999). No hay reglas ni recetas fijas para la construcción del modelo, así que la actividad misma de figurarse qué encaja con qué y cómo proporciona una oportunidad para aprender sobre el modelo. Una vez construido el modelo, no aprendemos sobre sus propiedades mirándolo; tenemos que usarlo y manipularlo para descubrir sus secretos.

Dependiendo del tipo de modelo que tratemos, la construcción y manipulación de un modelo equivalen a distintas actividades que requieren metodologías distintas. Los modelos materiales parecen sencillos, ya que se usan en contextos experimentales comunes (*e.g.*, colocamos el modelo de un auto en un túnel de viento y medimos su resistencia al aire). Por ello, en tanto nos concierne el aprendizaje sobre los modelos, los modelos materiales no plantean preguntas que vayan más allá de las que conciernen a la experimentación en general.

Este no es el caso con los modelos ficticios y abstractos. ¿Qué restricciones hay en su construcción y cómo los manipulamos? Una respuesta natural sería que lo hiciéramos nosotros al realizar un experimento mental. Distintos autores (e.g., Brown, 1991; Gendler, 2000; El Skaf y Stuart, 2024; Norton, 1991; Reiss, 2003; Sorensen, 1992; Sartori, 2023) han explorado esta línea argumental, pero han alcanzado conclusiones muy distintas – y, comúnmente, en conflicto – acerca de

cómo se realizan los experimentos mentales y cuál es el estatus de sus resultados (para más detalles, véase la entrada de Experimentos mentales).

Una importante clase de modelos es de naturaleza computacional. En algunos casos, es posible obtener resultados o resolver ecuaciones de un modelo matemático de forma analítica. No obstante, este no es el común de los casos. En este punto, las computadoras tienen un gran impacto porque nos permiten resolver problemas que, de otro modo, resultan intratables. Por lo tanto, los métodos computacionales nos proporcionan conocimiento sobre (las consecuencias de) los modelos donde los métodos analíticos permanecen en silencio. Muchas partes de las investigaciones actuales – tanto en las ciencias naturales como en las sociales – dependen de simulaciones computacionales, que ayudan a los científicos a explorar las consecuencias de modelos que no pueden investigarse de otros modos. Algunos ejemplos [de casos] que se exploran mediante simulaciones computacionales son: la formación y desarrollo de las estrellas y galaxias, la dinámica de las reacciones a de iones pesados a altas energías, la evolución de la vida, los estallidos bélicos, la progresión de una economía, el comportamiento moral y las consecuencias de los de los procesos en la toma de decisiones de una organización, por mencionar algunos.

Las simulaciones computacionales también tienen importancia heurística. Ellas pueden sugerir nuevas teorías, modelos e hipótesis basadas, por ejemplo, en una exploración sistemática del espacio de parámetros del modelo (Hartmann, 1996). Sin embargo, las simulaciones computacionales también conllevan peligros metodológicos. Por ejemplo, ellas pueden proporcionar resultados engañosos porque, dada la naturaleza discreta de los cálculos realizados en una computadora digital, sólo permiten explorar una parte [de la imagen] completa del espacio de parámetros, y este subespacio no refleja necesariamente todas las características importantes del modelo. La severidad de este problema es, de algún modo, mitigada por el aumento del poder de las computadoras modernas. No obstante, la disponibilidad de mayor potencia computacional también puede tener efectos adversos: puede

incentivar a los científicos a proponer rápidamente modelos cada vez más complejos, pero conceptualmente prematuros que involucran presuposiciones o mecanismos pobremente entendidos y demasiados parámetros ajustables adicionales (Para una discusión acerca de un problema relacionado en ciencias sociales, véase Braun y Saam, 2015, Ch. 3). Esto puede llevar a un incremento en la adecuación empírica – que puede ser bienvenido para ciertas tareas de previsión –, pero no necesariamente a un mejor entendimiento de los mecanismos subyacentes. Como resultado, el uso de simulaciones computacionales puede modificar el peso que les asignamos a varios objetivos en las ciencias. Finalmente, la disponibilidad de [mayor] potencia computacional puede seducir a los científicos para realizar cálculos que no tienen el grado de confiabilidad que uno esperaría. Esto ocurre, por ejemplo, cuando las computadoras se usan para propagar distribuciones de probabilidad a lo largo del tiempo, lo cual puede resultar engañoso (véase Frigg, 2014). Por consiguiente, es importante no dejarse llevar por los medios que ofrecen las nuevas y potentes computadoras, perdiendo de vista el objetivo genuino de la investigación. Para una discusión sobre más problemas relacionados con las simulaciones computacionales, referimos al lector a la entrada de Simulaciones computacionales en ciencias.

3.2 Aprendiendo acerca de los sistemas objetivo

Una vez que tenemos conocimiento sobre el modelo, este conocimiento debe "traducirse" en conocimiento sobre el sistema objetivo. Es en este punto en el que la función representacional de los modelos vuelve a ser importante: si un modelo representa, entonces puede instruirnos acerca de la realidad porque al menos algunas partes o aspectos del modelo tienen partes o aspectos correspondientes en la realidad. No obstante, si el aprendizaje está conectado a la representación y si hay distintos tipos de representaciones (analogías, idealizaciones, etc.), entonces también hay distintos tipos de aprendizaje. Por ejemplo, si tenemos un modelo que consideramos una descripción realista, la transferencia de

conocimiento del modelo al objeto de estudio se logra de manera distinta a cuando lidiamos con una analogía o con un modelo que implica presuposiciones idealizadas. Para una discusión sobre las distintas formas en que la función representacional de los modelos puede ser explotada para aprender sobre el objetivo, referimos al lector a la entrada de Representación científica.

3.3 Explicaciones con modelos

Algunos modelos explican; ahora bien, ¿cómo pueden cumplir esta función, dado que típicamente implican idealizaciones? ¿estos modelos explican *a pesar de* o *a raíz de* las idealizaciones que involucran? ¿el uso explicativo de los modelos presupone que representan, o los modelos no representacionales también pueden explicar? ¿Y qué tipo de explicación proporcionan los modelos?

Hay una extensa tradición que exige que los *explanans* de una explicación científica sean verdaderos. Encontramos este requisito tanto en el modelo nomológico-deductivo (Hempel, 1965) como en la literatura más reciente. Por ejemplo, Strevens (2008, p. 297) afirma que "ninguna aproximación causal de la explicación […] permite que los modelos no verídicos expliquen". Para una discusión más exhaustiva, véase también Colombo *et al.* (2015).

Los autores que trabajan en esta tradición niegan que las idealizaciones contribuyan positivamente a la explicación y exploran cómo los modelos pueden explicar a pesar de estar idealizados. McMullin (1968, 1985) argumenta que una explicación causal basada en un modelo idealizado sólo deja fuera las propiedades que resultan irrelevantes para la respectiva tarea explicativa (véase también Salmon, 1984; y Piccinini y Craver, 2011 para una discusión sobre bocetos de mecanismo). Friedman (1974) argumenta que un modelo más realista (y, por lo tanto, menos idealizado) explica mejor [en términos] de la teoría de la unificación. La idea es que las idealizaciones pueden (al menos en principio) ser deidealizadas (para una discusión crítica de esta afirmación en el contexto

del debate sobre explicaciones científicas, véase Batterman, 2002; Bokulich, 2011; Morrison, 2005, 2009; Jebeile y Kennedy, 2015; y Rice, 2015). Strevens (2008) argumenta que un modelo explicativo causal debe ofrecer una representación precisa de las relaciones o procesos causales relevantes que el modelo comparte con el sistema objetivo. Las suposiciones idealizadas de un modelo no hacen diferencia en el fenómeno en consideración y, por consiguiente, resultan explicativamente irrelevantes. En contraposición, Potochnik (2017) y Rice (2015) argumentan que los modelos explicativos pueden *distorsionar* directamente muchas de las causas que *hacen la diferencia*.

De acuerdo con la teoría de Woodward (2003), los modelos son herramientas para identificar las relaciones causales que se sostienen entre ciertos factores o procesos, y son estas relaciones las que realizan el trabajo explicativo. Más específicamente, las explicaciones proporcionan información sobre los patrones de dependencias contrafácticas entre el *explanans* y el *explanandum* que

> "Nos habilitan a ver qué clase de diferencias habría tenido lugar en el *explanandum* si los factores citados en el *explanans* hubiesen sido diferentes de diversas maneras posibles. (Woodward, 2003, p. 11)

Las caracterizaciones causales de la explicación también han suscitado varias afirmaciones sobre cómo los modelos idealizados pueden proporcionar explicaciones, explorando así hasta qué punto la idealización permite que el modelo explicativo haga una representación tergiversada de factores causales irrelevantes (Elgin y Sober, 2002; Strevens, 2004, 2008; Potochnik, 2007; Weisberg, 2007, 2013). Sin embargo, tener características causalmente relevantes en común con el sistema real sigue desempeñando un papel esencial al mostrar cómo los modelos idealizados pueden ser explicativos.

No obstante, ¿es realmente la veracidad del *explanans* lo que hace que el modelo sea explicativo? Otros autores han seguido una línea más radical y argumentan que los modelos falsos no explican sólo *a pesar de* su falsedad, sino, de hecho, a *raíz de* ella. Cartwright (1983, p. 44) sostiene

que "la verdad no explica mucho". En su denominada "teoría del simulacro explicativo", ella sugiere que nosotros explicamos un fenómeno al construir un modelo que ajusta al fenómeno en el marco básico de una gran teoría (1983, Ch. 8). De acuerdo con esta teoría, el modelo mismo es la explicación que buscamos. Esto cuadra bien con las intuiciones científicas básicas, pero nos deja con la pregunta de qué noción de la explicación estamos trabajando (véase también Elgin y Sober, 2002) y de qué funciones explicativas desempeñan las idealizaciones en los modelos explicativos (Rice, 2018, 2019). Wimsatt (2007, Ch. 6) destaca el papel de los modelos falsos como medios para llegar a teorías verdaderas. Batterman y Rice (2014) argumentan que los modelos explican, ya que los detalles que caracterizan a un sistema en específico no importan para la explicación. Bokulich (2008, 2009, 2011, 2012) sigue una línea de razonamiento similar y sostiene que el poder explicativo de los modelos está estrechamente ligado a su naturaleza ficticia. Bokulich (2009) y Kennedy (2012) presentan una caracterización no representacional de la explicación mediante modelos (véase también Jebeile y Kennedy, 2015). Reiss (2012) y Woody (2004) ofrecen una discusión general sobre la relación entre representación y explicación.

3.4 Entendimiento con modelos

Muchos autores han señalado que el entendimiento es uno de los objetivos centrales de las ciencias (véase, por ejemplo, de Regt, 2017; Elgin, 2017; Khalifa, 2017; Potochnik, 2017). En algunos casos, queremos entender un fenómeno en particular (*e.g.,* por qué el cielo es azul); en otros casos, queremos entender una teoría científica específica (*e.g.,* la mecánica cuántica) que da cuenta del fenómeno en cuestión. A veces, ganamos entendimiento de un fenómeno al entender la teoría o el modelo correspondiente. Por ejemplo, la teoría del electromagnetismo de Maxwell nos ayuda a entender por qué el cielo es azul. No obstante, es controvertido si entender un fenómeno *siempre* presupone un entendimiento de la teoría correspondiente (de Regt, 2009, p. 26).

A pesar de que hay muchas formas distintas de obtener entendimiento, los modelos y la actividad del modelamiento científico son particularmente importantes aquí (de Regt *et al.*, 2009; Morrison, 2009; Potochnik, 2017; Rice, 2016). Esta intuición puede rastrearse, al menos, hasta Lord Kelvin que, en sus famosas *Lecturas de Baltimore sobre Dinámica Molecular y la Teoría ondulatoria de la Luz* de 1884, sostuvo que "la evaluación de [la pregunta] '¿entendemos o no un tema particular en física?' es '¿podemos hacer un modelo mecánico al respecto?'" (Kelvin 1884 [1987, p. 111]; véase también Bailer-Jones 2009, Ch. 2; y de Regt 2017, Ch. 6).

Ahora bien, ¿por qué desempeñan los modelos tal papel crucial en el entendimiento de una materia? Elgin (2017) argumenta que esto no es *a pesar de*, sino *a raíz de* que los modelos son literalmente falsos. Ella considera a los modelos falsos como "falsedades felices" que ocupan un lugar central en la epistemología de las ciencias y menciona al modelo de gases ideales en mecánica estadística y al modelo de Hardy-Weinberg en genética como ejemplos de modelos literalmente falsos que son centrales para sus respectivas disciplinas. El entendimiento es holístico y concierne a un tópico, disciplina o materia en lugar de enunciados aislados o hechos. Entender un contexto significa tener

> "Un compromiso epistémico con un cuerpo de información comprensivo y sistemáticamente enlazado que está fundado en hechos, es debidamente receptivo a razones o evidencias y permite [generar] inferencias no triviales, argumentos y, quizás, acciones sobre el tópico al que pertenece la información" (Elgin, 2017, p. 44)

Los modelos desempeñan un papel crucial en la búsqueda de estos compromisos epistémicos. Para una discusión sobre la teoría de Elgin acerca de los modelos y el entendimiento, véase Baumberger y Brun (2017) y Frigg y Nguyen (2021).

Elgin (2017), Lipton (2009) y Rice (2016) todos argumentan que los modelos pueden usarse para entender, independientemente de su capacidad para ofrecer una explicación. Otros autores, entre ellos

Strevens (2008, 2013), sostienen que el entendimiento presupone una explicación científica y que

> "Un individuo tiene entendimiento científico de un fenómeno sólo en los casos en los que él aprehende una explicación científica correcta de aquel fenómeno (Strevens, 2013, p. 510; véase, sin embargo, Sullivan y Khalifa, 2019)

Desde esta perspectiva, el entendimiento consiste en una forma particular de acceso epistémico que un individuo científico tiene a una explicación. Para Strevens, este aspecto es la "aprehensión", mientras que para de Regt (2017) es la "inteligibilidad". Es importante señalar que tanto Strevens como de Regt sostienen que dichos aspectos "subjetivos" constituyen un tema valioso de investigación en la filosofía de las ciencias. Esto contrasta con la concepción tradicional (véase, e.g., Hempel, 1965) que delega [estos aspectos] al dominio de la psicología. Véase Friedman (1974), Trout (2002) y Reutlinger *et al.* (2018) para una discusión más exhaustiva sobre el entendimiento.

3.5 Otras funciones cognitivas

Además de las funciones ya mencionadas, se ha enfatizado de distintas formas que los modelos realizan una variedad de otras funciones cognitivas. Knuuttila (2005, 2011) argumenta que el valor epistémico de los modelos no se limita a su función representacional y desarrolla una teoría que concibe a los modelos como artefactos epistémicos que nos permiten recolectar conocimiento de diversas maneras. Nersessian (1999, 2010) destaca el papel de los modelos análogos en la formación de conceptos y [en] otros procesos cognitivos. Hartmann (1995) y Leplin (1980) discuten [sobre] los modelos como herramientas para la construcción de teorías y enfatizan su valor heurístico y pedagógico. Epstein (2008) enumera un conjunto de funciones específicas de los modelos en las ciencias sociales. Peschard (2011) investiga la forma en la que los modelos pueden usarse para construir otros modelos y generar

nuevos sistemas objetivos. Isaac (2013) discute los usos no explicativos de los modelos que no dependen de sus capacidades representacionales.

Las compensaciones [*trade-offs*] son una parte inherente del modelamiento científico, ya que los modelos deben mantener un equilibrio entre las desideratas en competencia, como la precisión, la generalidad y la simplicidad. Richard Levins (1966) articuló este desafío en su ensayo "La estrategia de construcción de modelos en biología poblacional", argumentando que no existe un modelo único que pueda maximizar simultáneamente las tres desideratas. En cambio, los científicos deben tomar decisiones estratégicas sobre qué aspectos del sistema priorizar en sus modelos. Por ejemplo, un modelo puede sacrificar algo de realismo para lograr una mayor generalidad, o bien simplificar dinámicas complejas para hacerlo computacionalmente manejable.

El marco teórico de Levins ha desatado extensas discusiones filosóficas; algunos han argumentado que las compensaciones son inevitables en el modelamiento (Weisberg, 2006; Odenbaugh, 2006), mientras que otros han propuesto aproximaciones integrativas que combinan múltiples modelos para lograr un entendimiento más comprensivo de sistemas complejos (Orzack y Sober, 1993; Matthewson, 2011). Yoshida (2021) ha enfatizado la importancia de la yuxtaposición de múltiples modelos, en la que los distintos modelos se comparan y contrastan para ganar un entendimiento más matizado del fenómeno en investigación.

Estas discusiones destacan los desafíos epistémicos de la construcción de modelos a lo largo de las ciencias. Sea en física, biología, economía o ciencias del clima, las compensaciones siempre están presentes y los científicos deben considerar con cuidado las fortalezas y limitaciones de sus modelos. Al reconocer estas compensaciones, los investigadores pueden tomar decisiones más informadas sobre cómo representar y estudiar sistemas complejos, y los filósofos pueden entender mejor el papel de los modelos en la investigación científica.

4. Modelos y teoría

Una pregunta importante concierne a la relación entre los modelos y las teorías. Hay un espectro completo de posiciones desde los modelos siendo subordinados a las teorías hasta los modelos siendo independientes de las teorías.

4.1 Modelos como subsidiarios de la teoría

Para discutir sobre la relación entre los modelos y las teorías en las ciencias, es útil recapitular brevemente las nociones de modelo y de teoría en la lógica. Una *teoría* se considera un conjunto (con frecuencia deductivamente cerrado) de enunciados en un lenguaje formal. Un *modelo* es una estructura (en el sentido introducido en la Sección 2.3) que hace verdaderos todos los enunciados de una teoría cuando sus símbolos se interpretan como si refiriesen a los objetos, relaciones y funciones de una estructura. La estructura es un *modelo de* la teoría, en el sentido de que es correctamente descrito por la teoría (véase Bell y Machover, 1977, o Hodges, 1997, para detalles). Los modelos lógicos también son referidos, a veces, como "modelos de la teoría" para indicar que son interpretaciones de un sistema formal abstracto.
A veces, los modelos en ciencias trasladan de la lógica la idea de ser interpretaciones de un cálculo abstracto (Hesse, 1967). Esto es sobresaliente en física, donde las leyes generales – tales como la ecuación del movimiento de Newton – yacen en el corazón de la teoría. Estas leyes se aplican a un sistema particular – e.g., un péndulo – mediante la elección de una función de fuerza especial, la toma de suposiciones sobre la distribución de masa del péndulo, etc. El modelo resultante, entonces, es una interpretación (o realización) de la ley general.

Es importante mantener separadas las nociones de un modelo lógico y de uno representacional (Thomson-Jones, 2006): son conceptos distintos. A veces, un modelo puede ser lógico sin ser representacional, y *viceversa*. No obstante, esto no significa que, a veces, no pueda haber un modelo de ambos sentidos en uno solo. De hecho, como señala Hesse

(1967), muchos modelos en las ciencias son tanto lógicos como representacionales. El modelo del movimiento planetario de Newton es un ejemplo de ello: el modelo – que consiste de dos esferas homogéneas perfectas, que se atraen gravitacionalmente la una a la otra y están localizadas en un espacio, por lo demás, vacío – es, simultáneamente, un modelo lógico (porque hace verdaderos los axiomas de la mecánica newtoniana cuando son interpretados como refiriendo al modelo) y un modelo representacional (porque representa al sol y la tierra reales).

Hay dos caracterizaciones centrales de las teorías científicas: la denominada concepción sintáctica de las teorías y la denominada concepción semántica de las teorías (véase la entrada de La estructura de las teorías científicas). En ambas concepciones, los modelos desempeñan un rol subsidiario frente a las teorías, aunque de maneras muy distintas. La concepción sintáctica de las teorías (véase la entrada en la sección de La concepción sintáctica) conserva las nociones lógicas de un modelo y de una teoría. Interpreta la teoría como un conjunto de enunciados en un sistema lógico axiomatizado y al modelo como una interpretación alternativa de un cálculo determinado (Braithwaite, 1953; Campbell, 1920 [1957]; Nagel, 1961; Spector, 1965). Por ejemplo, si tomamos la matemática usada en la teoría cinética de los gases y reinterpretamos los términos de su cálculo de tal forma que se refieran a bolas de billar, [entonces] las bolas de billar son un modelo de la teoría cinética de los gases en el sentido de que todos los enunciados de la teoría aparecen como verdaderos. El modelo se supone que es algo con lo que estamos familiarizados y sirve para el propósito de hacer más palpable el cálculo formal abstracto. Una teoría dada puede tener distintos modelos, y el modelo que elegimos depende tanto de nuestros objetivos como de nuestro conocimiento previo. Defensores de la concepción sintáctica discrepan sobre la importancia de los modelos. Carnap y Hempel pensaron que los modelos sólo servían para fines pedagógicos o estéticos y son, en último término, dispensables, ya que toda la información relevante está contenida en la teoría (Carnap, 1938; Hempel, 1965; véase también Bailer-Jones, 1999). Nagel (1961) y Braithwaite (1953), por otro

lado, enfatizan el papel *heurístico* de los modelos y Schaffner (1969) sugiere que los términos teóricos obtienen, al menos, parte de su sentido de los modelos.

La concepción semántica de las teorías (véase la entrada en la sección La concepción semántica) prescinde de los enunciados en un sistema lógico axiomatizado e interpreta la teoría como una familia de modelos. Desde esta perspectiva, una teoría es literalmente una clase, conglomerado o familia de modelos – los modelos son las piezas de construcción con las que se elaboran las teorías científicas. Distintas versiones de la concepción semántica trabajan con distintas nociones de modelo, pero, como se señaló en la Sección 2.3, en la concepción semántica los modelos se interpretan mayoritariamente como estructuras teóricas de conjuntos. Para una discusión sobre diferentes opciones, referimos al lector a la entrada relevante de esta enciclopedia (enlazada al inicio de este párrafo).

4.2 Los modelos como independientes de la teoría

Tanto en la concepción sintáctica como en la semántica de las teorías, los modelos se consideran subordinados a la teoría y no desempeñan un papel fuera del contexto de la teoría. Esta visión de los modelos ha sido desafiada de distintas maneras, con autores señalando que los modelos disponen de varios grados de libertad respecto de la teoría y funcionan de manera autónoma en muchos contextos. La independencia puede adoptar muchas formas y gran parte de la literatura sobre modelos concierne a la investigación de distintas formas de independencia.

Modelos completamente independientes de la teoría. La separación más radical del análisis de los modelos centrado en la teoría es el reconocimiento de que existen modelos completamente independientes de cualquier teoría. Un ejemplo de estos modelos es el modelo de Lotka-Volterra. El modelo describe la interacción entre dos poblaciones: una de animales depredadores y otra de presas (Weisberg, 2013). El modelo fue construido usando sólo la matemática de las ecuaciones diferenciales y

suposiciones relativas al sentido común sobre los depredadores y las presas. No se apela a una teoría de las interacciones depredador-presa ni a una teoría del crecimiento poblacional; el modelo es independiente de las teorías sobre su materia de estudio. Si un modelo se construye en un dominio donde no hay una teoría disponible, entonces el modelo a veces se refiere como un "modelo sustituto" (Groenewold, 1961) porque el modelo sustituye una teoría.

Modelos como medio para explorar la teoría. Los modelos también pueden usarse para explorar teorías (Morgan y Morrison, 1999). Una forma obvia en la que esto ocurre es cuando un modelo es un modelo lógico de la teoría (véase Sección 4.1). Un modelo lógico es un conjunto de objetos y propiedades que hacen verdadero un enunciado formal, de modo que uno ve en el modelo cómo se desarrollan los axiomas de la teoría en una configuración particular y qué clase de comportamientos ellos dictan. Sin embargo, no todos los modelos que son usados para explorar teorías son modelos lógicos y, [además,] los modelos pueden representar características de las teorías de otros modos. Considérese, por ejemplo, la teoría del caos. Las ecuaciones de sistemas no lineales, tales como las que describen el problema de los tres cuerpos, tienen soluciones muy complejas para estudiar con métodos rudimentarios, e incluso las simulaciones computacionales están limitadas en muchas formas. Consideraciones abstractas sobre los comportamientos cualitativos de las soluciones muestran que existe un mecanismo apodado "estirar y plegar" (véase la entrada de Caos). Para obtener una idea de la complejidad de las dinámicas exhibidas por [el mecanismo de] estirar y plegar, Smale propuso estudiar un modelo simple de fluido – ahora conocido como el "mapa de la herradura" (Tabor, 1989) – que nos proporciona una comprensión importante de la naturaleza del [mecanismo de] estirar y plegar. Otros ejemplos de modelos de este tipo son el modelo de anillo de Kac, usado para estudiar las propiedades de equilibrio de sistemas en mecánica estadística (Lavis, 2008), y el domo de Norton en mecánica newtoniana (Norton, 2003).

Modelos como complementos de las teorías. Una teoría puede estar incompletamente especificada en el sentido de que sólo impone ciertas restricciones generales, pero permanece en silencio sobre los detalles de las situaciones concretas, los cuales son proporcionados por un modelo (Redhead, 1980). Un caso especial de esta situación es cuando una teoría cualitativa ya es conocida y el modelo introduce mediciones cuantitativas (Apostel, 1961). El ejemplo de Redhead de una teoría indeterminada de esta forma es la teoría axiomática de la mecánica cuántica de campos, que solo impone ciertas restricciones generales a la mecánica cuántica de campos, pero no proporciona una caracterización de campos particulares. Harré (2004) señala que los modelos pueden complementar las teorías al proporcionar mecanismos para procesos que quedan sin especificar en la teoría, aunque estos sean responsables de provocar el fenómeno observado.

Las teorías pueden resultar muy complicadas de manejar. En tales casos, un modelo puede complementar una teoría al proporcionar una versión simplificada del escenario teórico, lo que permite [llegar a] una solución. La cromodinámica cuántica, por ejemplo, no puede usarse fácilmente para investigar la física de un núcleo atómico, aunque sea la teoría fundamental relevante. Para evitar esta dificultad, los físicos construyeron modelos fenomenológicos manejables (como el modelo del bolso del MIT) que efectivamente describen el nivel relevante de libertad del sistema en consideración (Hartmann, 1999, 2001). La ventaja de estos modelos es que producen resultados [en casos] en los que las teorías permanecen en silencio. Su inconveniente es que, normalmente, no es claro cómo se entiende la relación entre el modelo y la teoría, ya que ambos, estrictamente hablando, son contradictorios.

Modelos como teorías preliminares. La noción de un modelo como un sustituto de una teoría está estrechamente ligada a la noción de un *modelo de desarrollo.* Este término fue acuñado por Leplin (1980), quien señaló cómo de útiles eran los modelos en el desarrollo temprano de la teoría cuántica, y ahora se usa como una noción paraguas para cubrir los

casos en los que los modelos son una especie de ejercicio preliminar para [obtener una] teoría.

[Dicha noción] también está estrechamente ligada a la noción de un modelo probatorio (o "modelo de estudio"). Los modelos de este tipo no cumplen una función representacional y no se espera que nos instruyan sobre nada más allá del modelo mismo. El propósito de estos modelos es examinar nuevas herramientas teóricas que más tarde se emplean para construir modelos representacionales. En la teoría de campos, por ejemplo, el denominado modelo-φ^4 fue extensivamente estudiado no porque se creyese que representaba cualquier aspecto real, sino porque servía para distintas funciones heurísticas: la simplicidad del modelo-φ^4 permitió a los científicos "tener una sensación" de cómo es la teoría cuántica de campos y extraer algunas características generales que este modelo simple compartía con [otros modelos] más complicados. Los físicos pudieron estudiar técnicas complejas, tales como la renormalización en una configuración sencilla, y fue posible familiarizarse con mecanismos importantes – en este caso, la ruptura de simetría – que podrían usarse luego en diferentes contextos (Hartmann, 1995). Esto no aplica sólo en la física. Como Wimsatt (1987, 2007) señala, un modelo falso en genética puede realizar muchas funciones útiles, entre las que encontramos las siguientes: el modelo falso puede ayudar a responder preguntas sobre los modelos más realistas, proporcionar una arena para responder preguntas sobre las propiedades de modelos más complejos, "excluir" fenómenos que no hubieran sido observados de otro modo, servir como un caso límite de un modelo más general (o dos modelos falsos pueden definir los extremos de un *continuum* de casos en los que se supone que se encuentra el caso real), o guiar en la identificación de variables relevantes y en la estimación de sus valores.

Modelos interpretativos. Cartwright (1983, 1999) argumenta que los modelos no sólo ayudan en la aplicación de teorías de algún modo incompletas; ella afirma que los modelos también están involucrados *siempre* que una teoría con una estructura matemática general es aplicada. Las principales teorías de la física – mecánica clásica, electrodinámica,

mecánica cuántica y así sucesivamente – entran en esta categoría. Las teorías de este tipo están formuladas en términos de conceptos abstractos que deben concretizarse para que la teoría proporcione una descripción del sistema objetivo, y [para] concretizar los conceptos relevantes se introducen objetos y procesos idealizados. Por ejemplo, cuando aplicamos la mecánica clásica, el concepto abstracto de fuerza debe reemplazarse por una fuerza concreta, como la gravedad. Para obtener ecuaciones manejables, este proceso debe aplicarse a un escenario simplificado, por ejemplo, [uno] de dos planetas homogéneos y perfectamente esféricos en un espacio por lo demás vacío, en lugar de la realidad en toda su complejidad. El resultado es un modelo interpretativo que cimenta la aplicación de teorías matemáticas a objetivos del mundo real. Tales modelos son independientes de la teoría en tanto esta no determina su forma, y, con todo, son necesarios para la aplicación de la teoría a un problema concreto.

Modelos como mediadores. La relación entre modelos y teorías puede ser complicada y desordenada. Los contribuyentes de una colaboración programática de ensayos editada por Morgan y Morrison (1999) se reunieron en torno a la idea de que los modelos son instrumentos que median entre las teorías y el mundo. Los modelos son "agentes autónomos" en tanto son independientes de las teorías y de sus sistemas objetivos, y es esta independencia la que les permite mediar entre ambos. Las teorías no nos entregan algoritmos para la construcción de un modelo; ellas no son "máquinas expendedoras" en las que uno puede insertar un problema y obtener un modelo (Cartwright 1999). La construcción de un modelo normalmente requiere conocimiento detallado sobre materiales, esquemas de aproximación y configuraciones que la teoría correspondiente no proporciona. Además, el funcionamiento interno de un modelo suele estar dirigido por una variedad de teorías que trabajan de manera cooperativa. En el modelamiento climático contemporáneo, por ejemplo, los elementos de distintas teorías – entre [las que encontramos] la dinámica de fluidos, la termodinámica y el electromagnetismo – se ponen a trabajar de manera

cooperativa. Lo que entrega los resultados no es la rigurosa aplicación de una teoría, sino las voces de distintas teorías cuando se usan en coro, la una con la otra en [la forma de] un modelo.

En casos complejos, como el estudio de un sistema láser o del clima global, los modelos y las teorías pueden entrelazarse a tal punto que resulta poco claro dónde deba dibujarse la línea entre ambos: ¿dónde termina el modelo y comienza la teoría? Este no es sólo un problema para el análisis filosófico; también surge en la práctica científica. Bailer-Jones (2002) entrevistó a un grupo de físicos acerca de su entendimiento de los modelos y de su relación con las teorías, y reportó concepciones ampliamente diversas: 1) no hay una diferencia sustantiva entre el modelo y la teoría; 2) los modelos *se vuelven* teorías cuando se incrementa su grado de confirmación; 3) los modelos contienen simplificaciones y omisiones, mientras que las teorías son precisas y completas; 4) las teorías son más generales que los modelos, y el modelamiento trata sobre la aplicación de teorías generales a casos específicos. La primera sugerencia pareciera ser muy radical para hacer justicia a las muchas facetas de la práctica en la que claramente se distinguen modelos y teorías. La segunda concepción está alineada con el lenguaje común, en el que los términos "modelo" y "teoría" se emplean a veces para expresar la actitud de alguien hacia una hipótesis particular. La frase "es sólo un modelo" indica que la hipótesis en juego es afirmada sólo de modo tentativo, o incluso se sabe que es falsa, mientras que algo recibe la etiqueta de "teoría" si ha adquirido cierto grado de aceptación general. Sin embargo, este uso de "modelo" difiere de los usos observados en las Secciones 1 a 3 y, por lo tanto, no resulta útil si el objetivo es entender la relación entre las teorías y los modelos científicos (e, incidentemente, uno puede igualmente descartar afirmaciones especulativas como "sólo una teoría"). La tercera propuesta es correcta al asociar los modelos con las idealizaciones y simplificaciones, pero se excede al limitar esto a los modelos; de hecho, las teorías también pueden contener idealizaciones y simplificaciones. La cuarta concepción parece estar estrechamente alineada con los modelos interpretativos y con la idea de que los modelos son mediadores, pero ser

más general es una noción gradual y, por ello, no proporciona un criterio divisorio claro para distinguir entre teorías y modelos.

5. Modelos y otros debates en filosofía de las ciencias

El debate sobre los modelos científicos tiene repercusiones importantes para otros problemas en la filosofía de las ciencias (para una caracterización histórica de la discusión filosófica sobre los modelos, véase Bailer-Jones, 1999). Tradicionalmente, los debates sobre, digamos, el realismo científico, el reduccionismo y las leyes de la naturaleza se expresaron en términos de teorías, porque estas se consideraban las principales portadoras del conocimiento científico. Una vez reconocido el importante lugar que ocupan los modelos en el edificio de las ciencias, estos problemas tuvieron que ser reconsiderados con un enfoque en los modelos. La pregunta es si – y, de ser así, cómo – las discusiones acerca de estos problemas cambian cuando cambiamos el enfoque de las teorías a los modelos. Hasta ahora, no ha surgido ninguna caracterización comprensiva basada en modelos de alguno de estos problemas, pero los modelos han dejado una huella importante en las discusiones sobre estos tópicos.

5.1 Modelos, realismo y las leyes de la naturaleza

Como hemos visto en la Sección 1, los modelos típicamente ofrecen una representación distorsionada de sus objetivos. Si uno ve la ciencia como principalmente basada en modelos, [entonces] esto puede tomarse como sugerente de una interpretación antirrealista de las ciencias. No obstante, los realistas niegan que la presencia de idealizaciones en los modelos ocasione que una aproximación realista de las ciencias [sea] imposible y señalan que un buen modelo – aunque no sea literalmente verdadero – es, con frecuencia, al menos aproximadamente verdadero y/o puede ser mejorado mediante la deidealización (Laymon, 1985; McMullin, 1985; Nowak, 1979; Brzezinski y Nowak, 1992).

Aparte de las preocupaciones usuales sobre la elusividad de la noción de verdad aproximada (para una discusión, véase la entrada de Semejanza con la verdad), los antirrealistas han tenido problemas con esta réplica por dos razones (relacionadas). Primero, como señala Cartwright (1989), no hay razón para asumir que uno siempre puede mejorar un modelo al agregar correcciones deidealizadas. En segundo lugar, parece que la deidealización no concuerda con la práctica científica porque es inusual que los científicos inviertan esfuerzos en deidealizar repetidamente un modelo existente. Más bien, cambian a un marco de modelamiento distinto cuando los ajustes por realizar se vuelven muy complejos (Hartmann, 1998). Los distintos modelos del núcleo atómico son un ejemplo de ello: una vez se dieron cuenta de que los efectos de capas son importantes para entender varios fenómenos subatómicos, el modelo (colectivo) de gota líquida fue puesto a un lado y el modelo (partícula-única) de capas fue desarrollado para dar cuenta de los hallazgos correspondientes. Una dificultad adicional de la deidealización es que la mayoría de las idealizaciones no son "controladas". Por ejemplo, no es claro de qué modo uno podría deidealizar el modelo de bolsa del MIT para llegar, eventualmente, a la cromodinámica cuántica, la teoría subyacente supuestamente correcta.

Un argumento antirrealista adicional, el "argumento de los modelos incompatibles", toma como punto de partida la observación de que los científicos, con frecuencia, usan exitosamente muchos modelos incompatibles de *uno y el mismo* sistema objetivo para fines predictivos (Morrison, 2000). Estos modelos aparentemente se contradicen entre sí al atribuir propiedades distintas al mismo sistema objetivo. En física nuclear, por ejemplo, el modelo de gota líquida explora la analogía entre el núcleo atómico y una gota (cargada) de fluido, mientras que el modelo de capas nucleares describe propiedades nucleares en términos de las propiedades de los protones y neutrones que constituyen el núcleo atómico. Esta práctica parece causar un problema para el realismo científico: los realistas suelen sostener que existe una relación estrecha entre el éxito predictivo de una teoría y su aproximada veracidad. Ahora

bien, si muchos modelos del mismo sistema son predictivamente exitosos y si estos modelos son mutuamente inconsistentes, entonces es difícil sostener que todos ellos sean aproximadamente verdaderos.

Los realistas pueden reaccionar a este argumento de diversas formas. Primero, ellos pueden desafiar la afirmación de que los modelos en cuestión son, en efecto, predictivamente exitosos. Si los modelos no son buenos para predecir, entonces el argumento queda bloqueado. Segundo, ellos pueden defender una versión del "realismo perspectival" (Giere, 20006; Massimi, 2022; Rueger, 2005). Defensores de esta posición (que también se llama, a veces, "perspectivismo") la sitúan en algún lugar entre el realismo científico "estándar" y el antirrealismo; dónde yace exactamente esta posición intermedia es un tema activo de debate (Massimi, 2018a, b; Saatsi, 2016; Teller, 2018; y las contribuciones de Massimi y McCoy, 2019). Tercero, los realistas pueden negar que haya un problema en primer lugar, ya que los modelos científicos – que son siempre idealizados y, por tanto, falsos hablando en sentido estricto – son sólo el vehículo equivocado para hacer un punto sobre el realismo (que debería discutirse en términos de teorías).

Sober (2015) relaciona el debate sobre la selección de modelos en la estadística con el debate entre los realistas e instrumentalistas en la filosofía de las ciencias. Mientras que el realista considera que el objetivo de las ciencias es encontrar teorías verdaderas, el instrumentalismo de Sober considera que el objetivo de las ciencias es encontrar teorías predictivamente precisas. Al final, Sober defiende el eslogan "instrumentalismo para los modelos, realismo para los modelos ajustados" (2015, p. 144). Véase también la entrada de filosofía de la estadística.

Un punto particularmente focal del debate realista [concierne] a las leyes de la naturaleza, donde emergen las preguntas [relativas a] qué son las leyes y si son retratadas de modo fidedigno por nuestras representaciones científicas. De acuerdo con las dos teorías actualmente

dominantes, el enfoque del mejor sistema y el enfoque determinista[8], las leyes de la naturaleza se entienden como universales dentro de su alcance, lo que significa que son aplicables a todo lo que hay en el mundo (para una discusión sobre leyes, véase la entrada sobre leyes de la naturaleza). Esta aproximación a las leyes no parece encajar con una concepción que sitúa a los modelos en el centro de la investigación científica. ¿Qué papel desempeñan las leyes generales en las ciencias si son los modelos [aquello] que representan lo que ocurre en el mundo? ¿y cómo se relacionan los modelos y las leyes?

Una respuesta posible a estas preguntas es argumentar que las leyes de la naturaleza gobiernan entidades y procesos en un *modelo* en lugar de en el mundo. Las leyes fundamentales, desde esta aproximación, no establecen hechos sobre el *mundo*, sino que son verdaderas respecto de las entidades y procesos en el modelo. Esta concepción ha sido defendida desde distintas variantes: Cartwright (1983) argumenta que todas las leyes son leyes *ceteris paribus*. Cartwright (1999) utiliza las "capacidades" (que ella considera anteriores a las leyes) e introduce la noción de una "máquina nomológica". Esto es

> "Un arreglo (lo suficientemente) ajustado de componentes o factores con capacidades (lo suficientemente) estables que, en el tipo correcto de ambiente (lo suficientemente) estable, va a, con una operación repetida, dar lugar al tipo de comportamiento regular que representamos mediante nuestras leyes científicas."

[8] NT: original "*best-system approach*" y *"necessitarian approach"* son conceptos paraguas en la discusión sobre las leyes de la naturaleza que encapsulan dos posturas generales y comparten un núcleo que establece las leyes como universales dentro de su alcance. Por un lado, el enfoque del mejor sistema (*best-system approach*) sostiene que las leyes pertenecen a un sistema deductivo que es la mejor combinación de fuerza y simplicidad; por otro lado, el enfoque determinista (*necessitarian approach*) afirma que las leyes son necesariamente verdaderas. Para una exposición más amplia sobre estas posturas, véanse las secciones 2 y 8 de la entrada sobre Leyes de la naturaleza en la enciclopedia de Stanford (Carroll, John W., "Laws of Nature", *The Stanford Encyclopedia of Philosophy* (Winter 2025 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/win2025/entries/laws-of-nature/>).

(1999, p. 50; véase también la entrada acerca de leyes ceteris paribus)

Giere (1999) argumenta que las leyes de una teoría son concebidas de mejor modo no como la codificación de verdades generales sobre el mundo, sino, más bien, como enunciados abiertos que pueden completarse de diversas maneras en el proceso de construcción de modelos científicos más específicos. Posiciones similares también han sido defendidas por Teller (2001) y van Fraassen (1989).

5.2 Modelos y Reduccionismo

El problema de los modelos múltiples mencionado en la Sección 5.1 también plantea la pregunta de cómo se relacionan entre sí los diferentes modelos. Evidentemente, múltiples modelos del mismo sistema objetivo no suelen encontrarse en una relación deductiva, ya que normalmente se contradicen entre sí. Algunos [autores] (Cartwright, 1999; Hacking, 1983) han sugerido una imagen de las ciencias según la cual no hay relaciones sistemáticas que se sostengan entre los distintos modelos. Algunos modelos están enlazados porque representan el mismo sistema objetivo, pero esto no implica que estén en ninguna otra relación adicional (deductiva o de otro tipo). Nos encontramos con un mosaico de modelos y todos ellos se sostienen [de modo] *ceteris paribus* en sus respectivos dominios de aplicación.

Algunos [autores] argumentan que esta imagen es, al menos, parcialmente incorrecta, ya que hay varias relaciones de interés que se sostienen entre los distintos modelos y teorías. Estas relaciones abarcan desde minuciosas relaciones reductivas (Scheibe, 1997, 1999, 2001: esp. Chs. V.23 y V.24) y aproximaciones controladas sobre relaciones de límites singulares (Batterman, 2001 [2016]) a relaciones estructurales (Gähde, 1997) y relaciones más bien laxas denominadas “historias” (Hartmann, 1999; véase también Bokulich, 2003; Teller, 2002; y la colección de ensayos en la parte III de Hartmann *et al.*, 2008). Estas sugerencias se han realizado en la base de casos de estudio y [aún] queda

por ver si una caracterización más general de estas relaciones puede ser dada y si se puede proporcionar una justificación más detallada para ellas, por ejemplo, dentro de un marco bayesiano (los primeros pasos hacia un entendimiento bayesiano de las relaciones reductivas pueden ser encontrado en Dizadji-Bahmani *et al.*, 2011; Liefke y Hartmann, 2018; y Tešić, 2019).

Los modelos también figuran en el debate sobre la reducción y la emergencia en la física. Aquí, algunos autores argumentan que la aproximación moderna a la renormalización desafía el modelo de reducción de Nagel (1961) o la doctrina más amplia de reducciones (para una discusión crítica, véase, por ejemplo, Batterman, 2002, 2010, 2011; Morrison, 2012; y Saatsi y Reutlinger, 2018). Dizadji-Bahmani *et al.* (2010) ofrecen una defensa del modelo Nagel-Schaffner de reducción y Butterfield (2011a, b, 2014) argumenta que la renormalización es consistente con la reducción nageliana. Palacios (2019) muestra que las transiciones de fase son compatibles con el reduccionismo y Hartmann (2001) argumenta que el programa de investigación sobre teorías de campos efectivas es consistente con el reduccionismo (véase también Bain, 2013, y Franklin, 2020). Rosaler (2015) sostiene una forma "local" de reducción que considera que la relación fundamental de la reducción se sostiene entre modelos, no entre teorías, lo cual, no obstante, es compatible con el modelo Nagel-Schaffner de la reducción. Véase también la entrada sobre relaciones intrateóricas en física y sobre la reducción científica.

En las ciencias sociales, los modelos basados en agentes (MBAs) se utilizan cada vez más (Klein *et al.*, 2018). Estos modelos muestran cómo los patrones de comportamiento, sorprendentemente complejos a escala macro, pueden emerger de un número reducido de reglas simples de comportamiento de los agentes individuales y de sus interacciones. Esto plantea preguntas similares a las mencionadas anteriormente sobre la reducción y la emergencia en física, pero hasta ahora uno sólo puede encontrar comentarios dispersos sobre la reducción en la literatura. Véanse Weisberg y Muldoon (2009) y Zollman (2007) para la aplicación

de los MBAs a la epistemología y a las estructuras sociales en las ciencias, y Colyvan (2013) para una discusión sobre las preguntas metodológicas planteadas por los modelos normativos en general.

Bibliografía

Achinstein, Peter, 1968, *Concepts of Science: A Philosophical Analysis*, Baltimore, MD: Johns Hopkins Press.

Akerlof, George A., 1970, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism", *The Quarterly Journal of Economics*, 84(3): 488–500. doi:10.2307/1879431

Ankeny, Rachel A. and Sabina Leonelli, 2020, *Model Organisms*, Cambridge: Cambridge University Press.

Apostel, Leo, 1961, "Towards the Formal Study of Models in the Non-Formal Sciences", in Freudenthal 1961: 1–37. doi:10.1007/978-94-010-3667-2_1

Bailer-Jones, Daniela M., 1999, "Tracing the Development of Models in the Philosophy of Science", in Magnani, Nersessian, and Thagard 1999: 23–40. doi:10.1007/978-1-4615-4813-3_2

——, 2002, "Scientists' Thoughts on Scientific Models", *Perspectives on Science*, 10(3): 275–301. doi:10.1162/106361402321899069

——, 2009, *Scientific Models in Philosophy of Science*, Pittsburgh, PA: University of Pittsburgh Press.

Bailer-Jones, Daniela M. and Coryn A. L. Bailer-Jones, 2002, "Modeling Data: Analogies in Neural Networks, Simulated Annealing and Genetic Algorithms", in Magnani and Nersessian 2002: 147–165. doi:10.1007/978-1-4615-0605-8_9

Bain, Jonathan, 2013, "Emergence in Effective Field Theories", *European Journal for Philosophy of Science*, 3(3): 257–273. doi:10.1007/s13194-013-0067-0

Bandyopadhyay, Prasanta S. and Malcolm R. Forster (eds.), 2011, *Philosophy of Statistics* (Handbook of the Philosophy of Science 7), Amsterdam: Elsevier.

Barberousse, Anouk and Pascal Ludwig, 2009, "Fictions and Models", in Suárez 2009: 56–75.

Bartha, Paul, 2010, *By Parallel Reasoning: The Construction and Evaluation of Analogical Arguments*, New York: Oxford University Press. doi:10.1093/acprof:oso/9780195325539.001.0001

——, 2013 [2019], "Analogy and Analogical Reasoning", in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition). URL = <https://plato.stanford.edu/archives/spr2019/entries/reasoning-analogy/>

Batterman, Robert W., 2002, *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*, Oxford: Oxford University Press. doi:10.1093/0195146476.001.0001

——, 2001 [2016], "Intertheory Relations in Physics", in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition). URL = <https://plato.stanford.edu/archives/fall2016/entries/physics-interrelate>

——, 2010, "Reduction and Renormalization", in Gerhard Ernst and Andreas Hüttemann (eds.), *Time, Chance and Reduction: Philosophical Aspects of Statistical Mechanics*, Cambridge: Cambridge University Press, pp. 159–179.

——, 2011, "Emergence, Singularities, and Symmetry Breaking", *Foundations of Physics*, 41(6): 1031–1050. doi:10.1007/s10701-010-9493-4

Batterman, Robert W. and Collin C. Rice, 2014, "Minimal Model Explanations", *Philosophy of Science*, 81(3): 349–376. doi:10.1086/676677

Baumberger, Christoph and Georg Brun, 2017, "Dimensions of Objectual Understanding", in Stephen R. Grimm, Christoph Baumberger, and Sabine Ammon (eds.), *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*, New York: Routledge, pp. 165–189.

Bell, John and Moshé Machover, 1977, *A Course in Mathematical Logic*, Amsterdam: North-Holland.

Berry, Michael, 2002, "Singular Limits", *Physics Today*, 55(5): 10–11. doi:10.1063/1.1485555

Black, Max, 1962, *Models and Metaphors: Studies in Language and Philosophy*, Ithaca, NY: Cornell University Press.

Bogen, James and James Woodward, 1988, "Saving the Phenomena", *The Philosophical Review*, 97(3): 303–352. doi:10.2307/2185445

Bokulich, Alisa, 2003, "Horizontal Models: From Bakers to Cats", *Philosophy of Science*, 70(3): 609–627. doi:10.1086/376927

——, 2008, *Reexamining the Quantum–Classical Relation: Beyond Reductionism and Pluralism*, Cambridge: Cambridge University Press. doi:10.1017/CBO9780511751813

——, 2009, "Explanatory Fictions", in Suárez 2009: 91–109.

——, 2011, "How Scientific Models Can Explain", *Synthese*, 180(1): 33–45. doi:10.1007/s11229-009-9565-1

——, 2012, "Distinguishing Explanatory from Nonexplanatory Fictions", *Philosophy of Science*, 79(5): 725–737. doi:10.1086/667991

Bolker, Jessica A., 2009, "Exemplary and Surrogate Models: Two Modes of Representation in Biology", *Perspectives in Biology and Medicine*, 52: 485–499.

Braithwaite, Richard, 1953, *Scientific Explanation*, Cambridge: Cambridge University Press.

Braun, Norman and Nicole J. Saam (eds.), 2015, *Handbuch Modellbildung und Simulation in den Sozialwissenschaften*,

Wiesbaden: Springer Fachmedien. doi:10.1007/978-3-658-01164-2

Brewer, William F. and Clark A. Chinn, 1994, "Scientists' Responses to Anomalous Data: Evidence from Psychology, History, and Philosophy of Science", in *PSA 1994: Proceedings of the 1994 Biennial Meeting of the Philosophy of Science Association*, Vol. 1, pp. 304–313. doi:10.1086/psaprocbienmeetp.1994.1.193035

Brown, James, 1991, *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*, London: Routledge.

Brzezinski, Jerzy and Leszek Nowak (eds.), 1992, *Idealization III: Approximation and Truth*, Amsterdam: Rodopi.

Burian, Richard M., 1993, "How the Choice of Experimental Organism Matters: Epistemological Reflections on an Aspect of Biological Practice", *Journal of the History of Biology*, 26: 351–368.

Butterfield, Jeremy, 2011a, "Emergence, Reduction and Supervenience: A Varied Landscape", *Foundations of Physics*, 41(6): 920–959. doi:10.1007/s10701-011-9549-0

——, 2011b, "Less Is Different: Emergence and Reduction Reconciled", *Foundations of Physics*, 41(6): 1065–1135. doi:10.1007/s10701-010-9516-1

——, 2014, "Reduction, Emergence, and Renormalization", *Journal of Philosophy*, 111(1): 5–49. doi:10.5840/jphil201411111

Callender, Craig and Jonathan Cohen, 2006, "There Is No Special Problem about Scientific Representation", *Theoria*, 55(1): 67–85.

Campbell, Norman, 1920 [1957], *Physics: The Elements*, Cambridge: Cambridge University Press. Reprinted as *Foundations of Science*, New York: Dover, 1957.

Carnap, Rudolf, 1938, "Foundations of Logic and Mathematics", in Otto Neurath, Charles Morris, and Rudolf Carnap (eds.), *International Encyclopaedia of Unified Science*, Volume 1, Chicago, IL: University of Chicago Press, pp. 139–213.

Cartwright, Nancy, 1983, *How the Laws of Physics Lie*, Oxford: Oxford University Press. doi:10.1093/0198247044.001.0001

——, 1989, *Nature's Capacities and Their Measurement*, Oxford: Oxford University Press. doi:10.1093/0198235070.001.0001

——, 1999, *The Dappled World: A Study of the Boundaries of Science*, Cambridge: Cambridge University Press. doi:10.1017/CBO9781139167093

Colombo, Matteo, Stephan Hartmann, and Robert van Iersel, 2015, "Models, Mechanisms, and Coherence", *The British Journal for the Philosophy of Science*, 66(1): 181–212. doi:10.1093/bjps/axt043

Colyvan, Mark, 2013, "Idealisations in Normative Models", *Synthese*, 190(8): 1337–1350. doi:10.1007/s11229-012-0166-z

Contessa, Gabriele, 2010, "Scientific Models and Fictional Objects", *Synthese*, 172(2): 215–229. doi:10.1007/s11229-009-9503-2

Crowther, Karen, Niels S. Linnemann, and Christian Wüthrich, 2021, "What We Cannot Learn from Analogue Experiments", *Synthese*, 198(16): 3701–3726; first online: 4 May 2019. doi:10.1007/s11229-019-02190-0

Da Costa, Newton and Steven French, 2000, "Models, Theories, and Structures: Thirty Years On", *Philosophy of Science*, 67(supplement): S116–S127. doi:10.1086/392813

Dardashti, Radin, Stephan Hartmann, Karim Thébault, and Eric Winsberg, 2019, "Hawking Radiation and Analogue Experiments: A Bayesian Analysis", *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 67: 1–11. doi:10.1016/j.shpsb.2019.04.004

Dardashti, Radin, Karim P. Y. Thébault, and Eric Winsberg, 2017, "Confirmation via Analogue Simulation: What Dumb Holes Could Tell Us about Gravity", *The British Journal for the Philosophy of Science*, 68(1): 55–89. doi:10.1093/bjps/axv010

de Regt, Henk, 2009, "Understanding and Scientific Explanation", in de Regt, Leonelli, and Eigner 2009: 21–42.

——, 2017, *Understanding Scientific Understanding*, Oxford: Oxford University Press. doi:10.1093/oso/9780190652913.001.0001

de Regt, Henk, Sabina Leonelli, and Kai Eigner (eds.), 2009, *Scientific Understanding: Philosophical Perspectives*, Pittsburgh, PA: University of Pittsburgh Press.

Dietrich, Michael R., Rachel A. Ankeny, Nathan Crowe, Sara Green, and Sabina Leonelli, 2020, "How to Choose Your Research Organism", *Studies in History and Philosophy of Science Part C* (*Studies in History and Philosophy of Biological and Biomedical Sciences*), 80: 101–227.

Dizadji-Bahmani, Foad, Roman Frigg, and Stephan Hartmann, 2010, "Who's Afraid of Nagelian Reduction?", *Erkenntnis*, 73(3): 393–412. doi:10.1007/s10670-010-9239-x

——, 2011, "Confirmation and Reduction: A Bayesian Account", *Synthese*, 179(2): 321–338. doi:10.1007/s11229-010-9775-6

Downes, Stephen M., 1992, "The Importance of Models in Theorizing: A Deflationary Semantic View", *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1992(1): 142–153. doi:10.1086/psaprocbienmeetp.1992.1.192750

Elgin, Catherine Z., 2010, "Telling Instances", in Roman Frigg and Matthew Hunter (eds.), *Beyond Mimesis and Convention* (Boston Studies in the Philosophy of Science 262), Dordrecht: Springer Netherlands, pp. 1–17. doi:10.1007/978-90-481-3851-7_1

——, 2017, *True Enough*. Cambridge, MA, and London: MIT Press.

Elgin, Mehmet and Elliott Sober, 2002, "Cartwright on Explanation and Idealization", *Erkenntnis*, 57(3): 441–450. doi:10.1023/A:1021502932490

Elliott-Graves, Alkistis, 2020, "What Is a Target System?", *Biology & Philosophy*, 35, Article 28. doi:10.1007/s10539-020-09745-3

El Skaf, Rawad and Michael T. Stuart, 2024, "Scientific Models and Thought Experiments: Same Same but Different", in Knuuttila, Carrillo and Koskinen 2024: ch. 24.

Epstein, Joshua M., 2008, "Why Model?", *Journal of Artificial Societies and Social Simulation*, 11(4): 12. [Epstein 2008 available online]

Fisher, Grant, 2006, "The Autonomy of Models and Explanation: Anomalous Molecular Rearrangements in Early Twentieth-Century Physical Organic Chemistry", *Studies in History and Philosophy of Science Part A*, 37(4): 562–584. doi:10.1016/j.shpsa.2006.09.009

Franklin, Alexander, 2020, "Whence the Effectiveness of Effective Field Theories?", *The British Journal for the Philosophy of Science*, 71(4): 1235–1259. First online: 3 August 2018. doi:10.1093/bjps/axy050

Freudenthal, Hans (ed.), 1961, *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences*, Dordrecht: Reidel. doi:10.1007/978-94-010-3667-2

Friedman, Michael, 1974, "Explanation and Scientific Understanding", *Journal of Philosophy*, 71(1): 5–19. doi:10.2307/2024924

Frigg, Roman, 2010a, "Fiction in Science", in John Woods (ed.), *Fictions and Models: New Essays*, Munich: Philosophia Verlag, pp. 247–287.

——, 2010b, "Models and Fiction", *Synthese*, 172(2): 251–268. doi:10.1007/s11229-009-9505-0

——, 2023, *Models and Theories: A Philosophical Inquiry*, London: Routledge.

Frigg, Roman, Seamus Bradley, Hailiang Du, and Leonard A. Smith, 2014, "Laplace's Demon and the Adventures of His Apprentices", *Philosophy of Science*, 81(1): 31–59. doi:10.1086/674416

Frigg, Roman and James Nguyen, 2016, "The Fiction View of Models Reloaded", *The Monist*, 99(3): 225–242. doi:10.1093/monist/onw002 [Frigg and Nguyen 2016 available online]

——, 2021, "Mirrors without Warnings", *Synthese*, 198: 2427–2447. doi:10.1007/s11229-019-02222-9

Fumagalli, Roberto, 2016, "Why We Cannot Learn from Minimal Models", *Erkenntnis*, 81(3): 433–455. doi:10.1007/s10670-015-9749-7

Gähde, Ulrich, 1997, "Anomalies and the Revision of Theory-Elements: Notes on the Advance of Mercury's Perihelion", in Maria Luisa Dalla Chiara, Kees Doets, Daniele Mundici, and Johan van Benthem (eds.), *Structures and Norms in Science* (Synthese Library 260), Dordrecht: Springer Netherlands, pp. 89–104. doi:10.1007/978-94-017-0538-7_6

Galison, Peter, 1997, *Image and Logic: A Material Culture of Microphysics*, Chicago, IL: University of Chicago Press.

Gelfert, Axel, 2016, *How to Do Science with Models: A Philosophical Primer* (Springer Briefs in Philosophy), Cham: Springer International Publishing. doi:10.1007/978-3-319-27954-1

Gendler, Tamar Szabó, 2000, *Thought Experiment: On the Powers and Limits of Imaginary Cases*, New York and London: Garland.

Gibbard, Allan and Hal R. Varian, 1978, "Economic Models", *The Journal of Philosophy*, 75(11): 664–677. doi:10.5840/jphil1978751111

Giere, Ronald N., 1988, *Explaining Science: A Cognitive Approach*, Chicago, IL: University of Chicago Press.

——, 1999, *Science Without Laws*, Chicago, IL: University of Chicago Press.

——, 2006, *Scientific Perspectivism*, Chicago, IL: University of Chicago Press.

——, 2009, "Why Scientific Models Should Not be Regarded as Works of Fiction", in Suárez 2009: 248–258.

——, 2010, "An Agent-Based Conception of Models and Scientific Representation", *Synthese*, 172(2): 269–281. doi:10.1007/s11229-009-9506-z

Godfrey-Smith, Peter, 2007, "The Strategy of Model-Based Science", *Biology & Philosophy*, 21(5): 725–740. doi:10.1007/s10539-006-9054-6

——, 2009, "Abstractions, Idealizations, and Evolutionary Biology", in Anouk Barberousse, Michel Morange, and Thomas Pradeu (eds.), *Mapping the Future of Biology: Evolving Concepts and Theories* (Boston Studies in the Philosophy of Science 266), Dordrecht: Springer Netherlands, pp. 47–56. doi:10.1007/978-1-4020-9636-5_4

Groenewold, H. J., 1961, "The Model in Physics", in Freudenthal 1961: 98–103. doi:10.1007/978-94-010-3667-2_9

Grüne-Yanoff, Till, 2009, "Learning from Minimal Economic Models", *Erkenntnis*, 70(1): 81–99. doi:10.1007/s10670-008-9138-6

Hacking, Ian, 1983, *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge: Cambridge University Press. doi:10.1017/CBO9780511814563

Hale, Susan C., 1988, "Spacetime and the Abstract/Concrete Distinction", *Philosophical Studies*, 53(1): 85–102. doi:10.1007/BF00355677

Harré, Rom, 2004, *Modeling: Gateway to the Unknown* (Studies in Multidisciplinarity 1), ed. by Daniel Rothbart, Amsterdam etc.: Elsevier.

Harris, Todd, 2003, "Data Models and the Acquisition and Manipulation of Data", *Philosophy of Science*, 70(5): 1508–1517. doi:10.1086/377426

Hartmann, Stephan, 1995, "Models as a Tool for Theory Construction: Some Strategies of Preliminary Physics", in Herfel et al. 1995: 49–67.

——, 1996, "The World as a Process: Simulations in the Natural and Social Sciences", in Hegselmann, Mueller, and Troitzsch 1996: 77–100. doi:10.1007/978-94-015-8686-3_5

——, 1998, "Idealization in Quantum Field Theory", in Shanks 1998: 99–122.

——, 1999, "Models and Stories in Hadron Physics", in Morgan and Morrison 1999: 326–346. doi:10.1017/CBO9780511660108.012

——, 2001, "Effective Field Theories, Reductionism and Scientific Explanation", *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 32(2): 267–304. doi:10.1016/S1355-2198(01)00005-3

Hartmann, Stephan, Carl Hoefer, and Luc Bovens (eds.), 2008, *Nancy Cartwright's Philosophy of Science* (Routledge Studies in the Philosophy of Science), New York: Routledge.

Hegselmann, Rainer, Ulrich Mueller, and Klaus G. Troitzsch (eds.), 1996, *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View* (Theory and Decision Library 23), Dordrecht: Springer Netherlands. doi:10.1007/978-94-015-8686-3

Helman, David H. (ed.), 1988, *Analogical Reasoning: Perspectives of Artificial Intelligence, Cognitive Science, and Philosophy* (Synthese Library 197), Dordrecht: Springer Netherlands. doi:10.1007/978-94-015-7811-0

Hempel, Carl G., 1965, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free Press.

Herfel, William, Wladiysław Krajewski, Ilkka Niiniluoto, and Ryszard Wojcicki (eds.), 1995, *Theories and Models in Scientific Process* (Poznań Studies in the Philosophy of Science and the Humanities 44), Amsterdam: Rodopi.

Hesse, Mary, 1963, *Models and Analogies in Science*, London: Sheed and Ward.

——, 1967, "Models and Analogy in Science", in Paul Edwards (ed.), *Encyclopedia of Philosophy*, New York: Macmillan, pp. 354–359.

——, 1974, *The Structure of Scientific Inference*, London: Macmillan.

Hodges, Wilfrid, 1997, *A Shorter Model Theory*, Cambridge: Cambridge University Press.

Holyoak, Keith and Paul Thagard, 1995, *Mental Leaps: Analogy in Creative Thought*, Cambridge, MA: MIT Press.

Horowitz, Tamara and Gerald J. Massey (eds.), 1991, *Thought Experiments in Science and Philosophy*, Lanham, MD: Rowman & Littlefield.

Isaac, Alistair M. C., 2013, "Modeling without Representation", *Synthese*, 190(16): 3611–3623. doi:10.1007/s11229-012-0213-9

Jebeile, Julie and Ashley Graham Kennedy, 2015, "Explaining with Models: The Role of Idealizations", *International Studies in the Philosophy of Science*, 29(4): 383–392. doi:10.1080/02698595.2015.1195143

Jones, Martin R., 2005, "Idealization and Abstraction: A Framework", in Jones and Cartwright 2005: 173–217. doi:10.1163/9789401202732_010

Jones, Martin R. and Nancy Cartwright (eds.), 2005, *Idealization XII: Correcting the Model* (Poznań Studies in the Philosophy of the Sciences and the Humanities 86), Amsterdam and New York: Rodopi. doi:10.1163/9789401202732

Kelvin, William Thomson, Baron, 1884 [1987], *Notes of lectures on molecular dynamics and the wave theory of light. Delivered at the Johns Hopkins University, Baltimore* (aka Lord Kelvin's Baltimore Lectures), A. S. Hathaway (recorder). A revised version was published in 1904, London: C.J. Clay and Sons. Reprint of the 1884 version in Robert Kargon and Peter Achinstein (eds.), *Kelvin's Baltimore Lectures and Modern Theoretical Physics*, Cambridge, MA: MIT Press, 1987.

Khalifa, Kareem, 2017, *Understanding, Explanation, and Scientific Knowledge*, Cambridge: Cambridge University Press. doi:10.1017/9781108164276

Klein, Dominik, Johannes Marx, and Kai Fischbach, 2018, "Agent-Based Modeling in Social Science History and Philosophy: An Introduction", *Historical Social Research*, 43(1): 243–258.

Knuuttila, Tarja, 2005, "Models, Representation, and Mediation", *Philosophy of Science*, 72(5): 1260–1271. doi:10.1086/508124

——, 2011, "Modelling and Representing: An Artefactual Approach to Model-Based Representation", *Studies in History and Philosophy of Science Part A*, 42(2): 262–271. doi:10.1016/j.shpsa.2010.11.034

Knuuttila, Tarja, Natalia Carrillo, and Rami Koskinen (eds.), 2024, *The Routledge Handbook of Philosophy of Scientific Modeling*, London: Routledge.

Kroes, Peter, 1989, "Structural Analogies Between Physical Systems", *The British Journal for the Philosophy of Science*, 40(2): 145–154. doi:10.1093/bjps/40.2.145

Lange, Marc, 2015, "On 'Minimal Model Explanations': A Reply to Batterman and Rice", *Philosophy of Science*, 82(2): 292–305. doi:10.1086/680488

Lavis, David A., 2008, "Boltzmann, Gibbs, and the Concept of Equilibrium", *Philosophy of Science*, 75(5): 682–692. doi:10.1086/594514

Laymon, Ronald, 1982, "Scientific Realism and the Hierarchical Counterfactual Path from Data to Theory", *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1982(1): 107–121. doi:10.1086/psaprocbienmeetp.1982.1.192660

——, 1985, "Idealizations and the Testing of Theories by Experimentation", in Peter Achinstein and Owen Hannaway (eds.), *Observation, Experiment, and Hypothesis in Modern Physical Science*, Cambridge, MA: MIT Press, pp. 147–173.

——, 1991, "Thought Experiments by Stevin, Mach and Gouy: Thought Experiments as Ideal Limits and Semantic Domains", in Horowitz and Massey 1991: 167–191.

Leonelli, Sabina, 2010, "Packaging Small Facts for Re-Use: Databases in Model Organism Biology", in Peter Howlett and Mary S. Morgan (eds.), *How Well Do Facts Travel? The Dissemination of Reliable Knowledge*, Cambridge: Cambridge University Press, pp. 325–348. doi:10.1017/CBO9780511762154.017

——, 2016, *Data-Centric Biology: A Philosophical Study*, Chicago, IL, and London: University of Chicago Press.

——, 2019, "What Distinguishes Data from Models?", *European Journal for Philosophy of Science*, 9(2): article 22. doi:10.1007/s13194-018-0246-0

Leplin, Jarrett, 1980, "The Role of Models in Theory Construction", in Thomas Nickles (ed.), *Scientific Discovery, Logic, and Rationality* (Boston Studies in the Philosophy of Science 56), Dordrecht: Springer Netherlands, 267–283. doi:10.1007/978-94-009-8986-3_12

Levins, Richard, 1966, "The Strategy of Model Building in Population Biology", *American Scientist*, 54: 421–431.

Levy, Arnon, 2012, "Models, Fictions, and Realism: Two Packages", *Philosophy of Science*, 79(5): 738–748. doi:10.1086/667992

——, 2015, "Modeling without Models", *Philosophical Studies*, 172(3): 781–798. doi:10.1007/s11098-014-0333-9

Levy, Arnon and Adrian Currie, 2015, "Model Organisms Are Not (Theoretical) Models", *The British Journal for the Philosophy of Science*, 66(2): 327–348. doi:10.1093/bjps/axt055

Levy, Arnon and Peter Godfrey-Smith (eds.), 2020, *The Scientific Imagination: Philosophical and Psychological Perspectives*, New York: Oxford University Press.

Liefke, Kristina and Stephan Hartmann, 2018, "Intertheoretic Reduction, Confirmation, and Montague's Syntax–Semantics

Relation", *Journal of Logic, Language and Information*, 27(4): 313–341. doi:10.1007/s10849-018-9272-8

Lipton, Peter, 2009, "Understanding without Explanation", in de Regt, Leonelli, and Eigner 2009: 43–63.

Luczak, Joshua, 2017, "Talk about Toy Models", *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 57: 1–7. doi:10.1016/j.shpsb.2016.11.002

Magnani, Lorenzo, 2012, "Scientific Models Are Not Fictions: Model-Based Science as Epistemic Warfare", in Lorenzo Magnani and Ping Li (eds.), *Philosophy and Cognitive Science: Western & Eastern Studies* (Studies in Applied Philosophy, Epistemology and Rational Ethics 2), Berlin and Heidelberg: Springer, pp. 1–38. doi:10.1007/978-3-642-29928-5_1

Magnani, Lorenzo and Claudia Casadio (eds.), 2016, *Model-Based Reasoning in Science and Technology: Logical, Epistemological, and Cognitive Issues* (Studies in Applied Philosophy, Epistemology and Rational Ethics 27), Cham: Springer International Publishing. doi:10.1007/978-3-319-38983-7

Magnani, Lorenzo and Nancy J. Nersessian (eds.), 2002, *Model-Based Reasoning: Science, Technology, Values*, Boston, MA: Springer US. doi:10.1007/978-1-4615-0605-8

Magnani, Lorenzo, Nancy J. Nersessian, and Paul Thagard (eds.), 1999, *Model-Based Reasoning in Scientific Discovery*, Boston, MA: Springer US. doi:10.1007/978-1-4615-4813-3

Mäki, Uskali, 1994, "Isolation, Idealization and Truth in Economics", in Bert Hamminga and Neil B. De Marchi (eds.), *Idealization VI: Idealization in Economics* (Poznań Studies in the Philosophy of the Sciences and the Humanities 38), Amsterdam: Rodopi, pp. 147–168.

Massimi, Michela, 2018a, "Four Kinds of Perspectival Truth", *Philosophy and Phenomenological Research*, 96(2): 342–359. doi:10.1111/phpr.12300

——, 2018b, "Perspectival Modeling", *Philosophy of Science*, 85(3): 335–359. doi:10.1086/697745

——, 2019, "Two Kinds of Exploratory Models", *Philosophy of Science*, 86(5): 869–881. doi:10.1086/705494

——, 2022, *Perspectival Realism*, Oxford: Oxford University Press.

Massimi, Michela and Casey D. McCoy (eds.), 2019, *Understanding Perspectivism: Scientific Challenges and Methodological Prospects*, New York: Routledge. doi:10.4324/9781315145198

Matthewson, John, 2011, "Trade-offs in Model-building: A More Target-oriented Approach", *Studies in History and Philosophy of Science Part A*, 42(2): 324–333.

Mayo, Deborah, 1996, *Error and the Growth of Experimental Knowledge*, Chicago, IL: University of Chicago Press.

——, 2018, *Statistical Inference as Severe Testing: How to Get Beyond the Statistics Wars*, Cambridge: Cambridge University Press. doi:10.1017/9781107286184

McMullin, Ernan, 1968, "What Do Physical Models Tell Us?", in B. Van Rootselaar and J. Frits Staal (eds.), *Logic, Methodology and Philosophy of Science III* (Studies in Logic and the Foundations of Mathematics 52), Amsterdam: North Holland, pp. 385–396. doi:10.1016/S0049-237X(08)71206-0

——, 1985, "Galilean Idealization", *Studies in History and Philosophy of Science Part A*, 16(3): 247–273. doi:10.1016/0039-3681(85)90003-2

Morgan, Mary S., 1999, "Learning from Models", in Morgan and Morrison 1999: 347–388. doi:10.1017/CBO9780511660108.013

Morgan, Mary S. and Marcel J. Boumans, 2004, "Secrets Hidden by Two-Dimensionality: The Economy as a Hydraulic Machine", in Soraya de Chadarevian and Nick Hopwood (eds.), *Model: The Third Dimension of Science*, Stanford, CA: Stanford University Press, pp. 369–401.

Morgan, Mary S. and Margaret Morrison (eds.), 1999, *Models as Mediators: Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press. doi:10.1017/CBO9780511660108

Morrison, Margaret, 1999, "Models as Autonomous Agents", in Morgan and Morrison 1999: 38–65. doi:10.1017/CBO9780511660108.004

——, 2000, *Unifying Scientific Theories: Physical Concepts and Mathematical Structures*, Cambridge: Cambridge University Press. doi:10.1017/CBO9780511527333

——, 2005, "Approximating the Real: The Role of Idealizations in Physical Theory", in Jones and Cartwright 2005: 145–172. doi:10.1163/9789401202732_009

——, 2009, "Understanding in Physics and Biology: From the Abstract to the Concrete", in de Regt, Leonelli, and Eigner 2009: 123–145.

——, 2012, "Emergent Physics and Micro-Ontology", *Philosophy of Science*, 79(1): 141–166. doi:10.1086/663240

Musgrave, Alan, 1981, "'Unreal Assumptions' in Economic Theory: The F-Twist Untwisted", *Kyklos*, 34(3): 377–387. doi:10.1111/j.1467-6435.1981.tb01195.x

Nagel, Ernest, 1961, *The Structure of Science: Problems in the Logic of Scientific Explanation*, New York: Harcourt, Brace and World.

Nersessian, Nancy J., 1999, "Model-Based Reasoning in Conceptual Change", in Magnani, Nersessian, and Thagard 1999: 5–22. doi:10.1007/978-1-4615-4813-3_1

——, 2010, *Creating Scientific Concepts*, Cambridge, MA: MIT Press.

Nguyen, James, 2020, "It's Not a Game: Accurate Representation with Toy Models", *The British Journal for the Philosophy of Science*, 71(3): 1013–1041. doi:10.1093/bjps/axz010

Nguyen, James and Roman Frigg, 2021, "Mathematics Is Not the Only Language in the Book of Nature", *Synthese*, 198, 5941–5962. doi:10.1007/s11229-017-1526-5

——, 2022, *Scientific Representation*, Cambridge: Cambridge University Press.

Norton, John D., 1991, "Thought Experiments in Einstein's Work", in Horowitz and Massey 1991: 129–148.

——, 2003, "Causation as Folk Science", *Philosopher's Imprint*, 3: article 4. [Norton 2003 available online]

——, 2012, "Approximation and Idealization: Why the Difference Matters", *Philosophy of Science*, 79(2): 207–232. doi:10.1086/664746

Nowak, Leszek, 1979, *The Structure of Idealization: Towards a Systematic Interpretation of the Marxian Idea of Science*, Dordrecht: D. Reidel.

Odenbaugh, Jay, 2002, "Complex Systems, Trade-offs, and Theoretical Population Biology: Richard Levins's 'Strategy of Model Building in Population Biology' Revisited", *Philosophy of Science*, 70: 1496–1507.

——, 2006, "The Strategy of 'The Strategy of Model Building in Population Biology'", *Biology & Philosophy*, 21: 607–621.

Orzack, Steven Hecht and Elliott Sober, 1993, "A Critical Assessment of Levins's The Strategy of Model Building in Population Biology (1966)", *The Quarterly Review of Biology*, 68(4): 533–546.

Otsuka, Jun, 2023, *Thinking About Statistics: The Philosophical Foundations*, London: Routledge.

Palacios, Patricia, 2019, "Phase Transitions: A Challenge for Intertheoretic Reduction?", *Philosophy of Science*, 86(4): 612–640. doi:10.1086/704974

Peschard, Isabelle, 2011, "Making Sense of Modeling: Beyond Representation", *European Journal for Philosophy of Science*, 1(3): 335–352. doi:10.1007/s13194-011-0032-8

Piccinini, Gualtiero and Carl Craver, 2011, "Integrating Psychology and Neuroscience: Functional Analyses as Mechanism Sketches", *Synthese*, 183(3): 283–311. doi:10.1007/s11229-011-9898-4

Pincock, Christopher, 2012, *Mathematics and Scientific Representation*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199757107.001.0001

——, 2022, "Concrete Scale Models, Essential Idealization and Causal Explanation", *British Journal for the Philosophy of Science*, 73(2): 299–323.

Portides, Demetris P., 2007, "The Relation between Idealisation and Approximation in Scientific Model Construction", *Science & Education*, 16(7–8): 699–724. doi:10.1007/s11191-006-9001-6

——, 2014, "How Scientific Models Differ from Works of Fiction", in Lorenzo Magnani (ed.), *Model-Based Reasoning in Science and Technology* (Studies in Applied Philosophy, Epistemology and Rational Ethics 8), Berlin and Heidelberg: Springer, pp. 75–87. doi:10.1007/978-3-642-37428-9_5

Potochnik, Angela, 2007, "Optimality Modeling and Explanatory Generality", *Philosophy of Science*, 74(5): 680–691.

——, 2017, *Idealization and the Aims of Science*, Chicago, IL: University of Chicago Press.

Poznic, Michael, 2016, "Make-Believe and Model-Based Representation in Science: The Epistemology of Frigg's and Toon's Fictionalist Views of Modeling", *Teorema: Revista Internacional de Filosofía*, 35(3): 201–218.

Psillos, Stathis, 1995, "The Cognitive Interplay between Theories and Models: The Case of 19th Century Optics", in Herfel et al. 1995: 105–133.

Redhead, Michael, 1980, "Models in Physics", *The British Journal for the Philosophy of Science*, 31(2): 145–163. doi:10.1093/bjps/31.2.145

Reiss, Julian, 2003, "Causal Inference in the Abstract or Seven Myths about Thought Experiments", in *Causality: Metaphysics and Methods Research Project*, Technical Report 03/02. London: London School of Economics.

——, 2006, "Social Capacities", in Hartmann et al. 2006: 265–288.

——, 2012, "The Explanation Paradox", *Journal of Economic Methodology*, 19(1): 43–62. doi:10.1080/1350178X.2012.661069

Reutlinger, Alexander, 2017, "Do Renormalization Group Explanations Conform to the Commonality Strategy?", *Journal for General Philosophy of Science*, 48(1): 143–150. doi:10.1007/s10838-016-9339-7

Reutlinger, Alexander, Dominik Hangleiter, and Stephan Hartmann, 2018, "Understanding (with) Toy Models", *The British Journal for the Philosophy of Science*, 69(4): 1069–1099. doi:10.1093/bjps/axx005

Rice, Collin C., 2016, "Factive Scientific Understanding without Accurate Representation", *Biology & Philosophy*, 31(1): 81–102. doi:10.1007/s10539-015-9510-2

——, 2018, "Idealized Models, Holistic Distortions, and Universality", *Synthese*, 195(6): 2795–2819. doi:10.1007/s11229-017-1357-4

——, 2019, "Models Don't Decompose That Way: A Holistic View of Idealized Models", *The British Journal for the Philosophy of Science*, 70(1): 179–208. doi:10.1093/bjps/axx045

——, 2021, *Leveraging Distortions: Explanation, Idealization, and Universality in Science*. Boston, MA: MIT Press.

Rosaler, Joshua, 2015, "Local Reduction in Physics", *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 50: 54–69. doi:10.1016/j.shpsb.2015.02.004

Rueger, Alexander, 2005, "Perspectival Models and Theory Unification", *The British Journal for the Philosophy of Science*, 56(3): 579–594. doi:10.1093/bjps/axi128

Rueger, Alexander and David Sharp, 1998, "Idealization and Stability: A Perspective from Nonlinear Dynamics", in Shanks 1998: 201–216.

Saatsi, Juha, 2016, "Models, Idealisations, and Realism", in Emiliano Ippoliti, Fabio Sterpetti, and Thomas Nickles (eds.), *Models and*

Inferences in Science (Studies in Applied Philosophy, Epistemology and Rational Ethics 25), Cham: Springer International Publishing, pp. 173–189. doi:10.1007/978-3-319-28163-6_10

Saatsi, Juha and Alexander Reutlinger, 2018, "Taking Reductionism to the Limit: How to Rebut the Antireductionist Argument from Infinite Limits", *Philosophy of Science*, 85(3): 455–482. doi:10.1086/697735

Salis, Fiora, 2021, "The New Fiction View of Models", *The British Journal for the Philosophy of Science*, 72(3): 717–742. doi:10.1093/bjps/axz015

Salmon, Wesley C., 1984, *Scientific Explanation and the Causal Structure of the World*, Princeton, NJ: Princeton University Press.

Sartori, Lorenzo, 2023, "Putting the 'Experiment' Back Into the 'Thought Experiment'", *Synthese*, 201, article 34. doi:10.1007/s11229-022-04011-3

——, forthcoming, "Model Organisms as Scientific Representations", *The British Journal for the Philosophy of Science*. doi:10.1086/728259.

Schaffner, Kenneth F., 1969, "The Watson–Crick Model and Reductionism", *The British Journal for the Philosophy of Science*, 20(4): 325–348. doi:10.1093/bjps/20.4.325

Scheibe, Erhard, 1997, *Die Reduktion physikalischer Theorien: Ein Beitrag zur Einheit der Physik, Teil I: Grundlagen und elementare Theorie*, Berlin: Springer.

——, 1999, *Die Reduktion physikalischer Theorien: Ein Beitrag zur Einheit der Physik, Teil II: Inkommensurabilität und Grenzfallreduktion*, Berlin: Springer.

——, 2001, *Between Rationalism and Empiricism: Selected Papers in the Philosophy of Physics*, Brigitte Falkenburg (ed.), New York: Springer. doi:10.1007/978-1-4613-0183-7

Shanks, Niall (ed.), 1998, *Idealization in Contemporary Physics*, Amsterdam: Rodopi.

Shech, Elay, 2018, "Idealizations, Essential Self-Adjointness, and Minimal Model Explanation in the Aharonov–Bohm Effect", *Synthese*, 195(11): 4839–4863. doi:10.1007/s11229-017-1428-6

Sismondo, Sergio and Snait Gissis (eds.), 1999, *Modeling and Simulation*, Special Issue of *Science in Context*, 12(2).

Sjölin Wirling, Ylwa and Till Grüne-Yanoff, 2021, "The Epistemology of Modal Modelling", *Philosophy Compass*, e12775. doi:10.1111/phc3.12775

Sober, Elliott, 2015, *Ockham's Razors: A User's Manual*, Cambridge: Cambridge University Press.

Sorensen, Roy A., 1992, *Thought Experiments*, New York: Oxford University Press. doi:10.1093/019512913X.001.0001

Spector, Marshall, 1965, "Models and Theories", *The British Journal for the Philosophy of Science*, 16(62): 121–142. doi:10.1093/bjps/XVI.62.121

Sprenger, Jan and Stephan Hartmann, 2019, *Bayesian Philosophy of Science*, Oxford: Oxford University Press.

Staley, Kent W., 2004, *The Evidence for the Top Quark: Objectivity and Bias in Collaborative Experimentation*, Cambridge: Cambridge University Press.

Sterrett, Susan G., 2006, "Models of Machines and Models of Phenomena", *International Studies in the Philosophy of Science*, 20(1): 69–80. doi:10.1080/02698590600641024

——, 2020, "Scale Modeling", in Diane Michelfelder and Neelke Doorn (eds.), *The Routledge Handbook of Philosophy of Engineering*, New York: Routledge, pp. 394–407.

Strevens, Michael, 2004, "The Causal and Unification Approaches to Explanation Unified—Causally", *Noûs*, 38(1): 154–176. doi:10.1111/j.1468-0068.2004.00466.x

——, 2008, *Depth: An Account of Scientific Explanation*, Cambridge, MA, and London: Harvard University Press.

——, 2013, *Tychomancy: Inferring Probability from Causal Structure*, Cambridge, MA, and London: Harvard University Press.

Suárez, Mauricio, 2004, "An Inferential Conception of Scientific Representation", *Philosophy of Science*, 71(5): 767–779. doi:10.1086/421415

—— (ed.), 2009, *Fictions in Science: Philosophical Essays on Modeling and Idealization*, London: Routledge. doi:10.4324/9780203890103

——, 2024, *Inference and Representation: A Study in Modeling Science*, Chicago, IL: Chicago University Press.

Sugden, Robert, 2000, "Credible Worlds: The Status of Theoretical Models in Economics", *Journal of Economic Methodology*, 7(1): 1–31. doi:10.1080/135017800362220

Sullivan, Emily and Kareem Khalifa, 2019, "Idealizations and Understanding: Much Ado About Nothing?", *Australasian Journal of Philosophy*, 97(4): 673–689. doi:10.1080/00048402.2018.1564337

Suppe, Frederick, 2000, "Theory Identity", in William H. Newton-Smith (ed.), *A Companion to the Philosophy of Science*, Oxford: Wiley-Blackwell, pp. 525–527.

Suppes, Patrick, 1960, "A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences", *Synthese*, 12(2–3): 287–301. Reprinted in Freudenthal 1961: 163–177, and in Suppes 1969: 10–23. doi:10.1007/BF00485107 doi:10.1007/978-94-010-3667-2_16

——, 1962, "Models of Data", in Ernest Nagel, Patrick Suppes, and Alfred Tarski (eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*, Stanford, CA: Stanford University Press, pp. 252–261. Reprinted in Suppes 1969: 24–35.

——, 1969, *Studies in the Methodology and Foundations of Science: Selected Papers from 1951 to 1969*, Dordrecht: Reidel.

——, 2007, "Statistical Concepts in Philosophy of Science", *Synthese*, 154(3): 485–496. doi:10.1007/s11229-006-9122-0

Swoyer, Chris, 1991, "Structural Representation and Surrogative Reasoning", *Synthese*, 87(3): 449–508. doi:10.1007/BF00499820

Tabor, Michael, 1989, *Chaos and Integrability in Nonlinear Dynamics: An Introduction*, New York: John Wiley.

Teller, Paul, 2001, "Twilight of the Perfect Model", *Erkenntnis*, 55(3): 393–415. doi:10.1023/A:1013349314515

——, 2002, "Critical Study: Nancy Cartwright's *The Dappled World: A Study of the Boundaries of Science*", *Noûs*, 36(4): 699–725. doi:10.1111/1468-0068.t01-1-00408

——, 2009, "Fictions, Fictionalization, and Truth in Science", in Suárez 2009: 235–247.

——, 2018, "Referential and Perspectival Realism", *Spontaneous Generations: A Journal for the History and Philosophy of Science*, 9(1): 151–164. doi:10.4245/sponge.v9i1.26990

Tešić, Marko, 2019, "Confirmation and the Generalized Nagel–Schaffner Model of Reduction: A Bayesian Analysis", *Synthese*, 196(3): 1097–1129. doi:10.1007/s11229-017-1501-1

Thomasson, Amie L., 1999, *Fiction and Metaphysics*, New York: Cambridge University Press. doi:10.1017/CBO9780511527463

——, 2020, "If Models Were Fictions, Then What Would They Be?", in Levy and Godfrey-Smith 2020: 51–74.

Thomson-Jones, Martin, 2006, "Models and the Semantic View", *Philosophy of Science*, 73(5): 524–535. doi:10.1086/518322

——, 2020, "Realism about Missing Systems", in Levy and Godfrey-Smith 2020: 75–101.

Toon, Adam, 2012, *Models as Make-Believe: Imagination, Fiction and Scientific Representation*, Basingstoke: Palgrave Macmillan.

Trout, J.D., 2002, "Scientific Explanation and the Sense of Understanding", *Philosophy of Science*, 69(2): 212–233. doi:10.1086/341050

van Fraassen, Bas C., 1989, *Laws and Symmetry*, Oxford: Oxford University Press. doi:10.1093/0198248601.001.0001

Walton, Kendall L., 1990, *Mimesis as Make-Believe: On the Foundations of the Representational Arts*, Cambridge, MA: Harvard University Press.

Weisberg, Michael, 2006, "Forty Years of 'the Strategy': Levins on Model Building and Idealization", *Biology & Philosophy*, 21: 623–645.

——, 2007, "Three Kinds of Idealization", *Journal of Philosophy*, 104(12): 639–659. doi:10.5840/jphil20071041240

——, 2013, *Simulation and Similarity: Using Models to Understand the World*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199933662.001.0001

Weisberg, Michael and Ryan Muldoon, 2009, "Epistemic Landscapes and the Division of Cognitive Labor", *Philosophy of Science*, 76(2): 225–252. doi:10.1086/644786

Wimsatt, William, 1987, "False Models as Means to Truer Theories", in Matthew Nitecki and Antoni Hoffman (eds.), *Neutral Models in Biology*, Oxford: Oxford University Press, pp. 23–55.

——, 2007, *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*, Cambridge, MA: Harvard University Press.

Woodward, James, 2003, *Making Things Happen: A Theory of Causal Explanation*, Oxford: Oxford University Press. doi:10.1093/0195155270.001.0001

Woody, Andrea I., 2004, "More Telltale Signs: What Attention to Representation Reveals about Scientific Explanation", *Philosophy of Science*, 71(5): 780–793. doi:10.1086/421416

Yoshida, Yoshinari, 2021, "Multiple-Models Juxtaposition and Trade-offs among Modeling Desiderata", *Philosophy of Science*, 88(1): 103–123.

Zollman, Kevin J. S., 2007, "The Communication Structure of Epistemic Communities", *Philosophy of Science*, 74(5): 574–587. doi:10.1086/525605

Herramientas académicas

Cómo citar esta entrada [en INGLÉS].
Avance de la versión en PDF de esta entrada en los amigos de la sociedad SEP.
Véase los tópicos y pensadores relacionados con esta entrada en el Proyecto de ontología filosófica de internet (InPhO).
La bibliografía ampliada de esta entrada en PhilPapers, con los enlaces a su base de datos.

Otros recursos de internet

Discusiones conceptuales e históricas

- Enciclopedia de internet de artículos filosóficos sobre modelos
- Bibliografía (1450-2008), Mueller Science

Ejemplos de modelos

- Modelos interactivos de distintas ciencias (Phet, Universidad de Colorado, Boulder)
- Modelos del clima global (Climate.gov)
- Modelo de doble hélice del ADN (Proteopedia)
- Una guía biológica al modelamiento matemático en ecología y evolución (Sarah Otto y Troy Day)
- Modelo Lotka-Volterra (analyticphysics.com)
- Modelo Schelling de Segregación (Frank McCown)
- Modelamiento común (NetLogo)

- Redes sociales y económicas: Modelos y análisis (cursos de Stanford Online)
- Modelos de redes neurales (TensorFlow)

Entradas relacionadas

Objetos abstractos | Razonamiento análogo y analógico | Biología: experimentos en | Caos | Ficciones: entidades ficticias | Leyes de la naturaleza | Leyes de la naturaleza: *ceteris paribus* | Física: relaciones intrateóricas en | Reducción, científica | Ciencia: teoría y observación en | Ciencia: unidad de | Explicación científica | Realismo científico | Representación científica | Teorías científicas: estructuras de | Simulaciones en ciencia | Estadística, filosofía de | Experimentos mentales | Semejanza con la verdad

Agradecimiento

Nos gustaría agradecer a Jow Dewhurst, James Nguyen, Alexander Reutlinger, Collin Rice, Dunja Šešelja y Paul Teller por sus fructíferos comentarios en los borradores de la versión revisada de 2019. Cuando escribimos la versión original en 2006, nos beneficiamos de los comentarios y sugerencias de Nancy Cartwright, Paul Humphreys, Julian Reiss, Elliot Sober, Chris Swoyer y Paul Teller.

Copyright © 2025 por
Roman Frigg <r.p.frigg@lse.ac.uk>
Stephan Hartmann <stephan.hartmann@lrz.uni-muenchen.de>

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 310-354
ISSN 3087-260X
Santiago de Chile

UNA APOLOGÍA DE LOS CONFLICTOS ENTRE METAFÍSICA Y CIENCIA EN LA METAFÍSICA NATURALIZADA[1]

Rasmus Jaksland[2]
Traducido por Gabriel Donoso Umaña[3]

Resumen

De acuerdo con la metafísica naturalizada, la metafísica debe ser informada por nuestra mejor ciencia disponible y no apoyarse en el razonamiento *a priori*. En consecuencia, la metafísica naturalizada

[1] Esta es una traducción al español del artículo de Rasmus Jaksland, "An apology for conflicts between metaphysics and science in naturalized metaphysics" (2021), *European Journal for Philosophy of Science*, *11*(74). La obra original está licenciada bajo Creative Commons Attribution 4.0 International (CC BY 4.0). Salvo por ajustes estilísticos propios de la traducción, las únicas modificaciones introducidas consisten en la modificación de la nota al pie 2 (donde se agrega su información de contacto y su ORCID; contenido ya presente en el manuscrito original, pero no en la nota al pie), a las cursivas para expresiones no hispanas y en la adaptación (y corrección) de las referencias y citas al estilo APA 7. Siglas técnicas como QFT, QG o GR se han dejado en su formulación anglosajona.

[2] Departamento de Filosofía y Estudios Religiosos, Facultad de Humanidades, NTNU – Universidad Noruega de Ciencia y Tecnología, Trondheim, Noruega. Contacto: rasmus.jaksland@ntnu.no. ORCID: https://orcid.org/0000-0001-9531-9994.

[3] Profesor de Estado en Filosofía y Magíster en Filosofía de las Ciencias – Universidad de Santiago de Chile, Santiago, Chile. Contacto: gabriel.donoso.u@usach.cl. ORCID: https://orcid.org/0000-0002-4842-1387.

tiende a descartar los intentos de los metafísicos por disputar con la ciencia. Este artículo sostiene que, en cambio, la metafísica naturalizada debería acoger tales conflictos entre metafísica y ciencia. La metafísica naturalizada no es (y no debería ser) eliminativa respecto de la metafísica. Por lo tanto, si tales conflictos se originan en la ausencia inmediata, dentro de la ciencia, de una respuesta a una cuestión metafísica, entonces el conflicto no debe ser descartado, sino recibido como una ocasión para hacer (más) metafísica naturalizada. Que los conflictos entre metafísica y ciencia pueden ser beneficiosos para la metafísica naturalizada se ejemplifica mediante el caso de las teorías no espaciales de la gravedad cuántica. Estas teorías son criticadas por metafísicos que, a menudo siguiendo a David Lewis, argumentan que la distancia espacial es un elemento fundamental indispensable en cualquier metafísica coherente debido a su rol como la relación constructora del mundo (*world-making relation*). Sin embargo, se reconoce que el conflicto resultante está bien fundado, puesto que las teorías no espaciales de la gravedad cuántica no ofrecen ninguna relación constructora del mundo alternativa a la distancia espacial. En lugar de descartar este conflicto, la metafísica naturalizada debería, por lo tanto, recibir la resistencia lewisiana como un llamado a buscar una. Cómo se desarrolla esto al modo de una negociación entre la teoría científica y la cuestión metafísica se ejemplifica en la última parte del artículo, donde se propone el entrelazamiento (*entanglement*) como una relación constructora del mundo alternativa dentro de la gravedad cuántica de lazos (*loop quantum gravity*).

Palabras clave: *Metafísica naturalizada, Metodología metafísica, Gravedad cuántica, Fundamentalismo del entrelazamiento, Metafísica vs. ciencia, Superveniencia humeana.*

1. Introducción

Aunque no existe una definición unívoca de metafísica, a menudo se la caracteriza como el "estudio sistemático de la estructura más fundamental de la realidad" (Lowe, 1998, p. 2), como "el estudio de la realidad última" (van Inwagen, 2015, p. 1) o como "*la exploración de los rasgos más generales del mundo*" (Blackburn, 2002, p. 61, énfasis en el original).[4] Sin embargo, esta caracterización podría aplicarse igualmente a la física fundamental, al menos adoptando una voz realista. Es defendible que la física fundamental también esté interesada en la estructura y los rasgos de la realidad última. Por lo tanto, no es sorprendente que esta (aparente) superposición en el objeto de estudio entre la metafísica y la física fundamental sea una semilla de conflicto. Este artículo examina tales conflictos y, más generalmente, los conflictos entre metafísica y ciencia. El artículo sostiene que, en lugar de ser descartados como señales de un exceso de confianza por parte de los metafísicos, los conflictos entre metafísica y ciencia deberían ser bienvenidos como indicios de preguntas genuinas e importantes que deben ser respondidas por la metafísica naturalizada. Aunque la metafísica debería, en última instancia, ceder ante la ciencia —tal como sostiene la metafísica naturalizada—, los metafísicos solo deberían hacerlo una vez que la ciencia —a través de la metafísica naturalizada— provea una respuesta positiva (metafísica) a sus preguntas metafísicas.

[4] Tales caracterizaciones de la metafísica son disputadas por otros (por ejemplo, Bennett, 2016; Paul, 2012), quienes sostienen que las preguntas abordadas por la metafísica son diferentes de aquellas de la ciencia en general y de la física en particular. Sin embargo, este debate se dejará de lado aquí, puesto que el foco estará en los casos donde la metafísica —al menos tal como se usa el término aquí— entra efectivamente en conflicto directo con la ciencia, como se ejemplifica en el conflicto entre la metafísica lewisiana y la gravedad cuántica. Quizás esto muestre que "metafísica" está siendo mal empleada en el presente análisis, pero cualquiera que favorezca esta opinión puede reemplazar "metafísica" en lo que sigue por un término más adecuado que, a su juicio, represente mejor el tipo de preguntas que, por ejemplo, Lewis formula acerca de la construcción del mundo.

Como estudio de caso, el artículo utiliza las teorías de la gravedad cuántica —teorías que intentan reconciliar la mecánica cuántica y la relatividad general— en las cuales el espacio parece estar ausente en el nivel fundamental de la realidad descrita por dichas teorías. El espacio, la distancia y, en general, la extensión no parece figurar en la ontología fundamental implicada por teorías de gravedad cuántica como la gravedad cuántica de lazos (*loop quantum gravity*) y la teoría de conjuntos causales (*causal set theory*) (Huggett & Wüthrich, 2013). En consecuencia, tales teorías son recibidas con sospecha por algunos metafísicos y filósofos de la ciencia, quienes han puesto en duda la coherencia de ontologías no espaciales (por ejemplo, Esfeld, 2019; Hagar & Hemmo, 2013; Lam & Esfeld, 2013; Maudlin, 2007a). En lo que sigue, el foco estará en las preocupaciones metafísicas acerca de las teorías no espaciales de la gravedad cuántica que, de distintos modos, hacen eco de la tesis de David Lewis (1986a, 1986b, 1994) según la cual la distancia espacial es la relación fundamental en el mundo y, en particular, la relación en virtud de la cual los elementos de una ontología constituyen un mundo.[5] La desaparición del espacio en algunas teorías de gravedad cuántica da lugar, por lo tanto, a un conflicto entre estas teorías científicas y la metafísica lewisiana (Wüthrich, 2019): mientras los lewisianos[6] sostienen que la distancia espacial —dado su rol como relación constructora del mundo— es un elemento fundamental necesario en cualquier metafísica coherente, estas teorías de gravedad cuántica proponen que, en el mundo actual, el espacio no es fundamental.

[5] Para una exposición más detallada de la "relación de compañeros de mundo" (*worldmate relation*) de Lewis, véase Darby (2009).

[6] Usaremos "lewisianos" como un término colectivo para aquellos que expresan una preferencia por la fundamentalidad de la distancia espacial. Aunque los defensores de esta tesis no respalden, en general, todos los aspectos de la metafísica de Lewis, utilizaremos este término dado que, en particular, la tesis de la superveniencia humeana de Lewis es invocada de forma habitual en apoyo de la priorización ontológica de la distancia espacial (por ejemplo Esfeld & Deckert, 2017; Lam, 2016).

La metafísica naturalizada es conocida por criticar enérgicamente a la metafísica que no se encuentra suficientemente informada por los descubrimientos de nuestras mejores ciencias. Por ello, no es sorprendente que la metafísica naturalizada considere equivocado que la metafísica dispute con la ciencia: "Si hay una contradicción entre la física y la metafísica, entonces la metafísica debe ceder" (Bird, 2007, p. 7).[7] Este sentimiento es compartido incluso por metafísicos que no suelen ser considerados naturalistas. Jonathan Lowe, por ejemplo, aconseja a los metafísicos "abrirse a la posibilidad de que sus afirmaciones acerca de las características metafísicas de la actualidad puedan verse socavadas por desarrollos en la teoría científica empírica" (Lowe, 1998, p. 26). Incluso fuera de la metafísica naturalizada, son pocos los que se inclinarían a seguir a Parménides, quien —sobre bases *a priori*— argumentó célebremente en contra de la existencia del cambio, a pesar de la evidencia empírica en contra. En otras palabras, la visión heredada parece ser que la ciencia tiene prioridad sobre la teorización metafísica en casos de conflicto directo.

Esto promete resolver rápidamente el conflicto entre los lewisianos y las teorías no espaciales de la gravedad cuántica: dado que su resistencia —como la de Parménides— parece estar basada al menos en parte en razonamiento *a priori*, los lewisianos deberían simplemente ceder y abandonar la fundamentalidad de la distancia espacial.[8] Para la metafísica naturalizada, este conflicto es otra

[7] Evidentemente, este dictamen depende de una demarcación adecuada entre la física (y la ciencia en general) y la especulación descontrolada que puede llegar a involucrar algún formalismo matemático; donde algunos podrían preocuparse de que ciertas teorías de la gravedad cuántica sean ejemplos de esto último (véase, por ejemplo, Hedrich (2007) para una discusión de esta y cuestiones similares en relación con la teoría de cuerdas). Este asunto será dejado de lado aquí, por la razón discutida también más adelante en la nota al pie 20, a saber: que la presente defensa de los conflictos entre metafísica y ciencia no depende de la suposición de que las teorías científicas en cuestión sean especulativas o que aún no hayan recibido la adecuada confirmación empírica.

[8] Tal resistencia motivada a priori frente a la ciencia empírica será denominada "resistencia parmenídea" debido a la analogía metodológica con los argumentos *a*

ilustración de que la metafísica independiente de la ciencia no solo es fútil, sino también perjudicial cuando busca competir con la ciencia (por ejemplo, Bryant, 2020).[9] Sin embargo, este artículo propone que la metafísica naturalizada podría beneficiarse de preservar tales conflictos, al menos por un tiempo. Como observa Humphreys (2013): "[L]a ciencia contemporánea ha revelado un mundo mucho más sutil e interesante que los mundos a menudo simples de los ontólogos especulativos" (p. 75). En ocasiones, la ciencia entra en conflicto con la metafísica en su conjunto, en el sentido de que demuestra que ninguno de nuestros marcos metafísicos concebidos hasta ahora es viable. No obstante, con qué nueva metafísica reemplazará la ciencia a estas solo se revelará si recordamos examinar la metafísica de una teoría científica y no simplemente dar por terminada nuestra investigación —como a veces parece recomendar la metafísica naturalizada— cuando la ciencia demuestra que nuestras creencias metafísicas son erróneas. La presencia de conflictos entre metafísica y ciencia, sostengo, es a menudo la manera más vívida en que se manifiestan las sorpresas metafísicas de la ciencia, y por ello dichos conflictos son nuestra mejor señal de cuándo y dónde comenzar esa investigación. Esto no implica rechazar el veredicto según el cual la ciencia generalmente tiene prioridad sobre la metafísica. No obstante, el artículo propone la matización de que podría ser útil que la metafísica ofrezca un poco de resistencia parmenídea (adecuadamente motivada). Los conflictos resultantes fomentarán un compromiso con la cuestión metafísica subyacente y garantizarán así que la ciencia —a través de una metafísica naturalizada— no solo refute nuestra vieja metafísica, sino que

priori de Parménides contra la experiencia aparente del cambio. Esto, a pesar de la curiosa coincidencia de que la gravedad cuántica de lazos —especialmente con la aparente desaparición del tiempo (por ejemplo Anderson, 2012; Isham, 1993)— presenta similitudes en contenido con la metafísica de Parménides.

[9] Véase Ladyman & Ross (2007, pp. 19–20) para una crítica general de la metafísica de Lewis desde la perspectiva de la metafísica naturalizada.

también provea una nueva *metafísica* en su lugar. La resistencia parmenídea de los metafísicos podría, en otras palabras, resultar un recurso invaluable para la metafísica naturalizada. Más aún, el artículo sostiene que, al limitarse a plantear preguntas metafísicas aún sin respuesta, este tipo de resistencia parmenídea debería ser aceptable para la metafísica naturalizada incluso asumiendo todos sus postulados centrales.

Para ilustrar cómo los conflictos entre metafísica y ciencia pueden servir como generadores para la metafísica naturalizada, el artículo explora el conflicto entre las teorías no espaciales de la gravedad cuántica y la metafísica inspirada en Lewis. Los lewisianos sostienen que, sin el espacio en su ontología fundamental, tales teorías carecen de una relación constructora del mundo candidata y, en consecuencia, corren el riesgo de ser metafísicamente incoherentes. De acuerdo con la propuesta aquí presentada, esto es una ocasión para indagar en la cuestión metafísica que subyace al conflicto: ¿qué, si lo hay, podría reemplazar al espacio como la relación constructora del mundo en las teorías no espaciales de la gravedad cuántica?[10] En lugar de terminar con esta sugerencia metodológica, el artículo la ejemplifica revisando los detalles de la gravedad cuántica de lazos, sobre cuya base se argumenta que el entrelazamiento (*entanglement*) podría servir como la relación constructora del mundo en esta teoría no espacial. En pocas palabras, se constata que el entrelazamiento entre los nodos de la red de espines (*spin network*) de la gravedad cuántica de lazos es crucial para la emergencia del espacio en esta teoría. Esto se toma como indicio de que el entrelazamiento podría ser la relación en virtud de la cual los elementos de la red de espines constituyen un mundo y, por ende, la relación que cohesiona los mundos fundamentalmente no espaciales de la gravedad cuántica de

[10] Al formular esta pregunta con respecto a la gravedad cuántica, podríamos ser vistos como siguiendo la sugerencia de Norton (2020) de que "los metafísicos deberían utilizar las teorías cuánticas de la gravedad […] como incubadoras para una metafísica futura" (p. 1961).

lazos; una propuesta que, con razón, debería calificar como una alternativa "sutil e interesante" a las ontologías de la distancia, pero que, no obstante, está motivada por la ciencia, como requiere la metafísica naturalizada. Así, este planteamiento ejemplifica concretamente cómo la propuesta metodológica del artículo puede implementarse para contribuir a la metafísica naturalizada.

El artículo procede del siguiente modo: la Sección 2 introduce el estudio de caso: el conflicto entre los lewisianos y las teorías no espaciales de la gravedad cuántica. La Sección 3 discute luego cómo los filósofos que persiguen tales conflictos suelen ser ridiculizados en la metafísica naturalizada, pero argumenta que no deberían serlo si el conflicto está impulsado por la ausencia de una interpretación metafísica viable de la teoría científica. En tales casos, los conflictos pueden servir como una herramienta heurística para la metafísica naturalizada. La Sección 4 ejemplifica cómo esto se desarrolla en la propuesta de que el entrelazamiento es la relación constructora del mundo en la gravedad cuántica de lazos, respondiendo así a la pregunta que motivó el conflicto de los lewisianos con las teorías no espaciales de la gravedad cuántica. Finalmente, sigue una conclusión.

2. Un conflicto entre metafísica y ciencia

La tesis de la superveniencia humeana[11] "dice que en un mundo como el nuestro, las relaciones fundamentales son exactamente las relaciones espaciotemporales" (Lewis, 1994, p. 474).[12] Una relación fundamental es importante para el realismo modal de Lewis, puesto que explica por qué algo pertenece a uno y no a otro de todos los mundos posibles que existen realmente. Explica, en otras palabras, cómo dos entidades pueden ser reconocidas como *compañeros de*

[11] La tesis de la superveniencia humeana tiene más contenido que la afirmación sobre la fundamentalidad de las relaciones de distancia; véase, por ejemplo, Weatherson (2015) para una introducción.
[12] Es justo observar, como hace Darby (2009, p. 202), que Lewis consideraba esta tesis como falible.

mundo: "las cosas son compañeros de mundo si y solo si están relacionadas espaciotemporalmente. Entonces, un mundo está unificado por la interrelación espaciotemporal de sus partes" (Lewis, 1986a, p. 71). Por lo tanto, tanto la noción de mundo como la relación de compañeros de mundo están, según Lewis, fundamentadas en relaciones espaciotemporales.

El objetivo aquí no es una exégesis de Lewis, y el punto de partida para la discusión son, en cambio, propuestas más recientes, inspiradas por Lewis, que enfatizan la centralidad de la distancia *espacial* y que están dirigidas explícitamente contra las teorías no espaciales de la gravedad cuántica. Un ejemplo es la ontología de puntos materiales distanciados promovida por Esfeld & Deckert (2017). Ellos sostienen que la distancia espacial es la relación que hace el mundo; la "relación constructora del mundo" (*world-making relation*) en la terminología de Esfeld (2019). En su opinión, la necesidad de una relación constructora del mundo puede mantenerse independientemente del realismo modal. Con independencia de si existen los otros mundos posibles, hay un sentido plausible en el que nosotros —y las entidades con las que nos rodeamos— pertenecemos al mismo mundo. Somos compañeros de mundo y, como tales, alguna relación debe hacer verdadero este hecho: "dada una pluralidad de objetos, tiene que existir un cierto tipo de relaciones en virtud de las cuales estos objetos constituyan un mundo" (Esfeld & Deckert 2017, p. 3). La necesidad de una relación constructora del mundo proviene de la necesidad de pegamento ontológico, y esto se requiere tan pronto como nos comprometemos con la existencia de "una pluralidad de objetos" en el mundo actual. La tarea de la relación constructora del mundo es ser la relación que conecta los elementos de la ontología en configuraciones y, como tal, fundamenta si vivimos en un multiverso de partes más desconectadas o en un único universo conectado. Así, "¿qué hace que seamos compañeros de mundo?" se

sugiere como una cuestión metafísica genuina independientemente del realismo modal.[13]

La búsqueda de una teoría de la gravedad cuántica no comienza, sin embargo, con esta pregunta. Por el contrario, la investigación en gravedad cuántica está impulsada por una aspiración a la unificación de nuestra descripción de la naturaleza: "El problema de la gravedad cuántica (QG) es encontrar una teoría que describa los fenómenos en la intersección de la relatividad general (GR) y la teoría cuántica de campos (QFT)" (Crowther & Linnemann, 2017, p. 378). Este problema ha generado varios programas de investigación en competencia, siendo los más prominentes la teoría de cuerdas, la gravedad cuántica de lazos, la teoría de conjuntos causales y la triangulación dinámica causal (véanse las contribuciones en Oriti (2009) para una visión de conjunto de estos programas con enfoque en el espacio y el tiempo). En su intento por reconciliar la relatividad general y la mecánica cuántica, varios de estos parecen eliminar el espacio en el nivel fundamental de descripción.[14] La gravedad cuántica de lazos, que utilizaremos en lo que sigue para ilustración, ofrece una descripción fundamental del mundo en términos de una estructura de grafo abstracta compuesta por nodos y enlaces.[15] Tanto los nodos como los enlaces poseen una representación del espín (SU(2)) y juntos definen el espacio de Hilbert cinemático de estados. Los estados de este espacio de Hilbert son manifiestamente

[13] Véase Jaksland (2020) para un análisis más detallado de la necesidad de una relación constructora del mundo en ausencia del realismo modal.

[14] En la medida en que estas teorías aspiran a recuperar la relatividad general y la relatividad general describe la dinámica del espacio-tiempo, las teorías están comprometidas de manera ingenua con recuperar el espacio-tiempo como un fenómeno emergente. Si esto es suficiente para considerar el espacio-tiempo como real es objeto de debate (por ejemplo, Lam & Wüthrich, 2018, Le Bihan, 2018, 2021; Baron, 2019). Dejaremos de lado este debate aquí y nos limitaremos a observar que el espacio parece estar ausente en el nivel fundamental de la ontología, lo cual basta para generar el conflicto con la metafísica lewisiana.

[15] Esta es una de varias exposiciones posibles de la gravedad cuántica de lazos, véase Rovelli (2008) y las referencias allí incluidas para una visión general.

invariantes bajo transformaciones locales de *gauge* de la representación de espín, y esta presentación de la gravedad cuántica de lazos es por ello denominada como una red de espines. Puesto que las redes de espines son discretas, ya difieren de la estructura suave (*smooth*) del espacio. De hecho, la mayoría de los estados de redes de espines ni siquiera admitirán una aproximación como un espacio-tiempo suave del mismo modo en que una descripción termodinámica de equilibrio solo está disponible si el comportamiento macroscópico de un sistema es suficientemente robusto respecto de los cambios individuales entre los grados de libertad microscópicos. Cuando se dice que el espacio no es fundamental en la gravedad cuántica de lazos, esto se entiende como análogo al sentido en el que la temperatura no es fundamental en la mecánica estadística.[16]

Incluso si el espacio está ausente en el nivel fundamental según lo descrito por teorías no espaciales como la gravedad cuántica de lazos, Esfeld (2019) insiste en que la pregunta permanece acerca de qué hace que sea verdadero que seamos compañeros de mundo: "en este caso, necesitamos otra relación constructora del mundo distinta de la espacial o espaciotemporal" (p. 358). El problema, como lo formula Esfeld (2020) en otro lugar, es que "nadie ha elaborado hasta ahora una propuesta para otro tipo de relaciones que […] sea empíricamente adecuada" (p. 1892). Se podría objetar a Esfeld que si el espacio está ausente en el nivel fundamental de la descripción, entonces lo mismo podría ser el caso para los compañeros de mundo. En efecto, las personas, las mesas y las sillas no desempeñan ningún papel explícito, por ejemplo, en la descripción fundamental en términos de redes de espines en la gravedad cuántica de lazos. Esto,

[16] Dado que la termodinámica del equilibrio requiere que "todas los observables macroscópicos razonables tengan valores estacionarios" (Pitowsky 2006, p. 432), ninguna descripción macroscópica se aplica a muchos de los microestados cinemáticamente admisibles. Lo mismo se aplica a los estados genéricos de la red de espines: deben darse circunstancias especiales para que admitan una aproximación en donde algo sea reconocible como espacio. Véase Wüthrich (2017) para más detalles y discusión de la desaparición del espacio en la gravedad cuántica de lazos.

sin embargo, no muestra que no haya compañeros de mundo ni la necesidad de una relación constructora del mundo. Si dos elementos son compañeros de mundo en algún nivel de descripción, entonces, independientemente de a qué (complejas) correspondencias remitan al nivel fundamental de la descripción, cabe sostener que deben seguir siendo compañeros de mundo y que alguna relación debe hacer verdadero este hecho. Así, del mismo modo en que uno indaga cómo una teoría no espacial de la gravedad cuántica da cuenta del espacio, se puede indagar qué es lo que, en su nivel fundamental de descripción, hace que sea verdadero que somos compañeros de mundo. Como se discute con más detalle en la Sección 4, deberíamos mantenernos abiertos a la posibilidad de que estas teorías muestren que "compañero de mundo" es una categoría inadecuada, con la consecuencia de que no surge ninguna pregunta sobre la construcción del mundo; pero la razón sería más sutil que la ausencia, en el nivel fundamental, de las entidades que usualmente se supone que son compañeros de mundo. Más bien, la pregunta podría ser obsoleta si resulta que la categorización en términos de elementos y relaciones es problemática. En la situación actual, la necesidad de una relación constructora del mundo será en gran medida tomada como dada para fines de ilustración y, por ello, se deja abierta la posibilidad de que la física acabe por considerar que esta cuestión metafísica está mal planteada en lugar de proporcionar otra relación constructora del mundo;[17] aunque sostendremos que la gravedad cuántica de lazos sí parece sugerir el entrelazamiento como relación constructora del mundo alternativa.

Suponiendo que la pregunta acerca de la construcción del mundo está bien planteada, si la relación de distancia está ausente en

[17] Del mismo modo, no entraremos en el debate sobre si algunas visiones sobre la emergencia, la reducción ontológica y los niveles de realidad podrían permitir que entidades en algún nivel de descripción sean compañeros de mundo, mientras que entidades en otro nivel de descripción —como las de la base de superveniencia de las primeras— no lo sean.

el nivel fundamental de la ontología, entonces otra relación debe ocupar su lugar, pero, según Esfeld, no hay actualmente ninguna alternativa de este tipo disponible.[18] En consecuencia, Esfeld (2019) advierte contra la aparente importancia metafísica de estas teorías (aparentemente) no espaciales de la gravedad cuántica: "tal y como están las cosas, es razonable recomendar cautela respecto a proponer consecuencias ontológicas de gran alcance, como la desaparición del espacio-tiempo o de las relaciones espaciotemporales fundamentales" (p. 367). Estas preocupaciones acerca de mundos fundamentalmente no espaciales son razones para ser escépticos respecto a estas teorías científicas; especialmente de sus supuestas implicaciones metafísicas. En otras palabras, aunque Esfeld lo ofrece con un tono conciliador, las preocupaciones metafísicas se oponen a la teorización científica — en este caso, en la forma de teorías específicas de gravedad cuántica con implicaciones metafísicas preocupantes.

Mientras que Esfeld y los otros autores mencionados en la introducción son aquellos que expresan más explícitamente su preocupación acerca de la no-espacialidad de ciertas teorías de la gravedad cuántica, muchas de sus inquietudes son reflejadas por quienes prefieren la interpretación bohmiana de la mecánica cuántica debido a la preocupación por la ausencia de *beables* locales[19] en otras interpretaciones (por ejemplo, Bricmont, 2017; Maudlin, 2007a); algunos de estos expresando incluso una simpatía explícita por la superveniencia humeana (por ejemplo, Loewer, 1996; Miller, 2014). Asimismo, el programa de la ontología primitiva (también perseguido con mayor frecuencia en el contexto de la mecánica cuántica) requiere explícitamente que "cualquier teoría física fundamental satisfactoria […] contenga una hipótesis metafísica sobre lo que constituye los

[18] Dado que el espacio es emergente (en algún sentido) en las teorías no espaciales de la gravedad cuántica, uno podría proponer que la distancia podría seguir siendo la relación constructora del mundo. Véase, sin embargo, Wüthrich (2019, pp. 247-250) para un rechazo de este argumento.
[19] En el sentido de Bell (2001).

objetos físicos […] que reside en el espacio tridimensional o el espacio-tiempo y que constituye los bloques de construcción de todo lo demás" (Allori, 2015, p. 107). De acuerdo con la ontología primitiva (de este tipo), el espacio es una condición previa para cualquier metafísica satisfactoria y, por lo tanto, un componente esencial de la metafísica de una teoría física satisfactoria. En consecuencia, los defensores de la ontología primitiva y, más generalmente, aquellos que favorecen *beables* locales, muy probablemente compartirán el sentimiento de Esfeld de que las teorías sin espacio en el nivel fundamental de su ontología deberían ser tratadas con cautela, o quizás incluso considerar tales teorías como metafísicamente ilegítimas. Esto, entonces, genera una tensión con las teorías no espaciales de la gravedad cuántica y ofrece un ejemplo concreto de un conflicto entre metafísica y ciencia.

Antes de que procedamos a la actitud hacia tales conflictos en la metafísica naturalizada, uno podría preguntarse cómo este y otros conflictos similares pueden mantenerse el tiempo suficiente como para ser el objeto de una discusión metodológica. ¿Por qué, entonces, no se resuelven simplemente estos conflictos? Después de todo, las teorías científicas en cuestión están fácilmente disponibles en artículos de investigación, reseñas e incluso, a menudo, en libros de texto. Cuando Esfeld se preocupa por cuál es la relación constructora del mundo en las teorías no espaciales de la gravedad cuántica, la pregunta podría simplemente ser respondida con un "así es" — presentando la mejor descripción del mundo que ofrecen las teorías. Él podría simplemente echar un vistazo.[20]

[20] Como señala Esfeld (2019), "los enfoques de la gravedad cuántica que supuestamente implican que el espacio-tiempo o las relaciones espaciotemporales no son fundamentales son enfoques que, en el estado actual de las cosas, no producen ninguna predicción empírica" (pp. 365–366); se encuentran en un dominio muy especulativo de la física de altas energías. Esta es una nota justificada de precaución acerca de estas teorías, sin embargo, la defensa presentada aquí de los conflictos entre metafísica y ciencia no utilizará este hecho. Los metafísicos no

Sin embargo, la ciencia rara vez responde explícitamente las preguntas que interesan a los metafísicos. Como también se argumenta frecuentemente en relación con la subdeterminación de la metafísica por la ciencia (véanse, por ejemplo, Andersen & Becker Arenhart, 2016; French, 2011; Jones, 1991), la metafísica de una teoría científica rara vez se manifiesta en su formalismo. Así, tener la teoría fácilmente disponible por lo general no responderá inmediatamente a las indagaciones metafísicas que uno pueda dirigirle, y esto se debe a una buena razón: las preguntas que impulsan a las teorías científicas son diferentes de aquellas que impulsan la exploración metafísica de ellas.[21] Las teorías de gravedad cuántica y las preguntas metafísicas aquí dirigidas a ellas constituyen un buen ejemplo. Como se dijo anteriormente, la construcción de una teoría de gravedad cuántica está impulsada por el problema de reconciliar la relatividad general y la teoría cuántica de campos; un problema que ha demostrado ser difícil y que, por lo tanto, ha exigido desarrollos novedosos en la física teórica. Las teorías no espaciales están entre estos desarrollos, y su carácter sorprendente es consecuencia de estas dificultades. En otras palabras, estas teorías no fueron desarrolladas para cumplir una aspiración de explorar la posibilidad de una teoría física no espacial. Lo que estas teorías buscan poner de manifiesto es cómo podrían recuperar la relatividad general y la teoría cuántica de campos, y no qué reemplaza a la distancia como la relación constructora del mundo. Las teorías de gravedad cuántica están diseñadas para responder una

necesitan esto como excusa cuando plantean preguntas metafísicas acerca de teorías científicas.

[21] Esfeld (2019) se apoya en esta subdeterminación para cuestionar si el espacio está, de hecho, ausente en el nivel fundamental de estas teorías aparentemente no espaciales de la gravedad cuántica, aludiendo a la aparente no-localidad en la mecánica cuántica que, sin embargo, fue reinstaurada en la mecánica bohmiana (y así se demostró que era al menos consistente con el formalismo cuántico). Sin embargo, el argumento presentado aquí no se apoya en tal subdeterminación como parte de su defensa de los conflictos entre metafísica y ciencia, y se asumirá, por lo tanto, para fines del argumento, que el espacio-tiempo sí desaparece en estas teorías de la gravedad cuántica.

pregunta particular de la física; no diversas indagaciones metafísicas. Esto explica por qué un conflicto entre metafísica y ciencia puede perseverar: la ciencia rara vez, si acaso alguna vez, responde explícitamente las preguntas metafísicas, y esto —como se argumentará en la siguiente sección— es la razón por la que los conflictos entre metafísica y ciencia pueden resultar útiles.

3. Metafísica naturalizada y cuestiones metafísicas

La metafísica naturalizada presenta tanto un componente destructivo como un componente constructivo: el componente destructivo critica los métodos tradicionales de la metafísica —intuiciones, sentido común, análisis conceptual y razonamiento *a priori*—, mientras que el componente constructivo propone que la metafísica debería basarse, en cambio, en los hallazgos de nuestras mejores ciencias actuales: "El naturalismo exige que, dado que las instituciones científicas son los instrumentos mediante los cuales investigamos la realidad objetiva, sus resultados deberían motivar todas las afirmaciones acerca de esta realidad, incluidas las metafísicas" (Ladyman & Ross, 2007, p. 30).

Los defensores de la metafísica naturalizada sostienen que nuestras intuiciones, conceptos y patrones de razonamiento *a priori* son el resultado de la evolución biológica y, según Ladyman & Ross (2007), "no hay razón para imaginar que nuestras intuiciones habituales y respuestas inferenciales estén bien diseñadas para la ciencia o para la metafísica" (p. 3). Estas características cognitivas evolucionadas no proporcionan ninguna facultad que pueda ofrecer conocimientos acerca de la realidad y, por lo tanto, es problemático cuando han formado tradicionalmente la base metodológica para la metafísica. Esto es especialmente cierto cuando la metafísica se interesa por aquellos aspectos de la realidad con los que no nos encontramos en nuestro mundo de la vida, por ejemplo, el contenido y la estructura de la realidad fundamental. Aquí, cualquier dependencia de estas características evolucionadas es "ignorar el

hecho de que la ciencia, especialmente la física, nos ha mostrado que el universo es muy extraño para nuestra concepción heredada de cómo es" (Ladyman & Ross 2007, p. 10); sentimiento también expresado por Humphreys (2013, p. 75) más arriba. Según Ladyman y Ross, esta crítica también se aplica cuando la metafísica pasa de afirmaciones sobre el mundo actual a afirmaciones modales de posibilidad y necesidad. Al argumentar qué es metafísicamente posible o necesario, los métodos tradicionales de la metafísica han demostrado ser poco fiables.[22] Ladyman y Ross escriben:

> negamos que la indagación a priori pueda revelar qué es metafísicamente posible. Los filósofos han considerado a menudo como imposibles estados de cosas que la ciencia ha llegado a admitir. Por ejemplo, los metafísicos proclamaron con confianza que la geometría no euclidiana es imposible como modelo del espacio físico, que es imposible que no haya causalidad determinista, que el tiempo no absoluto es imposible, y así sucesivamente (Ladyman & Ross 2007, p. 16; véase también Maudlin 2007b, pp. 187–88).

Los métodos tradicionales de la metafísica no pueden emplearse ni para decir qué es el caso ni qué puede o no puede ser el caso. En particular, no pueden emplearse para informar qué rasgos son indispensables para mundos metafísicamente coherentes.

Mediante este componente destructivo de la metafísica naturalizada, no hay recursos dentro de la metafísica que puedan movilizarse para una disputa con la ciencia. De hecho, Ladyman y Ross rechazan enérgicamente este tipo de especulación:

> Los físicos no creen que existan fundamentos a priori lo suficientemente buenos como para mantener creencias acerca

[22] Dado que el objetivo aquí es una defensa de los conflictos entre metafísica y ciencia *dentro* de la metafísica naturalizada, se dará por sentada aquí esta supuesta falta de fiabilidad de los métodos tradicionales de la metafísica en cuestiones acerca de la posibilidad, aunque, como argumenta por ejemplo Morganti (2016, p. 87), es cuestionable si la ciencia ha sido más fiable en sus juicios sobre la posibilidad.

> de la constitución del mundo físico, y sugerimos que solo un filósofo temerario debería estar dispuesto a disputar con ellos sobre la base de sus intuiciones (Ladyman & Ross, 2007, p. 18).

Esto incluye las "corazonadas" (*hunches*) sobre la posibilidad y, por lo tanto, debería ser igualmente "temerario" disputar así si la ciencia en cuestión está empíricamente bien confirmada o no (aunque lo primero podría ser más descabellado que lo segundo). Incluso si una teoría científica no está actualizada, nada interno a la metafísica puede justificar la afirmación de que una teoría científica describe un estado de cosas imposible. Esto también está en buena concordancia con el componente constructivo de la metafísica naturalizada que requiere que la metafísica esté motivada por los resultados de la ciencia; estar en conflicto con una hipótesis científica —bien confirmada o no— parece ser exactamente lo opuesto. Pero, uno podría preguntarse, ¿cómo podemos decidir si esta teoría es una teoría científica — de modo que se trate de un conflicto entre la metafísica y la ciencia— y no una pieza de la metafísica misma, por la cual el conflicto sería interno a la metafísica? La metafísica naturalizada tiene dos respuestas independientes: primero, puede argumentarse que no hace ninguna diferencia si la teoría que se cuestiona desde la metafísica es científica o no. En ambas circunstancias, no hay una base epistémicamente legítima para el conflicto debido a los problemas con los métodos tradicionales de la metafísica. Segundo, Ladyman & Ross (2007) identifican la ciencia "utilizando factores institucionales como indicadores indirectos más que criterios directamente epistemológicos" (p. 37), de modo que "una 'hipótesis científica' se entiende como una hipótesis que es tomada en serio por una ciencia institucionalmente bona fide" (p. 30). Factores institucionales — como ser publicada en revistas científicas respetables y financiada por fondos de investigación científica (Ladyman & Ross, 2007, p. 36)— delimitan la ciencia de lo no científico en lugar de la confirmación empírica. Según Ladyman y Ross, la metafísica no puede cuestionar las hipótesis científicas identificadas institucionalmente, que incluyen

las de la mecánica cuántica y las teorías de la relatividad, pero también, posiblemente, las principales teorías en contienda de la gravedad cuántica, como la gravedad cuántica de lazos, la teoría de cuerdas, etc. Desde la perspectiva de la metafísica naturalizada, proclamar que los mundos sin espacio son imposibles parece diferir de la afirmación de que los mundos no euclidianos son imposibles, y uno podría, por lo tanto, especular si lo primero, como lo segundo, es un ejemplo de exceso de confianza por parte de los metafísicos.

Esta visión ha sido recientemente promovida por Lam & Wüthrich (2020). Suponiendo que la ciencia debería informar (o "guiar") a la metafísica, argumentan que

> desde el punto de vista de los enfoques de la QG [gravedad cuántica] que apuntan a la desaparición del espacio-tiempo […] asumir a priori un marco ontológico para la QG que dependa de algún trasfondo de espacio-tiempo estándar y liso (por ejemplo, asumir a priori una ontología de beables locales para la QG) no es ni físicamente ni metafísicamente legítimo (contrariamente a lo que a veces se afirma en la literatura, véase Esfeld, 2019). En efecto, en esta perspectiva, tal suposición metafísica es ilegítima, puesto que entra directamente en conflicto con ciertos principios físicos constitutivos en los que se basan los enfoques de QG considerados (Lam & Wüthrich, 2020, p. 346).

Dado que Esfeld y otros, en la construcción de Lam y Wüthrich, defienden la necesidad de que el espacio(-tiempo) esté en el nivel fundamental de la ontología sobre bases *a priori*, es *ilegítimo* que mantengan esta suposición metafísica cuando los enfoques considerados de la gravedad cuántica señalan la no-fundamentalidad del espacio(-tiempo). La teoría científica tiene prioridad sobre el razonamiento metafísico en casos de conflicto. Insistir en una suposición metafísica *a priori* en la interpretación de la teoría científica, esto es, "[p]ostular ontologías en algún espacio-tiempo fijo de trasfondo para estos enfoques de QG", concluyen Lam y

Wüthrich, "entra en tensión directa con el naturalismo que hemos adoptado y en particular con un enfoque naturalista de la metafísica" (Lam & Wüthrich, 2020, p. 347). Al enfatizar su foco en "enfoques de QG" específicos y al reconocer que estos todavía carecen de apoyo empírico, Lam y Wüthrich están abiertos a la posibilidad de que la teoría de la gravedad cuántica que finalmente sea vindicada sea una en la cual el espacio siga siendo fundamental. Sin embargo, insisten, y así hacen eco de la postura mencionada arriba, en que esto no excusa las objeciones metafísicas a estos enfoques no espaciales. Lam y Wüthrich apelan a un enfoque naturalista de la metafísica para argumentar que Esfeld y otros deben poner fin a su crítica ilegítima, metafísicamente motivada, de las teorías no espaciales de la gravedad cuántica.[23]

Pero, ¿por qué exactamente la metafísica naturalizada es crítica de los conflictos metafísicos (*a priori*) con la ciencia? El objetivo general de la metafísica naturalizada no es, y no debería ser, poner fin a la metafísica. La metafísica naturalizada implica una crítica de los métodos pero no del objeto de estudio de la metafísica tradicional. En contraste con los programas eliminativistas que abogan por el antirrealismo metafísico (por ejemplo, Carnap, 1950; Chalmers, 2009; Yablo, 1998), Ladyman (2017) insiste en que "la metafísica no debería ser abolida, sino reformada" (p. 143).[24] El componente constructivo

[23] Nótese que esto no implica que las teorías no espaciales de la gravedad cuántica consideradas por Lam y Wüthrich tengan prioridad sobre otras teorías de gravedad cuántica que son más hospitalarias con relaciones espaciales fundamentales (por ejemplo Goldstein & Teufel, 2001; Dürr et al., 2018). Así, aunque los metafísicos, según su punto de vista, deberían terminar su crítica a las teorías no espaciales, los metafísicos, por supuesto, son bienvenidos a dedicar su tiempo a las implicaciones metafísicas de otras teorías de la gravedad cuántica. Sin embargo, de acuerdo con la metafísica naturalizada, será la ciencia la que decida cuál de estas será vindicada y no algún argumento *a priori* que sostenga que una u otra teoría es metafísicamente imposible.

[24] La eliminación de los metafísicos tampoco es un rasgo característico de la metafísica naturalizada. La metafísica naturalizada no recomienda que los metafísicos se conviertan en científicos y adopten los métodos de la ciencia en metafísica; no se caracteriza por este tipo de naturalismo metodológico, a pesar de

de la metafísica naturalizada implica la introducción de nuevos enfoques de metafísica informados por la ciencia que puedan reemplazar aquellos métodos ilícitos tradicionalmente empleados en metafísica, preservando al mismo tiempo el objeto de estudio y, por consiguiente, las ambiciones de la metafísica.[25] Allí donde los antirrealistas metafísicos son críticos con las preguntas metafísicas —por ejemplo, describiéndolas como "carentes de contenido cognitivo" (Carnap, 1950, p. 28)—, la metafísica naturalizada es crítica de cómo procedemos al responder preguntas metafísicas. Para la metafísica naturalizada, las preguntas mismas no son el problema. Sin embargo, como matiza Ladyman (2017), "[e]so no significa que [los metafísicos naturalizados] aboguen por responder todas las mismas preguntas que plantean los metafísicos analíticos por medios diferentes", puesto que algunas de ellas "no hacen un contacto suficiente con la realidad como para que valga la pena tomarlas en consideración" (p. 143). Algunas —quizás incluso muchas— preguntas metafísicas son (actualmente) epistémicamente inseguras de responder, dado que el proceso de responderlas no estará lo suficientemente inspirado ni constreñido por la ciencia como para satisfacer los estándares de la metafísica naturalizada. No obstante, siguen siendo este tipo de preocupaciones *epistémicas* las que están —o al menos deberían estar— detrás de la actitud despectiva hacia los metafísicos que disputan con la ciencia.

Esto también debe aplicarse cuando los lewisianos preguntan por la relación constructora del mundo, proponen la distancia como

que algunos autores sostienen lo contrario (por ejemplo Esfeld, 2018; Hudson, 2016).

[25] Ney (2012), por ejemplo, considera que la tarea de la metafísica naturalizada es "establecer conclusiones acerca de la realidad última" (p. 76) y Ladyman & Ross (2007) sostienen que "ningún otro tipo de metafísica cuenta como investigación acerca de la naturaleza objetiva del mundo" (p. 9). En este sentido, la metafísica naturalizada difiere de otros intentos recientes de rescatar la metafísica que, en cambio, adoptan ambiciones más modestas respecto del contenido de la metafísica (por ejemplo Hofweber, 2016; Kraut, 2016; Thomasson, 2014).

respuesta y luego advierten contra las teorías no espaciales de la gravedad cuántica en ausencia de una alternativa a la distancia. Reconociendo que el objetivo de la *metafísica* naturalizada es la metafísica, el problema con este conflicto *no* es la pregunta metafísica que subyace a él: ¿cuál es la relación constructora del mundo? Más bien, la cuestión concierne a los métodos que se emplean para promover el conflicto, esto es, las intuiciones y el razonamiento *a priori* —las "corazonadas" de los metafísicos— que entran en el argumento en el sentido de que la distancia es indispensable. Ciertamente hay algún mérito (al menos desde una perspectiva naturalista) en la idea de que es imprudente otorgar demasiada importancia a nuestra aparente incapacidad de imaginar un mundo sin espacio; especialmente considerando el pobre historial de éxito de las proclamaciones sobre posibilidad metafísica. Sin embargo, los argumentos *a priori* para la indispensabilidad de la distancia no son los elementos decisivos en, por ejemplo, la advertencia de Esfeld contra las teorías no espaciales de la gravedad cuántica. Como se discutió en la Sección 2, la advertencia de Esfeld está motivada por la ausencia de una alternativa a la distancia como relación constructora del mundo y él reconoce explícitamente que "[l]a afirmación de que no hay relaciones espaciotemporales fundamentales podría ser verdadera" (Esfeld, 2019, p. 356). Este conflicto es, en otras palabras, no (solo) el resultado de una preferencia *a priori* por la distancia como relación constructora del mundo, como parecen sugerir Lam y Wüthrich, sino más bien impulsado principalmente por la aparente falta de cualquier otra respuesta a la pregunta metafísica. Mientras que la ausencia de espacio en el nivel fundamental en, por ejemplo, la gravedad cuántica de lazos implica que la distancia no puede ser la relación constructora del mundo, no está manifiesto en la presentación usual de la teoría qué otra relación hace que los elementos de la ontología sean compañeros de mundo. En consecuencia, los lewisianos, en su conflicto con estas teorías no espaciales de la gravedad cuántica, no son simplemente testarudos,

sino que mantienen su posición dado que estas teorías solo proporcionan una respuesta negativa a su indagación metafísica.

En este sentido, el conflicto, y la resistencia parmenídea de los lewisianos, es diferente de los conflictos en los que se encuentra disponible una interpretación metafísica de la teoría científica, pero algunos metafísicos simplemente la desprecian. Un ejemplo (admitido como controvertido) de esto último es cuando la teoría de la relatividad aparentemente se alinea con el eternalismo y la teoría-B del tiempo frente a la teoría-A del tiempo bajo la forma del presentismo o del universo creciente de bloque. El potencial conflicto entre los presentistas y la teoría de la relatividad es de un tipo diferente al del que se discute aquí y ejemplifica el conflicto entre los lewisianos y las teorías no espaciales de la gravedad cuántica. Con respecto a la cuestión de cómo concebir el tiempo, la teoría de la relatividad tiene una interpretación metafísica aparentemente consistente en el eternalismo, mientras que las teorías no espaciales de la gravedad cuántica no tienen tal interpretación respecto de la cuestión de la construcción del mundo. Podríamos decir que la ciencia, en casos como el de la teoría de la relatividad, no está realmente en conflicto con la metafísica, sino en conflicto con un lado dentro de un debate metafísico. Según los estándares de la metafísica naturalizada, los metafísicos que disputan con teorías científicas porque detestan la visión metafísica favorecida por la teoría probablemente deberían simplemente ceder (aunque este asunto es por supuesto mucho menos claro una vez que reconocemos que las implicaciones metafísicas podrían ser controvertidas y posiblemente subdeterminadas).

En comparación, la resistencia de los lewisianos a las teorías no espaciales no es simplemente una expresión de preferencia por ontologías basadas en el espacio (aunque esto, por supuesto, todavía puede tener un lugar como sugieren Lam y Wüthrich). Si una presentista abandonara su convicción en la existencia exclusiva del presente y solo mantuviera su pregunta sobre la naturaleza del tiempo, el conflicto con la teoría de la relatividad se disiparía inmediatamente

al responder la pregunta con el eternalismo.[26] No sería este el caso para los lewisianos. De no existir el conflicto mismo, entonces gran parte de la tensión entre metafísica y ciencia persistiría incluso si la controvertida propuesta positiva de que la distancia es la relación constructora del mundo quedara atrás. Los lewisianos moderados —entre los cuales bien podría estar Esfeld— podrían simplemente estar preguntando "¿qué es lo que hace que seamos compañeros de mundo en teorías no espaciales de la gravedad cuántica?", pero solo recibir la respuesta "no la distancia espacial". Dado que la metafísica naturalizada no renuncia en general a la significatividad ni a la aceptabilidad de las preguntas metafísicas, los lewisianos ciertamente pueden *esperar* una respuesta a su pregunta e insistir en que sigamos buscando una.

Para permanecer fiel al espíritu de la metafísica naturalizada, no obstante, esta búsqueda debe ser sensible a las señales provenientes de la teoría científica de que la pregunta no es epistémicamente segura de responder o que está planteada de tal manera que la teoría científica debe ser apropiada o "domesticada", como lo llaman Ladyman y Ross, para los propósitos metafísicos de nuestra pregunta:

> Un aspecto de dejar la ciencia indomesticada es reconocer que ella misma puede decirnos que hay preguntas que absolutamente no puede responder porque cualquier respuesta tentativa es tan probable como cualquier otra. Esto no implica que debamos buscar en una institución diferente de la ciencia para responder tales preguntas; en estos casos deberíamos olvidarnos de las preguntas (Ladyman & Ross, 2007, p. 30).

[26] Puede argumentarse que los teóricos-A podrían simplemente insistir en que la teoría de la relatividad con el eternalismo no proporciona una respuesta satisfactoria a su pregunta metafísica y que, en consecuencia, no hay una respuesta alternativa a dicha pregunta. Es, ciertamente, poco claro quién debería actuar como árbitro en tales casos. Existe, por lo tanto, el riesgo de que permitir conflictos entre ciencia y metafísica en general termine por sancionar cualquier conflicto entre metafísica y ciencia, en contra de lo que es la intención de esta propuesta.

Debemos aceptar que algunas preguntas metafísicas no son contestables y, al menos implícitamente, a partir de esta observación, que algunas preguntas metafísicas pueden, en su propia formulación, presumir supuestos metafísicos que podrían hacer que las preguntas sean inaplicables a la teoría científica considerada.

Las circunstancias pueden, en otras palabras, ser tales que una pregunta metafísica deba, en última instancia, quedar sin respuesta, pero insistir en la pregunta inicialmente promete ser una buena manera de descubrirlo. Así, la terquedad de los lewisianos impone una tensión —un eco de un conflicto— que asegura que se busque una respuesta positiva a la pregunta sobre la construcción del mundo. Aunque los lewisianos podrían no estar justificados en su preferencia por la distancia como la relación constructora del mundo debido a su origen en los problemáticos métodos tradicionales de la metafísica, su resistencia parmenídea asegura la reafirmación de la pregunta metafísica: ¿qué reemplaza a la distancia como la relación constructora del mundo en las teorías no espaciales de la gravedad cuántica? Más generalmente, están participando de una aspiración general según la cual las teorías científicas necesitan una interpretación metafísica coherente que vaya más allá de meras respuestas negativas a las indagaciones metafísicas; una aspiración que debería ser aceptable para y compartida por la metafísica naturalizada siempre que se mantenga alejada de los métodos disputados de la metafísica tradicional y sea cuidadosa de evitar domesticar las teorías científicas. A menos que permitamos cierta resistencia parmenídea, podríamos pasar por alto cuando las teorías científicas han revelado metafísicas hasta ahora inconcebibles.[27]

[27] Aunque lo presenta de manera algo diferente, Norton (2020, pp. 1966–1971) ofrece cuatro ejemplos históricos en los que la ciencia entró en conflicto con creencias metafísicas firmemente sostenidas y cómo estos conflictos conducen a lo que él describe como "revoluciones conceptuales". En nuestros términos, estos conflictos, tal como documenta Norton, fueron utilizados como heurísticas para el desarrollo de una nueva metafísica naturalizada. Norton (2020, pp. 1971–1980) también da cuenta de cómo las teorías no espaciales de la gravedad cuántica sugieren

Los conflictos abiertos de los metafísicos con las teorías científicas sirven para recordarnos que debemos buscar respuestas positivas a nuestras preguntas metafísicas y no conformarnos con respuestas negativas. En consecuencia, los conflictos entre ciencia y *metafísica en su conjunto* (y no solo un lado de un debate metafísico) son señales de cuándo la ciencia deja preguntas metafísicas sin responder. En la medida en que la metafísica naturalizada aspira a responder preguntas metafísicas, los conflictos entre metafísica y ciencia deberían, por lo tanto, ser bienvenidos como ocasiones para hacer (más) metafísica naturalizada. Esta apología de los *conflictos* entre metafísica y ciencia no ofrece ninguna justificación para priorizar los puntos de vista de la metafísica *a priori* por sobre los de la ciencia (o de la metafísica basada en la ciencia). Pero se argumenta que la metafísica a priori —mediante sus tensiones con la ciencia— sigue siendo importante y útil cuando la metafísica naturalizada intenta desarrollar una ontología informada por nuestras mejores teorías científicas. En este sentido, la propuesta presente puede considerar como una que añade otra función a la metafísica *a priori* dentro del enfoque de "caja de herramientas" (*toolbox-approach*) a la metafísica de French & McKenzie (2012), según el cual los métodos y marcos desarrollados dentro de la metafísica *a priori* tradicional se valoran por su utilidad como herramientas que los metafísicos naturalizados pueden emplear con diversos fines en su metafísica científicamente informada y constreñida (para una discusión más amplia, véase French & McKenzie, 2016; Ross, 2016; Le Bihan & Barton, 2018; French, 2018).

4. La relación constructora del mundo en gravedad cuántica

una reconfiguración de la distinción entre objetos concretos y abstractos, lo que podría verse como otro ejemplo más en el que un conflicto entre metafísica y gravedad cuántica impulsa desarrollos importantes en metafísica.

Los lewisianos pueden ser caracterizados como quienes indagan qué reemplaza a la distancia como la relación constructora del mundo en las teorías no espaciales de la gravedad cuántica; una pregunta legítima —incluso según los estándares de la metafísica naturalizada— sin una respuesta inmediata. En términos generales, la preocupación de los lewisianos por la coherencia metafísica de una teoría científica y el conflicto resultante —como otros conflictos entre metafísica y ciencia— puede ser recibida como la indicación de un posible problema abierto relacionado con el fundamento metafísico de la teoría en cuestión. Esta sección mostrará con mayor detalle —a través del ejemplo del conflicto entre los lewisianos y las teorías no espaciales de la gravedad cuántica— cómo tales preocupaciones metafísicas pueden utilizarse como una heurística en la metafísica naturalizada. Más precisamente, muestra cómo los lewisianos —y Esfeld en particular—, con su resistencia parmenídea, llaman la atención sobre la interesante pregunta metafísica: ¿qué reemplaza a la distancia como aquello que conecta los elementos de la ontología en las teorías no espaciales de la gravedad cuántica? Basándome (principalmente) en el razonamiento proveniente de la teoría de cuerdas, he mostrado en otra parte los detalles de cómo el entrelazamiento puede servir como una relación constructora del mundo alternativa (Jaksland 2020). El presente análisis retoma esta respuesta en términos de entrelazamiento, pero lo hace basándose en la gravedad cuántica de lazos, que ha sido el ejemplo principal de una teoría no espacial de la gravedad cuántica en el debate entre Esfeld (2019) y Lam & Wüthrich (2020). Aunque esta perspectiva desde la gravedad cuántica de lazos complementa bien otros argumentos a favor del entrelazamiento como relación constructora del mundo, el objetivo aquí es primero y ante todo mostrar cómo este trabajo se desarrolla como una metafísica naturalizada impulsada por el conflicto lewisiano con teorías no espaciales de la gravedad cuántica como la gravedad cuántica de lazos. En particular, el presente análisis indicará cómo responder a la pregunta lewisiana sobre la construcción del mundo —de acuerdo

con los estándares de la metafísica naturalizada— requiere negociar entre las restricciones provenientes de la teoría científica y los presupuestos que están implícitos en esta pregunta metafísica.

Para una ontología de objetos en el espacio, la distancia es una relación constructora del mundo ejemplar: cada objeto está a una distancia de cualquier otro objeto tal que la relación de distancia puede hacer que sea verdadero que dos objetos sean compañeros de mundo. Además, una relación constructora del mundo coherente debería, según Esfeld (2020), "(a) hacer el trabajo de individualizar objetos simples y (b) ser empíricamente adecuada" (p. 1892). Esto ya expone un dilema para nuestra investigación: ¿cuánto y qué aspectos de la pregunta metafísica deberían preservarse? Cuando se plantea una pregunta metafísica en el contexto de una teoría científica, siempre existe la posibilidad de que la pregunta sea descartada por considerarla mal planteada en lugar de responderla. La teoría científica podría simplemente exponer que la pregunta se formula sobre premisas falsas o depende de metáforas inapropiadas. En la medida en que la nueva metafísica interesante se debe a las respuestas, es importante que la pregunta metafísica esté bien planteada. En la formulación de la pregunta, Esfeld presupone una metafísica de objetos individuales; esto es parte de aquello acerca de lo cual la relación constructora del mundo debería dar cuenta. Si bien la adecuación empírica parece ser un requisito mínimo relevante, la individuación de objetos simples —especialmente dado que Esfeld (2020, p. 1893) exige discernibilidad absoluta[28]— conlleva el tipo de prejuicios metafísicos que corren el riesgo de hacer que nuestras preguntas estén mal planteadas. Más generalmente, si intentamos preservar demasiadas de nuestras intuiciones metafísicas al responder la pregunta metafísica que impulsa el conflicto, esto podría impedir las novedades metafísicas de la teoría y, además, empujarnos hacia la

[28] Véase Saunders (2006) para un análisis de los tipos de individuación en el contexto de la mecánica cuántica.

domesticación de la ciencia contra la que advierte la metafísica naturalizada. La ambición no debe ser satisfacer a metafísicos testarudos, sino utilizar los conflictos entre metafísica y ciencia como una ocasión para la exploración con la mente abierta.

Me parece que nunca hay garantía de que una pregunta metafísica esté bien planteada, ya que nunca puede estar completamente separada de cualquier supuesto metafísico de fondo. Incluso preguntar por una relación constructora del mundo sin supuestos sobre la naturaleza de los *relata* implica que hay una relación en algún sentido reconocible. El resultado de la exploración de las teorías no espaciales de la gravedad cuántica podría, por lo tanto, ser que también esta pregunta esté equivocada. Aun así, esta ausencia de relaciones sería un descubrimiento metafísico, especialmente si viene acompañada de indicaciones sobre cómo construir mundos sin relaciones, y la búsqueda de la pregunta metafísica detrás del conflicto entre metafísica y ciencia habría producido percepciones interesantes. Sí pienso, no obstante, que la pregunta de Esfeld y de los lewisianos acerca de la relación constructora del mundo en la gravedad cuántica puede ser respondida, al menos si se le despoja de sus prejuicios ontológicos sobre los objetos.

La propuesta de la distancia como relación constructora del mundo nos ofrece dos indicios sobre qué buscar. Primero, buscamos una relación que relacione cada par de elementos del mundo (aunque estos pueden no ser reconocibles como objetos en ningún sentido estricto). Segundo, la distancia es probablemente una relación constructora del mundo, lo cual sugiere que cualquier distancia que se derive de estas teorías no espaciales es una candidata probable. Siguiendo este segundo indicio, deberíamos observar cómo se supone que surge el espacio[29] a partir de las redes de espines de la gravedad

[29] Ninguno de estos enfoques es aún riguroso y, por lo tanto, no pueden afirmar con absoluta certeza cómo emerge el espacio.

cuántica de lazos.[30] Como se expuso en la Sección 2, las redes de espines son, de entrada, una estructura de grafo abstracta con una representación del espín (SU(2)) para los nodos del grafo y otra para los enlaces. A partir de esto, puede construirse un espacio de Hilbert que define los estados de la gravedad cuántica de lazos. La ruta hacia el espacio va a través del teorema del espín-geométrico de Penrose (1971), que implica que cada nodo puede asociarse a un poliedro; un objeto geométrico con caras poligonales descritas únicamente por las áreas y los ángulos entre sus caras. Más precisamente, para cada enlace puede construirse un escalar a partir de operadores de espín: cada nodo queda asociado con tantos escalares como enlaces tiene y, con el requisito adicional de invarianza *gauge* en la representación del espín, estos escalares de cada nodo determinan de manera única un poliedro, donde los escalares se identifican con las áreas de sus caras. Los ángulos entre las caras, junto con las áreas, pueden emplearse para definir una métrica tridimensional y el volumen del poliedro. Los ángulos, al igual que las demás propiedades, siguen siendo cuánticos y, por lo tanto, están asociados a operadores no conmutativos: "la forma de un poliedro cuántico es borrosa" (Bianchi, 2017, p. 112). Sin embargo, bajo condiciones adicionales de coherencia (para más detalles, véase Bianchi et al., 2011), las áreas toman valores definidos y el valor esperado de los operadores de ángulo aproxima los ángulos clásicos. Los poliedros se vuelven así semiclásicos y cada nodo, junto con sus enlaces, puede recibir de esta manera una interpretación geométrica como un fragmento de espacio cuyo volumen y métrica están determinados por la red de espines: "la estructura algebraica [de la representación del espín] determina la existencia de una métrica en cada nodo y, por lo tanto, equipa cada cuanto de espacio con una geometría" (Rovelli, 2011, p. 4). Cada nodo está enlazado con otros nodos y, desde el punto de vista geométrico, esto puede concebirse como poliedros adyacentes entre sí. De esta manera, un espacio

[30] Esta exposición está parcialmente basada en Rovelli (2011).

celular compuesto por muchos poliedros puede surgir de la red de espines. Es intrigante imaginar cómo este espacio, pese a su granularidad, puede aproximarse a un espacio liso, así como un dodecaedro regular —el sólido platónico consistente en doce pentágonos— puede parecer redondo desde lejos. Sin embargo, incluso desde lejos —y, por lo tanto, ignorando la granularidad debida a los poliedros— la métrica de este espacio es discontinua. En la red de espines, cualquier par de nodos enlazados tiene una interpretación geométrica como dos poliedros enfrentados. Dado que comparten el mismo enlace, las caras tienen el mismo área, pero como los nodos genéricos tienen un número diferente de enlaces con otros nodos, la forma y el ángulo entre las caras serán diferentes aunque tengan la misma área: no hay ajuste en su forma. En consecuencia, la métrica es generalmente discontinua (Bianchi et al., 2011, 11).

En un artículo con Antonio Vassallo, Esfeld especula sobre cómo tales redes de poliedros —o "átomos de espacio", como ellos los llaman— podrían conectarse entre sí para formar un espacio continuo. Sin embargo, ellos también reconocen implícitamente que la distancia espacial —o, en general, las "propiedades métricas"— está ausente en este nivel fundamental de redes de poliedros:

> agrupar los átomos de espacio juntos de una manera adecuada, tal como la representada por nodos y aristas en un grafo, hace posible que la configuración instancie propiedades métricas, mientras que los átomos individuales del espacio están conectados únicamente por una relación de contigüidad (Vassallo & Esfeld, 2014, p. 10).[31]

Así, Esfeld también parece estar de acuerdo en que la distancia no puede ser la relación constructora del mundo en la gravedad cuántica de lazos. La mencionada "relación de contigüidad" no es

[31] La ontología de Vassallo y Esfeld, compuesta por estos átomos de espacio, podría ser una candidata interesante para los elementos en una ontología con el entrelazamiento como la relación constructora del mundo, pero desarrollar esta propuesta se pospondrá para trabajos futuros.

mucho más que notar que los poliedros en la representación de la red de espines están conectados por enlaces en el grafo. Y si bien esto puede ser visualmente intrigante, no es más que otra forma de representar el formalismo. Proponer esta contigüidad como la relación constructora del mundo —lo cual ni Vassallo ni Esfeld hacen— parece equivalente a simplemente presentar el formalismo de la gravedad cuántica de lazos como respuesta.[32] Como ya anuncié, creo que podemos hacerlo mejor al considerar cómo el espacio emerge de la métrica inicialmente desconectada de poliedros contiguos.

Los espacios *continuos* (aunque todavía celulares), conocidos como geometrías de Regge, corresponden a estados especiales de redes de espines donde las caras están emparejadas por forma y alineadas. Investigaciones recientes sugieren que el entrelazamiento entre los nodos de la red de espines desempeña un papel importante en este efecto.[33] Se puede comprender el papel del entrelazamiento recordando que estamos tratando con poliedros cuánticos: al igual que un electrón puede encontrarse en una superposición de espín arriba y espín abajo, un poliedro cuántico puede estar en una superposición de diversas formas. Dos poliedros adyacentes, esto es, que comparten un enlace en el grafo, estando ambos en tal estado de superposición, podrían tener el mismo espectro de formas, pero al colapsar la superposición podrían colapsar en distintas formas: "su geometría tiene fluctuaciones no correlacionadas. A nivel clásico este comportamiento corresponde a una geometría retorcida —la geometría de una colección de poliedros con formas no correlacionadas" (Baytaş et al., 2018, p. 15). Considerando nuevamente a los electrones, el colapso del estado de superposición

[32] Cabe argumentar que puede y debe decirse más sobre este asunto, pero esto se pospondrá para trabajos futuros.
[33] Más precisamente, puede demostrarse que la geometría de twistores de estados genéricos de la red de espines se convierte en una geometría vectorial (de la cual las geometrías de Regge son un subconjunto) al entrelazar los nodos de la red con sus vecinos más cercanos (Baytaş et al., 2018).

de dos electrones puede estar correlacionado por el entrelazamiento, esto es, están correlacionados si son preparados en un estado inseparable, como el estado de Bell. Lo mismo ocurre en las redes de espines. Entrelazar nodos vecinos produce correlaciones entre los poliedros y, por lo tanto, parece ser una condición necesaria para el alineamiento y emparejamiento de formas de las caras. Baytaş et al. (2018) concluyen: "Los resultados presentados muestran claramente el papel del entrelazamiento en la unión de regiones cuánticas de espacio" (p. 16). Parece, en otras palabras, que en la gravedad cuántica de lazos el entrelazamiento es responsable de la emergencia del espacio celular continuo que "desde lejos" se verá como un espacio (semi-)clásico.

El entrelazamiento parece conectar las "regiones cuánticas del espacio" en forma de poliedros. Aunque estos poliedros o sus nodos relacionados se parezcan poco a los objetos tal como los conocemos, el entrelazamiento cumple así un papel similar al de la distancia al conectar los elementos de la ontología. El entrelazamiento es interesante a este respecto, puesto que comparte algunas de las características que hicieron de la distancia una probable relación constructora del mundo (Jaksland, 2020).[34] En primer lugar, el entrelazamiento es una propiedad extrínseca; algo está entrelazado con algo *más*. En segundo lugar, el entrelazamiento comparte la universalidad de la distancia: la distancia puede relacionar todo en el espacio con todo lo demás y, de modo análogo, el entrelazamiento puede darse entre cualesquiera grados de libertad cuánticos. Suponiendo que todos los grados de libertad sean cuánticos en gravedad *cuántica*, esto implica que todos los grados de libertad

[34] Lewis sostiene específicamente que, si el espacio ha de ser reemplazado, debe serlo por una relación que sea análoga (en un sentido especificado) a una relación espaciotemporal. Una discusión más detallada del grado en que el entrelazamiento satisface la condición de Lewis para una relación espaciotemporal análoga será pospuesta para trabajos futuros. Véase, sin embargo, Wüthrich (2019) para algunas observaciones preliminares acerca de relaciones espaciotemporales análogas en gravedad cuántica.

pueden estar entrelazados en tales teorías. Además, resultados provenientes de la teoría cuántica de campos algebraica relativista indican que todos los grados de libertad no solo pueden, sino que de hecho están, entrelazados, lo que atestigua la ubicuidad del entrelazamiento. Más precisamente, Redhead (1995) muestra que, en el estado de vacío, todos los subsistemas separados por intervalo espacial —todos los subsistemas que también están conectados por una distancia— están fuertemente entrelazados, y este resultado se generaliza a estados genéricos por Clifton y Halvorson (2001), quienes también muestran que ninguna operación local puede desenredar subsistemas separados por intervalo espacial (véanse Lam (2013) y Swanson (2020) para más detalles). En tercer lugar, el entrelazamiento es, vía la entropía de entrelazamiento y la información mutua, cuantificable como un escalar no negativo, igual que la distancia. Dados estos rasgos y su papel en la gravedad cuántica de lazos, el entrelazamiento es un reemplazo prometedor de la distancia como relación constructora del mundo. El entrelazamiento podría ser la relación en virtud de la cual los elementos de la red de espines constituyen un mundo y, por ende, la relación que pega los mundos fundamentalmente no espaciales de la gravedad cuántica de lazos. Si bien las teorías no espaciales de la gravedad cuántica solo habían proporcionado hasta ahora la respuesta negativa de que la distancia no es la relación constructora del mundo, la resistencia parmenídea de los lewisianos ha motivado el (preliminar) desarrollo de una nueva metafísica en la forma de un fundamentalismo del entrelazamiento hasta ahora no visto en la literatura metafísica.

La propuesta de que el entrelazamiento podría ser la relación constructora del mundo en la gravedad cuántica de lazos se afirma que califica como metafísica naturalizada. Es una propuesta que responde a una pregunta metafísica, pero cuya respuesta está motivada por una teoría científica. Al ser relativa a la gravedad cuántica de lazos, la propuesta no afirma nada acerca de lo que es y ciertamente no acerca de lo que debe ser el caso en la realidad efectiva.

Siguiendo la deferencia a la ciencia en la metafísica naturalizada, es la ciencia la que determina qué es y qué no es el caso. Por lo tanto, la afirmación es que, si la gravedad cuántica de lazos resulta finalmente vindicada, entonces el entrelazamiento *podría* ser la respuesta a la pregunta "¿qué hace que seamos compañeros de mundo?". Nuevamente, siguiendo el espíritu de la metafísica naturalizada, se enfatiza el "podría" porque la tesis debería considerarse falible y, además, futuras investigaciones podrían revelar que esta cuestión acerca de la construcción del mundo es, después de todo, epistémicamente insegura para responder según los estándares de la metafísica naturalizada, o que la pregunta puede ser propensa a la domesticación.

En conexión con esto último, se plantean dos advertencias: (1) Aunque el entrelazamiento se ofrece como respuesta a una pregunta metafísica, esta respuesta se origina en un compromiso serio con una teoría física y, por esta razón, el entrelazamiento no posee ninguna significación más allá de su función en la unión de poliedros en la gravedad cuántica de lazos. Aunque se recurra a él en esta pregunta metafísica, se lo hace como un elemento de la teoría bajo escrutinio que es seleccionado como candidato para responder a nuestra indagación metafísica, y su uso en metafísica debe mantenerse fiel a este origen para cumplir con el estándar de la metafísica naturalizada. (2) Esfeld podría insistir en que el entrelazamiento no es una relación constructora del mundo satisfactoria, puesto que no parece individualizar objetos simples y no puede proporcionar subsistemas separables. Ahora bien, podría ser que aún no hayamos encontrado la relación constructora del mundo correcta en la gravedad cuántica de lazos, pero parece más probable que la pregunta metafísica —y, especialmente, nuestra expectativa acerca de la respuesta— deba adaptarse a lo que la teoría proporciona. En este sentido, responder a las preguntas metafísicas planteadas a las teorías científicas debe adoptar la forma de una negociación entre nuestras aspiraciones metafísicas y los detalles de la teoría. La teoría podría, en última

instancia, indicar que partes de la pregunta metafísica estaban mal planteadas, de modo que estos aspectos se explican en lugar de responderse. Sin embargo, se aprenden lecciones metafísicas en ambos casos, pues el conflicto que impulsa la investigación ha resultado útil para la metafísica naturalizada.

5. Conclusión

La introducción afirmaba que pocos hoy se alinearían con Parménides y doctrinas metafísicas en contra de los sentidos y, por extensión, de la ciencia. Esto es probablemente para bien, pero quizás una resistencia parmenídea del tipo adecuado puede seguir siendo beneficiosa. La historia de la ciencia ha estado llena de sorpresas metafísicas. A menudo, la ciencia ha demostrado explícitamente que nuestras preconcepciones metafísicas eran erróneas, siendo la mecánica cuántica un ejemplo. La ciencia, sin embargo, rara vez es igualmente explícita acerca de qué teoría metafísica debe ocupar el lugar de nuestras preconcepciones; la industria de interpretar la mecánica cuántica ilustra muy bien esta diferencia. Si seguimos el principio de la metafísica naturalizada, los metafísicos no deberían iniciar disputas con la ciencia por estas preconcepciones perdidas: la metafísica debe ceder ante la ciencia. Sin embargo, si simplemente descartamos a los metafísicos así, perdemos un lado de la contribución que la ciencia hace a la metafísica: mientras que la ciencia seguirá demostrando que nuestras preconcepciones metafísicas son erróneas, corremos el riesgo de continuar sin proponer nuevas metafísicas en su lugar si silenciamos las preguntas metafísicas que subyacen a los conflictos entre metafísica y ciencia. Se debería mantener una resistencia parmenídea hasta que la ciencia tenga una historia metafísica *positiva* alternativa que contar.

Tal resistencia se aconsejaba específicamente en circunstancias donde no existe un marco metafísico conocido que sea consistente con la teoría científica. Mientras que algunas teorías científicas simplemente

se alinean con un lado de un debate metafísico, otros conflictos —como el de los lewisianos contra las teorías no espaciales de la gravedad cuántica— pueden interpretarse como un conflicto entre ciencia y metafísica en su conjunto. Es este tipo de conflictos el que es valioso como indicador de dónde la ciencia solo ha ofrecido una respuesta negativa y donde, en consecuencia, hay más trabajo por hacer para la metafísica naturalizada. Solo formulando las preguntas metafísicas revelaremos las metafísicas hasta ahora inconcebibles con las que la ciencia reemplaza nuestras preconcepciones metafísicas. Un poco de resistencia parmenídea asegura que lo hagamos así, y la metafísica naturalizada debería, por lo tanto, acoger los conflictos entre metafísica y ciencia como un recurso para el desarrollo metafísico. La Sección 4 mostró cómo llevar esto a cabo en el contexto de la gravedad cuántica de lazos —una teoría no espacial de la gravedad cuántica— con énfasis en qué reemplaza a la distancia como la relación constructora del mundo en esa teoría. La propuesta fue el entrelazamiento; una respuesta alcanzada como un equilibrio entre los detalles de la teoría científica y las aspiraciones metafísicas de la pregunta.

REFERENCIAS

Allori, V. (2015). Primitive ontology in a nutshell. *International Journal of Quantum Foundations, 1*(3), 107–122. https://ijqf.org/archives/2394

Andersen, F., & Becker Arenhart, J. R. (2016). Metaphysics within Science: Against radical naturalism. *Metaphilosophy, 47*(2), 159–180. https://doi.org/10.1111/meta.12175

Anderson, E. (2012). Problem of time in quantum gravity. *Annalen der Physik, 524*(12), 757–786. https://doi.org/10.1002/andp.201200147

Baron, S. (2019). The curious case of spacetime emergence. *Philosophical Studies*, *177*(8), 2207–2226. https://doi.org/10.1007/s11098-019-01306-z

Baytaş, B., Bianchi, E., & Yokomizo, N. (2018). Gluing Polyhedra with entanglement in loop quantum gravity. *Physical Review D,* *98*(2), 026001. https://doi.org/10.1103/PhysRevD.98.026001

Bell, J. S. (2001). The theory of local beables. In M. Bell, K. Gottfried, & M. Veltman (Eds.), *John S. Bell on the foundations of quantum mechanics* (pp. 50–60). World Scientific. https://doi.org/10.1142/4757

Bennett, K. (2016). There is no special problem with metaphysics. *Philosophical Studies,* *173*(1), 21–37. https://doi.org/10.1007/s11098-014-0439-0

Bianchi, E. (2017). Spinfoam gravity. In A. Ashtekar & J. Pullin (Eds.), *Loop quantum gravity: The First 30 Years* (pp. 97–124). World Scientific. https://doi-org/10.1142/10445

Bianchi, E., Doná, P., & Speziale, S. (2011). Polyhedra in loop quantum gravity. *Physical Review D,* *83*(4), 044035. https://doi.org/10.1103/PhysRevD.83.044035

Bird, A. (2007). *Nature's metaphysics: Laws and properties*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199227013.001.0001

Blackburn, S. (2002). Metaphysics. In N. Bunnin & E. P. Tsui-James (Eds.), *The Blackwell companion to philosophy* (pp. 61–89). Blackwell Publishing. https://doi.org/10.1002/9780470996362.ch4

Bricmont, J. (2017). The de Broglie–Bohm theory as a rational completion of quantum mechanics. *Canadian Journal of Physics,* *96*(4), 379–390. https://doi.org/10.1139/cjp-2017-0192

Bryant, A. (2020). Keep the chickens cooped: the epistemic inadequacy of free range metaphysics. *Synthese, 197*(5), 1867–1887. https://doi.org/10.1007/s11229-017-1398-8

Carnap, R. (1950). Empiricism, semantics, and ontology. *Revue Internationale de Philosophie, 4*(2), 20–40.

Chalmers, D. (2009). Ontological anti-realism. In D. Chalmers, D. Manley, & R. Wasserman (Eds.), *Metametaphysics: New essays on the foundations of ontology* (pp. 77–129). Oxford University Press. https://doi.org/10.1093/oso/9780199546046.003.0003

Clifton, R., & Halvorson, H. (2001). Entanglement and open systems in algebraic quantum field theory. *Studies in History and Philosophy of Modern Physics, 32*(1), 1–31. https://doi.org/10.1016/S1355-2198(00)00033-2

Crowther, K., & Linnemann, N. (2017). Renormalizability, fundamentality, and a final theory: The role of UV-completion in the search for quantum gravity. *The British Journal for the Philosophy of Science, 70*(2), 377–406. https://doi.org/10.1093/bjps/axx052

Darby, G. (2009). Lewis's worldmate relation and the apparent failure of Humean supervenience. *Dialectica, 63*(2), 195–204. https://doi.org/10.1111/j.1746-8361.2009.01192.x

Dürr, D., Goldstein, S., & Zanghi, N. (2018). *Quantum motion on shape space and the gauge dependent emergence of dynamics and probability in absolute space and time* (pp. 1–27). https://arxiv.org/abs/1808.06844 (Hay una versión posterior, publicada en una revista revisada por pares: Dürr, D., Goldstein, S. & Zanghí, N. Quantum Motion on Shape Space and the Gauge Dependent Emergence of Dynamics and Probability in Absolute Space and Time. Journal of Statistical Physics, 180(1-6), 92–134 (2020). https://doi.org/10.1007/s10955-019-02362-9).

Esfeld, M. (2018). Metaphysics of science as naturalized metaphysics. In A. Barberousse, D. Bonnay, & M. Cozic (Eds.), *The*

philosophy of science: A companion (pp. 142–170). Oxford University Press. https://doi.org/10.1093/oso/9780190690649.003.0004

Esfeld, M. (2019). Against the disappearance of spacetime in quantum gravity. *Synthese*, *199*(Suppl 2), pp. 355–369. https://doi.org/10.1007/s11229-019-02168-y

Esfeld, M. (2020). A proposal for a minimalist ontology. *Synthese, 197*(5), 1889–1905. https://doi.org/10.1007/s11229-017-1426-8

Esfeld, M., & Deckert, D. A. (2017). *A minimalist ontology of the natural world.* Routledge. https://doi.org/10.4324/9781315142272

French, S. (2011). Metaphysical underdetermination: why worry? *Synthese, 180*(2), 205–221. https://doi.org/10.1007/s11229-009-9598-5

French, S. (2018). Toying with the Toolbox: How Metaphysics Can Still Make a Contribution. *Journal for General Philosophy of Science, 49*(2), 211–230. https://doi.org/10.1007/s10838-018-9401-8

French, S., & McKenzie, K. (2012). Thinking outside the toolbox: Towards a more productive engagement between metaphysics and philosophy of physics. *European Journal of Analytic Philosophy, 8*(1), 42–59. https://eujap.uniri.hr/casopis/abstracts/8_1_4.pdf

French, S., & McKenzie, K. (2016). Rethinking outside the toolbox: Reflecting again on the relationship between philosophy of science and metaphysics. In T. Bigaj & C. Wüthrich (Eds.), *Metaphysics in contemporary physics* (pp. 25–54). Brill. https://doi.org/10.1163/9789004310827_003

Goldstein, S., & Teufel, S. (2001). Quantum spacetime without observers: Ontological clarity and the conceptual foundations of quantum gravity. In C. Callender & N. Huggett (Eds.), *Physics meets philosophy at the Planck scale* (pp. 275–289).

Cambridge University Press. https://doi.org/10.1017/CBO9780511612909.013

Hagar, A., & Hemmo, M. (2013). The primacy of geometry. *Studies in History and Philosophy of Modern Physics Part B: Studies in History and Philosophy of Modern Physics, 44*(3), 357–364. https://doi.org/10.1016/j.shpsb.2013.01.003

Hedrich, R. (2007). The internal and external problems of string theory: A philosophical view. *Journal for General Philosophy of Science, 38*(2), 261–278. https://doi.org/10.1007/s10838-007-9048-3

Hofweber, T. (2016). *Ontology and the ambitions of metaphysics*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198769835.001.0001

Hudson, H. (2016). Non-naturalistic metaphysics. In K. J. Clark (Ed.), *The Blackwell companion to naturalism* (pp. 168–181). Wiley-Blackwell. https://doi.org/10.1002/9781118657775.ch12

Huggett, N., & Wüthrich, C. (2013). Emergent spacetime and empirical (in)coherence. *Studies in History and Philosophy of Modern Physics Part B: Studies in History and Philosophy of Modern Physics, 44*(3), 276–285. https://doi.org/10.1016/j.shpsb.2012.11.003

Humphreys, P. (2013). Scientific ontology and speculative ontology. In D. Ross, J. Ladyman, & H. Kincaid (Eds.), *Scientific metaphysics* (pp. 51–78). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199696499.003.0003

Inwagen, P. van. (2015). *Metaphysics*. Westview Press. https://doi.org/10.4324/9780429495021

Isham, C. J. (1993). Canonical quantum gravity and the problem of time. In L. A. Ibort & M. A. Rodríguez (Eds.), *Integrable systems, quantum groups, and quantum field theories* (pp. 157–287). Springer. https://doi.org/10.1007/978-94-011-1980-1_6

Jaksland, R. (2020). Entanglement as the world-making relation: Distance from entanglement. *Synthese*, *198*(10), 9661–9693. https://doi.org/10.1007/s11229-020-02671-7

Jones, R. (1991). Realism about what? *Philosophy of Science, 58*(2), 185–202. https://doi.org/10.1086/289611

Kraut, R. (2016). Three Carnaps on ontology. In S. Blatti & S. Lapointe (Eds.), *Ontology after Carnap*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199661985.003.0003

Ladyman, J. (2017). An apology for naturalized metaphysics. In M. Slater & Z. Yudell (Eds.), *Metaphysics and the Philosophy of Science: New Essays* (pp. 141–162). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199363209.003.0008

Ladyman, J., & Ross, D. (2007). *Every thing must go: Metaphysics naturalized.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199276196.001.0001

Lam, V. (2013). The entanglement structure of quantum field systems. *International Studies in the Philosophy of Science, 27*(1), 59–72. https://doi.org/10.1080/02698595.2013.783976

Lam, V. (2016). Quantum structure and spacetime. In T. Bigaj & C. Wüthrich (Eds.), *Metaphysics in contemporary physics* (pp. 81–99). Brill. https://doi.org/10.1163/9789004310827_005

Lam, V., & Esfeld, M. (2013). A dilemma for the emergence of spacetime in canonical quantum gravity. *Studies in History and Philosophy of Modern Physics Part B: Studies in History and Philosophy of Modern Physics, 44*(3), 286–293. https://doi.org/10.1016/j.shpsb.2012.03.003

Lam, V., & Wüthrich, C. (2018). Spacetime is as spacetime does. *Studies in History and Philosophy of Modern Physics Part B: Studies in History and Philosophy of Modern Physics*, *64*, 39–51. https://doi.org/10.1016/j.shpsb.2018.04.003

Lam, V., & Wüthrich, C. (2020). Spacetime functionalism from a realist perspective. *Synthese*, *199*(Suppl 2), 335–353. https://doi.org/10.1007/s11229-020-02642-y

Le Bihan, B. (2018). Space emergence in contemporary physics: Why we do not need fundamentality, layers of reality and emergence. *Disputatio, 10*(49), 71–95. https://doi.org/10.2478/disp-2018-0004

Le Bihan, B. (2021). Spacetime emergence in quantum gravity: Functionalism and the hard problem. *Synthese*, *199*(Suppl 2), 371–393. https://doi.org/10.1007/s11229-019-02449-6

Le Bihan, B., & Barton, A. (2018). Analytic Metaphysics Versus Naturalized Metaphysics: The Relevance of Applied Ontology. *Erkenntnis*, *86*(1), 21–37. https://doi.org/10.1007/s10670-018-0091-8

Lewis, D. (1986a). *On the plurality of worlds*. Blackwell.

Lewis, D. (1986b). *Philosophical papers* (Vol. II). Oxford University Press.

Lewis, D. (1994). Symposium: Chance and Credence: Humean Supervenience Debugged. *Mind, 103*(412), 473–490. https://doi.org/10.1093/mind/103.412.473

Loewer, B. (1996). Humean supervenience. *Philosophical Topics, 24*(1), 101–127. https://doi.org/10.5840/philtopics199624112

Lowe, E. J. (1998). *The possibility of metaphysics: Substance, identity, and time*. Oxford University Press. https://doi.org/10.1093/0199244995.001.0001

Maudlin, T. (2007a). Completeness, supervenience and ontology. *Journal of Physics A: Mathematical and Theoretical, 40*(12), 3151. https://doi-org/10.1088/1751-8113/40/12/S16

Maudlin, T. (2007b). *The metaphysics within physics*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199218219.001.0001

Miller, E. (2014). Quantum entanglement, Bohmian mechanics, and Humean supervenience. *Australasian Journal of Philosophy,*

92(3), 567–583. https://doi.org/10.1080/00048402.2013.832786

Morganti, M. (2016). Naturalism and realism in philosophy of science. In K. J. Clark (Ed.), *The Blackwell companion to naturalism* (pp. 75–90). Wiley-Blackwell. https://doi.org/10.1002/9781118657775.ch6

Ney, A. (2012). Neo-positivist metaphysics. *Philosophical Studies, 160*(1), 53–78. https://doi.org/10.1007/s11098-012-9912-9

Norton, J. D. (2020). Incubating a future metaphysics: Quantum gravity. *Synthese, 197*(5), 1961–1982. https://doi.org/10.1007/s11229-017-1473-1

Oriti, D. (Ed.). (2009). *Approaches to quantum gravity: Toward a new understanding of space, time and matter*. Cambridge University Press. https://doi.org/10.1017/CBO9780511575549

Paul, L. A. (2012). Metaphysics as modeling: The handmaiden's tale. *Philosophical Studies, 160*(1), 1–29. https://doi.org/10.1007/s11098-012-9906-7

Penrose, R. (1971). Angular momentum: An approach to combinatorial spacetime. In T. Bastin (Ed.), *Quantum theory and beyond* (pp. 151–180). Cambridge University Press.

Pitowsky, I. (2006). On the definition of equilibrium. *Studies in History and Philosophy of Modern Physics Part B: Studies in History and Philosophy of Modern Physics, 37*(3), 431–438. https://doi.org/10.1016/j.shpsb.2006.03.001

Redhead, M. (1995). More ado about nothing. *Foundations of Physics, 25*(1), 123–137. https://doi.org/10.1007/BF02054660

Ross, D. (2016). Vikings or Normans? The Radicalism of Naturalized Metaphysics. *Metaphysica, 17*(2), 213–227. https://doi.org/10.1515/mp-2016-0018

Rovelli, C. (2008). Loop quantum gravity. *Living Reviews in Relativity, 11*(5), 1–69. https://doi.org/10.12942/lrr-2008-5

Rovelli, C. (2011). Loop quantum gravity: The first 25 years. *Classical and Quantum Gravity, 28*(15), 153002. https://doi-org/10.1088/0264-9381/28/15/153002

Saunders, S. (2006). Are quantum particles objects? *Analysis, 66*(289), 52–63. https://doi.org/10.1093/analys/66.1.52

Swanson, N. (2020). How to be a relativistic spacetime state realist. *The British Journal for the Philosophy of Science, 71*(3), 933–957. https://doi.org/10.1093/bjps/axy041

Thomasson, A. L. (2014). *Ontology made easy*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199385119.001.0001

Vassallo, A., & Esfeld, M. (2014). A proposal for a Bohmian ontology of quantum gravity. *Foundations of Physics, 44*(1), 1–18. https://doi.org/10.1007/s10701-013-9745-1

Weatherson, B. (2015). Humean supervenience. In B. Loewer & J. Schaffer (Eds.), *A companion to David Lewis* (pp. 99–115). Wiley-Blackwell. https://doi.org/10.1002/9781118398593.ch8

Wüthrich, C. (2017). Raiders of the lost spacetime. In D. Lehmkuhl, G. Schiemann, & E. Scholz (Eds.), *Towards a theory of spacetime theories* (pp. 297–335). Springer. https://doi.org/10.1007/978-1-4939-3210-8_11

Wüthrich, C. (2019). When the actual world is not even possible. In D. Glick, G. Darby, & A. Marmodoro (Eds.), *The Foundation of Reality: Fundamentality, Space, and Time* (pp. 233–254). Oxford University Press. https://doi.org/10.1093/oso/9780198831501.003.0014

Yablo, S. (1998). Does Ontology Rest on a Mistake? *Aristotelian Society Supplementary Volume, 72*(1), 229–261. https://doi.org/10.1111/1467-8349.00044

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 355-378
ISSN 3087-260X
Santiago de Chile

SCANLON SOBRE LA ONTOLOGÍA DE LAS RELACIONES DE RAZÓN[1]

Por John Skorupski[2]
Traducción por Cristóbal Andrade[3]

[1] NT: Traducción al español de "*Scanlon on the ontology of reason relations*" escrito por John Skorupski. Este texto se encuentra inédito en su idioma original; en una correspondencia personal (19/05/2025), el profesor Skorupski manifestó que no tenía planes de publicación, debido a su extensión y el presumible escaso interés de las revistas anglosajonas en virtud de su insistencia en tesis propuestas en su obra *The Domain of Reasons* (2010). No obstante lo anterior, en la misma comunicación autorizó expresamente su traducción y posterior publicación en español. La presente traducción se publica con su conocimiento y consentimiento, y tiene por objeto poner a disposición del público hispanohablante un texto central para la discusión metanormativa contemporánea en torno a la naturaleza de las razones normativas. Esta edición presenta una serie de modificaciones del original: las notas a pie de página que comienzan con "NT" (nota del traductor) y las notas sobre el autor y el traductor. A su vez, se modifica el sistema de referencias y citas al estilo APA 7.

[2] Profesor emérito de la cátedra de Filosofía Moral en la *School of Philosophical, Anthropological and Film Studies* de la *University of St. Andrews*, en Escocia, UK. Es *Fellow of the British Academy* (FBA) y *Fellow of the Royal Society of Edinburgh* (FRSE). Contacto: jms2@st-andrews.ac.uk.

[3] Cristobal Andrade es profesor de Filosofía por la Universidad de Valparaíso, Chile. Email: cristobal.andrade@estudiantes.uv.cl. ORCID: https://orcid.org/0009-0002-7660-1745.

§*Being Realistic about Reasons* de T. M. Scanlon (2014)[5] presenta una concepción general de las razones y la normatividad la cual ofrece, en particular en el capítulo 2, una concepción notable acerca de en qué consiste que los objetos normativos existan. Esta concepción es la que me gustaría discutir aquí.[6]

En su capítulo introductorio, Scanlon señala dos maneras en las cuales el debate meta-ético ha cambiado. Hasta finales de los años 70, cuando John Mackie publicó *Ethics: Inventing Right and Wrong* (1977), el foco estaba extendidamente centrado en la moralidad. A partir de entonces se ha ampliado a la noción general de la normatividad:

> Aunque la moralidad aún se mantiene muy discutida, una parte significativa del debate concierne al razonamiento práctico y a la normatividad más en general: las razones para la acción e incluso más ampliamente, las razones para la creencia y otras actitudes, las cuales son cada vez más reconocidas como normativas… (Scanlon, 2014, p. 1)

La segunda diferencia, dice él, es que el foco estaba en ese entonces en la motivación, especialmente en la pregunta de cómo los juicios morales podrían motivar (si lo hiciesen), mientras que ahora, aunque la importancia de las preguntas sobre la motivación no es disputada, el foco se ha movido hacia la idea de una *razón normativa* —para la acción, creencia u "otras actitudes". Scanlon observa más bien modestamente que esta afirmación sobre la meta-ética como un todo puede ser

---

[5] Mis agradecimientos a los miembros de la audiencia en la Conferencia de Metaética Rio-2015 quienes comentaron este artículo; especialmente a mi ponente de réplica, Rodrigo Gouvea.
Scanlon (2014) es una versión revisada de las *Locke lectures* las cuales Scanlon presentó en Oxford en 2009.
[6] NT: Para una comprensión cabal de este debate, es recomendable revisar trabajos como los de Blackburn (1984, 1993, 1998), Dummett (1996), Gibbard (1990, 2003) y Quine (1961).

tendenciosa, puesto que ambos virajes pueden ser virajes sólo en su propio pensamiento. Sin embargo, creo que la mayoría de las personas interesadas en el tema estarían de acuerdo en que estos dos virajes en el debate han sido, sin lugar a dudas, bastante generales, aunque de ningún modo universales. La ampliación de la atención ha tenido diversas fuentes, siendo no menos importante el desarrollo natural del tema; pero el pensamiento de Scanlon ha sido incuestionablemente una influencia muy importante en esto. Ahora pensamos cada vez más ampliamente en términos de meta-normatividad, antes que restringidamente en términos de solo meta-ética, y la idea de que el concepto de una razón es el concepto normativo clave está muy presente en el ambiente.

Las conferencias de Scanlon reflejan ambas tendencias. Se describe a sí mismo como un fundamentalista de razones. Este es el nombre que él le da a la visión

> de que las verdades sobre razones no son reducibles o identificables con verdades no-normativas, tales como verdades sobre el mundo natural de los objetos, causas y efectos físicos, ni tampoco pueden ser explicadas en términos de nociones de racionalidad o agencia racional que no sean en sí mismas afirmaciones sobre razones. (Scanlon, 2014, p. 2)

Señala que las razones podrían ser fundamentales en un sentido ulterior, a saber, que los otros conceptos normativos podrían ser todos analizables en términos de la noción de una razón. Esta es una visión hacia la cual él dice que está inclinado, pero la reconoce como controvertida y no propone defenderla en el libro (aunque luego la adopta como una "hipótesis de trabajo" (Scanlon, 2014, p. 34; véase también pp. 25–26)).

Estoy de acuerdo con el Fundamentalismo de Razones de Scanlon. También considero la hipótesis de que todos los conceptos normativos son analizables en términos de relaciones de razón altamente

plausible[7]. Es verdad, necesitamos más para establecer que el concepto de una relación de razón es el concepto normativo *fundamental*. Lo que se necesita es la tesis negativa de que el círculo de conceptos normativos no puede ser analizado igualmente bien en términos de algún concepto normativo distinto de aquel de una relación de razón, por ejemplo el concepto *bueno*. Puesto que si algo como eso puede ser realizado, resultaría que hay más de una manera de analizar algunos conceptos normativos en términos de otros; ninguno de ellos sería descriptible, por lo tanto, como fundamental. Yo mismo soy escéptico frente a si hay algún análisis de la normatividad que sea tan bueno o mejor que aquel en términos de razones. Sin embargo, no necesitamos perseguir el asunto aquí. La discusión ontológica de Scanlon —la cual él en efecto considera que es una discusión del dominio normativo en general— es conducida en términos de relaciones de razón y lo seguiré en ello. Pero el análisis funcionaría, si funciona, independientemente de lo que uno considere que son las propiedades normativas.

El principal objetivo de la discusión de Scanlon es ofrecer "una defensa calificada de un cognitivismo realista sobre las razones" (2014, p. 2). Por 'cognitivismo' Scanlon entiende la visión familiar de que las proposiciones sobre razones pueden ser verdaderas o falsas y, como creo que él aceptaría, verdaderas o falsas en un sentido genérico, perfectamente estándar, de 'verdadero' y 'falso' —no uno que sea especial para el caso normativo. 'Realista' lo usa de una manera más idiosincrática (Scanlon, 2014, p. 2). Su versión del cognitivismo es 'realista' en el sentido de que acepta que algunas afirmaciones normativas pueden no tener valores de verdad determinados —esto en contra de lo que Scanlon presumiblemente consideraría que es la insistencia poco realista de que cada afirmación normativa debe tener un valor de verdad determinado. Pienso que se refiere a que tal insistencia sería poco realista en un sentido familiar no-filosófico de la palabra. Claramente entonces, ser 'realista' en este sentido no implica un abrazo del *realismo normativo*, en cualquiera

[7] He argumentado que todos pueden ser analizados de esa manera en Skorupski (2010).

de los sentidos filosóficos comunes en los cuales el realismo sobre lo normativo es opuesto al anti-realismo, irrealismo, o cuasi-realismo. Por el contrario, me parece que Scanlon rechaza aquellas formas de realismo, mientras también rechaza el anti-realismo, irrealismo o cuasi-realismo. Como yo lo entiendo, él rechaza ambas alas de este debate como inapropiadamente metafísicas. Permitámonos perseguir esto un poco más allá.

§Como múltiples otros escritores sobre la normatividad, Scanlon considera que las razones son hechos los cuales están en una *relación de razón* con una persona y un tipo de respuesta de esa persona. Distingue además entre el hecho o hechos que constituyen una razón y las circunstancias en virtud de las cuales ellos constituyen una razón. (Por lo que uno podría decir, por ejemplo, que el hecho de que el edificio esté en llamas es una *razón* para que tú entres —dada la *circunstancia* de que eres un bombero quien ha sido puesto a cargo de salvar tantos edificios incendiándose como sea posible). Entonces

> ''Es una razón para'' es una relación de cuatro posiciones, R (*p*, *x*, *c*, *a*), que se sostiene entre un hecho, *p*, un agente, *x*, un conjunto de condiciones *c*, y una acción o actitud *a*. Esta es la relación que se sostiene sólo en caso de que *p* es una razón para una persona *x* en la situación *c* para hacer o sostener *a*. (Scanlon, 2014, p. 31)

Bajo esta concepción, como él señala, las razones, esto es, los hechos que son razones, no plantean ninguna pregunta ontológica especial que la noción de un hecho en general no plantee. En cuanto a esto, Scanlon entiende por hechos ''la reflexión de pensamientos verdaderos''; se distancia a sí mismo de una noción de correspondencia de un hecho como un *verificador* (*truth-maker*) para una proposición[8]. Su noción de hecho

[8] ''La idea de que los juicios normativos son correctos cuando se corresponden con los hechos normativos no es ninguna explicación de si estos 'hechos' son, como he sugerido,

es efectivamente la noción fregeana. Las razones son hechos en ese sentido especial de 'hecho' en el cual los hechos *son* verdades, contrario a lo que "hace" verdaderas a las proposiciones.

Esto aún mantiene abierta la pregunta del estatus ontológico de las relaciones de razón, contrario al de las razones. La respuesta de Scanlon en la Conferencia 2, aunque no completamente nueva, es desarrollada de manera innovadora. La idea básica es lo que él llama la "visión centrada en el dominio" (*domain-centered*) de existencia (Scanlon, 2014, p. 21). De acuerdo con esta, las preguntas acerca de lo que existe deberían ser pensadas como internas a los dominios de discurso tales como "las matemáticas, la ciencia, y el razonamiento práctico y moral" (Scanlon, 2014, p. 19). No hay

> condiciones generales, independientes al dominio, de 'existencia' tal que diversas afirmaciones existenciales hechas en cada dominio impliquen o presupongan que las entidades del tipo al que refieren satisfagan estas condiciones […] lo único común a estas afirmaciones existenciales a través de todos los dominios es la lógica puramente formal del cuantificador existencial […] las condiciones requeridas para que los objetos en diferentes dominios existan varían [*sic*] de dominio a dominio. (Scanlon, 2014, p. 25; nota al pie 12, p. 25)

En su conjunto, entonces, y a pesar de la elasticidad del término 'realismo' en los debates meta-normativos actuales, parecería incorrecto clasificar a Scanlon como un realista normativo. Ese término, creo, es mayormente usado con frecuencia para denotar la visión no sólo de que las afirmaciones normativas expresan proposiciones, sino que aquellas proposiciones son verdaderas sólo si tienen verificadores que las hagan verdaderas y, finalmente, que algunas de ellas son verdaderas. Dejemos la

---

meramente "la reflexión de pensamientos verdaderos" (Scanlon, 2014, p. 31). Véase también Scanlon (2014, nota al pie 9, p. 24; p. 62).

última tesis, puesto que, desde el punto de vista meta-normativo, añade una afirmación que debería ser tratada separadamente. El realismo sobre lo normativo, como usaré el término en esta conferencia, es simplemente la tesis de que las proposiciones normativas caen bajo una concepción de la verdad por correspondencia —dejando abierta la pregunta de si alguna de ellas es verdadera. Los típicos teóricos del error sobre lo normativo son, en este sentido, realistas: sostienen que las afirmaciones normativas caen efectivamente bajo una concepción de la verdad por correspondencia, pero creen que no son nunca verdaderas, porque no hay hechos que las hagan verdaderas. En contraste, Scanlon no es un realista en este sentido; como se acaba de señalar, él explícitamente niega que las verdades normativas deban tener verificadores. Él es un cognitivista, pero no es un realista.

§Scanlon también distancia su visión del expresivismo. Aquí, sin embargo, el punto de diferencia queda un tanto oscuro. Señala que los expresivistas como Allan Gibbard y Simon Blackburn dicen que su posición les permite aceptar que hay verdades normativas y, por lo tanto, también que hay hechos normativos, en el sentido fregeano de hecho. Y, por supuesto, ellos también estarán de acuerdo en que estas verdades normativas no tienen verificadores.

No es de hecho posible, me parece, establecer una diferencia entre el cognitivismo de Scanlon y el expresivismo complejo de Gibbard o el 'cuasi-realismo' de Blackburn sin desafiar directamente el derecho de los expresivistas a reclamar una noción (genérica, estándar) de la verdad para el caso de las proposiciones normativas —o, efectivamente, su derecho a reclamar la noción de una *proposición* normativa. Esto es algo que Scanlon no intenta, tal vez porque persigue el acuerdo antes que el desacuerdo. Como resultado, sin embargo, su rechazo del expresivismo parece ligeramente arbitrario.

Así que consideremos el problema un poco más de cerca. Tomemos, primero, la afirmación de que una expresión de la oración "el placer es bueno" en algunos contextos C constituye un acto de habla de

decir que el placer es bueno. Podemos asumir que el expresivista tendrá una concepción sobre cuándo es apropiado realizar tal acto de habla. Su siguiente afirmación es que una expresión de la oración ‘‘Es verdadero que el placer es bueno’’, en algunos contextos C, es una realización del mismo acto de habla. Llamemos a esta la afirmación de la identidad del acto de habla. Por supuesto que hay maneras obvias en las cuales los dos actos de habla no son exactamente el mismo —diferentes palabras son usadas. Pero concedamos, para los fines del argumento, que tenemos algún entendimiento de cómo esta afirmación acerca de la identidad del acto de habla es pretendida.

Contrastemos ahora la afirmación de que la proposición de que el placer es bueno es la misma proposición que la proposición de que es verdadero que el placer es bueno. Llamemos a esta identidad afirmada la identidad proposicional. Ella presupone una visión de la verdad, la cual uno podría llamar la visión simple, o la visión de la ‘‘no teoría’’, o (con algunas reservas) la visión minimalista. La visión simple tendrá algo que decir acerca de las ocurrencias de ‘‘es verdad’’ como predicado, así como de las ocurrencias de ‘‘es verdadero que’’ como un operador, pero no necesitamos perseguir eso aquí. De todas formas, algunos cognitivistas, entre los cuales yo soy uno, están de acuerdo con esta visión simple de la verdad. ¿Cómo difieren, entonces, estos cognitivistas de los expresivistas? La diferencia está en el orden de explicación. Para el cognitivista simple, la afirmación sobre la identidad del acto de habla, en el sentido en el que es verdadera, es verdadera en virtud de la identidad proposicional. Los actos de habla son idénticos en el sentido de que tienen el mismo contenido, esto es, son aserciones de la misma proposición. Esa explicación, no obstante, no puede funcionar para el expresivista, ya que el punto central del expresivismo respecto a expresiones tales como ‘‘el placer es bueno’’ es que, sea o no que pueden inocuamente, al final del análisis, decirse que son cuasi-aserciones de cuasi-proposiciones, las cuales pueden ser cuasi-verdaderas o cuasi-falsas, es imposible decir que son aserciones de proposiciones en algún sentido antecedente de proposición que un expresivista pueda dar por supuesto

como aplicable *ab initio* al discurso normativo. Aceptar, *ab initio*, que las afirmaciones normativas son aserciones de proposiciones eliminaría cualquier margen para una semántica expresivista distintiva para las oraciones normativas.

Por lo que el cognitivista presupone la noción de la *misma proposición* al entender la noción relevante del *mismo acto de habla*; él puede argumentar además que al entender la noción relevante de la identidad del acto de habla estamos implícitamente presuponiendo una comprensión de la proposición que el acto de habla asevera, y bajo la teoría simple de la verdad siendo aplicada a proposiciones. El expresivista tendrá que hacerlo hacia el otro lado, bajo la amenaza de desaparecer por completo. Tendrá que apelar a la identidad del acto de habla como una manera de construir la identidad proposicional, o cuasi-proposicional, y la verdad cuasi-proposicional, para el caso normativo. Si esto puede hacerse, sin, en algún paso u otro, cometer petición de principio mediante apelar implícitamente a una noción *ya comprendida* de la identidad proposicional, es cuestionable.

En contraste, los cognitivistas consideran evidente que las aserciones normativas son aserciones como cualquier otra. Pueden estar contentos con la semántica de condiciones de verdad, o cualquier forma de semántica para el discurso normativo que (i) sea uniformemente aplicada al discurso asertórico en general, así como al discurso normativo, y (ii) capture el punto básico de que la aserción es aserción de contenido judicativo (por usar la expresión fregeana). En la medida en que estos dos puntos sean respetados, la semántica no es, para el cognitivista, un gran problema.

Permítaseme llamar a esta posición *cognitivismo simple*. Pero antes de proceder, debería señalar la visión común de que el minimalismo sobre la verdad no es compatible con una concepción del significado basada en condiciones de verdad. Si esa visión fuese correcta, un cognitivista simple tendría que dar una concepción del significado que no fuera basada en condiciones de verdad. Tal vez sería una teoría del uso. Fuera cual fuera la teoría, el cognitivismo simple requeriría que se

preservara el punto de que las aserciones normativas expresan contenidos judicativos. Sin embargo, yo mismo no estoy de acuerdo con que esta visión común sea la correcta; me parece que adscribe un rol excesivamente ambicioso a la semántica basada en condiciones de verdad. Aquella semántica no debería ser vista, bajo mi visión, como ocupándose de elucidar filosóficamente qué *es* el significado o, correlativamente, el entendimiento. Tú puedes emplear la semántica de condiciones de verdad para un lenguaje particular sin tener una visión filosófica de la naturaleza del significado y del entendimiento. Si tu semántica funciona, proporciona una concepción sistemática de un lenguaje particular que tiene la propiedad de que conocerla, junto con conocer que es verdadera únicamente en virtud de las convenciones del lenguaje en cuestión, basta para comprender ese lenguaje.

Desde luego, hay también otra pregunta, y una buena, acerca de si lo que Michael Dummett ha llamado una concepción "robusta" (*full-blooded*) del entendimiento, o aprehensión de conceptos, debería descansar en la noción de verdad, o si debería más bien descansar en caracterizar conceptos en términos de normas para introducirlos y eliminarlos en el pensamiento. Puede ser que el minimalismo sobre la verdad lleve al cognitivista a la última opción. Pero mi punto presente es que eso, de ser verdadero, de ninguna manera impide que el cognitivista simple hable directamente sobre las condiciones de verdad para las oraciones normativas. La visión de la no-teoría de la verdad no es la visión de que no hay tal cosa como la verdad.

§Puesto que Scanlon no discute estos problemas sobre la verdad y el significado en sus conferencias, no estoy seguro de que pueda ser descrito como un cognitivista simple. Por otro lado, estoy bastante seguro de que eso es lo que debería ser. Esta posición es distinta del expresivismo, pero al mismo tiempo no es realista en el sentido meta-normativo previamente descrito. Y Scanlon no es ni un expresivista ni un realista en ese sentido.

Sería un paso ulterior describirlo como un irrealista, no obstante. Ahora, en vistas de lo que he dicho hasta ahora, esto puede parecer

meramente como una cuestión terminológica. Tal vez lo es. Sin embargo, me parece que necesitamos indagar más a fondo antes de decidir si lo es o no, porque los problemas son bastante esquivos.

Al comienzo de la Conferencia 2, la cual lidia con las objeciones "metafísicas" a su posición, Scanlon establece lo que dice que es "natural describir… como una objeción ontológica" a su forma de cognitivismo sobre las razones, a saber, "que la idea de que hay verdades normativas irreducibles tiene implicaciones que son incompatibles con visiones plausibles acerca de "lo que hay" (Scanlon, 2014, p. 17). Agrega inmediatamente que él mismo no cree que el "supuesto problema aquí sea uno de ontología", a pesar de que "sería útil considerar el asunto primero en esta forma ontológica". Luego introduce el criterio de Quine de "compromiso ontológico" —pero es notablemente cauteloso en su uso de esa frase. Claramente tiene alguna reserva, tal vez acerca de una ontologización excesivamente simplificada.

Al considerar el criterio de Quine, será útil distinguir dos tesis. La primera es que si tú refieres a un objeto *a*, entonces para algún *x*, *x* = a y estás refiriendo a *x*. Esto es simplemente una instancia del principio frecuentemente llamado generalización existencial —pero para no incurrir en petición de principio, seguiré a Graham Priest (2005) y referiré al cuantificador como el cuantificador particular y al principio como generalización particular. La segunda es la tesis de que cualquier objeto sobre el cual un cuantificador opera existe. Llamemos a este *el criterio cuantificacional de existencia*. Poniendo el criterio cuantificacional en combinación con la generalización particular, entonces da como resultado una idea antigua sobre la existencia, a saber, que a lo que sea que podamos referir en el pensamiento o discurso existe. Puesto que la generalización particular es un truismo lógico —si refieres a un objeto, refieres a algo—, la única manera de resistir esa conclusión es rechazar el criterio cuantificacional de existencia.

Hay otra idea antigua acerca de la existencia: que existir es tener estatus o posicionamiento causal. Llamemos a este *el criterio causal de la existencia*. Permítaseme enfatizar que en sí mismo este criterio no dice

nada sobre la naturaleza de la causación; es compatible con un amplio rango de visiones acerca de lo que la causación y la dependencia causal es. Como es familiar, no obstante, intentar combinar estos dos criterios hace surgir dificultades, tanto en meta-ética como en filosofía de la matemática. Ni las relaciones de razón ni los conjuntos parecen tener ningún tipo de posicionamiento causal, bajo cualquier interpretación razonable de posicionamiento causal.

Scanlon está comprometido con rechazar el principio causal como una condición general de la existencia, puesto que él niega que *haya* alguna condición general como tal: "lo único común a las afirmaciones existenciales a través de todos los dominios es la lógica puramente formal del cuantificador existencial". En contraste, acepta el criterio cuantificacional —pero lo acepta como un criterio para la existencia *dentro de un dominio*. Así añade que

> Nuestros compromisos ontológicos en este sentido general [el sentido capturado por el criterio cuantificacional] no representan una afirmación por nuestra parte acerca de lo que *el mundo* contiene, en cualquier sentido significativo de "el mundo" (Scanlon, 2014, p. 24).

Preguntar lo que "realmente hay", "en el mundo como tal", es, en su visión, hacer una pregunta sobre una noción vacía. Por lo demás, formar un concepto general disyuntivo de existencia a través de los dominios es, piensa él, una maniobra sin sentido. ¿Cuán seriamente "ontológica" es, entonces —si puedo ponerlo de esta manera— su aceptación del criterio cuantificacional?

§Para examinar la pregunta podemos consultar si hay un dominio de las ficciones, en el sentido de Scanlon de 'dominio'. La "lógica puramente formal" del cuantificador se aplica: podemos considerar, por ejemplo, la proposición de que no hay un detective ficticio más grande que Sherlock Holmes, e inferir que si es verdad entonces Sherlock Holmes es al menos

tan grande como Hércules Poirot. Podemos preguntar cuántos detectives ficticios ha habido. Así parecemos estar comprometidos, por el criterio de Scanlon, con la existencia de los objetos ficticios —por supuesto que en el dominio de las ficciones. Pues, si estamos inclinados a intentar parafrasear estas referencias aparentes a los detectives ficticios, eso implica que sí tenemos en mente, después de todo, alguna condición trans-dominio sobre la existencia diferente a, o más rica que, el criterio cuantificacional —tal vez el criterio causal. Si el criterio cuantificacional es *todo* lo que está en juego, no hay fundamento para cuestionar nuestro ''compromiso ontológico'' con los objetos ficticios.

Aunque ciertamente decimos cosas como

(1) Hay muchos personajes en *Guerra y Paz*, algunos los cuales (realmente) existieron y otros que no.

En esta oración parece *prima facie* obvio que 'existe' o 'existe realmente' está usado como un predicado ontológico, mientras el verbo 'haber' (*to be*) está usado en un sentido de cuantificador ontológicamente no-comprometido. Pero de acuerdo con el criterio cuantificacional, si cuantificamos sobre personajes en *Guerra y Paz*, estamos comprometidos con su existencia. Por lo que, si tomamos el ejemplo al pie de la letra, deberíamos rechazar el criterio cuantificacional.

En el ejemplo, he usado el verbo 'haber' para expresar el cuantificador particular. En el discurso ordinario, podemos usar ambos verbos 'existir' y 'haber' en aquel rol cuantificacional. Sería, por ejemplo, perfectamente inteligible decir

(2) Existen personajes en *Guerra y Paz* que no son reales y personajes que lo son.

O podemos no usar ninguno: podemos decir

(3) Algunos de estos personajes en *Guerra y Paz* realmente existieron y algunos no.

En contraste, 'existir' antes que 'haber' es el verbo más naturalmente usado para expresar un predicado ontológico, como en "El flogisto no existe". Hay alguna rareza, o implicación de especial significatividad, en usar 'haber' en una posición de predicado, como en "Dios lo hay", o "No hace nada, sólo lo hay".

El criterio cuantificacional de existencia obtiene alguna plausibilidad ilegítima del hecho de que la palabra 'existe' puede ser *doblemente* usada para expresar el cuantificador particular *y* en el rol de predicado ontológico. Por lo que en lo que sigue usaré 'existir', y solo 'existir', como el predicado ontológico y no lo usaré en absoluto en el rol cuantificacional.

Si podemos referir a lo no-existente, un predicado ontológico se necesita. Si asevero que el flogisto no existe, parezco referir al flogisto. Luego, por generalización particular, refiero a algún *x* tal que *x* = flogisto y *x* no existe.

Ahora, Scanlon no deja espacio para el uso irreduciblemente predicativo, no-cuantificacional, de 'existe'. Considera que 'existe' expresa un cuantificador, aunque, por supuesto, un cuantificador enfocado en el dominio. Presumiblemente, él podría decir que la oración de *Guerra y Paz* significa que algún *x* en el dominio de las ficciones es un personaje en esa novela y algún *x* en el dominio de las causas es un personaje en esa novela. ¿Qué sucede entonces con el enunciado de que el flogisto no existe? Presumiblemente, esto dice que algún *x* en el dominio de las entidades postuladas es flogisto, pero ningún *x* en el dominio de las causas es flogisto.

A pesar del elemento que este análisis tiene en común con el criterio de Quine del compromiso ontológico, me parece que sería mejor presentado como *rechazando* el criterio de Quine —en el sentido de rechazarlo como *vacío*: como pretendiendo decir algo donde no hay nada que decir. Una visión así podría ser puesta de esta manera. Sabemos cómo

establecer verdades que contienen cuantificadores en sus variados dominios. Hay detectives ficticios en el dominio de las ficciones, electrones en el dominio de las causas, y relaciones de razón en el dominio normativo. Todos estos juegos del lenguaje son jugados. Hablar de estas verdades como revelando nuestros "compromisos ontológicos" en variados dominios parece decir algo más, pero en realidad no hace más que repetir, de manera engañosa, que estas verdades son verdades.

Scanlon parece querer evitar llegar a esta visión radicalmente anti-metafísica. Rechaza la ontología cuando es entendida como una "idea dominio-trascendente como la idea de Quine del compromiso ontológico" —"estoy rechazando [dice él] esta idea general de existencia y argumentando que las preguntas ontológicas genuinas son todas dominio-específicas" (Scanlon, 2014, nota al pie 9, p. 24) . Pero, añade, "[d]ecir esto no es negar que hay importantes e interesantes preguntas metafísicas y ontológicas. Sólo es decir que estas preguntas son dominio-específicas (Scanlon, 2014, p. 25). Sin embargo, las preguntas que Scanlon tiene en mente no son, me parece, preguntas metafísicas. Consideremos las "preguntas ontológicas genuinas" primero. Puedes preguntar si las partículas de Higgs realmente existen, o si los detectives ficticios más grandes que Sherlock Holmes realmente existen. La primera pregunta es ontológica; bien podrías dudar si la segunda lo es —pero dada su visión de la existencia centrada en el dominio no queda claro que Scanlon debiera dudarlo. Sin embargo, creo que estaría de acuerdo con que ninguna es una pregunta metafísica. Si pueden ser resueltas son resueltas, respectivamente, por los métodos de la física y la discusión literaria sobre la ficción de detectives. Lo mismo aplica si preguntas si las relaciones de razón realmente existen. La respuesta es que sí, si es que hay razones. Y es la discusión normativa de primer orden la que resuelve si hay razones. Esto, como digo, es algo con lo que Scanlon estaría de acuerdo.

La posición es entonces que no hay ninguna pregunta distintivamente *metafísica* que hacer sobre la existencia o sobre lo que existe. Esto es reminiscente de las visiones de Carnap sobre las preguntas

internas y externas, como Scanlon señala (Scanlon, 2014, nota al pie 3, p. 19; p. 23). ¿Qué hay sobre las otras preguntas que Scanlon tiene en mente como "preguntas metafísicas" dominio-específicas? Me parece que son preguntas sobre la sintaxis lógica del dominio —sobre cómo regimentar de la mejor manera su discurso. Por lo que, por ejemplo, la pregunta de si deberíamos distinguir los hechos que son razones de otros hechos que las constituyen como razones me parece ser una pregunta sobre cuál es la manera más útil para regimentar el discurso sobre las razones —la respuesta puede variar dependiendo de la investigación. No considero útil llamar a tales preguntas metafísicas, pero acepto que esto sí es verdaderamente una cuestión terminológica.

A la luz de esta discusión, pienso que podemos concluir con Scanlon que sería simplemente engañoso tanto llamar a su visión irrealista como llamarla realista (en el sentido estipulado arriba). Puesto que estas etiquetas parecen presuponer que hay alguna pregunta metafísica, en el sentido de "trascendiendo al dominio", "externa", sobre la ontología de las relaciones de razón. Y Scanlon no piensa que la haya.

§La visión de Scanlon es elegantemente mínima. Pero no me parece que capte todo lo que pensamos. A mí me parece plausible que el ejemplo de *Guerra y Paz*, como en (1), utiliza un uso predicativo de 'existir' el cual es (i) ontológico en un sentido dominio-trascendente y (ii) privilegia el dominio de las causas: existir en la "vida real", "realmente existir", es tener un posicionamiento causal (*causal standing*). Eso es lo que Napoleón tiene y Pierre no. *Hay* una cosa tal como el mundo; es el dominio de las causas, y Napoleón está en él mientras Pierre no. Sin embargo, hay verdades sobre Pierre, tal como hay verdades sobre Napoleón. Puesto que podemos referir a lo no-existente, las referencias a Pierre no necesitan ser parafraseadas fuera de estas verdades para que sean consideradas como verdades.

Esto es el irrealismo cognitivo sobre los objetos ficticios. Rechaza el criterio cuantificacional de existencia. Y añadiré que acepta el criterio

causal de la existencia, aunque esta tesis ulterior exige mayor discusión de la que puedo dar aquí.

§La pregunta que podemos levantar ahora es si un cognitivismo irrealista similar es apropiado para los objetos normativos —específicamente, para las relaciones de razón. Yo creo que lo es. Pero debería decir de inmediato que eso no implica que las relaciones de razón sean objetos ficticios o postulados teóricos inexistentes como el flogisto. Tomar la primera visión es ser un ficcionalista, tomar la segunda es ser un teórico del error. Es importante ver que estas no son las únicas opciones. La visión que quiero defender es, de hecho, mucho más cercana a la de Scanlon que a estas, a pesar de las diferencias que han sido discutidas en este artículo.

El punto clave es que los enunciados sobre objetos ficticios y reales putativos pero inexistentes tienen condiciones de verdad dependientes de la mente, mientras que los enunciados sobre relaciones de razón no. Pero antes de desarrollarlo, necesito hacer un punto sobre el principio causal.

El principio causal es importante en diversos contextos. En particular, sin embargo, tiene una implicación para el conocimiento. Entraña que es perfectamente legítimo preguntar, de cualquier existente independiente de la mente conocido, mediante qué modo de receptividad, en términos de Kant, uno lo conoce. Es legítimo porque, puesto que la existencia es posicionamiento causal, el conocimiento de existentes independientes de la mente debe, en algún punto, involucrar una transmisión causal que termina en un receptor localizado como conocedor. Como ya he dicho, al utilizar aquí la noción de causa no estoy dirimiendo entre visiones de lo que puede estar involucrado en que A produzca B, y no estoy, en particular, asumiendo que lo que sea que es deba ser reducible en términos naturalistas. Incluso así, el punto esencial se mantiene: el conocimiento de existentes independientes de la mente debe involucrar alguna forma de receptividad conectando dos existentes distintos, como sea que la receptividad opere.

Supongamos que somos persuadidos, por ejemplo, por medio de la discusión de Scanlon sobre la epistemología de lo normativo, que el conocimiento de relaciones de razón o de conjuntos, los cuales también discute, no involucra ninguna receptividad —sólo espontaneidad en los conocedores, para usar el término que Kant contrasta con la receptividad. Entonces debemos concluir que las relaciones de razón y los conjuntos no son existentes independientes de la mente. Esta conclusión podría tentar a uno a pensar que son existentes dependientes de la mente. Pero también deja abierta la posibilidad de que las relaciones de razón y los conjuntos sean *inexistentes* independientes de la mente. A diferencia de los objetos ficticios, los cuales son objetos de la imaginación dependientes de la mente, y de los postulados teóricos no-existentes, los cuales son objetos de la cognición dependientes de la mente, las relaciones de razón son objetos de la cognición *independientes de la mente*.

¿Por qué, después de todo, debería ser imposible referir a lo que no existe? La única condición que puede ser legítimamente impuesta a la referencia es que deberíamos saber, y ser capaces de comunicarnos mutuamente, acerca de lo que estamos hablando —cuál es nuestro tema. Esa condición requiere que deberíamos ser capaces de anclar nuestra referencia de alguna manera inter-subjetivamente inteligible. Cuando aquello de lo que estamos hablando es Sherlock Holmes, la condición está ampliamente satisfecha. Él es el detective que Conan Doyle inventó, el que se dice que vive en Baker Street, y así. En el caso del flogisto, es una entidad teórica postulada por los científicos del siglo dieciocho para explicar fenómenos que ahora pensamos como oxidación.

Contrario a las intuiciones parmenídeas, la presentación como un objeto *de re* del pensamiento y el discurso no entraña en sí misma existencia. Es verdad de Sherlock Holmes que tú y yo estamos hablando sobre él, a pesar de que no existe. Por supuesto, puesto que no existe no tiene ninguna propiedad que presuponga existencia, pero que estén hablando de ti no es —a diferencia de ser un detective, vivir en Baker Street, o tocar el violín— una propiedad que presuponga existencia, más

de lo que ser un detective *ficticio,* o ser el detective ficticio más grande, es una propiedad que presuponga existencia.

§Pasamos ahora a las relaciones de razón y los conjuntos. Como con Sherlock Holmes y el flogisto, el hecho de que sean objetos legítimos de referencia y cuantificación, y de que haya verdades sobre ellos, no implica que existan. Crucialmente, sin embargo, no debemos saltar de esto a la conclusión de que son objetos ficticios como Sherlock Holmes o, como el flogisto, existentes putativos que resultaron no existir. Estos dos tipos de referentes no-existentes —las ficciones y los existentes putativos— son imaginados, inventados o postulados. Sólo en virtud de estos actos mentales se vuelven objetos posibles de referencia; si la cadena de referencia a ellos es rastreada, se debe llegar a algunos actos mentales de este tipo. La posibilidad de referir a ellos, y por ende de hacer enunciados verdaderos o falsos sobre ellos, presupone la ocurrencia de tales actos mentales. Eso no es el caso ya sea con los conjuntos o con las relaciones de razón. La posibilidad de referir al conjunto de los planetas, por ejemplo, no presupone que alguien haya imaginado, inventado o postulado tal conjunto. Ese conjunto es un objeto posible de referencia independiente de la mente. De la misma manera con las relaciones de razón: por ejemplo, la posibilidad de referir a la relación de razón que se sostiene entre el hecho de que una acción será placentera y esa acción, no presupone que cualquiera haya imaginado, inventado o postulado tal relación.

La posibilidad de referir a esta relación de razón, por supuesto, depende —como siempre— de la posibilidad de anclar la referencia, con una adecuación suficiente para que nosotros conozcamos de qué estamos hablando. En el caso de las relaciones de razón, eso es hecho por ascenso semántico. Podemos decir, por ejemplo, que el hecho de que alguna acción será placentera es siempre una razón para hacerla. Ascendiendo un orden, podemos decir, equivalentemente, que el hecho de que será placentero siempre mantiene una relación de razón para hacerlo. Podríamos decir, por ejemplo, que siempre se mantiene en una relación

de razón *pro tanto* para hacerlo, pero no siempre en una relación de razón suficiente. Cuando decimos estas cosas, referimos a relaciones de razón. Y puesto que tal enunciado de primer orden es objetivamente verdadero o falso —esto es, su valor de verdad no depende de lo que pensamos o sintamos o hayamos decidido como su valor de verdad— se sigue trivialmente que el enunciado de segundo orden es, en exactamente el mismo sentido, objetivamente verdadero o falso.

El enunciado de segundo orden cuantifica sobre relaciones de razón. Dice que alguna relación de razón se sostiene entre el hecho de que hacer algo será placentero y hacerlo. Desde un punto de vista puramente lógico-semántico, eso está bien. La semántica asigna valores semánticos; asignará valores a predicados de razón y a términos singulares denotando relaciones de razón. Para pasar de eso a la conclusión de que las relaciones de razón existen, necesitamos el criterio cuantificacional de la existencia. Esa es una tesis metafísica, no semántica. Nótese que esto va más allá de la semántica, de modo tal que la pregunta sobre si la partícula de Higgs existe *no* va más allá de la física. La física pretende lidiar con lo que existe. Esa es su misión. Por ende, hipotetizar una partícula es hacer un postulado sobre lo que existe. La semántica, en contraste, pretende lidiar con aquello sobre lo que hablamos cuando usamos diversos fragmentos del lenguaje. Esa es su misión. La idea adicional de que todo aquello de lo que hablamos existe es completamente ajena a ella.

El irrealismo sobre las relaciones de razón rechaza esta idea. Sostiene una tesis metafísica diferente, a saber, que lo real, lo existente, es aquello que tiene posicionamiento causal. Las relaciones de razón, entonces, son irreales, inexistentes, pero objetivas, independientes de la mente. Podría resumir esto por medio de una observación bien conocida, a veces atribuida a Kreisel: lo que le importa a la matemática es la objetividad de los enunciados matemáticos, no la existencia de los objetos matemáticos. De la misma manera con la normatividad: lo que le importa a la normatividad es la objetividad de los enunciados normativos, no la existencia de los objetos normativos. Los enunciados normativos son

objetivos: los objetos normativos son irreales. Son, aún así, objetos posibles de referencia.

Otra manera de iluminar la posición irrealista es señalar cómo nos trae de vuelta a lo que se ve hoy en día frecuentemente como una distinción pasada de moda y de petición de principio: la distinción entre las proposiciones factuales y las normativas.

La distinción es impopular porque choca con una idea popular: el factualismo —la tesis de que todo el contenido proposicional es contenido factual. La teoría de la verdad por correspondencia está comprometida con el factualismo, puesto que, bajo esta visión, lo que determina el contenido de una aserción es su verificador, y los verificadores son hechos. De otra manera, una persona que sostiene que la única noción sostenible de un hecho es la noción fregeana, de acuerdo con la cual un hecho es sólo una verdad, está también comprometida con ella. Por lo que puede ser que Scanlon esté comprometido con ella.

Hay otra, sin embargo, y muy común, noción de hecho. Bajo este entendimiento, un hecho es la instanciación de alguna propiedad o relación real, esto es, de una propiedad o relación que tiene posicionamiento causal. Esta es, por ejemplo, la noción de hecho que está en juego cuando hablamos, como lo hacemos habitualmente, sobre las relaciones causales entre hechos (el hecho de que la grúa se quede vertical es gracias al hecho de que han sido instalados contrapesos suficientes). Es obviamente un truismo que si aseveras que algún hecho se obtiene, en este sentido de hecho, tu enunciado es verdadero sólo si el hecho se obtiene. Puesto que esto es un truismo, debe ser respetado por cualquier concepción de la verdad; debería ser, en consecuencia, claramente distinguida de la concepción de la correspondencia. No entraña que, a diferencia de la concepción de la correspondencia, *toda* aserción es factual o que todo el contenido es contenido factual. Y, por supuesto, bajo la visión irrealista, una proposición puramente normativa sobre las relaciones de razón no tiene contenido factual en este sentido común de hecho, puesto que las relaciones de razón son irreales. Una proposición puramente normativa dice algo perfectamente; sólo no dice nada factual.

Así, aseverar que el hecho de que alguna acción será placentera es siempre una razón para hacerla dice algo normativo sobre tal hecho —no dice nada *factual* sobre ese hecho.

Este punto de vista, creo, captura nuestra visión preteórica sobre la normatividad más cercanamente y claramente que el punto de vista propuesto por Scanlon, el cual introduce una visión de la existencia centrada en el dominio mientras aún aplica el criterio cuantificacional. Pero alguien podría preguntar, ¿no es este punto de vista en realidad una mera variante terminológica de la visión de Scanlon? Veo el punto, pero mi inclinación tentativa es decir que no. Es verdad, es una cuestión de convención cómo usamos la *palabra* 'existir'. Y esa palabra ocurre en el discurso ordinario de una variedad de maneras (como lo hace la palabra 'hecho'). Algunas veces juega el rol cuantificacional y algunas veces juega el rol ontológico. Por lo que mi versión del cognitivismo sí hace alguna sistematización del habla ordinaria tal como lo hace la de Scanlon. Además, si Scanlon está dispuesto a aplicar su noción de existencia centrada en el dominio a los objetos ficticios y a los postulados teóricos, entonces las dos concepciones son, sin duda, muy cercanas. Ambas niegan que el cognitivismo simple sobre lo normativo requiera cualquier forma de realismo metafísico, natural o no-natural. Aún así, pienso que mi concepción clarifica de mejor manera por qué esto es así, mediante una mejor concepción de criterios ónticos, y por esa razón no pienso que las dos concepciones sean simplemente variantes terminológicas. A diferencia de Scanlon, pienso que sí tenemos una concepción del mundo, o de la realidad; a diferencia del realista, no creo que pensemos que contiene 'hechos normativos'. La realidad es el orden causal unificado mediante el cual hallamos nuestro camino. Estamos en ella de una manera en la cual las relaciones de razón no lo están.

REFERENCIAS

Blackburn, S. (1984). *Spreading the Word. Groundings in the Philosophy of Language*. Oxford University Press.

Blackburn, S. (1993). *Essays in Quasi-Realism*. Oxford University Press. https://doi.org/10.1093/oso/9780195080414.001.0001

Blackburn, S. (1998). *Ruling Passions. A Theory of Practical Reasoning*. Oxford University Press. https://doi.org/10.1093/oso/9780198247852.001.0001

Dummett, M. (1996). "What is a Theory of Meaning? (I)". In *The Seas of Language*. Oxford University Press. https://doi.org/10.1093/0198236212.001.0001

Gibbard, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Oxford University Press. https://doi.org/10.1093/oso/9780198249856.001.0001

Gibbard, A. (2003). *Thinking How to Live. Harvard University* Press. https://doi.org/10.4159/9780674037588

Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. Penguin. (Hay traducción al castellano: Mackie, J. L. (2000). *Ética: la invención de lo bueno y lo malo* (T. Fernández Auz, Trans). Editorial Gedisa).

Priest, G. (2005). *Towards Non-Being. The Logic and the Metaphysics of Intentionality*. Oxford University Press. https://doi.org/10.1093/0199262543.001.0001

Quine, W. V. O. (1961). *From a Logical Point of View* (Second ed.). Harvard University Press. (Obra original publicada en 1953. Hay traducción al castellano: Quine, W. V. O. (2002). *Desde un punto*

de vista lógico (Segunda ed.). (M. Sacristán, Trans) Barcelona: Ediciones Paidós).

Scanlon, T. M. (2014). *Being Realistic about Reasons*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199678488.001.0001

Skorupski, J. (2010). *The Domain of Reasons*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199587636.001.0001

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 379-438
ISSN 3087-260X
Santiago de Chile

ENSEÑANZA, MÉTODO Y PRÁCTICA DE LA FILOSOFÍA[1]

Entrevista al Dr. Luis Placencia[2]
Kuminak Lefio[3]— Entrevistador

Luis Placencia García es profesor asociado de la Facultad de Filosofía y Humanidades y de la Facultad de Derecho de la Universidad de Chile, Doctor en Filosofía por la Martin-Luther-Universität Halle-Wittenberg (2014, *summa cum laude*) y licenciado por la Pontificia Universidad Católica de Chile (2005, egresado con distinción máxima). Sus líneas de investigación se centran en historia de la filosofía moderna (autores como *Descartes, Spinoza, Kant, Hegel*), filosofía práctica y problemas de autoconciencia y autoengaño. Así mismo, ha realizado estancias

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Entrevistas de la Revista Homónima*. Cómo citar: Lefio, K. (2026). Enseñanza, método y práctica de la filosofía / Entrevista con el Dr. Luis Placencia. Revista de Filosofía Homónima 1(1), 379-438.
[2] Email: luis.placencia@uchile.cl. ORCID: https://orcid.org/0000-0003-1052-0846
[3] Licenciado en Filosofía de la Universidad de Chile, estudiante del diplomado "Fundamentos de la Física" de la misma casa de estudios. Actualmente se desempeña como jefe en Gestión Editorial de la Revista Homónima. Email: kuminak.lefio@ug.uchile.cl. ORCID: https://orcid.org/0009-0005-8121-4063.

de investigación en la Universität Leipzig (2015) y en la Rheinische Friedrich-Wilhelms-Universität Bonn (2017–2018).

I. Introducción

Kuminak Lefio

La entrevista consiste en aproximadamente 15 preguntas centrales, con sus seis núcleos temáticos correspondientes, por lo que espero que tenga tiempo de contestarlas y la facilidad de hablarlos por el interés que nos generaba tener su perspectiva. Las preguntas intentan mezclar elementos que les sean más familiares —de lo que pensamos que le puede ser más familiar para usted—, y otros que lo saquen de su zona de confort. Pensamos ente ellas, cuestiones relacionadas con su vida, tanto como estudiante como profesor de filosofía; cuestiones sobre el método de la filosofía su concepción de ella y su práctica; reflexiones en torno a la interdisciplinariedad; acerca de cuestiones de género en filosofía; y el futuro de nuestra disciplina dado el ascenso de nuevas tecnologías.

II. Como estudiante y como profesor

Kuminak Lefio

Un gusto poder tener esta entrevista con usted.

Me interesa partir, desde la siguiente idea: conversando con varios estudiantes de filosofía, tanto durante la carrera, en los congresos, etc. me di cuenta de que muchos de nosotros partíamos, al entrar en la carrera, con una idea bien distinta de lo que era la filosofía, en comparación con la idea con la que salimos de la carrera, o con la que fuimos evolucionando a medida que conocimos y estudiamos la filosofía. Entonces, por lo

mismo, me interesaba preguntarle ¿Cuál era la idea inicial que tenía de la filosofía, cuando entró a estudiarla, y cómo ella fue evolucionando, cambiando, a medida que fue estudiándola y desarrollándose en toda su carrera filosófica? ¿Cómo se fue desarrollando esa concepción que tenía de la filosofía?

Luis Placencia

Cuando yo entré a estudiar filosofía, tenía una idea muy vaga de lo que era. Para mí era una especie de ejercicio de la curiosidad sin restricciones metódicas, pero con una cierta rigurosidad —que yo no sabía muy bien cómo expresar en qué consistía— en torno a las cuestiones fundamentales de la vida humana, o lo que yo entendía que eran las cuestiones fundamentales de la vida humana. Entonces, cuando decidí estudiar filosofía, decidí, en parte, motivado por ese tipo de concepción.

A mí me gustaban mucho otras humanidades también, la historia en particular, también la literatura, pero me parecía que justamente la filosofía era, en algún sentido, más universal, por esta razón que les mencionaba. Y, en la medida que fui estudiando filosofía, empecé a darme cuenta de que cuando la filosofía era "bien hecha" —al menos me parecía a mí que lo era; todo lo estoy tratando de expresar del modo vago en el que yo lo percibía en ese momento—, cuando la filosofía era convincente, planteaba ideas que me parecían bien articuladas, y por lo mismo robustas, me di cuenta de que ese tipo de filosofía normalmente estaba vinculada más bien a unas constricciones metódicas bien definidas; a una cierta conciencia de por qué se hacía lo que se hacía del modo en que se hacía. Y eso para mí fue muy importante, porque, de hecho, en algún momento pensé en dejar la filosofía a causa de que cuando estuve el primer semestre de la universidad, experimenté algo —que creo que muchos estudiantes experimentan— que es no entender muy bien de qué se trataba esto.

Y me pasó como pasaba en ese entonces, y creo que todavía pasa en algunas facultades, que tenía mis cursos de lengua clásica simultáneamente. Me gustaba mucho el mundo clásico, probablemente como tenía esta especie de inquietud por las preguntas de las humanidades, entendida en términos muy generales, y dado que el mundo clásico es una especie de hummus del que surgen ese tipo de preguntas, me interesó por lo tanto mucho desde joven, desde antes de entrar a estudiar filosofía. Entonces, esos cursos de lengua clásica me satisfacían mucho porque yo veía que progresaba; uno hace esta especie de camino curioso en el que pasa de no saber nada, a después poder leer una línea, después poder leer un párrafo, poder leer un texto. Entonces, eso me gustó mucho. Y en la filosofía me sentía más bien desorientado. Veía que se trataba de una exposición de ideas diferentes, algunas de ellas muy interesantes, pero no lograba comprender cómo podía hacerse aquello de buena manera. Hasta que, bueno, el contacto con un profesor en particular, con el profesor Vigo, me generó un impacto muy profundo en ese sentido, porque me pareció que él respondía a unos estándares que a mí me satisfacían mucho. Yo consideraba muy admirable el modo en que él hacía la filosofía. Y en parte de él aprendí esta preocupación, en el fondo, por las cuestiones metódicas de la filosofía. Después me di cuenta de que eso fue decisivo para mí para quedarme en la filosofía, y no intentar hacer una especie de carrera en las lenguas clásicas, que es algo por lo que pensé en algún momento, sobre todo en el latín, que me fascinaba mucho.

Kuminak Lefio

Sí, justamente, relacionándolo con eso, hay algo que me interesaba comentarle, esta idea con la que muchos estudiantes lidian a lo largo de la carrera, de esta vaguedad, de estar un poco perdidos respecto a cómo entender la filosofía, y también cierto ánimo de decepción, que muchas

veces se da. ¿Cómo cree usted que se podría hoy en día, sabiendo que eso se sigue dando, apoyar a esos estudiantes? o tal vez ¿Cuál es su manera de dar un apoyo a esos estudiantes en encontrar una forma de entender la filosofía?

Luis Placencia

La respuesta que yo daría es una respuesta en alguna medida desilusionante, pero que tiene que ver con mi propia experiencia. Creo que la manera de tratar de proponerle a los estudiantes un camino en la filosofía es tratar uno mismo de hacer bien su trabajo. Tengo la impresión de que muchas veces en filosofía simplemente enseñamos cosas, decimos cosas, sin tener muy claro por qué lo estamos haciendo. Y creo que una parte esencial del hacer bien nuestro trabajo como profesores consiste justamente en hacer o intentar hacer, cada vez, esa reflexión previa; tratar de explicarnos por qué hacemos lo que hacemos. Por ejemplo, a mí me toca, como ustedes lo saben, cada cierto tiempo dar un curso de filosofía moderna, que es un curso obligatorio. Entonces, yo trato de iniciar siempre el curso explicando por qué yo creo que es importante estudiar eso y por qué es importante estudiar la historia de la filosofía en términos generales.

Eso me parece que es una cosa que uno debería tratar de hacer. Tengo la impresión de que no siempre se hace, que más bien se asume que lo que está dado está dado por alguna razón y entonces simplemente hay que ir a eso que sería la filosofía. Pero cuando uno opera de ese modo justamente oblitera uno de los aspectos esenciales de la filosofía, que es la necesidad de dar cuenta de sí misma. Es decir, trata la filosofía como si fuese una disciplina cualquiera, como cualquiera de las otras, que son todas ellas muy respetables, por supuesto, pero que no tienen esta característica distintiva. Así como la filosofía tiene también, frente a esas otras disciplinas, una serie de desventajas. Nosotros no podemos hacer

experimentos, no tenemos métodos formalizados de control, etc. Pero —eso en parte— está intrínsecamente relacionado con el hecho de que justamente no asumimos nada como dado, o intentamos asumir lo menos posible como dado. Lo cual significa entonces que no podemos simplemente llegar y enseñar las materias de las que se trata, como si fuese obvio y trivial el por qué esas cosas son importantes.

Creo que ese ejercicio —el ejercicio de tratar de hacer bien nuestro propio trabajo— obviamente es importante. Ahora, tengo la impresión también de que el abandono en las carreras de filosofía —si uno ve los números es una situación más o menos común también con el resto de las humanidades— obviamente tiene que ver con una cuestión epocal y también laboral. Hay muchas personas que a medida que va avanzando en la carrera se dan cuenta de que tienen que hacer un esfuerzo, que a veces es significativo, y el "premio" —por decirlo de algún modo— al final del esfuerzo es incierto. Y eso es un problema que, desde las mismas facultades de filosofía, nosotros no podemos solucionar. Lo que sí creo que podemos hacer es tratar de hacer un ejercicio razonado de lo que hacemos. Y eso sí sería muy importante.

Lo otro que creo que sería muy importante, especialmente en una facultad como la mía, es tener tasas de ingresos que sean razonables. Si uno admite 100 estudiantes a entrar a estudiar filosofía en un país de 18 millones de personas, está admitiendo unas cifras que creo que son irreales, sobre todo si hay 7 u 8 facultades más de filosofía en el país. Entonces, es muy natural que ante un escenario de esa naturaleza ciertos estudiantes se desanimen. Normalmente dicen los estudiantes: "Yo aquí estoy haciendo un esfuerzo enorme, aprendiendo unas cosas que son rarísimas, que no me queda muy claro cómo se vinculan con el mundo en el que yo vivo, y no termino de ver el beneficio inmediato de todo aquello". Esto es muy natural en la filosofía. La filosofía en alguna medida siempre consiste en desvinculación del mundo inmediato. Pero claro, eso tiene que tener algún tipo de explicación. Y eso es lo que yo a veces siento

que nos hace falta a nosotros los docentes comunicar de buena manera: El por qué esto que hacemos tiene en alguna medida algún sentido, aunque sea inmanente a la filosofía.

Kuminak Lefio

Usted mencionó brevemente aquello que parecía diferenciar a la filosofía de otras disciplinas. Justamente, desearía que profundizase un poco más en si es que cree que la filosofía tiene algún tipo de reflexión particular que la diferencia de otras disciplinas, un poco como lo que decía recién, y cuál sería ese tipo de reflexión que la caracteriza. Si es solo esto que mencionaba, el intentar dar lo menos por dado o lo menos por supuesto, o si hay algo más en el ejercicio filosófico que sea característico a diferencia de otras disciplinas.

Luis Placencia

Varias cosas. Por supuesto, una primera cosa que es relevante tener en cuenta es que esta es una idea de la filosofía. En ningún caso pienso que esta sea la única manera de concebir la filosofía que sea admisible. Hay muchas personas que hacen filosofía y lo hacen muy bien y no están de acuerdo conmigo en esto. Tengo unos colegas con los que he conversado esto y bueno, tienen dudas de estas cosas que yo digo y creo que con cierta razón. No lo doy por una especie de verdad apodíctica, pero creo que hay buenas razones para pensar lo que estoy diciendo.

Obviamente, de la idea de que uno asuma la menor cantidad de prejuicios, es decir, que uno tome como ya admitidas, o por ya sabidas, la menor cantidad de cosas, se desprenden una serie de consecuencias. Porque, fíjense ustedes, que esto vale incluso para la idea misma de la

filosofía. Entonces, se desprende el hecho de que la filosofía tiene que hacerse cargo de sí misma, pero tiene que hacerse cargo de sí misma —en mi parecer al menos— de un modo que implica tratar de reflexionar sobre cómo hacer filosofía de una manera que ella se aproxime tanto cuanto sea posible a algo que pudiéramos denominar saber.

La aspiración de la filosofía no es meramente ser una forma de reflexión, sino que ser una forma de reflexión que en alguna medida aspira tanto cuanto se pueda a presentarse como un modo de conocimiento intersubjetivamente válido. Reflexionar es una actividad que en general los seres humanos realizan espontáneamente, especialmente a partir de cierta edad. Hay una edad en que la gente se pone filósofa, se pone a pensar "¿Por qué estoy aquí? ¿Por qué nos morimos?" Ese tipo de cosas, u otras, como "¿Existe la Torre Entel cuando yo la estoy mirando?" Ese tipo de reflexiones. Pero evidentemente la filosofía lo que busca es, recogiendo esa tendencia espontánea, si no en todos los seres humanos, al menos en una cantidad significativa de ellos, intentar llevarla a una expresión que pudiésemos denominar científica. Es decir, una expresión que responda a ciertos controles metódicos.

Otro rasgo fundamental de la filosofía no es que ella carezca de método, a diferencia de otras formas de conocer que tienen un método dado, o unos métodos —pueden ser varios— sino que la filosofía es un tipo de —Heidegger lo dice así— lucha por el método. Es decir, de ejercicio constante de reflexión en torno a cuáles son los métodos más adecuados para realizar las tareas que la filosofía se da a sí misma. Ese es otro rasgo que es muy relevante; no tenemos métodos dados *ex-ante*, pero eso no nos libera de la necesidad de darnos ciertos métodos. Es decir, la relación de la filosofía con la metodología es una relación distinta a la que tienen otras disciplinas.

Ese es un segundo rasgo ya relevante, digamos, que diferencia a la filosofía de otro tipo de disciplina. Y obviamente de estos rasgos también

se desprende una cierta flexibilidad. La filosofía en principio, en virtud de estas cosas que les acabo de mencionar, puede pretender hacerse cargo de asuntos que son muy diferentes. Esta especie de flexibilidad está considerada como un rasgo propio de la filosofía ya desde los primeros filósofos. Ellos lo expresaban de otro modo, como si la filosofía fuera una especie de disciplina universal que se ocupa de todo.

Evidentemente, ya esos primeros filósofos, pienso en Aristóteles por ejemplo, veían que ese de todo no podía significar ocuparse de todo en el sentido de estrictamente todas las cosas en todos los sentidos posibles. Sino que se requería una especie de configuración metódica que permitiese ocuparse de todo. De todas las cosas, pero de un cierto modo. Aristóteles llamaba a eso, la disciplina que estudia lo que es en tanto que algo que es. Entonces, bueno, más allá de que uno comparta eso que dice Aristóteles o no, ahí hay una lección metódica también importante. Esa curiosidad que a mí me atraía tanto, que parece ser propia del filósofo, esa especie de interés universal, tiene que ser ejercido sobre la base de una cierta concepción; cuál es el tipo de encuadre bajo el cual vamos a ejercer esa curiosidad. No es una especie de curiosidad librada a sí misma sin ningún tipo de restricción, ¿no? No es la curiosidad del polímata que se ocupa de cuántas cosas puede.

Kuminak Lefio

Justamente usted mencionaba que le interesaba tratar este tipo de temas metodológicos y más bien, acerca de los fundamentos de por qué tratamos las cosas como las tramamos en filosofía y que le gustaba mostrar aquello, también, en sus clases. Por lo mismo, ya sea en las clases universitarias que usted da, no sé si ha tenido la oportunidad también de hacer clases a secundarios, me interesaba preguntarle si alguna vez le ha ocurrido que en esas clases o en otras instancias con estudiantes que usted

haya visto algún tipo de cambio en su forma de pensar respecto a la filosofía por lo que decían sus estudiantes.

Luis Placencia

Sí, pasa mucho. Yo nunca di clases en el colegio salvo unos talleres de lengua clásica que di en el Liceo 1, y que fue una experiencia muy grata por lo demás. De hecho fue lo primero que hice como docente, salvo algunas cosas que hice en el colegio como "docencia" en ayuda a compañeros, y cosas así. Pero como docente, propiamente tal, lo primero que hice fueron unos talleres de lengua clásica y de lógica que para mí fueron muy entretenidos; lo pasé muy bien y fueron muy gratos. Pero eran personas que iban ahí porque querían. Para mí era muy importante eso, entonces después apenas pude tomé la oportunidad de empezar a trabajar en la universidad y no salí nunca más de ahí, porque en general siempre me gustó mucho la idea de darle clases a personas que iban básicamente porque habían escogido hacer eso; que no era una especie de obligación.

Para mí eso es muy importante. Obviamente sin desconocer que la enseñanza primaria y secundaria es central y se hace ahí una tarea importantísima. Lo que pasa es que yo no creo tener la habilidad para ejercer esa tarea de un modo que no sea sumamente dependiente de cierta madurez de las personas. Ese es un ejercicio que requiere de una habilidad y unas capacidades que yo no tengo pero que son capacidades muy importantes, propias de una forma muy excelsa de la inteligencia, de la cual yo carezco.

Pero volviendo al punto, sí, frecuentemente ocurre —me parece— que en las clases uno recibe información de perspectivas que son interesantes. Ahora, para eso obviamente uno requiere que la clase funcione de un modo en que hoy es muy difícil hacerla funcionar. Se necesita que tenga la forma primero de un diálogo, pero también de un diálogo sobre la base de cierta información. Es cierto que a veces la ingenuidad permite adquirir una perspectiva fresca y novedosa y en ocasiones muy necesaria

para quienes están mucho tiempo ocupados con un asunto. Pero, también es cierto que si esa ingenuidad está en alguna medida informada, las probabilidades de que ocurra ese fenómeno, en el que quién está demasiado cercano a un asunto logra una cierta distancia, se maximizan.

Entonces, me da la impresión de que ocurre tendencialmente menos que se puedan hacer clases dialogadas y que se puedan hacer clases en que ese diálogo además esté fundado en una especie de preparación. Probablemente, porque tendemos a concebir, cada vez, más la clase como una especie de fenómeno vinculado a enseñar ciertos contenidos. Y eso en la filosofía creo que es muy complejo porque en general mi impresión es que en la filosofía lo que deberíamos tratar de enseñar más bien es cómo reflexionar filosóficamente.

Kant tiene una frase que es muy citada, aunque él ahí está diciendo una cosa ligeramente distinta a lo que se le hace decir cuando se cita la frase, pero no tiene importancia —tergiversemos a Kant un rato, sin problema—, que es esto de que "No se aprende filosofía sino sólo a filosofar". Esa observación, entendida de un modo distinto a como Kant la hace ver ahí, es una observación muy importante. Porque en mi opinión la enseñanza de la filosofía debería estar más orientada, no tanto a contenidos, sino a que el estudiante desarrolle por sí mismo las capacidades para interpretar un texto, hacerse preguntas adecuadas respecto de él, y por tanto entrenarse en la reflexión filosófica, en una especie de diálogo con otras mentes, a veces mentes ausentes, que son las mentes de los autores que escriben los textos, porque normalmente no están ahí de cuerpo presente, ora porque ya está muerto, ora porque son personas que pertenecen a otros mundos, a otros universos; viven, en Stanford, o en Europa, etc.

Entonces, me ha pasado muchas veces que he aprendido mucho de los estudiantes. De hecho, he trabajado en estudios, he publicado cosas con gente que ha estudiado conmigo. Y para mí en general, la experiencia con

esos estudiantes ha sido muy enriquecedora. Es probablemente una de las partes más hermosas del trabajo académico y la docencia, cuando uno tiene la oportunidad de ver crecer a un estudiante y ver cómo se transforma en un colega. Obviamente, es una parte que también puede ser potencialmente difícil de administrar —digamos—, especialmente si uno es inseguro. Pero es un proceso muy satisfactorio.

Kuminak Lefio

Para cerrar esta parte acerca de los estudiantes, también me gustaría preguntarle si pudiera darles un consejo importante a los estudiantes de filosofía hoy en día. ¿Cuál sería?

Luis Placencia

Carpe Diem, que aprovechen el tiempo. Yo siempre cuento una historia que a mí me pasó cuando me fui a doctorar. Yo tenía un profesor que falleció ya hace un tiempo largo, y lo extraño mucho, Don Alfonso Gómez Lobo, una extraordinaria persona. Yo fui ayudante de él varios años y para mí fue una linda experiencia conocerlo. Y digo linda no solo porque aprendí muchísimo, él era un señor admirable, tenía un conocimiento del mundo griego como el que tengo yo de esta cuadra. Era su mundo. Era una bellísima persona. Y me acuerdo de que cuando le fui a contar que me iba a doctorar a Alemania, que además era el país donde él se había doctorado, me sonrió con esa sonrisa plena que tenía. Y me miró con una mirada llena de felicidad. Y me dice, "Qué envidia; qué ganas de empezar de nuevo". Y él era una persona que vivió muy bien, o sea, yo sé que no me lo decía en términos de que él había vivido una mala vida. Sino más bien al revés; con ese deseo en el fondo autoafirmativo, que decía algo así como "Esta vida que yo he llevado ha sido muy linda y empezó con algo

muy lindo, que fue el periodo de estudiar". Y el periodo de estudiante, yo creo, que es muy lindo y no vuelve más. Es una cosa que uno normalmente, cuando es estudiante, no lo aquilata lo suficiente. Después viene otro período en el que uno finalmente no dispone de la cantidad de tiempo para sí mismo que tiene cuando uno es estudiante.

Obviamente ocurre por razones diferentes. Puede ser porque uno mismo o su entorno envejece y hace que uno tenga que disponer de una cantidad de tiempo para otros. Por supuesto, también está el trabajo, que es una carga significativa, y suele ser también tiempo que uno emplea en atender asuntos de otros. O también muchas veces las personas hacen familia, tienen hijos —como yo mismo—, y tienen que dedicarles tiempo a otros; un tiempo que a veces es muy gozoso, pero que no es tiempo dedicado al estudio.

Entonces, mi consejo sería que piensen que el tiempo que tienen ahora no lo van a volver a tener. Y es un tiempo muy lindo que permite hacer cosas que uno después no puede hacer. Creo que eso es bien importante. O sea, hay una cuestión, si uno quiere, *kairológica*[4] de la situación, que uno tiene que saber identificar. Tiempo para el "carrete", por ejemplo, uno tiene después; es verdad que a medida que uno se va haciendo más viejo, también se le va haciendo más difícil resistir ese tipo de instancias, digamos.

Pero el tiempo para estar sentado un buen rato en la biblioteca, para curiosear, para leer cosas simplemente porque sí, porque a uno le dan ganas, le llamó la atención. Ese tiempo lo tiene uno mientras estudia, sea en el pregrado o en el posgrado. Entonces es un tiempo muy bello que mi único consejo para un estudiante es que lo aproveche mucho, porque si a uno realmente le gusta lo que hace, —claramente si la filosofía es una tortura, el consejo sería otro, sería estudiar otra cosa—, si le gusta la filosofía, es muy lindo tener la ocasión de dedicarle un tiempo a algo. Y

[4] "Kairológica" de "kairós" (ocasión, tiempo preciso para algo, en griego clásico)

eso es algo que uno después no puede hacer. Tiene que andar robando tiempo a distintas cosas para poder tener esa intensidad de trabajo y de estudios, si acaso la puede llegar a tener.

Kuminak Lefio

Pasando a otro tema, de lo que yo he tenido la oportunidad de tener clases y discutir con usted, me consta la importancia que le otorga a los problemas filosóficos y cómo los formulamos en nuestra práctica, como parte de su concepción misma de la filosofía y de cómo la enseña también. Por lo mismo, me gustaría preguntarle más específicamente ¿Cuáles fueron los problemas filosóficos que han marcado su carrera filosófica hasta ahora y también cómo llegó a ellos? ¿Qué hizo que le importasen, le interesasen y le llamasen la atención?

Luis Placencia

Me han interesado muchas cosas. Muchas de ellas son cosas a las que yo después no me dediqué, en parte porque me di cuenta de que había gente que lo había hecho ya muy bien. En cualquier caso, lo había hecho de un modo en el que yo no lo iba a poder hacer, probablemente con una calidad y sofisticación que yo no iba a lograr.

Me interesaron preguntas que uno podría denominar filosóficas de antes de estudiar formalmente filosofía, pero probablemente de un modo que yo no admitiría como significativo después de haber estudiado filosofía. Probablemente una de las primeras preguntas que me causó una gran admiración y que hacía parte de estas preguntas que me inquietaban de antes, era la cuestión del tiempo. Es decir, ¿Qué clase de cosa podía ser eso? Algo que parece en alguna medida ser lo más real del mundo, que

marca nuestra vida por completo, sentimos su realidad efectiva en nuestro propio cuerpo, en su paso, pero por otro lado parece ser lo más irreal que hay. Si buscamos dónde está el tiempo, no lo encontramos en ninguna parte. De hecho, probablemente una de las formulaciones más brillantes de lo que es una pregunta filosófica es la formulación de Agustín, en torno a ese tipo de preguntas: ¿Qué es el tiempo? Si me preguntan, no lo sé. Si no me preguntan, lo sé. Esa especie de paradoja es indicio de que hay un problema filosófico y me parece que el tiempo es un problema filosófico en ese sentido.

Cuando yo estudié en el pregrado, en la facultad en la que yo estudiaba, Aristóteles era una especie de presencia constante, era una facultad muy aristotélica, en buena hora en mi opinión, y en particular uno de mis maestros, que ya mencioné, el profesor Vigo, había escrito trabajos extraordinarios sobre el tiempo, especialmente en Aristóteles, pero también con influjo de una serie de otros autores.

Entonces, me acuerdo de que en el seminario privado que tenía él, en el coloquio, leímos en esa época *La Física* de Aristóteles, en el libro cuarto, capítulos 10 a 14. Estaba también la profesora Paloma Baño en esa época, que está como una colega aquí con nosotros. Ella estaba muy interesada en esas cosas, de hecho escribía sobre el tiempo de Aristóteles en esa época.

Y estaba la profesora Gabriela Rossi, que ahora es profesora de la Universidad de Los Andes y fue profesora de la Universidad Adolfo Ibáñez y de la Católica de Valparaíso, que estaba también trabajando sobre los temas de la Física; estaba el profesor Boeri, también el profesor Vigo. Yo me acuerdo de que ahí me embelesé con ese asunto. Leí cuanto cayó en mis manos sobre el tema, estos textos de Aristóteles que leíamos en el coloquio; leímos también, me acuerdo, esos famosos *Problemas Fundamentales de la Fenomenología* de Heidegger, donde él trata esta cuestión en distintos autores, que son realmente fascinantes.

Y yo por mi cuenta leí también una cantidad de otros trabajos —Por ejemplo, el famoso artículo de MacTaggart. Eso fue un asunto que me fascinó mucho. Y en parte, fíjese usted, un poco en esa guisa, llegué a ocuparme con lo que fue mi primer tema, el primer tema que me interesó, que yo trabajé con cierta seriedad, que era el problema del espacio, un problema relativamente similar, digamos. Y ahí me ocupé con la cuestión del espacio en la filosofía de Kant. Escribí mi primer trabajo, que después se publicó como un libro. Me fascinó ese asunto.

Me fascinó el problema de qué tipo de entidad o de forma de ser tenía el espacio. Que parecía, por un lado, ser una especie también de entidad omnipresente, pero completamente imposible de localizar. Más bien era una especie de condición de posibilidad de localizar entidades. Entonces, esta especie de realidad curiosa, digamos, que tienen estos objetos, me llamó mucho la atención en su momento. Y creo yo que eso me abrió también, especialmente a partir del estudio del enfoque kantiano de estos asuntos, cierta sensibilidad para lo que yo creo que ha sido una ocupación muy constante, que es el intento de ocuparme con determinados objetos, que, justamente, sin tener el modo de ser de las cosas, son, no obstante, entidades muy reales, digamos. O entidades de enorme significación, cuya realidad es difícil de discutir.

De alguna manera fue un modo en que la filosofía me dio una especie de antídoto a una tendencia que yo tenía muy naturalmente. Esta especie de tendencia que creo, en general, tienen casi todas las personas hoy por hoy, de considerar como efectivamente reales solamente los objetos materiales. Considerar que vivimos en un mundo de cosas; de cosas empíricamente observables. Y bueno, Kant y el enfoque de algunos de mis maestros, hizo que me fascinara por este tipo de objetos que parecen tener esta especie de realidad sui generis. La realidad de lo que no es cósico, de lo que no es empíricamente observable de manera directa, pero, no obstante, parece ser en alguna medida indiscutible. En alguna medida por esa razón también, y esto fue de lo que me ocupé en el doctorado, me interesó

también el problema de la acción humana, que parece tener un estatus similar, y ahí me ocupé nuevamente también del asunto de la mano de Kant, que en realidad ha sido el filósofo del que más tiempo le he ocupado.

Y finalmente, me han ocupado también en los últimos años cuestiones vinculadas a algo que podríamos denominar en términos amplios como el conocimiento de sí. Creo yo que es otra dimensión de la filosofía que es muy significativa: parece ser que la filosofía, junto con los otros rasgos que mencioné anteriormente, tiene, al menos bajo una concepción de ella, esta idea de ser una especie de disciplina que busca que la racionalidad humana, los seres humanos se comprendan a sí mismos a la vez que comprendan al mundo. Y eso para mí es muy fascinante. Entonces, problemas como el conocimiento de sí, la autoconciencia y fundamentalmente el autoengaño, que me interesaba mucho también, son cuestiones a las que le he dedicado cierto esfuerzo, y esto último ya más en términos sistemáticos y no solamente en Kant. Pero, en general, si tuviese que presentar una especie de itinerario lo haría de esa manera.

Kuminak Lefio

En esa línea, deseába plantearle lo siguiente, ya que mencionó su interés por los problemas del autoconocimiento y también esta idea de la relación entre el mundo y cómo nos conocemos a nosotros mismos. Como se mencionó también recién, para la filosofía los posibles intereses son sumamente amplios; esto mismo que usted dijo, que la filosofía podría tratarse de todo, al menos en alguna forma, y algunas de esas formulaciones puede ser tan amplias como "la filosofía se trata del mundo y nuestro lugar en él". O, más cercano a lo que usted mismo ha mencionado, "el mundo y cómo nos conocemos a nosotros mismos". Pero de esto surge un cuestionamiento, y un cuestionamiento que me parece que es adecuado que también usted pueda contestar, ¿Es

adecuado concebirnos a nosotros como parte de este mundo que investigamos? ¿O hay algo particular acerca de la comprensión de nosotros mismos que lo distingue de los demás objetos que podemos tener en nuestra reflexión? Deseaba que pudiese profundizar acerca de su postura sobre esto, especialmente, por su interés en las preguntas del autoconocimiento.

Luis Placencia

Es una pregunta muy difícil, porque, por de pronto, supone una especie de paso posterior a los que ya hemos dado, que es tratar de reflexionar sobre qué es lo que llamamos mundo. Cuando hablamos del mundo, especialmente en un ámbito de reflexión pre-filosófica, o incluso en un ámbito ya no de reflexión, sino que simplemente correspondiente a lo que llamamos la vida cotidiana, parece ser que hablamos de una especie de ítem, como si se tratase de una cosa más o de un agregado de cosas. Esa es la manera más natural de hablar del mundo.

Vean esa serie tan extraordinaria de "Cosmos", de Sagan, comienza justamente con esa frase tremenda, "El Cosmos es todo lo que hay, todo lo que fue y todo lo que alguna vez habrá". Fíjense que la definición es extensional. El Cosmos es un conjunto de ítems que existen, existieron o van a existir. Con lo tremenda que es la frase, me acuerdo yo cuando la escuché la primera vez, era niño, hubo una ilusión tremenda. ¿Qué es eso?, cabe preguntarse si eso realmente podría ser lo que llamamos mundo. Creo que ahí falta algo, que es que, aquellos seres, aquellas entidades, que en el contexto de esos ítems que están dados, experimentan esos ítems de un modo tal que buscan hacer sentido de ellos.

Esos somos nosotros. Y parece ser que "mundo" tiene un sentido, que en alguna medida es ligeramente distinto de este otro sentido, de esta especie

de sentido que podríamos denominar extensional. Lo podríamos denominar el marco de significaciones bajo el cual los seres humanos hacemos sentido de aquello que experimentamos. Y el marco de significaciones a partir del cual también transformamos eso que experimentamos.

Los seres humanos tenemos además esa cualidad curiosa de que simplemente no nos conformamos con hacer lo que sería nuestra naturaleza y observar lo que hay, sino que tratamos de transformar esa naturaleza y a la vez transformarnos a nosotros mismos. Piénsenlo ustedes, estamos en una universidad. Una universidad no sería posible si no tuviésemos esa curiosa tendencia: de tratar de hacer con nosotros algo en alguna medida completamente diferente a lo que somos desde un comienzo. Si siguiéramos con nuestros impulsos naturales no estaríamos aquí. Estaríamos haciendo otras cosas. A lo mejor pueden ser cosas mejores, más interesantes, no tengo ni idea, pero en cualquier caso estamos aquí.

Entonces, claro, yo creo que si uno entiende el mundo en ese segundo sentido, el papel que tiene la filosofía es un papel en algún sentido diferente, ya no se trata meramente de dar cuenta de lo que hay, sino de cómo comprendemos aquello que hay y de cómo eso que hay adquiere cierta relevancia; algunas formas incluso adquieren el tipo de entidad que tiene a partir del modo como nosotros prestamos sentido a las cosas. Es decir, dicho de otra manera, parece ser que nosotros no experimentamos el mundo como un conjunto de simples cosas existentes dentro de las cuales nosotros somos una más, sino que más bien en un sentido derivado, no este sentido extensional, constituimos un mundo de significaciones cuyo foco somos nosotros.

III. Mujeres en la filosofía y la academia

Kuminak Lefio

Muchas gracias por la respuesta. Me interesa pasar a otro tema, y un tema que creo que es muy relevante para nuestro contexto actual como academia filosófica, y se trata de la cantidad significativamente mayor de hombres que de mujeres que se dedican a la práctica. Al menos, como un dato —pues, no hay estadísticas al respecto—, desde mi experiencia como estudiante y la de muchas compañeras en la carrera de filosofía entran a estudiar muchos más hombres que mujeres. Se nota este fenómeno en los cursos y también da la impresión, por lo que piensan también otros académicos, que eso se replica posteriormente también en el desarrollo académico.

Conversando con varios académicos, varios tienen intuiciones o teorías respecto a por qué se da este fenómeno, y por qué no se replica en otras carreras de nuestra misma Facultad. Nos interesaba mucho conocer más a profundidad las distintas perspectivas que hay al respecto, y en este caso la suya. ¿Por qué cree que se da esto en la filosofía? Sabiendo que no es único a la filosofía, pero que sí parece ser un caso destacable y que preeminentemente se da de tal modo.

Luis Placencia

Mire, el ¿por qué? es una pregunta que yo no puedo responder, digamos. Es decir, ahí estoy, como dicen los alemanes, *sobrepreguntado*. No tengo las competencias para responder eso, no sé si alguien las tiene por lo demás. En cualquier caso creo que es un problema, en eso estamos todos de acuerdo. Es un problema y hay que intentar solucionarlo, porque evidentemente el talento está repartido de manera relativamente equitativa. En ningún sentido es el caso que los hombres tengan más

talento, las mujeres menos, etc. Es un fenómeno que ocurre efectivamente en muchos ámbitos. Hay un ámbito, ustedes saben, con el que tengo una cierta relación, por mis pasiones personales, y en el que ocurre de manera muy dramática, que es el ajedrez. Y es muy curioso. Yo tengo una hija que juega ajedrez, y a veces le tengo que explicar a algunas personas por qué hay premios femeninos. Porque muchas personas dicen "el ajedrez es un deporte en el que la condición física, la fuerza, digamos, no juega ningún papel". Entonces, lo más natural sería pensar que no hay ninguna razón para que existan competencias separadas. Es razonable pensar que, por ejemplo, en el tenis, los hombres no compitan con las mujeres porque la fuerza física, por razones fisiológicas, es completamente distinta. *Sacar* más fuerte ya da una ventaja. Sacar desde más arriba también, qué sé yo. Y es razonable.

Es como la diferencia de peso en el boxeo. Competirán los que pesan 100 kilos con los que pesan 100 kilos; si me ponen a mí a competir con un tipo de 120 kilos, me noquea en el primer round, sin ningún problema. Entonces, no es un problema, digamos así, de capacidades para el deporte, sino de que cada cual ejerza dentro de un ámbito en el cual esas capacidades sean comparables, por ejemplo.

Pero, lo que pasa es que en el ajedrez ocurre un fenómeno muy curioso, que es que hasta aproximadamente los 10, 12 años, las niñas y los niños se desenvuelven de manera muy equitativa ¡Muy equitativa! De hecho, a mí me toca ir a torneos escolares muy frecuentemente, y yo veo cómo entre los 10, 12 años, muchas veces, si es que no siempre, las niñas ganan. Hay un par de niñas de 10, 12 años en Chile que son buenísimas. Pero en algún momento pasa algo —que no sé qué es— que hace que muchas veces las niñas dejen de competir. De hecho, hace poco estuve en un torneo *sub-18* donde no había niñas o mujeres. ¿Qué pasa? No tengo idea. Lo que sí creo, como les decía primero, que es un problema. Hay que tratar de entender por qué se produce el problema, porque tomar medidas para resolver un problema suele ser más efectivo si uno entiende

las causas. Obviamente, con esto no quiero decir que no se deba hacer nada hasta que no conozcamos las causas, porque uno va comprendiendo las causas en la medida en que uno hace cosas para solventar el problema. Si hace algo que no funciona, bueno, después intenta otra cosa. Pero, sí hay que tener la suficiente claridad para saber qué es lo que está funcionando, qué es lo que no. A veces hay distintos tipos de agendas que no tienen la claridad suficiente respecto a eso.

Ahora, lo que sí también creo que puedo decir, que es importante, es que, aunque se da esa situación que usted acaba de mencionar, el panorama ha cambiado mucho para mí. Cuando yo entré a estudiar filosofía —tengo la suficiente cantidad de años como para poder hacer un contraste— hace más de 20 años estudié en una facultad en la que había una sola profesora; había menos mujeres que las que hay hoy en filosofía aquí, y en general en el medio. Lo cual, supongo, quiere decir que algunas de las políticas que se han estado tomando han tenido un efecto positivo. Entonces, supongo yo que, más allá de cuál sea la causa de por qué ocurre eso, sí es un fenómeno que está cambiando. Probablemente, como se trata de algún tipo de fenómeno de difícil comprensión y cuyas causas son muy profundas, el cambio va a ser lento. Pero eso no significa que tengamos que dejar de insistir. Porque evidentemente, como decía anteriormente, creo que estamos todos de acuerdo, en que hay un problema. Pero es difícil saber cuáles son las causas. Es muy difícil. Obviamente, creo que hay ciertas cosas que uno intuitivamente podría decir que tienen cierta significación, aunque no sé si sean la causa o una de las causas, pero en cualquier caso son significativas. Por ejemplo, aunque esto está cambiando, la existencia de modelos, en alguna medida, también ayuda a que el cambio vaya profundizándose. Pero, para las mujeres de hace 20 años atrás, y para qué decir las de antes, probablemente era mucho más difícil insertarse en un mundo en el que todas las personas eran en algún sentido completamente distintas a ti, y eso es muy importante.

Como les mencioné anteriormente, yo tengo una hija chica y es lo primero que se da cuenta inmediatamente cuando va a un lugar en el que solo es de la ausencia de mujeres; se dice "¿Y por qué no hay una mujer? ¿Por qué solo hay jugadores varones? ¿Por qué hay solo profesores varones?", es una cosa que le llama inmediatamente la atención. No sé si ese es el caso de todas las mujeres, pero evidentemente hay una cantidad significativa de mujeres y de niñas para las que eso genera un impacto. Y evidentemente el mensaje que hay implícito tras la ausencia de personas como yo es "esto no es para mí". Entonces, yo creo que la presencia de modelos es relevante y en eso me parece que las políticas que coadyuven a que existan esos modelos son importantes.

Kuminak Lefio

¿Qué creería que son los mejores siguientes pasos al tomar esta situación e intentar mejorarla? ¿Cree que agregaría algo más? ¿Qué se le ocurre además de buscar dar modelos a las estudiantes de filosofía y tal vez resaltarlos en la filosofía?

Luis Placencia

Sí. Creo que hoy en día uno de los problemas es que, en general, empezamos a replicar dentro de la filosofía una especie de esquema: ocurre cuando la filosofía se percibe como algo exclusivamente masculino —y lo mismo pasa, por ejemplo, en la ingeniería—, y entonces surgen subdisciplinas consideradas "para mujeres". Eso es lo que veo que está ocurriendo; puedo estar equivocado. Entonces, determinados asuntos, son asuntos que tratan las colegas y hay otros asuntos que tratan los colegas, y creo que eso no está bien. Lo que deberíamos tratar de hacer es de generar incentivos, no sé cuáles, pero generarlos para que eso deje

de ser así. Entiendo que, por ejemplo, en las ingenierías ocurre otro tanto. En las ingenierías ocurre que una cantidad mayor de mujeres va o a la biotecnología o a la ingeniería industrial. Pero en la ingeniería, por ejemplo, eléctrica, matemática o química, en cambio, se empieza a replicar el fenómeno de que quienes las estudian son, en su mayoría, hombres.

Kuminak Lefio

Eso haría que la estadística en general se viese como que se está más equiparado la situación pero en realidad...

Luis Placencia

Es que en algún sentido lo está, digamos. Pero, se vuelve a generar esta especie de fenómeno de separación —o de sesgo de género— en determinadas subdisciplinas. Y obviamente una manera que puede tener uno de reaccionar ante eso es la misma manera como se reaccionaba hace 20 o 30 años atrás en relación a la filosofía, es decir: "que en realidad es una cosa de interés, dejemos que cada cual haga lo que quiera". Pero, verlo así, evidentemente implica soslayar el hecho de que los intereses también están, por ejemplo, vinculados a modelos. El hecho de que las personas sienten, en el fondo, que un lugar no les pertenece, es en la medida en que hay un modelo único que seguir.

Es muy importante tratar de diversificar o tratar de hacer que esta especie de "generación de compartimientos estancos", en el ámbito de la filosofía, deje de ocurrir. Ahí me parece que hay un punto importante, y las colegas también tienen una tarea relevante que hacer, porque hay ciertas subdisciplinas en las que obviamente es tentador involucrarse, porque

hay un punto de partida que evidentemente los varones no tienen: "la experiencia". Es decir, evidentemente, ser mujer —así como ser varón también, pero en este caso estamos hablando de ser mujer— provee una cantidad de experiencias que no son transmisibles de manera discursiva. Por decirlo así, hay un saber, que está constituido a partir de las propias vivencias. Pero, por otro lado, también es cierto que es relevante, por las razones que fueron mencionadas previamente, que las colegas también adquieran una relevancia significativa en otros ámbitos. Y ya lo han hecho algunas en algunos contextos. Pienso en Ruth Barcan Marcus; es difícil encontrar una persona de la relevancia de ella en la historia de la lógica del siglo XX. Entonces, ahí yo creo que hay una tarea también por hacer.

Kuminak Lefio

Conversando con compañeros, existen perspectivas bien variadas en torno a qué lugar ocupa la filosofía en la vida de las personas que se dedican a ella. En un extremo, hay quienes —en un esfuerzo de dividir la vida laboral y la vida personal— tienden a ver a la filosofía completamente como su vida profesional. En ese sentido, los aspectos de su vida diaria o personal son separados de su vida laboral, y en este segundo aspecto sitúan a la filosofía. Por otro lado, hay gente que —en el otro extremo, completamente contrario— les parece aquello imposible. La filosofía, en el fondo, invade y llena cada elemento de sus vidas diarias. En ejemplos extremos, hay gente que dice que "su relación con los autores es prácticamente igual de importante a la relación con su pareja"[5]. Me gustaría saber su perspectiva al respecto de su relación con la filosofía. Dentro de estas caracterizaciones, ¿Dónde cree que se encuentra usted? ¿Qué relación cree que tiene la filosofía con su vida diaria?

[5] Dicho por Ernesto Castro

Luis Placencia

Tiendo más a lo segundo, espontáneamente. Ahora, no en la variante *venérea*[6], digamos. Por supuesto, amo mucho más a mi mujer que a *La crítica de la razón pura*, por razones que creo que son más o menos evidentes. *La crítica de la razón pura* no es un ser vivo; el amor uno lo experimenta, al menos en cierta variante, entre seres vivos. Es una cosa, podríamos llamarlo así, de mamíferos. Uno necesita sentir el calor de otro ser humano, o de otro ser vivo. Tiendo más naturalmente a lo segundo, aunque no sé si eso es bueno. Además, no me parece que haya en todo esto nada que sea bueno de suyo. Por rasgos más bien personales, tiendo a embelesarme con los asuntos, pero eso puede tener también formas *paróxicas*.

Entonces, no creo que haya una sola manera de relacionarse con la filosofía que sea razonable. Entre otras cosas también porque las circunstancias son muy distintas. Creo que cuando uno es estudiante, por ejemplo, es muy positivo tener una especie de relación muy profunda con el asunto —entre otras cosas por las razones que ya mencioné— pero cuando uno es más grande, las cosas cambian. Imagínese usted que cuando uno está a cargo —por ejemplo— de otros seres humanos y no solo de sí mismo la situación es muy distinta. Yo creo que es un error categorial pensar que unas ideas son más importantes que los seres humanos más próximos y amados. Es decir, yo no puedo comparar a mi hija con Kant. No solo porque pienso que mi hija es más importante, sino porque esos dos ítems pertenecen a órdenes completamente diferentes. Es como comparar a *Colo Colo*[7] con mi hija o a *Colo Colo* con Kant.

Pero sí, creo que en general de joven es muy positivo tener una relación en la que uno le da tanto cuanto pueda a la filosofía. Porque primero es una ocupación muy peculiar. Segundo es una ocupación en la que

[6] "Venérea" de "Venus" (atractivo, encanto, deseo, en latín).
[7] Equipo de Futbol Chileno

normalmente tiene sentido ingresar o dedicarse a ella porque a uno realmente le apasiona. Porque si es por ganar dinero o escalar en la cadena trófica social, no es el lugar; para eso hay que ir a otra carrera.

Entonces, si es así, es bueno dedicarle mucho tiempo y tratar de involucrarse. Pasa también que las situaciones cambian. A veces también ocurre que uno está en contextos —por ejemplo— laborales que no son tan favorables. Entonces, en ese sentido, esa es una situación en la que se torna muy positivo hacer una especie de corte entre la vida privada y la profesional.

En suma, no creo que haya una sola manera. Depende siempre de quién es uno, en qué contexto está, cuáles son las circunstancias y en qué fase de su vida está. No hay receta en esto.

Kuminak Lefio

Me interesa hacerle una pregunta teórica, pasando a un punto más complejo tal vez; más filosófico. Cuando investigamos problemas filosóficos, como hemos venido mencionando antes, ¿Cómo podemos distinguir una genuina aporía, una situación realmente paradójica, que se arraiga en algún elemento fundamental de nuestra condición humana y sus limitaciones, de algo que sea meramente una disonancia cognitiva, que se nos presenta en la experiencia como algo aporético, pero que realmente no lo es? Y, en la filosofía, ¿Llegamos a conocer definitivamente cuáles son estas aporías reales? ¿O es un error también pensar esta diferenciación?

Luis Placencia

Creo que hay una inquietud muy profunda tras la pregunta: que es que no todo lo que parece ser un problema filosófico, lo es. Y la pregunta es profunda porque además considera uno de los rasgos formales que nos permite distinguir un problema filosófico de algo que no es un problema filosófico, que es justamente el carácter aporético del problema. Muchas veces tenemos un indicio de eso a partir de una paradoja, o de una situación paradojal. Es decir, del hecho de que tengamos muy buenas razones para sostener dos tesis que son incompatibles no se sigue que la paradoja no presente un problema meramente aparente, o sea un pseuproblema. Creo que eso efectivamente puede ocurrir, y de hecho hay ejemplos de eso.

Ahora, creo que en ese sentido la historia de la filosofía sí provee un insumo que es muy interesante. Porque de alguna manera nos provee o nos da cierta información relativa a —por ejemplo— aporías o problemas que son relativamente permanentes. Hay ciertos problemas de la historia de la filosofía que parece ser que son recurrentes, para no decir permanentes. Es decir, que, cada cierto tiempo, surgen. Mencionábamos anteriormente el problema del tiempo, y un ejemplo era Agustín de la antigüedad tardía, hace muchísimos años. Pero, nadie diría que el problema acerca de la cuestión de la naturaleza del tiempo está resuelto.

Por supuesto tenemos una cantidad de información de la cual Agustín no disponía. No había el desarrollo de la física que hay hoy en la época de Agustín. Pero eso hace que el problema, lejos de estar resuelto, sea más difícil de resolver, porque lo que dice la física sobre el tiempo, no se relaciona de manera tan directa y evidente con lo que las personas piensan comúnmente que es el tiempo. Y creo que quienes sostienen que uno simplemente tiene que darle razón a la física, están operando a partir de un prejuicio filosófico. No digo que sea falsa esa tesis. Puede ser verdadera. Yo tiendo a pensar que no lo es —pero esa es una opinión mía, podemos hablar después de eso. En cualquier caso, esa es una tesis filosófica que requiere algún tipo de acreditación o de argumento. No es

evidente que la visión científica del mundo —si es que la hay— tenga que tener algún tipo de prioridad —salvo en el ámbito de la ciencia— sobre la visión de los seres humanos comunes y corrientes, que en algún sentido tienen una cierta idea compartida de lo que es el tiempo y que es distante de la idea que sostiene la física. A lo mejor con el tiempo se van a aproximar ambas visiones, eso es una posibilidad, pero hasta el momento es distante.

Entonces, lo que yo diría es que ahí tiene uno un criterio —no una regla que permite discernir de manera perfecta, pero sí un criterio— que le permite a uno reflexionar sobre el carácter genuino o aparente de una paradoja. Y yo creo que esa es una de las razones por las cuales, de hecho, la filosofía se ocupa con su propia historia, porque la historia en ese sentido sirve como una especie de método de control de la propia reflexión.

IV. Reflexiones sobre la interdisciplinariedad

Kuminak Lefio

En su experiencia como profesor, ya no solo a estudiantes de filosofía, sino también a estudiantes, por ejemplo, de derecho, también nos interesaba mucho preguntarle si cree que el pensamiento y la enseñanza filosófica tienen un impacto real en la formación y desempeño de otros estudiantes y profesionales de otras disciplinas, o si acaso en ellos la filosofía no llega a permear de la forma que uno tal vez quisiera.

Luis Placencia

Es difícil para mí responder esa pregunta porque supone varias cosas. Supone, por ejemplo, una distancia de mi parte, con lo que yo mismo realizo que no sé si tengo. Espero tenerla, pero no sé si la tengo. Por eso tengo que ser honesto. Evidentemente todos nosotros tendemos a pensar que lo que hacemos es importante. Y cuando pensamos que no lo es, eso nos genera un problema. Podríamos incluso llegar a decir que es un problema existencial.

Entonces, tengo el sesgo de pensar que lo que hago es importante y trato de hacer también las cosas que hago de manera tal que sean significativas para los estudiantes que les toca estar conmigo. Ahora, eso por una parte hace difícil responder la pregunta, y por otra parte es difícil responder la pregunta también porque yo no tengo los estándares de un abogado. De manera tal que me es difícil saber si para la formación de un abogado es relevante lo que nosotros hacemos.

Lo que sí tengo son reflexiones de colegas que sí son abogados que van en esa dirección. Hay un excelente libro —que creo haberles mencionado alguna vez en otro contexto— del profesor Rodrigo Valenzuela Cori, que es un importante abogado chileno que hasta hace poco enseñó en la Facultad de Derecho de la Universidad de Chile. El profesor Valenzuela sostiene en ese libro una cosa que creo que tiene mucho sentido —y la dice él como abogado, no yo como filósofo— que en general la enseñanza del derecho en Chile está excesivamente concentrada en el aprendizaje de códigos y doctrinas, entrenándose muy poco las capacidades hermenéuticas de los estudiantes. Estas capacidades son relevantes porque lo propio del derecho —piensa él y creo que tiene razón— en lo que se refiere a su ejercicio, es justamente el contexto adversarial. Dicho de otra manera, los abogados son personas que se desempeñan en conflictos. El derecho es un modo de resolver conflictos a partir de ciertas reglas dadas que son asumidas como válidas para las partes, y a las cuales las partes se atienen para la resolución de ese conflicto. Pero fíjense ustedes que eso significa entonces que el mero conocimiento de las reglas

y de la interpretación de las mismas no puede servir para resolver el conflicto. Porque si así fuese, o bien no habría conflicto, porque una de las partes se daría cuenta de que no tiene la razón, evidentemente, y simplemente no aspiraría a nada, o bien la resolución sería simplemente algorítmica —usted toma el conflicto, ve la regla, resuelve el problema.

Entonces, ¿qué es lo que hace un buen abogado? Él es capaz de interpretar la regla de manera tal que esa regla, en algún sentido, recoja los intereses de su parte. Y eso es lo que un buen abogado siempre tiene que poder hacer, lo cual obviamente supone que el conflicto se genere en un escenario —fundamentalmente el tipo de conflicto que es más interesante para un abogado— en el cual las reglas pueden, en alguna medida, adaptarse al interés de las partes. Pero eso supone una cierta capacidad de interpretación de las reglas que no viene dada con el mero aprendizaje de las mismas y de la doctrina. Por supuesto es muy importante esto último —sin eso no hay nada. El aprendizaje de las reglas y de la doctrina es el suelo que permite la resolución del conflicto justamente de un modo que no es el de la violencia, el del mero ejercicio del poder o el de la corrupción, que es el tipo de mecanismo que se supone soslaya o evita el derecho. De alguna manera lo que buscan los abogados es que los conflictos no se resuelvan de esa manera —por ¿Quién es más fuerte? ¿Quién tiene más plata? ¿Quién tiene más influencia? ¿Quién es más poderoso en la escala social? o ¿Quién tiene más gente a su lado? — sino que sobre la base de un conjunto de reglas, que además tienen origen legítimo y que surgen del poder delegado del pueblo a los representantes.

Ahora, los conflictos se generan, como decía anteriormente, ahí donde las reglas son susceptibles de interpretación, entonces es muy importante que los aprendan a interpretar. Creo que en ese sentido la filosofía sí puede ser importante para ellos —que es básicamente lo que dice Valenzuela. Pero no sólo la filosofía, también el arte, por ejemplo. Si hay un lugar en el que la hermenéutica o el ejercicio interpretativo se ejerce como en

ningún otro contexto, es en el arte. El arte es un ámbito en el cual, por un lado, hay una gran libertad interpretativa, pero parece ser que esa libertad no es carente de límite. Ahora, los límites son difusos, son mucho más difusos que en otros ámbitos.

Entonces, el mismo Valenzuela plantea el arte justamente como un caso también relevante. Por algo también a la filosofía hermenéutica —piénsese en Gadamer— le interesó muchísimo el arte, como prácticamente a ninguna otra forma de filosofía. El arte es el reino de la interpretación.

Kuminak Lefio

Tal vez ahora siguiendo esta idea, pero desde una perspectiva un poco más general, quisiera preguntarle, ¿cree que todos pueden comprender y sacarle provecho a la filosofía, cuando han intentado atenderla? ¿O es una posibilidad que solo está limitada a algunas personas? Y también, por ende, ¿Tiene sentido como una regla general fomentar el pensamiento filosófico en otras personas y en otras disciplinas, o es una cosa más bien de casos particulares donde cumple rédito o se le saca provecho? ¿Qué piensa usted al respecto?

Luis Placencia

Lo que pasa es que, planteado en ese nivel de generalidad, evidentemente tendría que inclinarme con la segunda opción, pero en un sentido muy amplio. Es decir, evidentemente la filosofía no puede ser significativa para todos los seres humanos en todo contexto. Por ejemplo, para un niño de un año, no puede tener ningún tipo de significación —básicamente porque no está en condiciones de cumplir con una serie de requisitos,

pero que por otro lado son muy elementales. Los requisitos que se deben cumplir para que la filosofía pueda tener algún tipo de significación o de relevancia son menores que los requisitos que se tienen que cumplir para que la astrofísica, por ejemplo, tenga algún tipo de significación.

O la econometría. La econometría a la mayor parte de los seres humanos les interesa muy poco. En cambio, las preguntas de naturaleza filosófica probablemente les interesan a más seres humanos que las preguntas económicas. Esto no quiere decir que la filosofía sea más importante que la econometría. Depende acerca de qué estamos hablando, por ejemplo, si se trata de fenómenos como "calcular el costo de un crédito", me imagino que ahí la filosofía deja de ser tan importante; su importancia tiende a cero en esos contextos. Pero en los contextos apropiados, me parece que la filosofía puede ser muy significativa para muchos seres humanos. Y probablemente puede ser muy significativa, o más significativa, para una serie de disciplinas muy especializadas —eso creo que sí es así.

Esto no quita que se requieran ciertas condiciones de base. El dominio del lenguaje, la capacidad para leer textos y una cierta disposición reflexiva son probablemente indispensables, ya que con alguien que simplemente está obstinado en defender una cierta idea de las cosas no se puede filosofar. Es algo sobre lo que Platón ya reflexionó largamente, a su manera, en los *Diálogos*. En varios *Diálogos*, Sócrates dice: "Con este no se puede conversar". Obviamente esas condiciones se requieren. Asimismo, se requiere estar interesado en descubrir la verdad y compartir una serie de premisas sumamente elementales respecto de la racionalidad humana. Por ejemplo, el principio de contradicción o la idea de que la racionalidad humana nos permite, a partir de una conversación, ponernos de acuerdo en ciertas cosas.

Dadas esas condiciones básicas —y algunas otras más— la filosofía es significativa para personas que no estudian la carrera —el fondo de la

pregunta probablemente es ese— o para personas que no van a estudiar filosofía. Es decir, para personas que digamos están en el colegio, universidades, etc. De hecho, creo que eso que yo llamaría filosofía, en esos contextos, es lo que a veces se confunde con lo que hoy por hoy se llama "pensamiento crítico", pero que es el menos crítico de todos los pensamientos. Lo que se entiende por pensamiento crítico es el aprendizaje de una serie de dogmas, que serían la crítica. Pero, no hay nada menos crítico que transformar la crítica en un conjunto de dogmas. Más bien la crítica consiste justamente en tratar de tener la menor cantidad de dogmas posible.

Y en ese sentido, creo que la filosofía entendida como "pensamiento crítico" en esa significación específica, es decir, el tipo de pensamiento que busca asumir la menor cantidad de dogmas posibles, y reflexionar desde ese punto de partida, por supuesto que puede ser significativa para seres humanos que no estudian filosofía y que no van a estudiar filosofía profesionalmente.

Kuminak Lefio

Deseaba llevar esta misma pregunta a una perspectiva más ética, en el aspecto de la posibilidad de la filosofía, ya no sólo que sea significativa, como usted lo mencionaba, sino que sea beneficiosa y que aporte a vivir una vida buena. Hay situaciones que ponen en duda esta idea, si es obvio que la filosofía sea siempre un bien. Por ejemplo, teniendo en cuenta casos de estudiantes que se enfrentan a diversos problemas de salud mental en el proceso de estudiar la carrera, o casos en los que las reflexiones filosóficas nos pueden llevar a crisis cognitivas de tal magnitud que nos incapaciten. Pensando en estos ejemplos u otros que a usted se le pueden ocurrir ¿Cree que la filosofía o el estudio de la misma son un bien en sí mismo? ¿Es un beneficio para todos los que deciden estudiarla? O, en cambio, ¿cree que hay contextos en que la filosofía puede volverse

dañina o nociva? ¿O cree que en los ejemplos que le mencionaba no hay relación causal o una relación real entre ellos y la filosofía?

Luis Placencia

Yo veo una diferencia, por lo pronto, en tres niveles. Una cosa es la filosofía o la filosofía propiamente tal, otra cosa es el estudio de la filosofía y otra es la carrera de filosofía y el estudio de la filosofía en la carrera de filosofía. Son tres cosas distintas. Hay formas de hacer la carrera de filosofía que yo pienso que son profundamente antifilosóficas, en el segundo sentido del término. Y la tarea de los que estamos enseñando filosofía, en la carrera de filosofía, es que esas dos cosas se acerquen, tanto cuanto se pueda. Por supuesto, ahí nos topamos con una serie de dificultades muy serias, como que las ideas que hay de lo que sería la filosofía son muy distintas. Hay colegas que simplemente piensan que hacer filosofía es hacer lo que hacen los filósofos profesionales y tratar de surfear esa ola.. A mí eso no me parece, no estoy de acuerdo con eso, pero entiendo que hay colegas que tienen esa idea y que son muy capaces, son personas excelentes en muchos sentidos, en tanto especialistas —por ejemplo— y como filósofos también, en alguna medida… para qué decir como profesionales. Sin embargo, no la comparto.

En cualquier caso, creo que —en parte por eso y por otras razones— es relevante hacer esa distinción. La carrera de filosofía no enseña a ser filósofo. Tampoco enseña filosofía. En el mejor de los casos, enseña a filosofar. Y creo que muchas veces no lo logra.

Entonces, si uno toma la filosofía como esta especie de ejercicio de autoesclarecimiento realizado a partir de un punto de arranque que es máximamente distante de dogmas, me parece que la filosofía —entendida así— es probablemente una tarea que es positiva, si no para todos los seres humanos, para una cantidad muy grande de ellos. Creo que todos

nosotros en alguna medida requerimos de entender en algún nivel qué es lo que somos, por qué hacemos lo que hacemos, por qué nuestra vida es como es, por qué le damos significación a las cosas que le damos significación, por qué hay otras cosas que nos parecen irrelevantes, por qué nos hemos equivocado respecto de cosas que nos han parecido relevantes —que en realidad a lo mejor no lo son—, por qué otros le dan relevancia a cosas que nosotros mismos no le damos relevancia, y así. Ese es un tipo de ejercicio que es muy importante para la vida humana en términos generales, y que, realizado de un cierto modo, es un ejercicio que llamaría filosófico. Es decir, ejercido a partir de ciertas concepciones racionales.

Ahora, no estoy seguro de si el ejercicio profesional necesariamente es positivo para todo el mundo, porque lo que hace el ejercicio profesional de ese tipo —el ejercicio profesional de la persecución de ese tipo de preguntas— es algo así como maximizar la tendencia, llevarla a otro nivel, transformarla en una especie de ocupación —llamémoslo así— científica, rigurosa, etc. Y eso no sé si en general es bueno para todo el mundo. De hecho, es más, no estoy seguro de si es bueno incluso para los que nos dedicamos a eso. Pero, probablemente si lo hiciese todo el mundo, la sociedad colapsaría —no tendríamos disponible una cantidad de bienes que necesitamos mucho más para la vida cotidiana que este tipo de reflexión. Entonces, ahí distinguiría entre esos dos niveles.

Y la carrera de filosofía —por supuesto— puede ser terrible para mucha gente. Creo que uno no debería andar por el mundo entusiasmando a la gente que estudie la carrera de filosofía, sino que más bien dejar que llegue aquí la gente que está muy convencida de algo así. Esto no lo digo porque la carrera sea terrible, sino porque si uno realmente no tiene una inclinación y un interés muy fuerte por esto, puede ser una cosa muy desagradable. Con el agravante, de que tiene un "premio" que es muy reducido. En las Facultades de Derecho un montón de gente — estimo que un tercio de los estudiantes que ingresan cada año— son estudiantes

que en realidad no tienen un interés muy marcado por ese tipo de carrera, como por la justicia o por los códigos, o por lo que sea. Simplemente descubren que es lo que se les da más fácilmente y dentro de lo que se les da más fácilmente les permite vivir de un modo que ellos consideran cercano —o que satisface— ciertos estándares. Eso en la filosofía es más difícil, y no hay razones para pensar que vaya a dejar de serlo. Lo cual no significa que uno estudiando filosofía vaya a vivir mal. Significa simplemente que el esfuerzo que hay que hacer para lograr vivir de esto es significativo y es mucho más significativo que el que hay que hacer en otros ámbitos. Y a veces, haciendo unos esfuerzos enormes, muy significativos, uno recibe un "premio" que es menor al que recibe haciendo un esfuerzo similar en otro contexto. Y "premio" no quiere decir solamente aquí el dinero que uno gana, también el reconocimiento, el prestigio social, etc.

Si uno está dispuesto a correr todos estos riesgos porque tiene un amor por las preguntas filosóficas y las discusiones en torno a ese tipo de asuntos, y por la escritura o la lectura o ambas cosas, puede ser muy positivo estudiar filosofía. Si no es ese el caso, lo desaconsejaría —sin duda.

V. La Filosofía y las nuevas tecnologías

Kuminak Lefio

Llegando al último eslabón de nuestra entrevista, nos gustaría partir preguntándole ¿Usted cree que la filosofía requiere de la tecnología para desarrollarse? En su opinión, ¿A medida que se desarrollan diferentes tecnologías, la filosofía se estudia siempre un poco "a la antigua" o se mueve con los nuevos desarrollos?

Luis Placencia

Por la razón mencionada anteriormente, es decir, aquellas que indican que la filosofía se da a sí misma o tiene que buscar desde sí misma sus métodos, en principio, no hay una dependencia esencial como la que podría haber entre la química y las nuevas tecnologías. Las ciencias que trabajan con métodos que van "evolucionando", evidentemente se van enriqueciendo a partir del desarrollo tecnológico.

En la filosofía es distinto, porque nosotros no tenemos una especie de método que se vaya enriqueciendo a partir del desarrollo tecnológico. Lo cual no quiere decir que la tecnología sea irrelevante para nosotros desde todo punto de vista. Es decir, no veo nada negativo, y se me ocurren cosas positivas que podrían tener lugar empleando la tecnología, pero no creo que sea esencial. Más bien, no me situaría en la posición de que el método de la filosofía es el mismo de hace 2.000 años atrás o 2.500 años atrás, en ningún caso. Tampoco me situaría entre quienes estiman que es una especie de imperativo el emplear nuevas tecnologías. No veo por qué. Al revés, me parece que eso da lugar a tendencias perniciosas.

Un ejemplo que puedo poner es esto, no tiene que ver con el empleo de tecnología en el método filosófico, pero sí en la exposición, y eso revierte, por lo menos, en el método de enseñanza. Hoy por hoy es muy común sostener que una buena clase debe emplear ciertos métodos o ciertos recursos electrónicos o digitales, como por ejemplo una exposición en *PowerPoint*. Creo que una exposición en PowerPoint puede ser muy útil, bajo ciertas circunstancias, y puede ser tremendamente perniciosa, bajo otras. Si lo que yo quiero es conversar con usted y tener un diálogo sobre la base de un texto, el PowerPoint puede ser terrible. Porque de antemano ya le estoy diciendo a usted lo que yo creo que es importante. Cuando la idea, en realidad, es la inversa: que usted, a partir de su propia reflexión, sea capaz de descubrir cuáles son los aspectos relevantes, o

eventualmente descubra aspectos que son relevantes que no estaban anticipados en la preparación de la sesión.

Imagínese una conversación donde llego con un PowerPoint, y digo mis puntos son este, este y el otro. Esa es una conversación imposible, tiránica. Entonces es curioso, pero se toma como una especie de rasgo de horizontalidad, lo que en realidad es todo lo contrario: No hay verticalidad mayor que esa. Para una clase expositiva, un PowerPoint puede ser muy útil porque básicamente sirve para destacar aquellos aspectos que el expositor quiere que el auditor comprenda como esenciales, pero para una clase dialógica —para un seminario— el PowerPoint puede —no digo que lo sea— pero puede ser fatal.

Esa es la impresión que yo tengo al menos. Ha sido un ejemplo, y esta tecnología en realidad no tiene nada de nuevo, es muy antigua, pero creo que uno puede hacer extensivo esta consideración con una serie de otros ejemplos.

Kuminak Lefio

De hecho nos gustaría llevar la discusión a nuestro contexto más actual. En el cual cada vez se hacen más presentes las herramientas de inteligencia artificial en nuestra práctica filosófica. Tanto en el estudio como en la investigación. Al mismo tiempo estas herramientas mejoran su efectividad y se vuelven cada vez más indistinguibles del trabajo hecho sin inteligencia artificial. A propósito, vale la pena hacer el *disclaimer* de que estas preguntas de la entrevista fueron hechas sin inteligencia artificial.

Profesor Luis Placencia

Y yo tampoco estoy respondiendo con uso de IA.

Kuminak Lefio

Acerca de este tema, me gustaría preguntarle ¿Cómo cree que el uso de la inteligencia artificial afecta a la adquisición de pensamiento filosófico en el estudio y la investigación? Y ¿A qué efectos positivos y negativos cree que hay que prestarle atención con respecto a esta nueva tecnología?

Luis Placencia

No tengo cómo saber todavía cómo afecta directamente porque es algo muy nuevo, pero hay ciertas ideas que puedo tener como hipótesis, y que creo que son hipótesis bien fundadas. Tengo la impresión de que si la inteligencia artificial se torna una especie de herramienta empleada masivamente para sustituir ciertas tareas que son esenciales para el desarrollo de la capacidad de reflexionar filosóficamente, el efecto que va a tener es muy pernicioso. Porque, por supuesto, es muy útil tener instrumentos que permitan descargarnos de ciertas tareas que pueden ser tediosas. Es muy útil, por ejemplo, tener un instrumento que nos permite corregir rápidamente la ortografía de un texto, pero eso no sustituye la tarea de comprensión y de conocimiento del idioma que implica el aprendizaje de la ortografía.

Entonces, creo que como instrumentos de suyo, como todos los instrumentos, no son ni negativos ni positivos. Todo depende del uso que hagamos de ellos. Si hacemos de ellos un uso que implique que vamos a perder ciertas habilidades —que el desarrollo de la historia humana nos ha mostrado que son relevantes— me parece que puede ser negativo. Por ejemplo, el aprendizaje de idiomas. Si alguien cree que la inteligencia

artificial puede sustituir el aprendizaje del idioma por uno mismo, me parece que está equivocado. Puede ser un instrumento muy útil, pero ¿cómo uno va a discernir si es que el instrumento está ejerciendo bien la tarea o no si uno mismo no conoce de alguna manera la lengua? No se puede. Por supuesto, es maravilloso poder hacerse una idea de lo que dice un texto escrito en macedonio, porque ni yo probablemente, ni ustedes tampoco, nos vamos a dar la tarea de aprender macedonio para leer un texto, pero es importante tener las capacidades que permitan discernir entre lo que significa hacerse una idea muy general de lo que dice un texto y haberlo leído en serio: Para leerlo en serio, uno lo que tiene que hacer es conocer la lengua.

Entonces, si la inteligencia artificial se va a transformar en una especie de instrumento que nos va a hacer pensar que esas habilidades son soslayables, como se quejan algunas personas que ha ocurrido con ciertas formas de la calculadora… La calculadora es fantástica, pero uno tiene que también ser capaz de hacer el cálculo por sí mismo. De lo contrario se pueden producir problemas muy serios. Es decir, uno tiene que ser capaz de entender los procesos que eventualmente son sustituidos por la máquina. Porque, de lo contrario, vamos no solo a depender de manera completamente excesiva de las máquinas, sino que en alguna medida vamos a ser incapaces de corregir nuestros errores. Las máquinas también pueden errar. Hay ciertas operaciones matemáticas que una calculadora de ordenador, por de pronto, las hace mal. Por ejemplo, las reglas del paréntesis; algunas calculadoras de móvil no las logran aplicar o no las aplican adecuadamente.

Entonces, me parece que requerimos de no perder al menos ciertas habilidades. Con esto no quiero decir que todas las habilidades sean relevantes, pero tengo la impresión de que además no estamos en condiciones de discriminar con tanta facilidad cuáles son las habilidades que son relevantes y cuáles no. Recién el paso del tiempo nos va enseñando cuáles son esas habilidades relevantes. Probablemente el

hecho de que haya algunas que hayan permanecido —siendo enseñadas, etc.— durante siglos es una muestra de su relevancia. Como, por ejemplo, escribir a mano. La cantidad de evidencia que hay ahora de que la escritura a mano maximiza la comprensión, frente a la escritura en teclado, es muy grande. Pero eso es una habilidad que se está perdiendo. Si perdemos eso, vamos a perder una cosa, en mi opinión, valiosa, cuyo valor incluso probablemente no hemos terminado de comprender.

Es lo que ha pasado también con la memoria. Durante décadas hemos sido entrenados en la idea de que el aprendizaje y el entrenamiento de la memoria era una cosa simplemente tonta, que eso podía ser sustituido por las máquinas. Es cierto, es muy valioso poder sustituir ciertos aspectos de la memoria por las máquinas, pero eso no implica que el ejercicio y el entrenamiento de la memoria por nuestra parte sea irrelevante; un ser humano sin memoria es un tonto.

Kuminak Lefio

Una pregunta tal vez un poco más entretenida. ¿Cómo habría cambiado su valorización o su disposición frente a la entrevista —y a esta conversación que hemos estado teniendo hasta ahora— si hubiese estado escrita con IA?

Luis Placencia

¿Usted dice las preguntas?

Kuminak Lefio

Si, las preguntas, no las respuestas.

Profesor Luis Placencia

Bueno, depende, porque, por de pronto ustedes están aquí de cuerpo presente, de manera tal que no tengo solo unas preguntas, sino que tengo a otras personas que —a veces sobre la base de lo que digo— reaccionan espontáneamente, me preguntan, etc. Además, los estoy observando a ustedes; eso tiene unos efectos que son muy significativos, propios de la conversación cara a cara.

Y además, en ese caso tendría que suponer que aunque ustedes hayan formulado o hubiesen formulado las preguntas así, en alguna medida las controlaron. Creo que no hay nada problemático en —por ejemplo— hacer un cuestionario de esa manera. Uno tiene que controlar si eso está bien hecho o está mal hecho, entre otras cosas. Y para eso uno tiene que tener habilidad; si uno no tiene la habilidad no puede controlar nada.

En cualquier caso, si eso hubiese sido así, me hubiese parecido preferible que ustedes hubiesen formulado las preguntas por ustedes mismos. Porque entiendo que la formulación de una lista de preguntas en alguna medida responde a una especie de idea de anticipación de lo que uno espera que sea una conversación. Y al menos, lo que la experiencia me indica, es que es mucho más interesante conversar con seres humanos que con Chat GPT. Más allá de que Chat GPT sea sorprendente y en muchas ocasiones parezca como si fuese un ser humano —todo eso es cierto, no lo niego—, hay algo acerca de la imprevisibilidad de la inteligencia humana, que probablemente la inteligencia artificial no tenga y sea muy difícil que alguna vez lo tenga —básicamente porque está muerta, no tiene intereses ni sentimientos. Más bien, aquí podría decir cosas que a ustedes los puede entusiasmar o herir o interesar, nada de eso es algo que puede tener lugar en el caso de una máquina.

Y eso vale también para el set inicial de preguntas. Yo supongo que el set inicial de preguntas —por ejemplo— está generado a partir de ciertos intereses, de una cierta visión del mundo, etc. que es algo muy difícil de esperar en una máquina.

De hecho, mis respuestas, por supuesto, en algún sentido suponen que yo anticipo en cierta medida cuáles son esos intereses y esa visión del mundo. Entre otras cosas porque también los conozco a ustedes. Entonces, no estoy en una situación en la que converso con dos perfectos desconocidos, sino con dos personas que sé en alguna medida quiénes son, qué es lo que piensan, por qué están haciendo esto. Y creo que toda esa información es muy significativa en un diálogo, en algún sentido lo orienta. Probablemente si yo hubiese sabido que esas preguntas están formuladas por una máquina, podría haber respondido o hubiese respondido de manera distinta. Incluso a lo mejor hasta con una chanza.

Kuminak Lefio

Me interesa hablar un poco más sobre sus prácticas más comunes o rutinarias en el estudio de la filosofía. ¿Cómo suele estudiar filosofía hoy en día? Y ¿Hay alguna herramienta digital o tecnológica que le gustaría que existiese para apoyar ese estudio de la filosofía pero que todavía no existe?

Luis Placencia

Mire, interesante. A mí me pasa en esto —probablemente porque soy un viejo en muchas otras cosas— que yo trato de depender lo menos de la tecnología, ya que a mí me tocó vivir el cambio. Yo empecé a estudiar filosofía en una época en que uno prácticamente no usaba instrumentos

digitales; estaba el computador —ya existía— y uno escribía en el computador. Yo alcancé a escribir trabajos a máquina en el colegio. Y uno escribía a máquina —cuando tenía máquina—, a veces lo escribía a mano cuando era más chico. Pero alcancé a escribir muchas cosas a máquina de escribir.

Entonces, cuando llegó la revolución del Internet eso fue tremendo, porque empezaron a aparecer una cantidad de instrumentos que además cambiaron con el paso del tiempo de manera rapidísima. En un comienzo fueron fascinantes.

Cuando yo estudiaba, experimenté —imagínense ustedes— la dificultad de conseguir un *paper*: uno tenía que ir a la biblioteca y el paper debía de estar ahí. Tenía que estar en el papel, ahí. Si no estaba ahí, había unas pocas opciones. Existía *JSTOR* ya y esas cosas; había unas pocas bibliotecas digitales, pero uno tenía que estar suscrito; no había estas páginas que permitan el pirateo y eso. Entonces, conseguir un paper era una cosa tremenda. Yo tuve profesores que me trajeron un paper fotocopiado de viajes que hacían. Le pedía al profesor: "Por favor, fotocópieme este paper". Y él iba, fotocopiaba el paper y me lo traía muy amablemente —algunos eran lo suficientemente amables como para hacer eso—, y cuando llegaba el paper era una cosa que uno estaba esperando desde hace meses. Entonces, no era como ahora que uno hace "clic" y descarga una cantidad de información impresionante. Me pasó que conocí eso.

También, conocí la *Stanford encyclopedia of philosophy* después de cuando empecé a estudiar —si existía antes, estaba en pañales y no era una cosa de uso recurrente—; conocí las bibliotecas digitales; conocí *Zotero*, que era una cosa extraordinaria, lo usé muchísimo tiempo; conocí todos los dispositivos o programas de escritura que son extraordinarios — Scrivener, por ejemplo, que es una maravilla.

Entonces, pasó que yo mismo noté que mi propia manera de estudiar empezó a cambiar mucho y empecé a perder cosas que tenía antes y que creo que eran más valiosas: Que era simplemente estar sentado con el libro durante muchas horas sin distraerme con nada más. Entonces, el esfuerzo que he estado más bien intentando hacer es el esfuerzo inverso, el esfuerzo de —a lo mejor es un error, pero ustedes ven que yo tengo fichitas de papel en mi oficina, y en mi casa tengo miles— que es el ejercicio simplemente de sentarse con el libro, escribir a mano y tomar apuntes en fichas como esas, que son básicamente un papel A4 partido en dos, y que corto con una regla. Y eso me sirve. Es cierto que a veces tiene un componente un poco desesperante, porque tengo que acordarme de dónde está la ficha que escribí, y cuándo la escribí. Tengo un montón de libretas también en la casa —como esta— que después tengo que sacar las hojas y ficharlas. Pero ¿sabe qué?, prefiero eso que ponerlo todo en un dispositivo digital donde después a veces ni siquiera me acuerdo qué escribí. Aunque, con las cosas que escribí en papel, me pasa también. A veces tengo que revisar mi ficha y cuando reviso la ficha me acuerdo, y a veces me acuerdo de que leí cosas que ya no me acordaba que había leído. Imagínese usted después de veintitantos años estudiando filosofía, pasa que uno no se acuerda de que leyó algo que en realidad había leído.

Entonces, en realidad, si usted me dijera qué quisiera yo que se desarrollara que no existe, le diría que no sé. Porque no he explorado, de hecho, lo que ya existe en su totalidad y no sé si alguien lo habrá hecho. La cantidad de cosas que hay es impresionante. Es muy fascinante. Y lo que hay, trato en alguna medida de no ser tan dependiente de ello. Ahora, lo que sí creo que sería muy deseable, pero no sé si depende del desarrollo de alguna herramienta tecnológica específica, sí creo que sería muy deseable el que dispusiésemos de un acceso más prístino y libre a la producción y el trabajo científico. En general, creo que todo lo que se pueda hacer en favor de la ciencia abierta y el desarrollo de plataformas de esa naturaleza —de software de esa naturaleza, software libre, etc.—, en general es muy positivo.

Por supuesto, no es agradable tener que depender de medios que son ilícitos para trabajar. En países como el nuestro, dependemos enormemente de eso. En otros países, donde la cantidad de recursos que hay es mucho mayor, no es así. Me gustaría mucho que pudiésemos disponer de bases de datos lo más centralizadas posible —creo que no las hay—, y que sean bases de datos que no sean ilegales, es decir, que uno acceda al trabajo de otros a través, ojalá, de una sola plataforma, y no tener que estar buscando en cuatro o cinco de ellas dónde puede estar el paper, y que además esa plataforma no esté desarrollada de manera ilícita. No porque tenga yo un problema con que sea ilícita esa plataforma, me parece que en realidad el problema es el sistema de reglas que ha incentivado que exista ese tipo de plataformas. Es decir, en realidad lo que hacen las editoriales es estafarnos a todos nosotros, porque, publican con cargo normalmente a proyectos que son financiados a partir de las rentas generales; trabajo que es hecho por otras personas; que es revisado por otras personas; y cobran una cantidad enorme a esas mismas personas para que lean esos trabajos. Entonces, eso es absurdo; no tiene ninguna lógica desde ningún punto de vista. El trabajo, en mi opinión, que hace la editorial es simplemente poner eso en un cierto medio en el que esté a disposición de todo el mundo. Eso es algo que puede hacer cualquier ser humano relativamente bien entrenado hoy por hoy, que esté —al menos— en el colegio.

Creo que eso es lo que yo esperaría, pero eso depende no de cambios tecnológicos sino de cambios legales. A lo mejor me falta imaginación, para pensar cosas. Pero es natural, porque crecí en un mundo en el que ninguno de estos dispositivos existía y lo curioso es que creo que podíamos estudiar bien. Si usted me apura, de hecho, en algún sentido podíamos estudiar mejor. Porque, ¿qué es lo que ha cambiado? Lo que ha cambiado hoy es que tenemos acceso a una cantidad enorme de información. Pero, por otro lado, es una cantidad de información que nunca vamos a poder procesar. Imagínese, que incluso con los libros de papel que ya tengo, y he comprado, tendría para leer toda mi vida. Y no

me sentiría, por otro lado, feliz delegando la tarea de la lectura de eso en un resumen que me haga una máquina porque tengo la impresión que —al menos lo que a mí me interesa genuinamente— lo realmente valioso, es leer y entender eso, y que está dado por el proceso de lectura y de comprensión, no simplemente por el resumen de la información. Por supuesto es útil disponer de eso. Con esto no estoy, de nuevo, queriendo hacer una especie de discurso antitecnológico. Es útil tener un resumen; uno puede hacerse una idea más próxima y eventualmente descartar alguna lectura o favorecer otra. Ayuda en alguna medida en el mare magnum de la cantidad de material que hay el tener esa herramienta.

VI. El Futuro

Kuminak Lefio

Me pareció muy interesante y valiosa la respuesta respecto al acceso abierto y voy a seguir intentando un poco que aplique la imaginación para orientarse hacia el futuro. Me gustaría preguntarle cómo se imagina que será nuestro futuro en unos 50 años más. Y en concordancia con ello, más específicamente, ¿cómo cree que se estudiará y desarrollará la filosofía en 50 años más? ¿Cree que en ese futuro se seguirá haciendo filosofía o se imagina que podría cambiar tan radicalmente, por el acervo de estas nuevas tecnologías, que en su opinión dejase de poder considerarse filosofía? Entendiendo que la filosofía tiene una historia más larga en comparación con estos 50 años que le propongo, pero que también estamos experimentando como los cambios que experimentamos son cada vez más radicales.

Luis Placencia

A ver, imaginar el futuro en 50 años más depende del ámbito; hay ámbitos que me lo imagino y otros que no sé muy bien cómo van a ser. Tampoco tengo el conocimiento como para hacer ese tipo de ejercicios de imaginación. Tomo como ejercicio de imaginación en general estas anticipaciones que hacen lo que hoy día llaman futurólogos, es decir, no le asigno otro valor que no sea ese. Porque el ejercicio, digámoslo así, anticipativo ha existido siempre y siempre ha existido gente en ese ámbito que ha acertado y gente que se ha equivocado, desde Julio Verne, y probablemente de antes. Entonces, hoy día —justamente tomando en cuenta la cantidad de tecnologías nuevas y de conocimientos que ellas implican, el desarrollo de las mismas y comprenderlas— no estoy en condiciones de imaginar qué tecnologías se van a desarrollar, cuáles son las más próximas, etc. Simplemente tengo ahí conocimientos de segunda mano; cosas que dicen otros, que además no siempre son consistentemente entre sí.

Ahora, sobre la filosofía propiamente tal, tengo la impresión de que es difícil que la filosofía desaparezca del todo, digamos, en el sentido de que no reflexionemos más sobre nuestra propia naturaleza; sobre el mundo y nuestro entorno; sobre el modo en que nos relacionamos con él; sobre nuestro propio destino, etc. No creo que eso sea posible, y no creo que la tarea de la reflexión en general sea sustituible por insumos tecnológicos, porque en general, además, no es una tarea ni ancilar, ni mecánica. De hecho, es más bien el tipo de tarea que uno emprende cuando se libera de ese tipo de ejercicios; cuando uno ya cocinó; hizo el aseo; ordenó la pieza y se procuró algo para comer, recién ahí puede tener tiempo para este tipo de tareas. Entonces, no lo veo como algo que vaya a desaparecer de manera tan simple. A lo mejor va a cambiar el modo como se ejerce. De hecho, vemos que hay cosas que están cambiando, por ejemplo, el modo como se ejerce la divulgación filosófica, que evidentemente es muy distinto a cómo se hacía antes. Y en eso la tecnología sí ha incidido. Hoy en día hay una cantidad de divulgación de tipo audiovisual, que es enorme, y que por supuesto es muy distinta a la divulgación "antigua" —

la divulgación se hace desde hace mucho tiempo. El famoso libro de Jaspers, por ejemplo, son programas de radio. La famosa disputa de Bertrand Russell con Copleston sobre la existencia de Dios, también fue un programa de radio. Entonces, divulgación, se ha hecho desde antaño, en distintos niveles y de distinta calidad, y hoy día evidentemente, por ejemplo, los medios audiovisuales —y para aquí decir, las plataformas de redes sociales; Instagram; esas cosas— sirven muchísimo. Hay gente que hace divulgación en ese plano, y la divulgación es muy importante, es crucial.

Hay gente que hace divulgación en ese plano y utiliza mucho esas plataformas, o los podcasts, etc. Yo creo que allí hay todo un mundo que probablemente va a cambiar mucho, por de pronto el primer acceso a la filosofía. Ocurría antiguamente que el primer acceso a la filosofía era dado por la compra de un libro en una cuneta —bueno, ese fue mi primer acceso a la filosofía—; un libro de dudosa procedencia; de dudosa calidad la traducción… uno leía eso.

Esto lo digo también porque hay mucha gente que se queja, que dicen "No, esto que dicen estos divulgadores es terrible". Bueno, puede ser. Hay divulgadores de todo tipo, unos son mejores que otros, pero en cualquier caso hacen algo —¡Hacen algo!— que yo por de pronto no hago, y que muchas otras personas no hacen, que provoca que la gente eventualmente desarrolle interés en hacer filosofía. Entonces lo encuentro positivo, y ahí creo que ese primer acceso probablemente está cambiando mucho. Probablemente va a cambiar en el futuro. Me imagino que el nivel de diferenciación con el que se va a llegar al público va a ser mucho mayor. Y eso evidentemente va a preparar una dificultad para el ejercicio más genuino de la filosofía, porque intentar integrar todo ese "mundo de diferenciación" va a ser más difícil que lo que ya es hoy, que ya es sumamente difícil, en cualquier caso, mucho más difícil que lo que era en el siglo XIX, o en el siglo IV a.C. De manera tal que evidentemente eso va a generar un desafío enorme —que creo— hace que sea muy factible

pensar que la filosofía como tal no necesariamente va a tender a desaparecer, sino que va a tender a tener más bien una mayor dificultad en el ejercicio de su tarea. De hecho en algún sentido —esto que les mencionaba— presenta ese tipo de dificultad: ¿Cómo alguien que se dedica a la filosofía, a la filosofía académica o lo que sea, integra el tipo de inquietudes que muchas veces se manifiestan a través de la divulgación, a través de medios audiovisuales? Que algunas de ellas son muy filosóficas, otras están a lo mejor más alejadas, pero tienen algún tipo de sustrato filosófico, y evidentemente eso genera una dificultad, pero a la vez un desafío para el ejercicio de la tarea de la filosofía académica.

Entonces, sí. Me imagino, por lo menos, un ejercicio mucho más diferenciado y localizado, por ejemplo, de la llegada de la filosofía en las formas más simples o más de divulgación al gran público. Es decir, si usted quiere, puesto de otra manera, a través de diferentes plataformas y de diferentes medios, va a dejar de existir, si es que ya no dejó de existir, eso que antiguamente se llamaba "el gran público", como ha dejado de existir eso que antiguamente se llamaba "la opinión pública". El uso del artículo determinado, de hecho, ya no tiene ningún sentido en ese plano. Usted tiene en las redes sociales, e incluso en los medios de difusión masiva, más bien fenómenos de segmentación: la búsqueda de públicos que parecen tener cada vez menos cosas en común entre sí, lo cual de nuevo genera una dificultad, que en algún sentido es de índole filosófica. Piense usted, por ejemplo, que en el plano político, ese tipo de fenómenos generan la dificultad de ¿cómo entonces nos concebimos, por ejemplo, como una nación, todavía como una unidad? Ahí donde parece que no tenemos nada en común entre aquellos individuos que formamos parte de una nación. Si no es ello mismo una dificultad de naturaleza filosófica, genera dificultades de naturaleza filosófica, o se vincula con problemas de naturaleza filosófica, por ejemplo, el problema de por qué obedecemos a un determinado orden político. Ya Hume decía: ¿no hay algo misterioso en eso, de que obedezcamos a un poder que en realidad solo está fundado en nuestro poder, pero que en algún sentido se opone muchas veces a

nuestra propia visión? Sin embargo, parece ser que por alguna razón lo hacemos. No estoy queriendo decir que eso sea estúpido, lo que estoy queriendo decir, más bien, es que por misterioso que sea, ocurre eso: Que el poder de la palabra de un individuo, o de dos o tres individuos —o de 120 individuos— termina resultando mandatorio para millones de individuos, cuya mayoría, por lo demás, suele pensar que esos individuos, son unos truhanes; gente que lo único que hace es no hacer nada, perder el tiempo, robarse el dinero, etc. Es cosa de ver cualquier encuesta, digamos. Pero curiosamente, más allá de que pensemos eso, o de que muchos piensen eso, cuando llega el día de la elección, vamos todos ahí en filita, nos ponemos en la filita, y ponemos el papelito en la urna.

No quiero decir con esto que hacemos una estupidez, lo que quiero decir es que opera ahí un tipo de fenómeno de obediencia política, que parece tener razones muy de fondo, y que uno —creo yo— tiene que preguntarse cuán sostenible es a partir de los cambios que les estoy mencionando. Parece ser que en alguna medida ese fenómeno está vinculado al hecho de que tenemos una cierta idea de que nos gobernamos a nosotros mismos entre un conjunto de iguales. ¿Cuán sostenible es eso si a medida que pasa el tiempo nos vamos considerando cada vez más desiguales, como cada vez más lejanos a lo que nos gobierna? Es un problema que evidentemente la nueva tecnología también genera, porque "la opinión pública" servía de modo de construir esa idea de una "identidad común". Por ejemplo, cuando era chico, o era joven incluso, uno llegaba a la casa y las familias veían el noticiario, muchas veces después de la teleserie —había la teleserie de la tarde y después venía el noticiario, eran tres, cuatro noticiarios nada más, en los cuales básicamente se pasaban las mismas noticias. Eso no existe ahora. Entonces, el noticiario en alguna medida modelaba la opinión común.

Hoy día la experiencia que tenemos del mundo común está mediada con nuestras distintas "cajas de resonancia". Si ustedes me preguntan a mí, por ejemplo, ¿Quién es famoso hoy? No tengo idea, pero no porque no

vea medios. Veo medios, pero los medios me dan a mí lo que ellos suponen que yo quiero ver, y con cierto acierto. No me envían noticias de "Balbi El Niño"[8], el que expulsaron del comando de Jeanette Jara, que yo supongo que debe ser una persona muy conocida, pero de cuya existencia yo me enteré el día de ayer o antes de ayer, por un fenómeno, digamos, electoral.

Kuminak Lefio

Sí, yo igual. Yo no sabía quién era.

Luis Placencia

Bueno, eso es índice de que ya tienen sus años, ¿ven? Claro, porque ese es el modo en que la política busca llegar al público joven.

Yo me enteré, por ejemplo, de la existencia de Dua Lipa hace un par de años atrás. Y me enteré de que era conocidísima, y no tenía ni idea. También me di cuenta de que, de hecho, había escuchado canciones de Dua Lipa, sin saber que eran de Dua Lipa.

Kuminak Lefio

Como síntesis de toda esta sección, planeé contarle acerca de una situación de un evento, que tal vez ya conozca, pero se lo contaré de todas formas, que ocurrió hace un par de años en el "Congreso Futuro". La introducción dice así: ¿Imagina ir a tomar una taza de café con su autor favorito fallecido hace más de 100 años? Eso ya no es algo imposible.

[8] El nombre real es "Balbi El Chamako"

Porque este lunes, sobre el escenario del Congreso Futuro, se revivió a nada menos que a Charles Darwin, quien fue entrevistado en el llamado Metaverso.

Y efectivamente, se trataba de un intento de replicar a Charles Darwin por medio de una inteligencia artificial que había sido entrenada en sus escritos, en sus intérpretes de sus escritos y otro tipo de textos relevantes. Y se le preguntaban a Charles Darwin, bajo esta simulación de inteligencia artificial —que podía responder, igual que se puede hacer con Chat GPT—, preguntas de todo tipo, incluyendo preguntas muy ajenas a su época, por ejemplo sobre qué pensaba sobre el cambio climático o cosas de ese estilo.

Si a usted le propusiesen ahora —o en un futuro con inteligencia artificiales aún más potentes— este tipo de conversaciones, pero con filósofos importantes que hace mucho tiempo fallecieron. Por dar algunos ejemplos, una discusión de los temas que usted quiera con Kant, Husserl u otro filósofo que se le venga a la mente. ¿Cree plausible la idea de que estas tecnologías de alguna forma logren recrear con fidelidad esta idea contrafáctica de "lo que diría el autor si estuviese vivo"? ¿Cree que este tipo de herramientas podrían llegar a reemplazar la necesidad de leer a los autores en el futuro? ¿O sería un error asumir aquello? Y más en general, ¿qué opina de esta forma de mirar "la historia de las ideas humanas" que adopta, por ejemplo, pero no únicamente, instancias como el "Congreso Futuro"?

Luis Placencia

Me parece que es un ejercicio muy divertido. La idea es divertida por sí misma, ¿no? y me imagino que sería divertido participar de la conversación. Ahora, en ningún caso lo creería un sustituto del ejercicio de la interpretación de los textos de los autores, por una razón bien

simple: porque que hay una diferencia entre leer por uno mismo un texto —además en el orden en el que el texto está dispuesto, si uno quiere, eventualmente por el autor, etc.—, que el de depender de una mediación. Ustedes sabrán que yo mismo, pero no solo yo, sino que mucha gente, desde el tipo de aprendizaje al menos en el que yo me formé, tratamos de hacer énfasis más bien en lo contrario; en la idea de que uno tiene que tratar de depender de la menor cantidad de mediaciones.

Siempre existen las mediaciones; las mediaciones no son sustituibles. Por ejemplo, nosotros leemos los textos de unas ediciones que cambian una serie de cosas. A veces presentan como libros cosas que no son libros —digamos, cosas así—, pero, uno tiene que saber eso y eventualmente tratar de ser consciente de aquellos aspectos. Ahora, con eso no quiero decir que este tipo de ejercicio sea completamente inútil, o que no tenga ninguna relevancia. Más bien al revés, creo que puede arrojar mucha información, por de pronto, información también de lo que es capaz de hacer la misma inteligencia artificial —por eso es interesante. Alguien que realmente sepa de *Darwin*, podría eventualmente comparar las respuestas con el conocimiento que él tiene, etc. Ahora, lo que creo que es insoslayable, es el hecho de que *Charles Darwin* está muerto. De manera tal que los procesos que hubiesen sido necesarios para que él se hubiese formado una opinión sobre los fenómenos del presente no pueden tener lugar —hay un *gap*, digamos así, un vacío—, y que habría implicado sin duda alguna una cantidad de aprendizajes, para un individuo como Darwin que ya no ocurrió. Uno no puede simplemente hacer que un individuo supuestamente se forme una opinión del presente a partir de opiniones que fueron emitidas hace 150-200 años atrás. Porque eso supone sostener que ese individuo no hubiese aprendido nada entre medio; revisado ninguna posición o no hubiese adquirido ninguna opinión nueva respecto a fenómenos que, por ejemplo, para él no eran anticipables.

El ejercicio me parece, desde ese punto de vista, un poco ficticio y creo que así no funciona. Pero por otro lado es muy interesante la idea de que uno pueda, en algún sentido, lúdicamente, imaginarse ese ejercicio de conversación. Me parece un instrumento curioso, divertido, eventualmente valioso, pero en ningún caso sustituto del ejercicio filosófico, más histórico a la vez —digamos, más propio—, que es tratar de situarse y de aproximarse, tanto cuanto sea posible, a la mirada del individuo concreto en ese momento y a lo que esa obra significa, no solo para ese momento, sino que para el desarrollo posterior de la obra.

Entonces, de hecho, fíjese usted que pienso que en la lectura de textos del pasado, mucho más importante que tratar de entender la opinión del autor del pasado, es entender el mensaje de la obra. El mensaje de la obra es, en algún sentido, independiente de la opinión del autor. Pero ese mensaje tampoco es accesible trivialmente a partir de la letra —llamémoslo así. Solo es reconstruible a partir del ejercicio de la inteligencia de hacerse preguntas con el texto.

Y, por supuesto, ahí las máquinas pueden ayudar mucho. Uno puede ayudarse con las máquinas buscando pasajes, dónde puede ser relevante leer para encontrar esto u otro. O, a lo mejor, incluso —en este modo que decía usted— pidiéndole a la máquina que haga como si fuese el autor. Pero yo no tomaría eso como la última palabra, sino que más bien simplemente como un insumo. Ahora, la literatura de intérpretes también es eso, en algún sentido: Un insumo a partir del cual uno se nutre de la visión de otros. Pero eso, en ningún caso, sustituye el propio ejercicio.

Eso es lo más importante. Lo que uno haga por uno mismo. Esto vale también para las habilidades, las capacidades, etc., y no hay sustitución.

Kuminak Lefio

En este aspecto más lúdico que usted menciona, ya que mencionó a Kant como uno de los autores más relevantes, tal vez, para su trayectoria, ¿qué cree que le llamaría la atención preguntarle? Entendiéndolo desde un ámbito más lúdico

Luis Placencia

Es una buena pregunta. No sé si tengo una pregunta que fuese "la pregunta que yo le haría". Probablemente lo que más me interesaría en ese caso, supongamos que yo tengo una máquina que trata de emular el pensamiento de Kant, me surgiría instintivamente hacerle muchas preguntas para tratar de ver cómo piensa. Cuán cerca está del Kant que yo mismo me imagino. Eso probablemente sería mi manera proceder, y las preguntas que haría, en cada caso, dependerían del tipo de respuestas que vaya recibiendo.

Ese ejercicio, es un ejercicio distinto al que usted me dijese: "¿qué haría usted si se encontrara con Kant?" Son situaciones distintas. Y además, de hecho, para esa última pregunta también es variable, porque una cosa es "¿qué haría si me encontrara con Kant y yo supiera que Kant pudo experimentar en algún nivel, o en alguna medida, los sucesos del mundo que transcurrieron desde que él murió y el mundo como es hoy?". Y otra cosa es si supiese que se ha quedado suspendido en 1804. Hay una serie de opiniones de Kant que a mí me parecen aberrantes, y me interesaría mucho —por ejemplo— saber qué pensaría él respecto de esas cosas teniendo una cantidad de información que era imposible que tuviese en ese momento. Las opiniones, por ejemplo, que Kant tiene sobre las "razas" y ese tipo de cosas, son opiniones que eran muy comunes en ese momento. O, por ejemplo, sobre el género, y que evidentemente dependían —si usted quiere— de prejuicios compartidos de la época. De manera tal que, más allá de que a uno le parezca muy lamentable y muy desagradable alguien que sostiene ese tipo de opiniones, uno puede en

alguna medida dispensar el hecho de que esas opiniones hayan sido sostenidas por el simple hecho de que eran opiniones bastante comunes —no digo universales, porque evidentemente no lo eran, pero sí muy comunes.

Lo interesante, por ejemplo, es saber cómo vería esas opiniones un autor como Kant si es que supiese todo lo que hemos aprendido. Eso es un punto probablemente interesante. Pero, eso es una pregunta que yo no le haría a una IA, que simplemente lo que hace es replicar —algo así— como el pensamiento de Kant y después aplicarlo a casos presentes.

Kuminak Lefio

Y sería sumamente interesante, especialmente en el caso de Kant por su orientación antidogmática, averiguar si es que él sería capaz de dejar esos prejuicios a la luz de nueva información.

Luis Placencia

Por supuesto. Por ejemplo, en el caso de las opiniones sobre el género, hay una evidencia masiva, ya existente, que haría que probablemente la mayor cantidad de autores que expresaron esas opiniones—de maneras que implicaban, por ejemplo, sostener que las mujeres eran incapaces de cierta forma de pensamiento, de cierta forma de liderazgo, etc.—, supongo, que la evidencia masiva de la que disponemos hoy, haría que cualquier autor relativamente reflexivo tuviese que al menos reconsiderar la opinión. Por supuesto, estoy seguro que había muchos casos en ese momento que sirven de ejemplo de lo contrario, pero en una cantidad mucho menor por razones estructurales.

Kuminak Lefio

Y menos valorado...

Luis Placencia

Por supuesto. En un contexto en el que el talento femenino no podía florecer como florece hoy, y probablemente —estamos todos de acuerdo— en que no florece en la forma máxima en que podría hacerlo, como tampoco florece el talento de muchos otros seres humanos que están en situaciones de desventaja frente a otros —creo que nadie puede pensar que en Chile el talento solamente florece en las comunas de las Condes, Vitacura y Lo Barnechea, o en Santiago, y que no florece en La Pintana, o en Puente Alto… Yo como "Puente Altino" me sentiría muy ofendido por ese tipo de razonamiento.

Kuminak Lefio

Llegando a la pregunta final de esta entrevista, y agradeciéndole mucho por el tiempo que nos ha dedicado, me complacería el tomarme la libertad de hacer una pregunta aún más fantasiosa. Si en una cantidad indefinida de años en el futuro, entre las nuevas estructuras de una sociedad tan distinta a la nuestra, que nos resultaría prácticamente imposible reconocernos en aquellos humanos que la habitan, alguien encontrase este pedazo impreso o digital de nuestra revista ¿Qué mensaje le dejaría a los humanos de dicho futuro que lo leerían?

Luis Placencia

Difícil pregunta. Porque supone dejarle un mensaje a alguien del cual yo no sé nada, salvo el hecho de que es máximamente lejano a mí… Probablemente le dejaría el mensaje de que traten de saber de nosotros, que hay cosas interesantes para descubrir.

Kuminak Lefio

Muchas gracias, profesor.

Luis Placencia

Gracias a ustedes.

25 de noviembre de 2025
Facultad de Filosofía y Humanidades
Universidad de Chile

HOMÓNIMA
Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 439-452
ISSN 3087-260X
Santiago de Chile

MÉTODOS LÓGICOS Y MATEMÁTICOS EN FILOSOFÍA: PERSPECTIVAS Y MOTIVACIONES[1]

Por Francisco Castro Richter[2]

La precisión suele considerarse una característica excesivamente cautelosa. Sin embargo, es precisamente lo contrario. Las afirmaciones vagas son las más difíciles de refutar. La ambigüedad es la defensa propia del oráculo. Ser preciso es facilitar al máximo que otros demuestren que uno está equivocado. Eso es lo que requiere el coraje. Pero la comunidad puede reducir el coste de la precisión teniendo en cuenta que los errores precisos suelen contribuir más al progreso científico que las verdades vagas.

Timothy Williamson

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima.* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Castro, F. (2026) Métodos lógicos y matemáticos en filosofía: perspectivas y motivaciones. *Revista de Filosofía Homónima, 1*(1), pp. 439-452. https://doi.org/xxxxxxx[Por asignar]

[2] Francisco Castro Richter es egresado de Licenciatura en Filosofía de la Universidad de Chile. Email: fcastror@fen.uchile.cl. ORCID: https://orcid.org/0009-0003-3167-6513.

1. Filosofía formal: antecedentes históricos contemporáneos

La lógica ha sido clásicamente considerada como una rama de la filosofía. Contemporáneamente se ha llegado a considerar, a la vez, como una rama de la matemática. Con antecedentes en nociones leibnizianas, estos desarrollos se asocian al temprano trabajo de George Boole en el siglo XIX, Gottlob Frege y Bertrand Russell en el mismo siglo y en los albores del siguiente, y quienes siguieran esas reflexiones o reflexiones asociadas (Alan Turing, Kurt Gödel, Alonzo Church, etc.).

Estos pensadores, inspirados en los desarrollos de la época en matemáticas, avanzaron en el tratamiento formal y axiomático de la lógica. Esto permitió poner la lógica y al razonamiento sobre la lógica al servicio del desarrollo de las matemáticas mismas en general, además de las ciencias empíricas y la filosofía. Esto incluyó el desarrollo de aproximaciones específicas a la discusión sobre fundamentos de la matemática y la atención a problemas filosóficos tradicionales con ayuda de los nuevos formalismos lógicos. La motivación de fondo, patente en clásicos como Frege y Russell, es la construcción de lenguajes artificiales capaces de purgar nuestras expresiones de los equívocos y ambigüedades, vaguedades e inconsistencias en los usos típicos del lenguaje natural. Mediante esto —se pensaba— podríamos alcanzar la precisión necesaria para someter a escrutinio las ideas de nuestro interés.

Estas discusiones sobre los fundamentos de la matemática, a su vez, sacudieron la conciencia teórica de los pensadores filosóficos y científicos de la época, quienes en su momento estuvieran continuando intensas tentativas por alcanzar fundamentos o justificación sólida para nuestros proyectos epistémicos en general (con la matemática tradicionalmente considerada como "reina" de las ciencias y ejemplar del razonamiento sólido y pulcro, deductivamente válido). Como se

subentiende también de lo antedicho, esto se daba a la par de discusiones sobre los fundamentos de las ciencias y el conocimiento en general y, sin duda además, de la filosofía y su propia naturaleza.

2. Matemáticas en filosofía: ¿Y a mí qué?

Respecto a lo que nos convoca en este primer número de *Homónima*, ¿Qué, en específico, relaciona la lógica y la matemática con la filosofía y sus métodos? ¿Cómo ha aportado y pudiera seguir aportando la existencia y desarrollo de estas áreas al desarrollo de la filosofía? ¿Qué tan importantes para la filosofía serían estos aportes? Sin duda, podría dar la impresión (tanto a quienes están entrando en filosofía o a quienes ya trabajan en ella) de que su relevancia es más bien modesta, sino incluso muy baja o —más dramáticamente— nula (al menos para la filosofía *propiamente tal*, por oposición a lo que vendrían a ser (en contraste, digamos) las ciencias. Eso, al menos en lo que se referiría a los problemas o aspectos más apremiantes, relevantes, urgentes o exaltados de la condición humana (dijera, quizás legítimamente, un crítico). Pero esto, creemos, está lejos de ser así y esperamos que tanto esta introducción, como el contenido presente y futuro de la sección de "Lógica y Matemática" de la Revista Homónima persuada a los más escépticos de ver las virtudes que estos estudios formales pueden ofrecer a la filosofía.

2.1 Filosofía teórica, filosofía práctica y filosofía matemática

La filosofía que implementa métodos matemáticos a menudo se encuentra domiciliada en la así llamada tradición "analítica" de la filosofía contemporánea (y a otras tradiciones asociadas con la misma, como lo son la filosofía "naturalista" o la filosofía —pretendidamente— "científica", además de la filosofía de la ciencia[3]). La impresión recién aludida que

[3] A propósito de la diferencia que existe entre filosofía analítica y filosofía de la ciencia propiamente tal, puede visitarse el artículo de Pablo Contreras Kallens y sus colegas

pudiera dar a algunos la filosofía matemática podría verse extendida a estas tradiciones filosóficas más amplias. La filosofía "anglo-analítica" y sus parientes podrían verse como algo sumamente tedioso, árido y sin un propósito significativo muy claro (¿A quién le importa, al final del día, digamos, la semántica de expresiones como "El Rey de Francia es calvo", o la semántica de los nombres propios, o de los nombres ficticios?).

Pero esto puede ser fácilmente engañoso. Problemas muy específicos en estas tradiciones contemporáneas terminan siendo, a menudo y desde el principio (o posteriormente a su desarrollo inicial), de impacto en asuntos de "mayor calibre" (o relevantes para otras áreas del conocimiento), de modo que eventualmente pueden impactar a los practicantes del área o al público como más apremiantes, urgentes o centrales desde una perspectiva vital, ética o "existencial".

Piénsese, por ejemplo, en las discusiones sobre filosofía del lenguaje y metafísica en la tradición "analítica" con aplicaciones en el ámbito social y político, en las aplicaciones de la epistemología para el diseño de esquemas deliberativos para toma de decisiones político-militares, en las aplicación de la ontología para tecnología informática en este mismo ámbito, en la presencia de investigadores filosóficos justamente en estas áreas por el interés de fuerzas militares en eventuales aplicaciones computacionales de la lógica, el uso de teoría de juegos y de la decisión en el estudio de los procesos de relevancia ética y normativa en general, entre otros ejemplos de altísimo interés humano. Ciertos filósofos muy destacados, tanto asociados como más lejanos a la tradición ``analítica'', han hecho aportes y/o colaboraciones por fuera de la filosofía académica tradicional, impactando (hasta cierto punto al menos) ya sea en ámbitos científicos, de política pública (o política en general) o en

Daniel J. Hicks y C. D. Jennings sobre redes y empleo en la filosofía académica (Contreras Kallens, Hicks, & Jennings, 2022).

discusiones de relevancia moral en general. Resulta muy fácil, ante casos así y otros, pensar que la filosofía *no da lo mismo*.

Esto mismo puede ser perfectamente el caso para cualquier investigación filosófica formal, en principio. Esto no es sorprendentemente si simplemente sopesamos, por ejemplo, la importancia social cardinal de temas como —por ejemplo— el conocimiento, la ética o nuestras concepciones de lo mental, temáticas donde, como puede apreciarse en una revisión siquiera superficial de la literatura, los métodos formales hacen también aparición.

2.2 Algunas gracias de los métodos formales en filosofía

Y entonces, como practicantes de la filosofía y aspirantes al gremio, ¿Qué hay de la importancia de la lógica y los modelos formales en general para lo que hacemos? Sabemos que ni siquiera en la llamada ``filosofía analítica'' la lógica (ni, más ampliamente, la matemática en general) es tan preponderante como lo fue en el trabajo de algunos de sus pioneros o inspiraciones intelectuales iniciales (como en Frege, Russell, Carnap, etc.). Sabemos que la mayoría de los filósofos no procede ni, a menudo, necesita proceder aplicando formalismos lógicos o siquiera esquemas de argumento con premisas y conclusiones enlistadas.[4]

Aun así, la formalización puede tener una importancia doble:

1. Un entrenamiento básico en lógica formal facilita la entrada a la práctica de la lógica informal y habilidades de pensamiento crítico

[4] En cualquier caso, sí es el caso que el razonamiento filosófico ordinario es constreñido por reglas, y estas podrían corresponder (o no hacerlo) a reglas lógicas. A modo de ejemplo, el razonamiento por casos está fundado en cierta clase de estructura lógica que lo respalda (aunque sin que sea necesario apelar a formalismo al emplearlo). Agradezco al Dr. Felipe Morales Carbonell por esta acotación.

que son fundamentales en la práctica argumental de la filosofía. Esto, sin embargo, es más bien un lugar común y no reivindica de suyo el uso extensivo de estas metodologías en la filosofía (a diferencia de lo que describimos en el siguiente punto).

2. Más al punto central, la lógica aplicada a la filosofía —agrupada en lo que a veces se llama *lógica filosófica*— (y los métodos matemáticos en general) pueden tener un rol positivo en la consecución de nuestras pretensiones de precisión conceptual y argumental. Si bien uno puede apreciar que ciertas formalizaciones en lógica pueden ser a menudo prescindibles en principio, el uso de métodos formales puede disminuir la probabilidad de dejar pasar premisas implícitas cuestionables, de generar algún grado de confusión en la comunicación entre investigadores, y puede habilitar a los investigadores inferir con precisión las consecuencias que se siguen de la asunción de supuestos básicos de origen propiamente filosófico (no, propiamente hablando, matemático).

En la siguiente sección elaboraremos más sobre este segundo punto y daremos algunos ejemplos de interés para el caso, y en lo cual repasaremos brevísimamente un número reducido de hitos en la historia de los métodos formales en filosofía contemporánea. Seguimos para esto de cerca, en lo siguiente, la exposición de Leon Horsten y Richard Pettigrew (Horsten & Pettigrew, 2014) en su artículo sobre métodos matemáticos en filosofía.

3. ¡Embárquense, filósofos matemáticos! Algunos hitos en la filosofía formal contemporánea y una invitación (cautelosa) a seguir explorándolos

La formalización matemática, para efectos de ser aplicada en ciencias empíricas o filosofía, no tiene mucho sentido si se hace como fin en sí mismo. Tampoco la formalización matemática es una manera de,

realmente y de suyo, *solucionar* un problema filosófico. Una formalización dependerá de intuiciones y nociones filosóficas previas que por fuerza no son de naturaleza matemática, sino (sorpresa) de tipo estrictamente filosófico. Y estas intuiciones y nociones no cuentan con un estatus incuestionable, por cierto. Por eso es natural que los procesos de formalización de problemas y posturas filosóficas no zanjen, de suyo, realmente estos temas (aunque puedan arrojar luz sobre ellos y contribuir a la investigación de todos modos).

Las primeras etapas de la filosofía asistida por métodos formales en el siglo XX estuvieron ampliamente marcadas por el uso de la lógica, muy notoriamente la lógica de primer orden y sistemas lógicos desarrollados en detalle posteriormente, a lo largo del mismo siglo (lógica modal, temporal, deóntica, erotética, etc.).

Probablemente el ejemplo clásico (y, en expresión de Frank Ramsey, sin duda, ``paradigmático'' al menos para este estilo de filosofía de análisis lógico y conceptual) es el tratamiento que hace Russell del problema de la semántica de los nombres propios y las descripciones definidas[5] en filosofía del lenguaje. Russell propuso una formalización (con lógica de primer orden con identidad) aún aludida de manera estándar para dar cuenta de lo que dice una descripción definida (y defendió la idea, mucho más controversial, de que los nombres propios son descripciones definidas ``disfrazadas'', con lo cual elabora posibles soluciones a algunos aspectos problemáticos o misteriosos de la referencia en el lenguaje[6]).

[5] Aquellas que sirven para identificar objetos o artículos individuales, como la expresión "El autor de *La República* y maestro de Aristóteles". Una marca notoria y obvia de estas expresiones es que suelen comenzar con artículos definidos (el, la).
[6] Como lo vendría siendo la semántica de los nombres ficticios y el problema de la informatividad de los juicios de identidad.

La aproximación y la formalización russelliana ayudó a eliminar ciertas ambigüedades que con (quizás demasiada) facilidad se ponen en el camino de la discusión y la comunicación entre investigadores. El alcance de las negaciones y de la cuantificación se precisó y la explicitación de los supuestos que irían, según Russell, asociados al uso de una descripción o nombre propio se hicieron más patentes. Otros filósofos (notoriamente Peter Strawson) estuvieron tajantemente en desacuerdo e incluso desafiaron la relativamente complicada formalización de las descripciones que propuso Russell (reemplazándola por la más simple posible: la asignación directa de un predicado a una constante), confrontando así además supuestos básicos y claves de la aproximación russelliana.

En los tiempos del auge del empirismo lógico se observó el insistente intento de elucidar y dar cuenta de la relación de confirmación entre teoría y evidencia (como tópico de filosofía y fundamentos de la ciencia empírica en general), donde se hicieron esfuerzos sostenidos (mediante análisis lógico) por dar cuenta de esta relación en términos sintácticos. Cuando se volvió patente que un análisis tal no podía ser provisto de manera satisfactoria, los filósofos comenzaron a tratar de dar cuenta de la relación de confirmación entre teoría y evidencia en términos probabilísticos (adelantando la relativa popularidad que han adquirido las aproximaciones probabilísticas en tópicos epistemológicos.

El legendario filósofo Rudolf Carnap comenzó su investigación con métodos formales en este tema creyendo con confianza en una tesis de tipo empirista (bastante radical) que puede nombrarse como "fenomenalismo lingüístico" (la idea de que las oraciones sobre *cosas* materiales pueden ser acabadamente 'traducidas' a oraciones estrictamente sobre *datos de los sentidos* preservando los valores de verdad en juego).

Carnap consideró que se podía hacer un estudio (de orden conceptual pero propiamente científico) de esto. En el relato de Hilary Putnam, es así como nació la disciplina que fuese luego llamada *lógica de la ciencia*. Pero, de acuerdo a la reconstrucción de la historia que hace el mismo Putnam, es también ahí donde murió también la postura fenomenalista en cuestión: treinta años de investigación lógica terminaron por establecer que un mapeo-traducción como el postulado no podría realizarse (incluso posiblemente por razones de principio). El lenguaje de cosas no puede, al parecer, ser traducido a un lenguaje puro de la experiencia (que refiera solo a "*sense data*"), con lo que al menos algunas nociones "cósicas" deben ser tomadas como "primitivos" en cualquier lenguaje apropiado para la empresa científica real (Putnam, 1997, pp. 19-20).

Es a propósito también de esta (extremadamente relevante) etapa de la filosofía contemporánea que Horsten y Pettigrew comentan aspectos importantes de lo que puede aportar la formalización lógica en filosofía:

> La formalización lógica obliga al investigador a precisar los conceptos filosóficos centrales. También puede mostrar cómo algunos conceptos y objetos filosóficos pueden definirse en términos de otros. Si se descubre que ciertos objetos se «construyen» como clases de otros objetos, se logra una clarificación ontológica. Además, insistir en una derivación lógicamente válida obliga al investigador a hacer totalmente explícitas todas las suposiciones necesarias. (Horsten & Pettigrew, 2014, p. 16)

Un hito notable en el desarrollo posterior de la filosofía formal fue dado por el trabajo de Alfred Tarski, quien fuese capaz de dar con un tratamiento lógico del concepto de modelo y de la noción de verdad en un modelo (lo que a su vez trajo una manera precisa de especificar la concepción clásica de verdad como correspondencia). Un modelo (en

lógica) corresponde a un conjunto de relaciones y funciones que son definidas en él de modo que nos permita especificar la referencia del vocabulario extra-lógico. Con ello y mediante cláusulas recursivas podemos dar cuenta de los valores de verdad de oraciones complejas como estando determinados por los valores de verdad de las partes que componen estas oraciones. Esto permitió dar cuenta, a su vez y de manera más precisa, de la noción misma de consecuencia lógica, ahora puesta en términos "modelo-teoréticos".

La investigación abierta por Tarski trajo nuevas formas de dar cuenta de las relaciones que hay entre conceptos filosóficos y las oraciones que expresan ideas o tesis filosóficas. Este enfoque modelo-teorético estuvo en primera instancia más bien asociado, fuertemente, a la investigación del concepto filosófico de verdad (algo particularmente impulsado, también, por la búsqueda de soluciones coherentes a la llamada "paradoja del mentiroso", famosamente atendida por Tarski apelando a una distinción entre niveles de lenguaje: a nivel básico, la distinción entre "lenguajes-objeto" y "meta-lenguajes"). Con este tipo de tratamiento pudo darse cuenta, al menos, de la consistencia o coherencia de ciertas teorías formales de la verdad. A su vez, la aproximación modelo-teorética facilitó la evaluación de teorías en mereología (el estudio de las relaciones entre parte y todo).

Estos modelos fueron expandidos para dar tratamiento a nociones intensionales (como las asociadas a la posibilidad, a la obligación moral, a la temporalidad o a nociones epistémicas, entre otras). Los modelos tarskianos pueden ser concebidos, en cierta forma, como estados posibles de cosas; necesitaríamos, en este contexto, dar con un concepto de modelo en el cual estados de cosas posibles (o mundos posibles) son representados como tales. Estos "modelos de mundos posibles" fueron desarrollados notablemente en los '50 desde el trabajo de Saul Kripke y otros. Esta clase de investigación tuvo aplicaciones notorias en epistemología, filosofía del lenguaje y metafísica (y fue, justamente, más o

menos en este periodo en que la expresión "lógica filosófica" comenzó a entrar en uso común, y en lo cual proliferaron elaboraciones de diversas áreas como lo son, precisamente, la lógica modal, la lógica deóntica, la lógica temporal, lógica epistémica, entre otras).

Décadas posteriores, no obstante la proliferación de enfoques específicamente lógicos, permitió la amplicación del set de herramientas disponibles para la filosofía formal. La teoría de juegos y la teoría de la decisión comenzaron a ser ocupadas para dar tratamiento a tópicos de filosofía práctica (piénsese, notablemente, en el trabajo de Ken Binmore y el subtexto de inspiración matemática en los tópicos "rawlsianos" de la filosofía política). Se incorporaron modelos algebraicos en mereología, teoría de grafos para dar cuenta de asuntos relativos a la noción de cualidad secundaria y la de (in)discriminabilidad perceptual, e incluso el uso del análisis funcional para dar cuenta de aspectos relevantes en el ámbito de las normas epistémicas.

Como ya adelantamos, comenzaron en su momento también a proliferar aproximaciones probabilísticas a la epistemología (hoy también con bajadas en la investigación metafísica). Filósofos de la ciencia contribuyeron con la incorporación de modelos y simulaciones computacionales, los que también tienen hoy en día aplicaciones notorias en áreas como la epistemología social (en efecto, ya se habla hace tiempo de la existencia de la "filosofía computacional", que vendría a ser justamente la realización de investigación filosófica con el uso de herramientas computacionales).

A pesar de lo interesante y útil que pudieran ser los modelos formales en el área, la filosofía formal no va sin sus riesgos. Particularmente, es fácil apresurarse a tomar de manera literal los modelos mismos (lo cual puede ser problemático y disputable, análogamente a la manera en que se disputa la literalidad que podemos atribuirle a los modelos y teorías en ciencias empíricas —por oposición a

una interpretación puramente "instrumentalista" u "operacionalista" de los modelos mismos). Puesto que además la filosofía (particular y notoriamente los modelos formales en filosofía) suele requerir a menudo un grado importante de idealización, se corre el riesgo de caer en una simplificación excesiva de los conceptos y temas tratados. No menos problemático es el hecho de que un formalismo por sí solo no va a purgar, necesariamente, la vaguedad de un concepto y ese hecho puede a veces ser oscurecido por la elegancia y el carácter formal estricto del modelo.

Considero que, sumando y restando, explorar el uso de modelos y formalismos en filosofía puede ser relevante y prometedor (aunque fuese únicamente para dar cuenta con mayor precisión de sus límites en el área, o para dar con resultados negativos y sacar provecho de lo que podríamos entender como "errores bien logrados").

Es por esto también que el uso de técnicas formales no purga de conflicto, disenso y creatividad al área. Citando a Horsten y Pettigrew, que cita a su vez a van Benthem:

> El modelado formal puede abrir posibilidades conceptuales. Al aislar una clase de modelos para una subdisciplina de la filosofía, se delimita un espacio de posibilidades. Estos espacios suelen ser mucho más ricos de lo que uno esperaría. Proporcionan un marco fértil y a la vez riguroso para explorar ideas creativas:
>
> > (…) a medida que desentrañamos nuestros conceptos, se revela su verdadera riqueza y se abren posibilidades hasta ahora no exploradas. Por lo tanto la precisión formal es un estímulo para la fantasía creativa (…). La oposición de moda entre 'libertad creativa' y 'armadura lógica' no hace justicia a la lógica (ni, temo, a la creatividad). ([van Benthem, 1982, p. 459]) (Horsten & Pettigrew, 2014, p. 21)

Estas consideraciones pueden analogarse a la idea más general según la cual los sistemas normativos que crean los humanos y bajo los cuales regimos nuestra conducta nos habilitan para realizar nuevas posibilidades no accesibles previamente a la implementación de esas normas. El lenguaje y la lógica cuentan como caso especial de este tipo de restricciones normativas auto-impuestas que nos habilitan crear de maneras completamente novedosas en relación a lo que podríamos hacer sin ellas.[7]

3. *Homónima* y los métodos formales en filosofía

La proyección, por parte de *Homónima*, es generar, en serie y para los distintos números de la revista, artículos que expongan aplicaciones filosóficas de la lógica y las matemáticas (así como material que facilite la comprensión de estas técnicas a nuestros lectores interesados pero con menos formación en estas temáticas). El objetivo es facilitar el acceso a nuestros colegas y a estudiantes del área interesados en expandir su repertorio de métodos y herramientas disponibles. Esperamos que esto contribuya una pieza relevante y de interés para la exploración metafilosófica en general y, con ello, al continuo perfeccionamiento colectivo de nuestro oficio.

[7] Agradezco a Kuminak Lefio por la sugerencia de agregar esta idea. Mi colega tuvo en mente esta consideración a propósito del trabajo de Jaroslav Peregrin y su "inferencialismo naturalizado" (Peregrin, 2024).

REFERENCIAS

Horsten, L., & Pettigrew, R. (2014). Mathematical Methods in Philosophy. En L. Horsten, & R. (. Pettigrew, *The Bloomsbury Companion to Philosophical Logic* (págs. 14-26). Bloomsbury.

Networks in philosophy: Social networks and employment in academic philosophy. *Metaphilosophy*, 653-684.

Peregrin, J. (2023). *Normative Species How Naturalized Inferentialism Explains Us.* Routledge. https://doi.org/10.4324/9781003388876

Peregrin, J. (2024). *Normative Species: How Naturalized Inferentialism Explains Us.* Nueva York: Routledge.

Putnam, H. (1997). Language and philosophy. En H. Putnam, *Philosophical Papers, Volume 2* (págs. 1-32). Cambridge University Press.

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 453-464
ISSN 3087-260X
Santiago de Chile

ALGUNAS OBSERVACIONES SOBRE GRÁFICOS PEIRCEANOS MODALES[1]

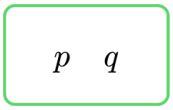
Felipe Morales Carbonell[2]

El filósofo pragmatista C.S. Peirce (1839-1914) dedicó gran parte de sus esfuerzos en su último periodo al desarrollo de una serie de sistemas lógicos diagramáticos, que denominó *gráficos existenciales*. Peirce desarrolló sistemas diagramáticos para lógica proposicional (clásica) (sus gráficos «alfa»), lógica de primer orden (sus gráficos «beta»), y lógica modal (sus gráficos «gamma»). Lo distintivo de estos sistemas es que en ellos las relaciones lógicas son representadas de manera gráfica o visual. En este artículo, quiero dar una breve presentación de las ideas centrales de estas lógicas, y hacer algunas observaciones sobre ciertas características de la manera en que capturan la noción modal de posibilidad.

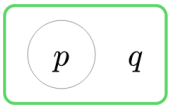
[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima.* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Morales, F. (2026) Algunas observaciones sobre gráficos peirceanos modales. *Revista de Filosofía Homónima, 1*(1), pp. 453-464. https://doi.org/xxxxxxx[Por asignar]

[2] Felipe Morales Carbonell tiene un Doctorado (PhD) y Máster en Filosofía (MPhil) por el Instituto de Filosofía de la KU Leuven, Bélgica. Es investigador postdoctoral en el Departamento de Filosofía de la Facultad de Filosofía y Humanidades de la Universidad de Chile. Email: ef.em.carbonell@gmail.com. ORCID: 0000-0001-5492-0759

La idea fundamental es que podemos representar conjuntos de afirmaciones como inscripciones sobre una superficie, que Peirce llama la "hoja de aserción". Asumamos que podemos representar proposiciones mediante letras minúsculas (p, q, …). Así, que y son verdaderas (lo que en la notación contemporánea escribiríamos como p∧q) correspondería a

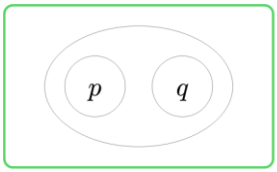


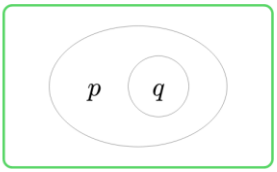
Para expresar que algo *no* es verdadero (es decir, es falso), lo inscribimos dentro de una línea cerrada que llamamos *corte*. Así, ¬p∧q correspondería a



Vamos a llamar a cualquier parte del diagrama (un conjunto de proposiciones y cortes) un *subgrafo*.

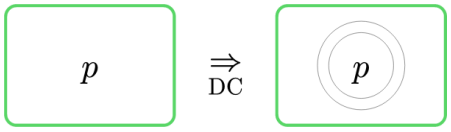
Teniendo un dispositivo para expresar conjunciones y negaciones, podemos expresar disyunciones (p∨q = ¬(¬p ∧¬q)) y condicionales (p → q = ¬(p∧¬q)):



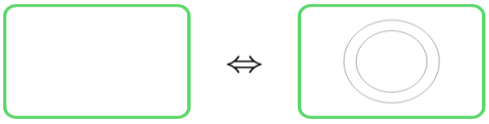


Dejo como tarea al lector pensar cómo podría representarse el bicondicional.

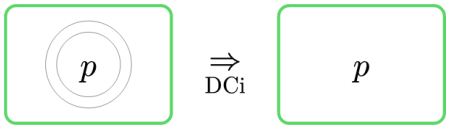
En los sistemas lógicos tradicionales, definimos una relación de consecuencia lógica entre conjuntos de fórmulas y fórmulas: decimos que de un conjunto de premisas $P_1 \ldots P_n$ se sigue la conclusión C. En el sistema de Peirce, la relación de consecuencia lógica ($\Rightarrow$) se define en términos de transiciones de gráficos a otros gráficos. Esta relación es reflexiva: así como de p se sigue p, el gráfico cuya única inscripción es p es consecuencia lógica de sí mismo. Además de esto, tenemos otras reglas que nos permiten transformar los gráficos inscribiendo proposiciones, borrándolas, trazando cortes y quitándolos. Por ejemplo, es claro (al menos desde la perspectiva de la lógica clásica) que si tenemos, también tenemos ¬¬p. Esto corresponde a la siguiente transformación:



Es claro que esto es válido para cualquier subgrafo, no solamente para inscripciones de proposiciones en la hoja de aserciones (¬(p∧q) es equivalente a ¬¬¬(¬¬p∧q), por ejemplo). En esta forma esquemática, podemos llamar a esta regla *corte doble* (DC). Como caso límite, la hoja vacía es un subgrafo, de modo que podemos introducir o eliminar un corte doble arbitrariamente:



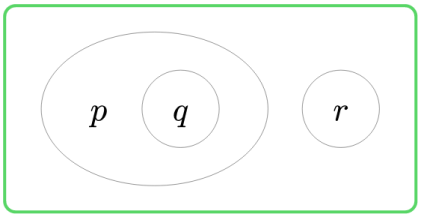
La transformación inversa, de ¬¬p a p, es igualmente válida (DCi):



Esto es equivalente a la regla usual de eliminación de la doble negación. Hay lógicas (las así llamadas lógicas *intuicionistas*) donde la regla de eliminación de la doble negación no es válida. La idea (a grandes rasgos) es que de que no haya una prueba de que no hay una prueba de una tesis, no se sigue que esa tesis sea verdadera. Vamos a volver a este punto después.

En efecto, podemos entender la hoja vacía como algo que es irrestrictamente verdadero (análogo a ⊤), y a un cierre vacío sobre una hoja vacía como la negación de lo irrestrictamente verdadero (análogo a ⊥). Esto significa que las tautologías son equivalentes a la hoja vacía–en otras palabras, para demostrar que un grafo es tautológico es suficiente mostrar que puede transformarse la hoja vacía en él, y para mostrar que es inconsistente, es suficiente mostrar que puede transformarse el corte vacío en él.

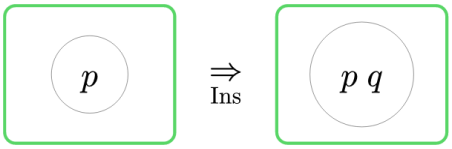
El sistema contempla solamente 3 reglas más. Para entenderlas tenemos que introducir la noción de nivel de un subgrafo, que es el número de cortes que lo rodean. Por ejemplo, en



el nivel de q es 2, el de p y r es 1, y la hoja de aserción tiene nivel 0. Aunque p y r tienen el mismo nivel, no están anidados en un mismo subgrafo que no sea la hoja entera misma.

Las reglas son:

- *Inserción*: se puede insertar cualquier gráfico en un nivel impar.

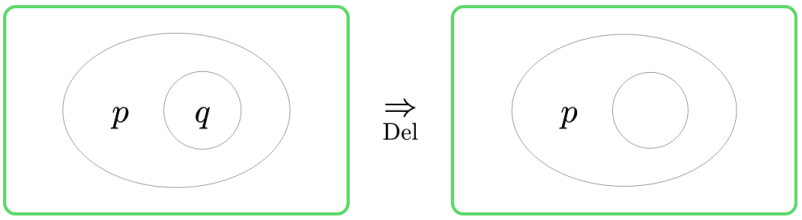


Por ejemplo:

Es decir, si p es falso, p∧q también es falso.

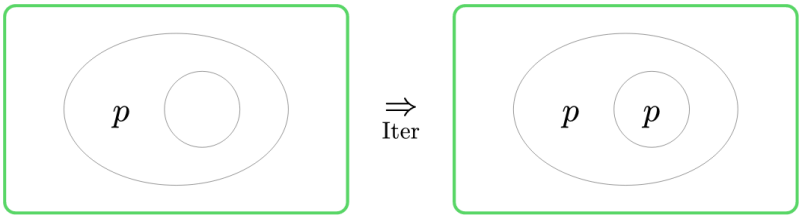
- *Borrado*: se puede borrar cualquier subgrafo en un nivel par.

Por ejemplo:



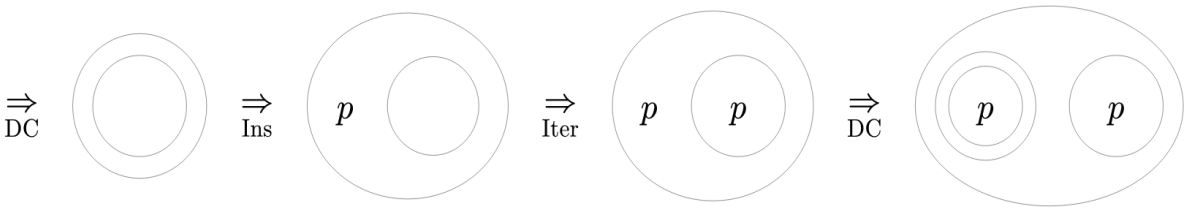
Vale la pena observar que aquí el grafo que resulta tiene que leerse como p→⊤: el corte vacío es de nivel 2, por lo que es la negación de un corte vacío de nivel 1, que como vimos, es ⊥. Como solo se puede borrar en un nivel par, la operación de borrado siempre será equivalente a Introducir ⊤.

- *Iteración/Deiteración*: se puede insertar o borrar una copia de un subgráfo a cualquier nivel anidado en relación con ese subgráfico.



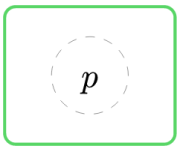
Es decir, si p→⊤ entonces p→p. En este ejemplo, el corte de nivel 2 está anidado en uno de nivel 1 que contiene, de modo que puede iterarse p en el corte de nivel 2.

Para ilustrar, considérese una prueba de p∨¬p como una tautología.



Nótese que en cada paso probamos una tautología: en el tercer paso (tras Ins), probamos p→⊤ como tautológica, y en el penúltimo paso (tras Iter) probamos p → p.

Mi interés en los gráficos Peirceanos tiene que ver con la extensión que hace Peirce para capturar el comportamiento de ciertas modalidades, el sistema de los así-llamados gráficos «gamma».[3] La idea básica es que a veces interesa razonar considerando enunciados que no han sido probados como verdaderos o falsos–algo que podría ser verdadero o falso. Para capturar esto Peirce introduce un segundo tipo de corte, que voy a llamar (siguiendo a Peirce) *corte roto*:



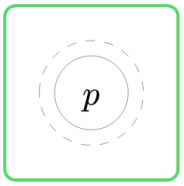
[3] Un poco más de contexto: me interesa el problema de si Peirce tiene los recursos para evitar paradojas del tipo de Fitch, que aparecen en lenguajes que tienen recursos modales y epistémicos. En algunos de estos sistemas, es demostrable que la cognoscibilidad colapsa en el conocimiento. En Grigoryan & Carbonell (2025) examinamos el problema desde una perspectiva general. Ahora me interesa examinarlo desde la perspectiva de la lógica diagramática Peirceana, pero no hablaré más en detalle de eso por ahora.

Nos imaginaríamos que p rodeado por un corte de este tipo significaría que p *es posible* (es decir, que este grafo sería equivalente a una afirmación de $\Diamond p$ en la notación contemporánea usual para la lógica modal), pero no es así. Peirce propone que deberíamos leer esto como $\neg p$ es posible, es decir, $\Diamond\neg p$. Veamos qué consecuencias trae esto–quizás examinándolas podamos entender mejor por qué Peirce adopta esta convención.

La negación de este grafo es:

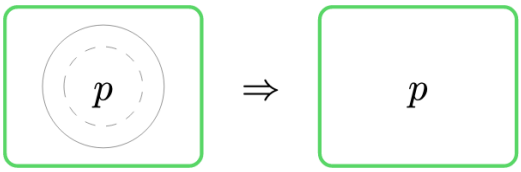
Esto es equivalente a decir que es necesario: no es el caso que sea posible ($\neg\Diamond\neg p$).

Del mismo modo, negar p dentro del corte roto es:

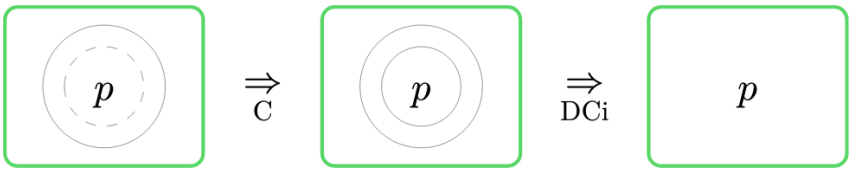


Esto es equivalente a decir que es posible: es posible que no sea el caso que no sea el caso ($\Diamond\neg\neg p$).

Los axiomas modales corresponden a ciertos patrones de transformación de estos grafos. Así, por ejemplo, el axioma T ($\Box p \rightarrow p$) corresponde a la transformación



Es fácil ver qué transformaciones corresponden al converso de T ($p \rightarrow \Diamond p$), al axioma K ($\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$), 4 ($\Box p \rightarrow \Box\Box p$) y 5 ($\Diamond p \rightarrow \Box\Diamond p$) (dejamos esto como ejercicio para el lector).[4] La validez de las inferencias modales se justifica en ciertas reglas, similares a las de la porción proposicional que vimos arriba, que tienen que ver con la posibilidad de convertir cortes continuos por cortes rotos, y otras manipulaciones de este tipo. Por ejemplo, T se puede justificar apelando a una regla (C) que dice que si un corte roto está anidado en un corte continuo a nivel par, puede convertirse en un corte continuo. Así:



Ahora que tenemos una idea de cómo funciona la lógica diagramática proposicional y modal de Peirce, me gustaría hacer algunas observaciones que ilustran una clase de problemas que el uso filosófico de la lógica presenta a menudo, que tienen que ver con la manera en que interpretamos un formalismo.

[4] Ma & Pietarinen (2017) describe una serie de sistemas modales empleando el marco de los gráficos existenciales Peirceanos.

A diferencia de lo que sucede en la lógica modal usual, aquí las modalidades de la necesidad y la posibilidad no son primitivas, sino derivadas: en el simbolismo, estas nociones emergen de la aplicación de cortes de distintos tipos. Esto presenta un problema: para que digamos que $\Diamond\neg\neg p$ es equivalente a $\Diamond p$, necesitamos que $\neg\neg p$ sea equivalente a p (de forma similar, para que $\neg\neg\Diamond p$ sea equivalente a $\Diamond p$, debemos tener la capacidad de eliminar la doble negación al comienzo). Si bien esto es clásicamente válido (y en efecto, es una de las asunciones centrales de los gráficos proposicionales Peirceanos), el hecho de que tengamos que asumir una lógica clásica (o al menos, una regla que permita la introducción y eliminación de dobles cortes alrededor de operadores modales) para poder derivar un concepto de posibilidad positivo (que no tenga que referir a la negación) es extraño. Un lógico intuicionista, que rechace la regla de doble negación, no obstante esto podría querer decir que hay un concepto de posibilidad que no es equivalente a $\Diamond\neg\neg p$ (para este lógico, podría distinguirse entre $\Diamond p$, $\Diamond\neg\neg p$ y $\neg\neg\Diamond p$).[5] La lógica de Peirce, en la forma presente, parece no poseer los recursos *conceptuales* para capturar estas distinciones.

Podría contraargumentarse que la preocupación que acabo de esbozar se debe a una suposición errónea, a saber, la idea que los primitivos notacionales de una lógica deben corresponden a los primitivos conceptuales de la misma. Esto no es necesariamente así; la idea sería que si bien en la lógica modal de Peirce la notación para la posibilidad es derivada, el concepto de posibilidad que queda expresado de esa manera derivada podría ser primitivo. Sin embargo, esta respuesta

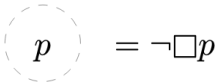
[5] Consideren cómo se puede expresar $\neg\neg\Diamond p$ en la lógica de los gráficos gamma. Todo corte (tanto continuo como cortado) introduce una negación. De modo que no puede haber un grafo que represente $\neg\neg\Diamond p$ con p dentro de un único corte. Sabemos que en la lógica modal tradicional $\Diamond = \neg\Box\neg$ y $\Box = \neg\Diamond\neg$, de modo que podemos convertir a $\neg\neg\Diamond$p primero en $\neg\neg\neg\Box\neg p$ (en este paso ya logramos introducir una negación antes de p) y luego a $\neg\neg\neg\Diamond\neg\neg p$. Podemos construir el grafo correspondiente, y aplicando la regla de doble corte iteradamente vemos que el grafo es equivalente al grafo de $\Diamond\neg\neg p$.

es insatisfactoria porque usualmente se cree que la notación lógica debe ser un medio conspicuo para expresar los conceptos lógicos que vienen al caso. Esto no es ajeno al pensamiento de Peirce: en efecto, la idea de que la modalidad pueda expresarse en términos de algo similar al corte con que expresa la negación es sugerente respecto a la manera en que Peirce concebía el papel del juicio modal en términos de posibilidades concernientes a lo que representa la hoja de aserciones. Si esta representa el contenido evaluado durante una sesión de pensamiento, la inscripción modal mediante el dispositivo del corte roto representa una forma alternativa de evaluar contenidos, distinta a la aserción y la negación. Esto es claro. Lo que no es claro es qué exactamente debería representar el corte como notación–en otras palabras, cómo deberíamos interpretarlo. La decisión de Peirce de adoptar la convención de que el corte roto representa la posibilidad de que algo no sea el caso es simplemente una solución posible al problema, entre otras.

Inicialmente, parecería sencillo reemplazar la interpretación del corte roto como $\Diamond\neg$ por una interpretación del corte roto como $\Diamond$, en cuyo caso la noción de posibilidad quedaría como primitiva.[6] Sin embargo, al hacer esto perdemos, al menos de manera preliminar, el aparataje inferencial que describimos arriba, por ejemplo cuando derivamos la validez del axioma T. ¿Qué reglas son válidas para esta interpretación del corte roto? Dejo eso como otro ejercicio para el lector.

Ha habido intentos de construir sistemas intuicionistas de grafos Peirceanos, y vale la pena examinar un poco cómo se resuelven estos problemas allí. Oostra (2022) desarrolla una lógica diagramática modal intuicionista. En esta lógica, el corte roto se interpreta como una inscripción de *contingencia*, es decir,

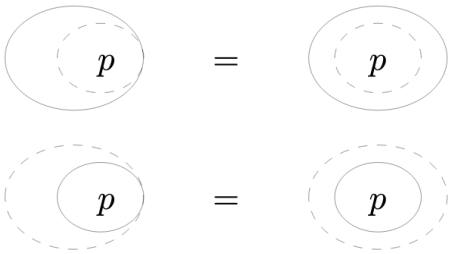
[6] Eventualmente, Peirce mismo adoptó un sistema lógico donde podía representar la noción de posibilidad directamente. Véase Schmidt (2025) para una discusión.



Así, p inscrito en un corte roto es equivalente a ¬□*p* en la lógica modal tradicional, no a ◇¬*p*. En este caso, pareciera que el problema que esbozamos no surge en el caso de la posibilidad, que queda como equivalente a ¬□¬*p*, sino en el caso de la necesidad, que queda equivalente a ¬¬□*p*. ¿No podría querer el intuicionista aceptar un concepto de necesidad que no sea equivalente a esto? En efecto, en lógicas modales intuicionistas las modalidades de necesidad y posibilidad no son necesariamente duales (es decir, las equivalencias ◇ = ¬□¬ y □ = ¬◇¬ pueden no ser válidas).[7] En el sistema de Oostra, esto se resuelve en que cada modalidad corresponde a un tipo distinto de corte primitivo:

necesidad posibilidad

Dado esto, y desde un punto de vista conceptual, ¿siguen siendo la necesidad la negación de la contingencia, y la posibilidad la contingencia de una negación? Es decir, ¿serían válidas las siguientes equivalencias?



[7] Cf. Wolter & Zakharyaschev (1999).

Esto va a verse reflejado en las reglas de transformación que adoptemos. Desde un punto de vista intuicionista estricto, la primera quizás solo podría ser aceptable en el sentido de izquierda a derecha (pasamos de un gráfico positivo a uno negativo, lo que está permitido), pero no de derecha a izquierda (porque pasamos de un gráfico negativo a uno positivo). En el segundo caso, todo depende de si tomamos a la posibilidad o a la contingencia como positivas o negativas. Para decidir eso, no bastan las consideraciones formales–hace falta un argumento filosófico.

COMENTARIO a "*Algunas observaciones sobre gráficos peirceanos modales*" de Felipe Morales Carbonell[8]

Kuminak Lefio Zamorano[9]

La relación, que queda patente en el texto, entre las investigaciones lógicas y filosóficas se muestra como una compleja interacción, con un

[8] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima.* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Lefio, K. (2026) Cometario a "Algunas observaciones sobre gráficos peirceanos modales" de Felipe Morales Carbonell. *Revista de Filosofía Homónima, 1*(1), pp. 464-472 https://doi.org/xxxxxxx[Por asignar]
[9] Licenciado en Filosofía de la Universidad de Chile, estudiante del diplomado "Fundamentos de la Física" de la misma casa de estudios. Actualmente se desempeña como jefe en Gestión Editorial de la Revista Homónima. Email: kuminak.lefio@ug.uchile.cl. ORCID: https://orcid.org/0009-0005-8121-4063.

potencial significativo de ser complementarias entre sí. Esto podría parecer obvio en tanto que entendamos a la lógica como una parte o rama de la filosofía, sin embargo, y a pesar de que lo anterior es verdadero, estaríamos perdiendo el matiz de una distinción más sutil entre lógica y filosofía, el cual no se trata de una diferenciación radical entre ambas como disciplinas, sino que es acerca del tipo de razonamiento que se dedicaría al análisis meramente formal y aquel que busca responder preguntas sustantivas acerca de los conceptos abordados en dicho sistema formal, las cuales parecen no ser decidibles meramente por este mismo.

En este caso particular, se ha mostrado como el estudio de los gráficos Pierceanos, como un sistema lógico particular, suscita cuestionamientos y dudas en ambos niveles. Por un lado, en su estructura formal es cuestionable si es que se trata de un sistema lógico completamente independiente del clásico o si parece tener que acudir a este para resolver algunas de las problemáticas planteadas. Por otro lado, nos motiva a reflexionar sobre el concepto de posibilidad y como juega un rol en nuestro razonamiento, a la vez que, en una pregunta en cierto sentido de segundo orden, acerca de cómo el sistema lógico que adoptemos debería ser un modelo adecuado de la forma normativamente correcta de desarrollar nuestro razonamiento.

Es por esto, que estos dos dominios de análisis, el formal y el sustantivo/conceptual, no parecen ser completamente independientes el uno del otro y funcionan mejor desarrollándose de forma juiciosa en conjunto. No solo de tal forma que nuestras propuestas sustantivas acerca de los conceptos relevantes guíen nuestra formulación de sistemas lógicos, sino también en tanto que el estudio lógico nos permite identificar los pilares fundamentales de nuestro razonamiento que requieren ser abordados con una postura filosófica robusta. Esto último, nos otorga una idea destacable acerca de la importancia del estudio de la lógica para nuestra teorización de otros problemas filosóficos. No me refiero únicamente a la idea comúnmente difundida de que estudiar la

lógica nos ayuda a ordenar nuestro pensamiento y ser claros en la formulación de nuestros argumentos. En cambio, de forma más específica quiero señalar cómo estudiar a fondo sistemas lógicos nos permite llegar a los conceptos fundamentales que utilizaremos en nuestra argumentación acerca de los problemas filosóficos que abordamos, y nos motiva a tomar una postura explícita respecto a cómo entender cada uno de ellos, así como las relaciones que tienen entre sí. Esto plausiblemente nos ayudará a ser más claros y ordenados en la argumentación, pero el rédito filosófico de identificar y vernos obligados a abordar estos pilares conceptuales no se limita a dicha claridad.

Bajo esta lectura, lo expuesto acerca de los gráficos Peirceanos no es solo relevante para la formulación formal o la postura filosófica que apoya este sistema lógico particular, sino que además nos suscita a tomar una postura sobre preguntas más generales: ¿Cómo se entiende la posibilidad? ¿Qué relevancia y potencial de modelación tiene esta forma diagramática de estructurar el razonamiento? ¿Cómo debería interactuar en nuestro razonamiento el que consideremos algo como plausible con el resto de proposiciones que pueden ser relevantes para dicho caso?, entre otros cuestionamientos que a alguien más que a mí se le ocurrirán.

Para desarrollar una de estas líneas de cuestionamiento, me extenderé sobre una idea que está de fondo en la lógica de gráficos Peirceanos: la posibilidad de modelar nuestro razonamiento lógico de forma diagramática. Esta es una idea sumamente llamativa, ya que promete ver el proceso de pensar lógicamente desde una perspectiva sustancialmente diferente a la forma tradicional. Y por lo mismo, no sería extraño suponer que nos permitiese llegar a nuevas conclusiones acerca de la estructura o las posibilidades del razonamiento lógico. Sin embargo, en el caso de los gráficos Peirceanos a pesar de que se presentan en una estructura diagramática y pretenden este cambio sustancial de perspectiva frente a la lógica, parece que no nos terminan entregando aquello que prometen.

Para argumentar lo anterior partamos desde la siguiente idea. Si nuestro objetivo fuese elaborar una perspectiva del razonamiento lógico que intentase modelarlo y comprenderlo diagramáticamente, lo que necesitaríamos no sería meramente una estructura construida por medio de dibujos cuyas reglas sean correctas para nuestros estándares ya establecidos acerca de cómo se debe conducir el razonamiento. Necesitaríamos además que aquellos dibujos y las reglas adecuadas para transformarlos representasen por sí mismos algo acerca del razonamiento lógico. Es decir, si quisiéramos llevar a sus últimas consecuencias el proyecto de modelar diagramáticamente el pensamiento lógico, debería poder ser explicable como cada aspecto del diagrama y sus transformaciones reflejan lo que normativamente entendemos como el razonamiento lógico. Esto es especialmente importante para que una lógica diagramática pueda llevarnos a conclusiones distintas acerca del razonamiento lógico y sus posibilidades, allí se juega la capacidad de distinguir sustancialmente este sistema lógico de otros y la posibilidad de defender que adoptarlo no es meramente un cambio superficial, sino que implica cambiar sustancialmente de perspectiva frente a la lógica.

Teniendo en mente esto, los gráficos Peirceanos si plantean algunas ideas que parecen atender a la exigencia recién planteada, mencionemos algunos ejemplos prominentes:

Por un lado, tenemos la intuición de que en nuestro razonamiento separamos aquello que consideramos verdadero de lo que consideramos falso. Esto se ve representado en los gráficos por medio del corte que aísla algunos elementos de otros y señala la falsedad de estos, separándolo de lo que se considera como verdadero. Es una idea interesante desde la que partir y que sí parece ser una manera plausible, entre muchas otras, de relacionar estas estructuras diagramáticas con aspectos de nuestro razonamiento lógico. No solo guarda una evidente relación con la forma en que razonamos lógicamente, sino que como toda lógica debería, estipula ideas normativas acerca de cómo conducir el razonamiento, en

este caso, se norma el deber mantener separado y distinguible aquello que se considera verdadero de lo que se considera falso.

Por otro lado, el querer construir un tipo de corte distinto para aquello que consideramos posible en vez de tajantemente verdadero o falso, intenta claramente reflejar un aspecto de nuestro razonamiento y normar el tipo de relaciones lógicamente válidas que se pueden establecer con esta posibilidad. Además, esto se condice con el interés que demuestra Peirce por el rol que juega lo plausible en nuestro razonamiento lógico y en la investigación científica.

No obstante, otros elementos, como las reglas de transformación de los gráficos, la idea de distintos niveles según los cortes alrededor de algo, diferenciar entre niveles pares e impares, entre otros, no parecen cumplir con dicha exigencia de la misma forma. Deberíamos ser capaces de explicar ¿Cómo el que se pueda insertar cualquier gráfico en un nivel impar refleja algo del modo en que eso funciona en el razonamiento lógico? Y aquí, al menos planteándonos desde la perspectiva exigente que ya definimos, no podemos solo responder "de la misma forma que lo hace la lógica clásica" o "ya que refleja un paso que ya consideramos como válido", lo que en cambio buscamos es que la estructura de gráficos, por su forma, ubicación u organización visual y estructural sea aquello que nos pueda otorgar la respuesta a la pregunta. En el caso de estos elementos, a pesar de cuan fructíferos son en la tarea formal de establecer este sistema lógico de manera adecuada, se vuelve difuso cómo nos *dicen algo* sobre el pensamiento lógico, especialmente si intentamos buscar algo distinto de lo que no dirían otros sistemas lógicos que no sean diagramáticos. En síntesis, la estructura y forma de estos elementos de los gráficos Peirceanos no parecen tener una relación clara con el razonamiento que modelan[10], es decir, es justamente lo que le otorga el

[10] Vease, Priest (2014) para una perspectiva más profunda y detallada acerca de lo que nuestros sistemas lógicos intentan modelar. Aquí el concepto de *lógica utens* nos permite

carácter diagramático a esta lógica aquello que carece de esta relación clara y directa.

Esta problemática parece darse por el método según el que se establecen estas reglas de transformación de grafos. Este parece proceder partiendo desde una serie de inferencias ya consideradas válidas en la lógica, para luego identificar el tipo de reglas de transformación de grafos que se adecuan para cumplir estas inferencias lógicas. Esto implica que el criterio, al menos el principal sino el único, para seleccionar este conjunto de reglas en lugar de otros, es justamente esta coincidencia con lo que ya hemos aceptado por válido en la lógica. Y no hay razón para suponer que este criterio por sí solo nos hará obtener una estructura diagramática que refleje nuestro razonamiento lógico, por el tipo de diagrama que es. La relación solo está establecida en la capacidad de producir inferencias correctas por medio de unos simbolismos que resultan tener un carácter diagramático, tal como también lo hacen simbolismos clásicos, y no entre la forma de ser de estos diagramas y nuestro razonamiento lógico; nunca se da esta relación directa en el método por el cual construimos las reglas de transformación de grafos.[11]

---

apuntar directamente a aquello con lo que se le exige a los gráficos Peirceanos que establezcan una relación clara.

[11] Teniendo en cuenta la importancia de la teoría semiótica de Peirce, vale la pena señalar que aquí el término "símbolos" se está utilizando de una forma más general y no apelando a la distinción planteada por el autor entre símbolos, índices e iconos. Sin demérito de lo anterior, aprovechare de considerar una posible objeción que se desprende de considerar esta distinción. En Legg (2011), donde se desarrolla una interesante discusión respecto a cómo llegamos a conocer *la necesidad lógica*, se sugiere que los gráficos Peirceanos serían principalmente iconos, en vez de símbolos o índices. Resumidamente, esto implica que como signo en sí mismo representan su objeto, y que las partes y relaciones de su estructura son representativas en sí mismas de la estructura de su objeto. No pretendo discutir profundamente esta idea en tan poco espacio, pero si planteare que creo que la problematización desarrollada no se disuelve por meramente señalar que se trata de un icono. En realidad, parece que ocurre lo opuesto, comprender a los gráficos como iconos que deberían permitir esta correspondencia entre las estructuras (partes y relaciones) del signo y el objeto, motivan a que nos preguntemos y respondamos como exactamente se da esta representación directa. Para defender que

El panorama para construir una lógica diagramática (con las aspiraciones antes mencionadas) parece ser algo más complejo, en este sabemos no solo que tenemos múltiples posibilidades de estructuras de gráficos y reglas de transformación, sino que además sabemos que el mero criterio de adecuación con las inferencias de la lógica tradicional no es suficiente. Incluso, desde esta perspectiva no sería extraño elucubrar que existan múltiples sistemas de gráficos y reglas capaces de acomodar dichas inferencias lógicas, pero que cada uno de ellos difiera de los demás. En cambio, debemos defender que un sistema diagramático particular con sus reglas propias, refleja en su estructura y normatividad (reglas) el funcionamiento del razonamiento lógico, he aquí el criterio que realmente nos permitiría elegir un sistema diagramático sobre otro. Este segundo método considero que no está desarrollado en el caso de los gráficos Peirceanos y, por ende, tampoco la argumentación necesaria para defender que esta lógica diagramática llegue a realizar un potencial mayor que ser meramente una traducción de unos símbolos a otros de la lógica tradicional.[12]

Lo anterior no implica que no exista potencial en los gráficos Peirceanos para realizar estas aspiraciones, antes mencionábamos algunos de los aspectos que sí parecían apuntar en esta dirección, pero sí es claro que hace falta recorrer este complicado camino de argumentar filosóficamente el entrelazamiento entre ciertos diagramas y el pensamiento lógico para lograrlo.

---

sean iconos efectivos en lo que buscan representar tenemos que dar cuenta de que se cumple la representación en los detalles del modelo, especialmente en los elementos diagramáticos de esta lógica que la distinguen de otras, y no solo en la generalidad de que se cumplen inferencias lógicas. Muchas gracias a Felipe Morales por hacer que me fijara en esta posible objeción.

[12] Obviamente acepto la falibilidad de esta afirmación y espero que este comentario sirva para desarrollar debates acerca de la posibilidad de una lógica diagramática, siendo yo el primero en alegrarme si se demuestra que los gráficos Peirceanos si cumplen con las exigencias que aquí se sugieren.

La discusión de la lógica de gráficos Peirceanos abre muchas otras incógnitas que no abordaré en profundidad aquí, pero que sí han tenido la posibilidad de ser sugeridas a lo largo del texto. Siguiendo esta línea, cerraré el texto dejando algunas interrogantes puntuales en las que el lector pueda vagar, hasta hallar o no una respuesta que lo satisfaga: ¿Qué significa entender la posibilidad de forma positiva o negativa? Es decir, como la afirmación o la negación de algo ¿Acaso no es la naturaleza misma de considerar algo posible, el considerar a la vez su afirmación y su negación? ¿O hay una manera correcta en la que se pueda expresar un sentido puramente positivo de la posibilidad? Finalmente ¿Hay alguna diferencia entre lo que entendemos por posible y lo plausible que es de tanto interés para autores como Peirce? ¿Ocupan ambas cosas el mismo lugar en nuestro razonamiento y en nuestras empresas de búsqueda de conocimiento?

REFERENCIAS

Grigoryan, K., & Carbonell, F. M. (2025). Peircean knowability and Fitch-like paradoxes. Res Philosophica, 102(2), 163-190. https://doi.org/10.5840/resphilosophica2664

Legg, C. (2011). The Hardness of the Iconic Must: Can Peirce's Existential Graphs Assist Modal Epistemology?. Philosophia Mathematica, Volume 20, Issue 1, February 2012, Pages 1–24, https://doi.org/10.1093/philmat/nkr005

Ma, M., & Pietarinen, A.-V. (2017). Gamma graph calculi for modal logics. Synthese, 195(8), 3621-3650. https://doi.org/10.1007/s11229-017-1390-3

Oostra, A. (2022). Intuitionistic and geometrical extensions of Peirce's existential graphs. En Advances in Peircean Mathematics (pp. 105-180). De Gruyter. https://doi.org/10.1515/9783110717631-003

Priest, G. (2014). Revising logic. En P. Rush (Ed.), The Metaphysics of Logic. (pp. 211-223). Cambridge University Press. https://doi.org/10.1017/CBO9781139626279.016

Schmidt, J. A. (2025). Peirce and modal logic: delta existential graphs and pragmaticism. Cognitio: Revista de Filosofia, 26(1), e60449. https://doi.org/10.23925/2316-5278.2025v26i1:e60449

Wolter, F., & Zakharyaschev, M. (1999). Intuitionistic modal logic. En Logic and Foundations of Mathematics (pp. 227-238). Springer Netherlands. https://doi.org/10.1007/978-94-017-2109-7_17

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 473-484
ISSN 3087-260X
Santiago de Chile

EVITANDO MALINTERPRETAR EL MÍTICO TEOREMA DE INCOMPLETUD DE GÖDEL: QUÉ ES LO QUE EL TEOREMA DICE EN REALIDAD[1]

J. Sebastián Castillo[2]

1. Introducción:

El teorema de incompletud de Gödel es uno de los pocos teoremas pertenecientes a las matemáticas superiores cuya fama ha trascendido mucho más allá de las fronteras de esta disciplina. En Filosofía ha sido comentado y discutido, y en ocasiones se lo ha intentado utilizar en contextos distintos del que le es propio. A ese respecto, es atingente

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima.* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Castillo, S. (2026) Evitando malinterpretar el mítico teorema de incompletud de Gödel: qué es lo que el teorema dice en realidad. *Revista de Filosofía Homónima, 1*(1), pp. 453-484. https://doi.org/xxxxxxx[Por asignar]

[2] Sebastián Castillo es Doctor en Ciencias, con mención en Matemáticas de la Universidad de Chile; Profesor de lógica y lógica matemática; Profesor asociado del departamento de Filosofía y Humanidades de la Universidad de Chile; y Profesor adjunto del Departamento de Matemática de la Pontificia Universidad Católica de Chile. Email: castillo@uchile.cl.

mencionar el hecho de que su enunciado es notoriamente técnico, por lo que se lo suele citar en versiones muy simplificadas, que cambian su verdadero sentido; a veces, incluso, se lo trata como si simplemente dijese "hay verdades indemostrables". La intención del presente artículo es explicar lo que realmente dice el teorema, de un modo no muy técnico.
La organización que adoptaremos será la siguiente (y mencionamos desde ya que la complejidad técnica irá de menos a más):

— La sección 2 es una contextualización del tema, completamente no técnica.
— En la sección 3 enunciaremos el teorema en una versión simplificada con respecto a su (bastante técnico) enunciado original, y realizaremos comentarios que pretenden ayudar a comprender lo que realmente dice el teorema (tal es nuestro principal objetivo en el presente artículo). Aquí ya aparecen algunos tecnicismos, pero intentaremos reducirlos a un mínimo, y explicarlos brevemente.
— Finalmente, en la sección 4 esbozaremos someramente algunas de las ideas que utilizó Gödel para construir su demostración. Esta es la parte más complicada de seguir si no se tiene familiaridad previa con los temas en ella aludidos, e inevitablemente habrá que entrar en ciertos tecnicismos, pero es también "optativa", al estar después de la sección 3, que es la que cumple nuestro objetivo principal de explicar lo mejor posible el verdadero significado del teorema de Gödel.

Importante: Haremos el intento de reducir al mínimo los términos técnicos, de modo que la explicación sea accesible sin necesidad de conocimientos avanzados, pero debemos admitir que seguramente será necesaria una familiaridad básica con la lógica formal para comprender las explicaciones contenidas en la parte 3, y aún más en la parte 4: lamentablemente, si se quiere dar una idea más o menos aproximada del verdadero enunciado del teorema, es imposible evitar un mínimo de

términos propios de la lógica formal; mencionémoslos desde ya, para que los tenga usted en mente en caso de que les sean familiares: lenguaje formal, conjunto consistente (de afirmaciones), demostración formal, sistema axiomático. En caso de que no le sean familiares, mencionamos desde ya que incluiremos aquí las definiciones de algunos de estos términos.

2. El teorema de incompletud, su trascendencia y sus dificultades

Los usualmente llamados *teoremas de incompletud* (o *incompletitud*) del eminente lógico y matemático Kurt Gödel fueron publicados en 1931, y constituyen un hito trascendental en la historia de la lógica matemática y de la matemática en general. Además, el primero de estos dos teoremas (que es el que suele ser aludido como *el* teorema de incompletud de Gödel) es sin duda uno de los resultados pertenecientes a las matemáticas superiores que más ha sido aludido y comentado fuera de esta disciplina (¡probablemente el que más!).

En particular, en filosofía tiene evidente interés, y ha sido objeto de numerosos comentarios, discusiones, y algunos intentos de aplicación fuera de su contexto original. Al hacer esto último (intentar aplicarlo fuera de su contexto original), a veces se lo ha interpretado casi como si el teorema dijera "hay verdades que son indemostrables", y desde ahí se da el paso a que "por el mismo teorema, en tal o cual área (distinta de las matemáticas) tendremos verdades indemostrables". Pero lo que el teorema afirma no es simplemente que "hay verdades indemostrables"; lo que afirma es que en un cierto contexto *concretamente delimitado* existen *determinados tipos* de afirmaciones que no son demostrables ni refutables *dentro de los límites correspondientes a ciertas herramientas.*

Pretender que el teorema puede aplicarse directamente en otros ámbitos implica que no se comprende (o no se conoce) lo que en realidad dice. En relación con ello, destaquemos el hecho de que su enunciado es bastante técnico y, consecuentemente, las personas que se aproximan a él desde otras disciplinas lo hacen, casi siempre, a través de versiones

simplificadas —o ultrasimplificadas—, que pueden conducir fácilmente a una interpretación errada de este teorema. El aludido intento que en ocasiones se ha hecho de aplicarlo fuera del ámbito que le es propio no sólo es "riesgoso", sino que en propiedad es automáticamente equivocado: un teorema matemático dice *exactamente* lo que dice, ni un ápice más, y ello (lo que dice) se refiere estricta y específicamente a los términos técnicos que aparecen en él, los que a su vez han sido previamente definidos de manera precisa, en un contexto matemático específico.

Lo que haremos a continuación, entonces, será intentar describir el enunciado del primer teorema de incompletud de Gödel de manera accesible, pero al mismo tiempo lo más apegada que sea posible a su verdadero significado; y a continuación, en la última parte, esbozaremos algunas de las ideas principales de su demostración.

3. Lo que *realmente* dice el teorema de incompletud

Comenzaremos esta parte mencionando, sólo a título ilustrativo, el enunciado con el que aparece en el artículo original de 1931: "Para cada clase recursiva primitiva y ω-consistente K de fórmulas, existe un signo de clase r tal que ni $v\ Gen\ r$ ni Neg($v\ Gen\ r$) pertenecen a Flg(K), donde v es la variable libre de r".

> (La lectura del enunciado original hace ciertamente comprensible el hecho, que antes hemos mencionado, de que las personas que se aproximan al teorema desde otras disciplinas lo hacen, casi siempre, a través de versiones sumamente simplificadas.)

Enunciemos ahora, entonces, nuestra versión informal del teorema. Para ello, recordemos que, como dijimos en la introducción, si bien haremos el intento de reducir al mínimo los términos técnicos, es imposible evitar un mínimo de expresiones propias de la lógica formal; por ese motivo, a continuación de nuestro enunciado informal del teorema incluiremos un

recuadro que resume las definiciones involucradas en él. Y destaquemos, también, que *es muy importante tener en mente que ninguna versión informal del teorema expresará exactamente lo que él dice, siempre será sólo una aproximación* (ello es válido, de hecho, para cualquier teorema matemático de cierto grado de complejidad técnica, y también es válido para el esbozo que realizaremos en la sección 4 de algunas de las ideas de la demostración).

Primer teorema de incompletud de Gödel (versión informal): Si se utiliza un lenguaje formal para establecer un sistema axiomático que es consistente y que permite formar un sistema de números naturales, entonces existe al menos una afirmación referente a dicho sistema axiomático que es verdadera pero que no se puede demostrar dentro del mismo sistema axiomático.

NOTA: Las siguientes las definiciones han aparecido en nuestra versión informal del teorema de incompletud (las plantearemos de manera más bien informal: no son definiciones realmente rigurosas — o derechamente son no rigurosas):
— Un lenguaje formal es un conjunto de símbolos, que posteriormente se usará para expresar simbólicamente determinados tipos de afirmaciones. A diferencia de los lenguajes naturales, en un lenguaje formal se fija desde el principio la totalidad de los símbolos que se usarán: no se pueden ir incorporando otros símbolos posteriormente.
— Un axioma es, en matemáticas, una afirmación que se da por aceptada sin demostración. El motivo por el que se la acepta puede ser que se la considera evidente, pero no necesariamente ha de ser por ello: en este contexto técnico, lo que importa es el hecho en sí de que se la acepta sin demostración; el motivo por el que no se pidió demostración alguna, es irrelevante. Al elegir un grupo de axiomas para proceder a realizar deducciones a partir de ellos, se ha establecido un sistema axiomático.

— Un conjunto inconsistente (de afirmaciones) es un grupo de afirmaciones a partir de las cuales es posible deducir una contradicción (es posible deducir a partir de ellas que una cierta afirmación es verdadera y que es falsa, cosa que *en principio* va contra lo aceptable). Lo contrario es un conjunto consistente (de afirmaciones): un grupo de afirmaciones a partir de las cuales es imposible deducir una contradicción.
—Los números naturales son "los que sirven para contar": 1, 2, 3, 4, 5, etc. Así que el número -3 no es un número natural, ni tampoco lo es el número ½, ni el número π. Mencionaremos también, por si hay dudas al respecto, que en algunos contextos matemáticos —como en teoría de conjuntos— se cuenta al cero como número natural. En general, en matemáticas, al trabajar en un contexto específico, es totalmente optativo elegir contar el cero como número natural o no hacerlo, sólo hay que asegurarse de explicitar la opción elegida.

Es importante mencionar que la "afirmación indemostrable" a la que se refiere el teorema no puede ser demostrada *dentro* del sistema axiomático con que se está trabajando, pero en principio es perfectamente factible que por fuera de dicho sistema sea posible construir una demostración de esta afirmación. Pero entonces, ¿por qué este teorema es considerado tan *demoledor* dentro de las matemáticas?: porque en matemáticas *siempre* se está trabajando dentro de *algún* sistema axiomático (a veces no explícitamente declarado), y una vez que nos hemos instalado ahí, por el teorema no podremos demostrar algunos resultados referentes a entidades de este mismo sistema (si el sistema cumple con las hipótesis del teorema). El tan extenso alcance que este teorema tiene en matemáticas se relaciona directamente con el hecho de que él no sólo se aplica a un sistema axiomático específico, ni a un grupo pequeño de ellos, sino a una cantidad infinita de sistemas axiomáticos, incluyendo a muchísimos de los que se utilizan en las matemáticas *reales*.

Observación: Según indica su enunciado, el teorema no es válido si el sistema adoptado no es lo suficientemente potente como para que en él podamos formar un sistema de números naturales; pero virtualmente todas las áreas de las matemáticas en las que en la práctica se trabaja requieren de la capacidad de contar, y por lo tanto necesitan disponer de un sistema de números naturales. Así que en la práctica el teorema de Gödel es aplicable a la mayor parte de la matemática *real*. Es atingente mencionar al respecto que en muchas áreas de la matemática los axiomas que permiten formar a los números naturales se suelen dar por subentendidos; es frecuente no explicitarlos, salvo en algunas áreas.

Observación: La otra excepción posible a la aplicabilidad del teorema, según indica su enunciado, es que el sistema axiomático adoptado sea inconsistente; pero un sistema tal es matemáticamente inútil: si de un sistema axiomático sabemos que es inconsistente, no lo utilizaremos. En propiedad, hay muchísimo que decir respecto a la consistencia e inconsistencia de sistemas axiomáticos, sobre todo a partir del *segundo* teorema de incompletud de Gödel; pero ello escapa por completo al alcance del presente artículo.

4. Esbozo de algunas ideas de la demostración del teorema

Como se mencionó en la introducción, esta sección final es la parte más complicada de este artículo si no se tiene familiaridad previa con los temas en ella aludidos, y aquí ya no podremos evitar entrar en algunos tecnicismos; pero esta parte es también "optativa", en el sentido de que ya hemos ejecutado nuestro principal objetivo: dar una idea informal del enunciado del teorema, de una manera lo menos técnica posible, pero al mismo tiempo lo más apegada posible a su verdadero significado.

Vamos entonces con nuestro esbozo de algunas de las ideas sobre las que se construyó la demostración del teorema. Para lograr esta demostración, Gödel procedió a asignar un número natural a cada símbolo del lenguaje

formal, y a partir de ello asignar un número natural a cada *afirmación* construida con el lenguaje formal (afirmaciones técnicamente llamadas "fórmulas bien formadas"), y a partir de ello asignar un número natural a cada *demostración* formal dentro del sistema. Una demostración consiste en una sucesión de afirmaciones relacionadas de ciertos modos, pero aquí lo importante es el hecho en sí de que es una sucesión de afirmaciones; dados números naturales asignados a dichas afirmaciones, el procedimiento de Gödel asigna, en base a ellos, un número a la sucesión de afirmaciones que es la demostración. Tras ello, el procedimiento que ideó Gödel permite, además de lo anterior, hacer algo notabilísimo: "codificar aritméticamente" ciertas *afirmaciones referentes a demostraciones realizadas dentro del sistema axiomático.* Compréndase bien: no sólo hablamos de asignar números a demostraciones, sino de codificar aritméticamente *afirmaciones referentes a demostraciones.* En términos más técnicos, el procedimiento de Gödel "codifica aritméticamente" ciertos conceptos metamatemáticos referentes al sistema axiomático y lógico utilizado, como el concepto de "sustituir una variable en una fórmula" o el de "ser demostrable dentro del sistema". Esta es, por mucho, la parte más sofisticada (y más técnica, y más extensa) de la demostración, y debe considerarse sencillamente brillante: es difícil sobreestimar la genialidad de Gödel, tanto al haber concebido la idea de "codificar aritméticamente" conceptos metamatemáticos, como al haberlo logrado, pues idear un procedimiento que concrete esta idea es descomunalmente no trivial.

El hecho de que es posible hacer las mencionadas asignaciones de números tiene relación directa con un resultado de la teoría de conjuntos que indica que la unión numerable de conjuntos numerables produce un conjunto numerable. Sin entrar en detalles al respecto, digamos que, aunque los símbolos de un lenguaje formal son tantos como los números naturales, y las afirmaciones formadas dentro del sistema son tantas como los números naturales, y las demostraciones formales son tantas como los números naturales, por las particulares propiedades de los conjuntos

infinitos resulta que es posible asignar un número natural distinto a cada símbolo del lenguaje y a cada afirmación y a cada demostración formal. La idea de poner en correspondencia hechos aritméticos con propiedades metamatemáticas de sistemas axiomáticos está en el núcleo de la demostración del teorema de incompletud, y se logra tras el trabajo previo de "aritmetización" de símbolos, afirmaciones y demostraciones. Lo único que habremos de agregar, antes de finalizar este artículo, será un muy somero esbozo de las bases de este proceso de "aritmetización". Nos conformaremos, eso sí, con dar sólo una idea muy vaga (*tenemos* que conformarnos con ello, por el carácter no *tan* técnico dentro del que estamos limitándonos).

Queremos entonces visualizar cómo es posible asignar números a cada símbolo del lenguaje formal, y a cada afirmación (fórmula bien formada), y a cada demostración formal, de modo que la asignación sea "bidireccional": no sólo debe asignarse un número, sino que el número asignado a un símbolo, afirmación o demostración, no puede haber sido asignado a ningún otro símbolo, afirmación, o demostración. Hay más de un modo de realizar esta asignación, pero lo natural (y así lo hizo Gödel) es basarla en un muy importante teorema referente a los números naturales conocido como *Teorema Fundamental de la Aritmética* (¡su nombre refleja por sí mismo su importancia!). En virtud de este teorema, cada número natural puede escribirse como multiplicación de números primos, y de una manera única salvo por el orden en que se escriben los factores. Por ejemplo, el número 12 se puede escribir como el producto de factores primos $12 = 2 \times 2 \times 3$, o lo que es lo mismo, $12 = 2^2 \times 3^1$. Note que hay otros modos de escribir el número 12 como multiplicación de números naturales, por ejemplo $12 = 4 \times 3$, o también $12 = 2 \times 6$; pero estas otras representaciones no constan de sólo números primos (el número 4 no es primo, el número 6 no es primo). Para escribirlo como producto de primos, no hay otro modo que multiplicando dos números 2 y un número 3; sólo se podría variar el orden en que se escriben.

Para los efectos que aquí nos interesan, lo importante es que el Teorema Fundamental de la Aritmética establece una relación biunívoca

entre cada número natural y un determinado conjunto de números primos elevados a determinados exponentes, y como resultado de ello, vemos que especificar un determinado conjunto de números primos y un determinado exponente para cada uno de ellos es equivalente a especificar un determinado número natural. Como dijimos, no entraremos en los detalles del proceso utilizado por Gödel (si no se tiene familiaridad previa con procedimientos de ese tipo, resultan bastante complicados de seguir), pero la idea es que después de asignar un número a cada símbolo del lenguaje formal, al ahora tomar una afirmación (una fórmula bien formada), los números correspondientes a los símbolos que aparecen en ella se usan para asignar exponentes a los primeros números naturales primos (2, 3, 5, 7, 11, ...). Por ejemplo, supongamos que al símbolo de igualdad se le asignó el número 2 y que al símbolo del número cero se le asignó el número 5; entonces a la afirmación 0=0 se le asignaría el número $2^5 \times 3^2 \times 5^5 = 900.000$. El punto es que el Teorema Fundamental de la Aritmética garantiza que con este procedimiento ningún número natural puede corresponder a más de una afirmación (fórmula bien formada), porque la descomposición en factores primos de cualquier número natural es única. En el ejemplo anterior, la única manera de descomponer el número 900.000 en factores primos es $2^5 \times 3^2 \times 5^5$. Y de modo análogo se lleva a cabo la codificación numérica de las demostraciones. En cuanto a las afirmaciones metamatemáticas, no podemos aspirar a describir aquí ni siquiera una idea aproximada de su construcción, pues ella requiere conceptos y procedimientos muchísimo más complejos.

Si bien es difícil que la muy breve descripción que hemos realizado permita hacerse una idea realmente clara del procedimiento concreto de la demostración (salvo si desde antes se contaba con conocimientos al respecto), al menos dará una cierta noción del tipo de ideas que permiten hacer corresponder hechos aritméticos con entidades relacionadas con sistemas axiomáticos. Concluiremos este artículo con dos observaciones finales:

— La descripción que hemos realizado de algunas ideas de la demostración del teorema permite comprender la necesidad de la hipótesis, en su enunciado, referente a que el sistema axiomático permita formar un sistema de números naturales: la demostración requiere, intrínsecamente, disponer de los números naturales.

— Repitamos algo ya mencionado, pero que es importantísimo de comprender y recordar: *ninguna versión informal del teorema expresará exactamente lo que dice el teorema, siempre será sólo una aproximación* (y mientras menos técnica es la versión, más alejada estará del verdadero significado del teorema). Lo mismo es válido para cualquier teorema matemático de cierto grado de complejidad técnica, así como para el esbozo que hemos realizado de algunas de las ideas de la demostración del teorema de Gödel, en que inevitablemente nos hemos visto en la obligación de incurrir en omisiones e inexactitudes, en vista de los límites dentro de los que aquí nos hemos situado.

BIBLIOGRAFÍA RECOMENDADA

1. En internet, en https://users.dcc.uchile.cl/~cgutierr/otros/godel.pdf se encuentra disponible el artículo *El Teorema de Incompletitud de Gödel (Versión para no iniciados)* de Claudio Gutiérrez, académico del Departamento de Ciencias de la Computación de la Universidad de Chile. En este artículo (originalmente publicado en Revista Cubo Mat. Educ., Universidad de la Frontera, Vol. 1, 1999, páginas 68-75) se entrega una resumida contextualización histórica y filosófica de los teoremas de incompletud de Gödel, y se resumen también los pasos principales de la demostración de estos dos teoremas.

2. En la dirección web: https://www-quantamagazine-org.translate.goog/how-godels-proof-works-

20200714/?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc se entregan varios detalles sobre el modo en que, en la demostración del teorema, se asignan números naturales y se hacen corresponder hechos aritméticos con propiedades de sistemas axiomáticos. El texto está en inglés, pero si se accede vía Google se ofrece opción de traducción.

3. En lengua hispana tenemos el privilegio de contar con una traducción que abarca casi todos los trabajos y escritos de Gödel, realizada por Jesús Mosterín, quien también elaboró comentarios introductorios para muchos de los artículos. El libro es K. Gödel, *Obras Completas*, Alianza Editorial, Madrid, 1981, 1989, 2006, 2018. En él se encuentra, en particular, "Sobre sentencias formalmente indecidibles de Principia Mathematica y sistemas afines", traducción a nuestro idioma del artículo original de Gödel publicado en 1931: "Über formal unentscheidbare Sätze der Principa Mathematica und verwandter Systeme".

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 485-503
ISSN 3087-260X
Santiago de Chile

BREVE GUÍA DE CÓMO ESCRIBIR UN ENSAYO FILOSÓFICO[1]

Alan Martin[2]
Kuminak Lefio[3]
Manuel Rodríguez[4]
Oscar Morales[5]

[1] Publicado: 27 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Cómo citar: Martin, A., Lefio, K., Rodríguez, M., & Morales, O. (2026). Breve guía de cómo escribir un ensayo filosófico. *Revista de Filosofía. Homónima, 1*(1), pp. 485-503. https://doi.org/xxxxxxx[Por designar]

[2] Doctor en Filosofía de la Universidad de Chile. Prof. Asistente de la Universidad de Chile. Director de la Escuela de Postgrado de la misma casa de estudios. Email: amartin@uchile.cl.

[3] Licenciado en Filosofía de la Universidad de Chile, estudiante del diplomado "Fundamentos de la Física" de la misma casa de estudios. Actualmente se desempeña como jefe en Gestión Editorial de la Revista Homónima. Email: kuminak.lefio@ug.uchile.cl. ORCID: https://orcid.org/0009-0005-8121-4063.

[4] Doctor en Filosofía. Prof. Asistente de la Universidad de Chile. Coordinador del Magíster en Estudios Cognitivos de la Universidad de Chile. Email: marodrig@uchile.cl.

[5] Oscar Morales es Licenciado en Filosofía, Universidad de Chile y estudiante de Magíster de la misma casa de estudios. Email: moralesbravo58@outlook.com. ORCID: https://orcid.org/0009-0002-3996-0028

¿Qué es un ensayo filosófico?

Un ensayo filosófico es una forma de argumentación racional cuyo propósito no consiste en persuadir mediante recursos retóricos ni en apelar a la autoridad del autor, sino en establecer la validez lógica, la solidez y la coherencia de los argumentos que se emplean para sostener o refutar una proposición filosófica. Se trata de un género dentro del amplio espectro de la literatura filosófica, en el que conviven múltiples estilos: desde los diálogos platónicos y los aforismos de Nietzsche hasta las meditaciones fenomenológicas, la prosa poética de Heidegger o Sartre y los artículos académicos contemporáneos. El ensayo constituye una entre diversas formas legítimas de hacer filosofía, y su valor distintivo radica en articular la originalidad interpretativa (y propositiva) con la disciplina argumentativa.

Aunque no existe un modelo único que prescriba cómo debe estructurarse un ensayo filosófico, sí se reconocen ciertos criterios de rigor indispensables: coherencia interna, claridad conceptual, fundamentación de las tesis, diálogo crítico con la tradición y con la bibliografía pertinente, así como un orden expositivo que facilite la comprensión del argumento. En este sentido, el ensayo filosófico es un ejercicio intelectual que exige precisión terminológica y un desarrollo argumentativo transparente, de modo que tanto las inferencias como los supuestos queden explícitos, permitiendo al lector evaluar de manera autónoma la plausibilidad de las conclusiones.[6]

En lo anterior, nos encontramos con una primera disyuntiva inherente a nuestra práctica filosófica, a la que tanto académicos como

[6] Esto no implica que los textos que fallan en cumplir estos criterios no pueden ser filosóficos. Debido a que no se trata de criterios que definan lo que *es* un ensayo filosófico, sino que nos guían normativamente acerca de cómo *deberíamos* practicar la escritura filosófica. La claridad puede ser una virtud intelectual deseable en la filosofía, pero no necesariamente está presente en todos los textos que catalogamos como filosóficos.

estudiantes, de una u otra forma, tenemos que hacer frente. Por un lado, el carácter permanentemente inconcluso de la filosofía hace que no tengamos una única metodología y estándar ya prescritos para cómo debemos escribir (y más ampliamente, reflexionar) acerca de problemas filosóficos. En otras palabras, es razonable que nos cuestionemos quién tiene la justificación última acerca de las formas correctas en las que se debe escribir filosóficamente. Y el estudio profuso por su cuenta parece no resolver este asunto ya que, la fuerza de dicho cuestionamiento sólo aparenta aumentar a medida que nos familiarizamos con la variedad inmensa de tradiciones filosóficas diversas y divergentes entre sí, que históricamente se han posicionado respecto a cómo hacer filosofía.

La otra cara de la disyuntiva se presenta al notar que la filosofía como práctica misma, se fundamenta en la capacidad de comunidades humanas de sostener estándares comunes. Esto es lo que permite que aquello que se piense y se escriba no sea sólo comunicable, sino que además evaluable y discutible de forma intersubjetiva por otros que compartan nuestra inquietud filosófica. Es decir, tampoco se puede dar la práctica filosófica como la conocemos, sin acuerdos acerca de cómo se debe o es recomendable pensar, escribir y discutir los problemas filosóficos.[7] Y por suerte, cuando cada uno de nosotros se enfrenta a este dilema, ya existen estándares dados, estos vienen de nuestro contexto cultural particular y de una multitud de líneas del pensamiento que han intentado hacer lo mismo a través de la historia de la filosofía.

[7] Estos acuerdos pueden ser tácitos, a veces pueden evidenciarse en la práctica de los filósofos y en ciertas ocasiones, cuando un filósofo o una corriente ponen el foco en los aspectos metodológicos de la filosofía, se discuten y defienden explícitamente. Muchas veces podemos ignorar estos acuerdos e identificar lo filosófico por un contenido particular o inclusive por lo que intuitivamente nos parece profundo, pero eso no niega la necesidad de que estos criterios acordados estén allí, les pongamos atención o no (y como buenos filósofos, en algún momento nos causara curiosidad indagar sistemáticamente sobre ellos).

Así la invitación que parece hacernos esta primera disyuntiva es a conservar nuestro espíritu crítico, evaluando profundamente la metodología y objetivos que adoptamos en nuestra escritura filosófica, y en el proceso, desarrollar una perspectiva filosófica propia con la que afrontar la escritura filosófica. Al mismo tiempo, que intentamos cumplir con los estándares intersubjetivos de evaluación que posibilitan el desarrollo de la filosofía en comunidad y que nos permiten aprender de nuestros maestros y compañeros, tanto como de las lecturas de quienes antes y ahora se enfrentan a la tarea de escribir filosofía.

El ensayo filosófico no tiene como propósito exponer temáticas generales de la disciplina, sino examinar debates, proposiciones o conceptos específicos con el fin de aportar críticamente al estado del arte. Su objetivo es investigar y no divulgar. Solo una propuesta filosófica controvertida es capaz de tener un impacto en el debate filosófico, y no se trata meramente de tener un nombre llamativo, sino de aquella propuesta que es capaz de mostrar un problema, mostrar su importancia y abordarlo con fuerza y rigor argumentativo. Por esta razón, todo ensayo debe responder a una pregunta filosófica clara y bien delimitada, inscrita dentro de un debate ya establecido en la tradición[8], de la cual se desprende la tesis que se defiende y la estrategia argumentativa que le da sustento. Al igual que un organismo vivo, el ensayo se compone de partes que cumplen funciones particulares y que lo estructuran en un inicio, un desarrollo y un cierre.

Ahora bien, las partes de un ensayo filosófico no dependen de la contrastación empírica ni de demostraciones matemáticas como ocurre en las ciencias para comprobar la verdad de una hipótesis. Sin embargo, esto no exime al ensayista de ofrecer razones justificadas. El juicio crítico en filosofía se apoya en la consistencia lógica, en la claridad de las

[8] Pueden existir ensayos que desencadenan en nuevas tradiciones, estos no son mayoritarios pero existen. Sobre estos es importante notar que aun así tienen un diálogo con su contexto filosófico y tradiciones previas.

definiciones y en la capacidad de situar la propia aportación en relación con posiciones previas. El motor de la filosofía es, en este sentido, la argumentación racional, entendida como el esfuerzo constante por fundamentar las ideas, someterlas a examen y dialogar con tradiciones ya existentes sin renunciar a la posibilidad de innovar en ellas. De forma complementaria, es importante considerar cómo la filosofía se encuentra inmersa en un medio académico e investigativo donde otras disciplinas con sus propios supuestos y metodologías se encuentran en una constante teorización que roza o directamente se funde con las temáticas que nosotros problematizamos. Y sin poner en riesgo la independencia de la filosofía, este intercambio de perspectivas, a veces más crítico y otras más acogedor, formará parte de nuestro medio de argumentación racional.

Así, dado el creciente interés suscitado en la comunidad filosófica por participar en la nueva *Revista de Estudiantes de Filosofía Homónima*, el equipo de ensayos sobre lectura y escritura filosófica ha considerado pertinente ofrecer una breve guía acerca de cómo elaborar un ensayo filosófico. Más que proponer una fórmula rígida, estas orientaciones son más bien sugerencias generales dirigidas a quienes deseen publicar, con el objetivo de brindar una visión clara sobre los elementos fundamentales que conforman un ensayo en filosofía. También, es nuestro objetivo suscitar el tipo de inquietudes metafilosóficas que son propias de nuestro quehacer filosófico, especialmente cuando intentamos tomar decisiones respecto a cómo acercarnos a la filosofía, y en este caso particular, cómo escribir al respecto.

1. Estructura de un ensayo

1.1. Título

El título constituye el rostro de toda investigación y actúa como el primer filtro a través del cual los lectores se aproximan a su contenido. De ahí que no solo debe guardar coherencia con el desarrollo del texto, sino

también ofrecer una síntesis clara y precisa de su propósito. En este sentido, el título cumple la función de condensar en pocas palabras el problema central que la obra aborda, orientando al lector respecto de sus objetivos y alcances. Por ello, es aconsejable evitar fórmulas enigmáticas, excesivamente generales o vagas, que dificulten la comprensión inmediata del tema tratado. Un buen título, en cambio, debe mantener el equilibrio entre concisión, rigor conceptual y claridad expositiva, de manera que logre captar el interés sin sacrificar exactitud.

1.2. Introducción

La introducción cumple un papel decisivo en todo escrito filosófico, pues orienta al lector en el objetivo del texto, al igual que el debate en el que se inscribe la investigación y define de manera explícita la perspectiva desde la cual será abordado. En este apartado debe formularse con claridad la hipótesis de investigación, utilizando un lenguaje preciso y evitando expresiones vagas o ambiguas. Una declaración inicial como "en este trabajo sostendré que…" no solo aporta transparencia, sino que también ofrece al lector una guía inmediata sobre la tesis que se defenderá a lo largo del texto. Además, es conveniente asumir que el lector posee cierto grado de familiaridad con la temática, aunque no necesariamente sea un experto, por ello es importante explicar los conceptos técnicos que resulten imprescindibles para la comprensión del argumento, cuidando siempre que dichas aclaraciones no interrumpan el flujo de la exposición ni desvíen la atención de la idea principal.

Además de exponer la hipótesis y los conceptos clave, la introducción debe situar al lector en el contexto del problema y mostrar de qué manera se aborda dentro de la extensa tradición filosófica en la que se inscribe. Es importante señalar si el enfoque adoptado corresponde a una reconstrucción fenomenológica, un análisis conceptual, un contraste crítico o una reconstrucción hermenéutica, por

mencionar algunos. Esto permite al lector comprender la posición de la obra dentro del espectro filosófico y la forma particular en que enfrenta el problema.

En muchos casos, la introducción también cumple la función de justificar la relevancia filosófica del problema planteado. Aunque en filosofía no existen debates clausurados ni temas que no puedan ser legítimamente tratados, a menudo resulta necesario explicar por qué la cuestión seleccionada merece atención y cuál es su aporte al desarrollo del pensamiento dentro del campo analizado. Así, un debate que fue relevante hace mucho tiempo puede ser nuevamente examinado, siempre que la discusión propuesta consiga aportar algo significativo a la reflexión filosófica en curso.

En este apartado inicial del ensayo, presentar con claridad el problema filosófico y la importancia del mismo, no tiene por que ser solo un paso rutinario que nos vemos obligados a cumplir como una mera formalidad, sino que puede ser una parte esencial de la investigación. Se trata de un espacio en el que podemos entender en profundidad los detalles del problema al que nos enfrentamos y los fundamentos que nos permiten formularlo como un problema filosófico (propio de la filosofía y abordable por sus metodologías). Como comprendamos, y seamos capaces de expresar, el problema será nuestra *primera guía* en el camino que emprendemos al intentar darle una respuesta satisfactoria en el resto de nuestro ensayo filosófico. Alternativamente, si nos encontramos en una situación tal que solo contamos con una idea vaga del problema que queremos abordar, es importante no paralizarse y dejar de investigar acerca de estos tópicos que nos interesan. En este caso, llegar a conocer y plantear de forma clara el problema de nuestra investigación será un *objetivo fundamental* que nos permitirá progresar en la investigación, motivándonos a estudiar rigurosamente el tema y esclarecer a cuál aspecto específico del mismo apuntan nuestras ideas. De una u otra forma, los problemas de nuestros ensayos son importantes.

Así mismo, la introducción cumple la función de trazar una línea de argumentación que anticipe al lector el desarrollo del ensayo. Para ello, resulta útil emplear expresiones como "la primera parte de este trabajo tiene como propósito…", "a continuación se desarrolla…" o "con el fin de demostrar…", que señalan de manera clara qué se abordará en cada sección y cómo se conectan entre sí los distintos apartados del texto. Este recurso no solo orienta al lector, sino que también obliga al autor a reflexionar sobre la organización de su trabajo, asegurando que cada sección cumpla un propósito específico dentro del conjunto y que la progresión argumentativa sea coherente. Así, enunciar explícitamente las partes que componen el ensayo permite establecer un mapa estructural que facilita la comprensión del texto y guía la lectura de manera efectiva.

Si bien es cierto que en la mayoría de las ocasiones no contamos con un conocimiento completo de todos los puntos que abordaremos ni de la forma en que los desarrollaremos, es común que, a medida que escribimos, surjan nuevas ideas que modifiquen el rumbo inicial del texto. Esto forma parte natural del proceso de escritura, pero para no perder claridad resulta fundamental precisar ciertos aspectos antes de comenzar. Incluso cuando no tengamos definida la conclusión a la que esperamos llegar, toda introducción debería responder a tres preguntas generales, con independencia del contenido específico del texto, que orientarán el desarrollo del ensayo: ¿qué pretende el autor?, ¿en qué debate se inscribe el problema que se abordará? y ¿cuáles son los argumentos que se propone desarrollar? Estas preguntas, aunque preliminares, ofrecen un marco de referencia indispensable que ayuda a dotar de orden, dirección y límites a la investigación.

1.3. Desarrollo argumentativo

Si bien, como mencionamos antes, la filosofía no recurre a métodos empíricos ni a demostraciones matemáticas, la manera en que somete a

prueba una hipótesis de investigación es mediante la construcción de argumentos estructurados con lógica[9] y claridad. La tarea central de un ensayo filosófico consiste en articular un esquema argumentativo en el que un conjunto de premisas conduzca de manera justificada a una conclusión. Un argumento no se reduce a la mera enumeración de afirmaciones, sino que constituye una cadena inferencial en la que cada premisa desempeña una función justificadora específica y explícita. En este sentido, un argumento puede definirse como un conjunto de proposiciones en el cual algunas de ellas —las premisas— se ofrecen en apoyo de otra —la conclusión—. La exigencia fundamental de todo autor consiste en mostrar la pertinencia de cada premisa y su conexión con la conclusión, explicando con rigor por qué la aceptación de aquellas debe llevar razonablemente a aceptar esta última.

Ciertamente, la argumentación filosófica posee sus particularidades y su contenido puede articularse de distintas maneras. Sin embargo, en términos formales todo debate racional debe ajustarse a ciertos criterios que permitan determinar la validez de los razonamientos. No todo razonamiento es aceptable, por lo que conviene distinguir entre validez y solidez. Un argumento es válido cuando la conclusión se sigue lógicamente de las premisas, independientemente de que estas sean verdaderas o falsas. En cambio, un argumento es sólido cuando, además de ser válido, las premisas resultan efectivamente verdaderas.

Un punto importante en un argumento es que las premisas sean suficientes para aceptar la conclusión. Sin embargo, el concepto de "suficiencia" no es del todo claro. Dammer (2009, p. 8) por ejemplo señala que las premisas deben ser suficientes en cantidad y calidad, pero también se puede entender suficiencia en términos de la lógica

[9] Esto no siempre implica recurrir a sistemas formales de la lógica, más generalmente se trata de la validez, relevancia, aceptabilidad y suficiencia de nuestros argumentos. Para lograr esto la lógica nos puede ser sumamente útil, pero no por ello está siempre presente de forma explícita.

proposicional, en particular en la relación que se da en el condicional desde el antecedente hacia el consecuente. Con el objetivo de aportar claridad en este respecto, es aconsejable determinar qué tipo de razonamiento es el que se está utilizando en el ensayo, y cuáles son las reglas formales de cada uno:

En el caso de un razonamiento deductivo, la verdad de las premisas implica necesariamente la verdad de la conclusión. En ese sentido, si bien la verdad de la conclusión no está contenida en ninguna de las premisas por separado, la adecuada articulación de las premisas debe contener a la conclusión. usualmente en este tipo de razonamiento hay una premisa de carácter general, una ley o principio, y lo que se intenta demostrar es que el fenómeno particular bajo escrutinio es un caso de esa premisa general. Por ejemplo, en un debate ético, se intentará demostrar que una opción particular responde a un principio ético universal (o bien que es incompatible con dicho principio, para refutarla). En estos casos, la conclusión se sigue necesariamente de las premisas, sin excepción, por lo que basta con mostrar un contraejemplo para refutar la tesis, es decir, es suficiente mostrar la posibilidad (conceptual o empírica) para refutar el argumento.

El razonamiento inductivo puede utilizarse para justificar la adopción de algún principio o regla general que no sea autoevidente. En este tipo de argumento las premisas intentan hacer probable la conclusión. El procedimiento consiste en enumerar casos que puedan contarse como del mismo tipo, para evidenciar que hay una conexión entre el antecedente y el consecuente, o, como se expresaría normalmente, entre la causa y el efecto. Aquí la noción de causalidad se puede entender en dos sentidos: como correlación entre dos eventos (los eventos del tipo A van seguidos de eventos del tipo B), o en términos contrafácticos (si A no hubiera sucedido, entonces B no habría sucedido, o bien que si B no hubiera sucedido tampoco A habría sucedido). En ambos casos es importante fijarse en dos aspectos: que los eventos puedan

ser caracterizados como siendo del mismo tipo, y que la muestra sea representativa para establecer la probabilidad, lo que rescata las anteriores nociones de cantidad y calidad, por un lado, y la idea de que en un condicional el antecedente es una condición suficiente para el consecuente y que el consecuente es una condición necesaria para el antecedente. Por ejemplo, si se pretende defender la tesis de que el existencialismo es una consecuencia de las condiciones de la Europa de entreguerras, habrá que mostrar, al menos, que hay una conexión conceptual entre lo significativo de ese periodo y ese tipo de pensamiento, y/o que hay otros casos históricos que son semejantes en los aspectos relevantes y donde se dio un caso similar de conexión.

Finalmente, cuando no contamos con principios generales ni tampoco tenemos una guía para caracterizar tipos de casos como semejantes, es decir en un caso problemático completamente novedoso, es necesario formular una hipótesis que pueda ser testeada inductivamente, y en caso de éxito poder contar con una regla general. El tipo de razonamiento que conduce a la formación de hipótesis es llamado por la tradición como "abductivo", reconociendo la formulación de C. S. Peirce (1955), que se caracteriza por esta necesidad de dar un salto creativo (pero no por ende invalido) al intentar explicar aquello que nos sorprende de la realidad y para lo cual los conceptos, esquemas y principios con los que ya contamos no dan abasto. Si bien no hay una formalización clara de este tipo de razonamiento, pues justamente es una discusión actual el cómo llegamos (ámbito descriptivo: psicología, neurociencia, ciencia cognitiva, filosofía de la mente, etc.) y cómo deberíamos llegar a nuestras ideas (ámbito normativo: epistemología), si hay al menos dos casos claros: la inferencia a la mejor explicación (IME) y el razonamiento metafórico. En el caso del razonamiento metafórico se intenta comprender un fenómeno nuevo en términos de un fenómeno ya conocido, distinto pero con semejanzas que se estiman relevantes, y se evalúan las consecuencias que se siguen de aplicar la metáfora. Por ejemplo, si la mente fuera al cuerpo como el software es al hardware, ¿que

se seguiría para el estudio de lo mental? si la hipótesis resulta fructífera entonces se acepta como guía de investigación. En el caso de IME, la pregunta se plantea entre hipótesis en competencia, y en este caso se evalúan tanto el potencial explicativo como la coherencia con el resto de lo que aceptamos como verdadero. Por ejemplo, si aceptamos la metáfora del computador para investigar lo mental, tenemos dos opciones: que la mente sea un procesador de información general que aprende de la experiencia o bien que la mente venga "pre-cargada" con programas específicos para cada dominio cognitivo. Dado que la evolución, a través de la selección natural, sólo puede funcionar en base al éxito (supervivencia, *fitness*), y que el éxito siempre es específico a alguna tarea, entonces parece ser una mejor hipótesis asumir que la mente no es una tabula rasa sino un conjunto de mecanismos específicos de procesamiento, al menos mientras no se cuestionen algunos de los supuestos del razonamiento.

Teniendo en cuenta, la cantidad de aristas que tienen los problemas en filosofía y la naturaleza misma de esta, es inevitable que nuestro ensayo filosófico tenga supuestos, por lo tanto, no podemos aspirar a que absolutamente todo lo que digamos quede profundamente justificado por otras premisas, también ellas igual de justificadas. Esto es verdadero para la filosofía en general, pero es aún más patente en textos que no pueden extenderse tanto, como lo es el ensayo filosófico. Esto no puede implicar que nos neguemos a justificar nuestras ideas, ni tampoco que intentemos ocultar estos supuestos al lector, por el contrario, nosotros debemos ser los primeros en captar nuestros supuestos y en mencionarlos transparentemente como tales. Como una buena regla general, allí donde nos parezca que no estamos suponiendo absolutamente nada, es mejor volver a revisar nuestro ensayo. Y allí cuando notemos los supuestos con los que nos comprometemos, revisar que sean lo suficientemente básicos y fáciles de aceptar cognitivamente, para elegir esos y no otros. A veces simplemente es práctico acotar nuestra investigación por medio de esta delimitación de supuestos, darlos por

sentados puede permitirnos desarrollar discusión hasta puntos significativamente más valiosos que a lo que llegaríamos si no lo hacemos.[10]

De igual forma, las premisas deben cumplir requisitos mínimos para ser presentadas válidamente en un argumento. Deben ser pertinentes al problema tratado, evitar trivialidades y estar libres de ambigüedades terminológicas. Tal como mencionamos, es recomendable que el autor explicite sus supuestos centrales y defina de manera operativa los conceptos controvertidos. Cuando una premisa sea debatible, debe ser defendida con razones sólidas o respaldada mediante referencias bibliográficas pertinentes, en lugar de dejarla como un supuesto implícito. También resulta fundamental evitar la circularidad entre premisas y conclusión, así como la inclusión de antecedentes irrelevantes, el diálogo con la bibliografía debe ser continuo. Situar la pregunta y la tesis en relación con posiciones previas exige mostrar a qué debates se responde, qué planteamientos se matizan y cuáles se rechazan. En la práctica, conviene priorizar las fuentes directamente relevantes y evitar listados extensos sin conexión argumentativa. Cada referencia debe aportar al planteamiento central, ya sea apoyándolo, problematizándolo o sirviendo de objeto de crítica.

La construcción de un buen argumento exige considerar distintas perspectivas y no limitarse a la propia. Un autor riguroso se adelanta a las objeciones más probables en su versión más robusta y ofrece respuestas bien fundamentadas, lo que fortalece la consistencia de su propuesta y

[10] De esto se desprenden un par de recomendaciones interesantes, por un lado, es deseable mantener nuestros supuestos estáticos y rígidos al interior de un ensayo y su correspondiente proceso de investigación. Sería muy difícil evaluar la validez de un ensayo que cambie a cada página los supuestos que sostienen sus posturas. Por otro lado, es virtuoso cambiar de supuestos entre distintas investigaciones, cuando la adopción de nuevos supuestos nos permite evaluar y justificar los anteriores. Esto es siempre y cuando no terminemos adoptando, a la vez y en el mismo sentido, presupuestos incompatibles entre sí.

evidencia que la tesis ha pasado por un examen crítico. En la práctica, resulta útil reservar un espacio del escrito para discutir objeciones y ofrecer réplicas, así como recurrir a contraejemplos o experimentos mentales cuando ayudan a poner a prueba la solidez de las premisas. Saber incorporar las críticas sin distorsionarlas ni debilitarlas de forma artificiosa es, en efecto, un rasgo que distingue la madurez filosófica. Por lo mismo, esto último es, a la vez, un ejercicio que debería fortalecer nuestra postura y entrenarnos en ser caritativo con las de los demás.

Para lograr lo anterior es una buena práctica (y comúnmente difundida) apoyarse en las objeciones que ya existen en la literatura sobre el tópico en discusión para nuestra posición o aquellas que son cercanas. Esto nos permite entrar en diálogo con lo que realmente sostienen investigadores que no opinan como nosotros, que comúnmente son posturas más complejas, matizadas y difíciles de confrontar que lo que nuestra imaginación ofrece, pero eso es justamente lo que le da riqueza al debate (y por ende a nuestro ensayo).[11]

Al considerar estos contraargumentos siempre hay que cuidar no caer en la "Falacia del hombre de paja", es decir, en lugar de probar si los argumentos de la proposición filosófica contrincante son válidos y hacerse cargo de ellos desde su mejor versión posible, se simplifican hasta distorsionar el sentido del argumento, solo con el fin de desacreditar la postura del oponente.

Ej.: A: para el utilitarismo lo correcto consiste en aquello que produce mayor placer para el mayor número de personas

[11] Esto nos exige que en el proceso de investigación salgamos de nuestra postura e indaguemos más allá de aquello con lo que estamos de acuerdo. Debemos apuntar a estar informados y actualizados del debate que nos compete en general, no solo de aquello que resuene amablemente con nuestras premisas, argumentos y conclusiones.

B: el utilitarismo es una doctrina inmoral que no considera el sufrimiento de quienes no conforman la mayoría

Conclusión

La conclusión constituye el momento de cierre del ensayo y cumple la función de integrar de manera coherente los elementos desarrollados en el cuerpo del texto. Su redacción exige un ejercicio de síntesis que no debe confundirse con una mera repetición de lo ya expuesto. Antes bien, se trata de condensar los aspectos más relevantes de la argumentación, de modo que el lector pueda reconocer con claridad cuál ha sido el itinerario seguido y qué resultados se han alcanzado. En este sentido, la conclusión debe, en primer lugar, ofrecer una respuesta explícita a la pregunta filosófica formulada en la introducción. El criterio decisivo es que dicha respuesta sea una consecuencia de las premisas y argumentos presentados a lo largo del trabajo. La introducción de consideraciones que no han sido previamente desarrolladas constituye un error metodológico, pues equivale a elaborar una conclusión sin fundamento, lo que ha sido metafóricamente denominado como la "falacia del conejo en el sombrero". En segundo lugar, una conclusión bien estructurada reconoce los límites de la investigación. El ensayo filosófico no pretende clausurar de manera definitiva los problemas tratados, sino contribuir a su esclarecimiento desde una perspectiva particular. Por ello, resulta pertinente señalar los aspectos que han quedado fuera del análisis o que requieren un examen más detallado, a fin de evitar la pretensión de haber agotado el tema en su totalidad. Por último, la conclusión no se reduce a un cierre retrospectivo, sino que cumple también una función prospectiva. A partir de las implicancias de lo desarrollado, puede abrir nuevas líneas de reflexión y plantear interrogantes que orienten investigaciones futuras. De este modo, el ensayo no solo concluye, sino que se proyecta hacia horizontes de discusión más amplios

Epílogo: *Algunas consideraciones importantes a tener en cuenta*

La elaboración de textos de investigación filosófica es un proceso público y riguroso que exige del autor un compromiso con la integridad ética de su obra. Por ello, es fundamental considerar ciertos aspectos éticos y formales.

Al autor debe velar por la integridad ética de su obra en al menos tres sentidos: 1) aunque la filosofía no suele involucrar investigaciones con seres sintientes, con frecuencia aborda temas sensibles que pueden afectar a las personas, por lo que cada análisis debe realizarse con el debido respeto; 2) es esencial no tergiversar las ideas de otros autores para favorecer el propio argumento, ya que ello constituye una práctica inapropiada; 3) la honestidad académica requiere reconocer de manera explícita las ideas o palabras ajenas, evitando cualquier forma de plagio, incluido el autoplagio o el plagio involuntario. Asimismo, cualquier uso de herramientas de inteligencia artificial debe declararse de forma transparente y limitarse a funciones de corrección, búsqueda o apoyo técnico, sin reemplazar el trabajo crítico y reflexivo del investigador. Esto no implica negar el posible aporte que significan estas herramientas si se ocupan de manera adecuada, en cambio, el énfasis está puesto en que no reemplacen aquello que es esencial para que tengamos una discusión filosófica crítica y ética a través de la redacción de ensayos.

La escritura de un texto filosófico es un proceso lento, sin atajos, que exige madurar las ideas con paciencia y rigor. Esto implica reescribir y corregir el borrador cuantas veces sea necesario, no solo en lo formal, sino también en lo conceptual, con el fin de lograr una versión final clara y

coherente. A su vez, es fundamental integrar diversas perspectivas mediante una recopilación bibliográfica cuidadosa y un diálogo crítico con las fuentes. Cada fuente debe ser revisada con detención, asegurando que las referencias y citas estén correctamente consignadas y en plena correspondencia con los argumentos desarrollados en la obra.

Un buen escrito, antes de ser enviado a revisión por parte de una revista, debe haber pasado por un diálogo previo con nuestros pares más cercanos, quienes, en la medida de lo posible, han de brindar una retroalimentación honesta y crítica que nos permita reconocer qué aspectos conviene mejorar de cara al envío final. Este proceso no solo contribuye a detectar eventuales debilidades argumentativas, sino que también facilita la evaluación de la claridad y coherencia del texto, de la solidez metodológica y conceptual de los planteamientos, así como de la pertinencia de la contribución respecto de la literatura existente. Además existen instancias sumamente importantes, como los congresos, coloquios, simposios y "workshops" donde se pueden presentar dichas investigaciones más allá de nuestro grupo cercano, lo cual aumenta la diversidad de puntos de vista que pueden ofrecernos críticas y comentarios a evaluar.

La extensión de un ensayo no puede determinarse de manera exacta, pues lo esencial es que el texto responda con claridad a la pregunta que se propone abordar. Los problemas complejos requieren respuestas igualmente complejas, por lo que desarrollar un tema profundo en unas pocas páginas puede resultar insuficiente. De manera orientativa, un ensayo académico suele situarse entre cinco y treinta páginas; sin embargo, esta medida no es estricta, ya que existen textos breves que se han consolidado como referentes importantes en la filosofía, así como extensos ensayos del tamaño de tratados que también han marcado la tradición académica. Lo relevante es que la extensión del ensayo se ajuste a la complejidad del tema y permita argumentar lo suficiente para que nuestra conclusión sea sólida.

Con independencia de la solidez lógica de los argumentos, si un texto carece de una redacción cuidada, los errores formales distraen al lector y pueden incluso llevarlo a desestimar la obra. Por ello, la claridad y el rigor estilístico son tan importantes como la coherencia argumentativa en la escritura filosófica. Una buena redacción exige precisión terminológica, frases concisas que eviten ambigüedades, un orden expositivo que guíe al lector en el desarrollo de la argumentación y la supresión de repeticiones innecesarias. La falta de claridad en la expresión suele ser reflejo de una falta de claridad en el pensamiento. En este sentido, los aspectos formales no constituyen un adorno accesorio, sino una condición indispensable para que las ideas puedan ser comprendidas y valoradas.

Aunque existen diversas maneras de estructurar la redacción de un ensayo, resulta fundamental establecer un orden específico para organizar las ideas. Algunos autores prefieren elaborar el resumen o la introducción una vez finalizado el texto, mientras que otros siguen el orden natural del desarrollo del ensayo desde el inicio. Cada estrategia tiene ventajas y dificultades; por ejemplo, redactar la introducción al final permite mayor claridad sobre los objetivos y la estructura, pero seguir el orden lineal del texto puede ayudar a mantener un flujo argumentativo coherente. No obstante, ninguna de estas opciones garantiza resultados uniformes para todos los autores, por lo que cada quien debe encontrar el método que mejor se adapte a su estilo de escritura y al tipo de ensayo que está elaborando.

REFERENCIAS

Damer, T. E. (2009). *Attacking faulty reasoning: a practical guide to fallacy-free arguments.* Wadsworth Cengage Learning.

Peirce, C. S. (1955). *Philosophical Writings of Peirce: Selected Writings.* In J. Buchler (Ed.). (pp. 150-156). Dover Publications.

BIBLIOGRAFÍA RECOMENDADA

Candia, C. y Flores, T. (2025). Manual de buenas prácticas en la investigación. CEDEA, DICP, Dirección Escuela de Postgrado, Facultad de Filosofía y Humanidades, Universidad de Chile.

Chudnoff, E. (s.f.). A guide to philosophical writing. Harvard University. Recuperado de: https://hwpi.harvard.edu/files/hwp/files/philosophical_writing.pdf

Feinberg, J., y Shafer-Landau, R. (2001). Doing philosophy: A guide to the writing of philosophy papers (2nd ed.). Wadsworth.

Gorovitz, S., et al. (1979). Philosophical analysis: An introduction to its language and techniques (3rd ed.). Random House.

Martinich, A. P. (1997). Philosophical writing (2nd ed.). Blackwell.

Seech, Z. (2003). Writing philosophy papers (4th ed.). Wadsworth.

Solomon, R. (2001). Writing philosophy [Appendix]. In The big questions: A short introduction to philosophy (6th ed.). Wadsworth.

Homónima • Revista de Filosofía
Número 1, Vol. 1 (2026), pp. 504-509
ISSN 3087-260X
Santiago de Chile

Lenguaje, mente y cultura: RESEÑA "Cartografías Lingüísticas: un abordaje desde y hacia la interdisciplinariedad"[1]

Rodrigo Manríquez[2]
Universidad de Chile

El libro "Cartografías lingüísticas" reúne ensayos organizados en cuatro secciones: enfoque histórico-cultural, ciencias sociales, filosofía y ciencias cognitivas. La propuesta de sus autores es clara: promover la convergencia interdisciplinaria y el trabajo conjunto para abordar problemas contemporáneos, a partir de la premisa de que el lenguaje constituye un eje epistémico central para comprender la vida social y reflexionar sobre el conocimiento.

En esta reseña considero pertinente explicitar mi formación en filosofía, ciencias cognitivas e historia, ya que, al tratarse de un libro que

[1] Publicado: 29 / 01 / 2026. Revista Open Access 4.0. *Artículos de la Revista Homónima* (ISSN 3087-260X), Departamento de Educación y Ciencias Sociales de la Universidad Andrés Bello. Manríquez, R. (2026). Lenguaje, mente y cultura [Reseña del libro "*Cartografías Lingüísticas: un abordaje desde y hacia la interdisciplinariedad*", por Albornoz, N. et al (Eds.)]. *Revista de Filosofía Homónima, 1*(1), pp. 504-509. https://doi.org/xxxxxxx[Por designar]

[2] Rodrigo Manríquez es Magíster en filosofía de la Universidad de Chile. Email: rodrigo.manriquez@ug.uchile.cl. ORCID: 0009-0004-1529-9671

sitúa a la lingüística en el centro y la reivindica como una disciplina de gran potencia teórica, mi posición como no lingüista resulta significativa para dar cuenta de ello. Esta perspectiva, además, permite iluminar uno de los objetivos del libro: mostrar la relevancia de la lingüística en ámbitos diversos e invitando a lectores para quienes estos estudios suelen pasar desapercibidos. "Cartografías lingüísticas" demuestra no solo la potencia propia de la lingüística, sino también su valor como eje para una producción genuinamente interdisciplinaria, especialmente frente a problemáticas que difícilmente se comprenden desde una sola disciplina.

El libro me resulta especialmente relevante para debates clásicos de la filosofía analítica, donde se ha planteado una disputa central sobre el conocimiento: si debe abordarse prioritariamente desde el estudio del lenguaje o desde el análisis de la mente y su uso del lenguaje (Putnam, 1975; Hacker, 1998). A mi juicio, esta distinción ha sido contraproducente, particularmente en el ámbito académico, como lo muestran las investigaciones reunidas en este volumen, cuyos resultados desbordan dicha oposición. "Cartografías lingüísticas" rompe esa barrera y constituye un ejemplo claro de cómo el trabajo interdisciplinario puede generar argumentos, herramientas conceptuales y metodológicas eficaces en casos concretos. En lugar de jerarquizar disciplinas, el libro propone un eje común —la lingüística— que articula enfoques diversos y permite producir un conocimiento relevante, genuino y capaz de enfrentar los desafíos contemporáneos.

La primera sección, dedicada a la relación entre historia y lenguaje, reúne ensayos sobre lenguas tradicionalmente marginadas en Occidente y subraya el lenguaje como punto de partida clave para la investigación cultural e historiográfica. Esto se ejemplifica en "Desplazamiento y revitalización de las lenguas huarpes a partir de las gramáticas misioneras de Luis de Valdivia", de Vázquez, que muestra cómo el lenguaje, junto con prácticas culturales y sociales, puede contribuir a la revitalización de una cultura afectada por el colonialismo. Asimismo, "Filologías y tipología lingüística: sobre prefijación y

morfemas no ligados en copto", de Yáñez, evidencia la fecundidad de la convergencia entre filología y lingüística en el estudio del copto.

En particular, el ensayo "El diccionario Seiwa jiten en la tradición lexicográfica hispanojaponesa: recorrido histórico y cultural", de Asato, muestra cómo el lenguaje revela no solo significados y usos, sino también dinámicas sociales y políticas. El estudio sitúa el temprano contacto entre España, Portugal y Japón en el siglo XVI, cuando la llegada europea alteró un contexto de relativa clausura cultural, cambio que se reflejó en diccionarios y en nuevas formas de relación mediadas por el idioma. A diferencia de la conquista americana, en Japón predominó una estrategia de persuasión orientada a la conversión cristiana, favorecida por la inestabilidad política y acompañada por la introducción de la imprenta y de las fábulas de Esopo.

El aumento de los intercambios dio lugar a diccionarios trilingües (español-portugués-japonés), hoy testimonio de un encuentro cultural que excede lo lingüístico y evidencia procesos de mediación cultural y religiosa. Asato vincula este legado con el siglo XX, cuando, tras la persecución de cristianos y el retorno de las órdenes religiosas, surgió el Seiwa jiten, mostrando que incluso un diccionario contemporáneo está siempre inscrito en un contexto histórico cambiante atravesado por factores culturales, políticos y religiosos.

La segunda sección se centra en la convergencia entre lenguaje y ciencias sociales. En este marco, destaca el ensayo "Voces glotopolíticas: la dimensión conversacional de la injusticia", de Espinoza, que muestra cómo el análisis del lenguaje en distintos contextos sociales permite identificar formas de vulneración e injusticia y, a la vez, ofrece una base para promover una mayor equidad. Por su parte, el ensayo "entre el código y la palabra: la interacción comunicativa en tiempos de inteligencia artificial", de Patiño y Valenzuela, articula una tríada entre lingüística, ciencias sociales e inteligencia artificial. La propuesta muestra cómo la IA puede emplearse para analizar interacciones en contextos rurales, enriqueciendo la investigación social y aportando, a la vez, al desarrollo de nuevos modelos de inteligencia artificial.

El ensayo "Control y consenso en tiempos de crisis: ideología de la armonía", de Alarcón y colaboradores, aborda problemáticas contemporáneas a partir del estallido social chileno de 2019, analizando el uso del concepto de *armonía* en las campañas del plebiscito. Como muestran los autores, esta noción —al igual que la de paz— se presenta en oposición al desorden y opera como herramienta de control social: la élite no solo ejerce poder, sino que también define los valores que deben regir la sociedad, apoyándose en mecanismos que inhiben transformaciones del *statu quo*. En este contexto, medios de comunicación y fuerzas políticas conservadoras apelaron a la armonía como horizonte deseable para el Estado chileno, en contraste con las demandas de cambios estructurales. El ensayo evidencia así cómo el análisis lingüístico permite revelar estrategias políticas y el carácter operativo de ciertos conceptos en escenarios de conflicto y violencia.

La tercera sección aborda la intersección entre filosofía, literatura y lingüística. En "Creatividad y productividad lingüística", Albornoz propone una discusión conceptual sobre la creatividad en contextos comunicativos desde la gramática de construcciones. Desde una perspectiva pragmática, sostiene que las construcciones no son estructuras fijas, sino dinámicas y situadas en la práctica comunicativa, en estrecha relación con las capacidades cognitivas de los hablantes, lo que habilitaría posibilidades de uso prácticamente infinitas en diversos contextos.

Silva y Ramírez, en "El centro no se sostiene: una aproximación lingüístico-enactiva sobre el uso de metáforas en la conceptualización de la depresión en los diarios de Sylvia Plath", presentan un estudio novedoso y pertinente de su obra. El ensayo articula herramientas del enfoque enactivo —que integra experiencia, biología, cognición y dimensión social— con la psiquiatría, ofreciendo un análisis sólido de los posibles trastornos mentales presentes en los diarios de Plath. Así, sus escritos dejan de leerse como mera producción literaria para entenderse también como testimonios de su modo de habitar el mundo, donde la

literatura encarna en la vida, el cuerpo y el lenguaje, tendiendo un puente fecundo entre la salud mental y los estudios lingüísticos.

De manera similar, Parra, en "Las estrategias narrativas de María Luisa Bombal y Luz de Viana: un estudio desde la poética cognitiva", analiza la obra de ambas autoras desde este enfoque. El ensayo muestra cómo las metáforas conceptuales permiten dar cuenta de los estados de ánimo en las novelas, al funcionar como herramientas cognitivas para abordar conceptos abstractos como la felicidad o la justicia.

Así, por ejemplo, el amor puede concebirse como un viaje, atribuyéndole rasgos que facilitan la comprensión de experiencias no directamente corporales. Desde esta perspectiva, la autora evidencia el uso sistemático de metáforas en ambas escritoras, siendo emblemático el caso de la niebla en Bombal, a la que se le confieren cualidades humanas.

La cuarta sección está dedicada a ciencias cognitivas. En "¿Cómo estudiamos y aportamos a las lenguas indígenas desde la computación?", de Anaís M. Almendra C., muestra cómo la inteligencia artificial y las herramientas computacionales pueden contribuir a la enseñanza y preservación de lenguas indígenas.

En "¿Cuántas palabras saben los adolescentes? Mitos, datos y desafíos del léxico juvenil", Cortés amplía el enfoque más allá del recuento léxico para enriquecer el lenguaje de los adolescentes en Chile. El trabajo sostiene que fomentar la reflexión sobre el lenguaje, dentro y fuera del aula, fortalece la conciencia léxica, amplía el vocabulario y mejora la comprensión y la competencia lectora y comunicativa.

En "Desarrollo del lenguaje en contextos atípicos: más allá del paradigma tradicional", González presenta un caso de aprendizaje lingüístico no convencional. El ensayo analiza a un niño autista, nacido y criado en Chile, que aprendió inglés de forma autodidacta, alcanzando mayor dominio que en español. Según la autora, rasgos característicos del autismo —como la sobrecarga ante la interacción social y los períodos de hiperfoco— fueron decisivos en este proceso. El acceso a internet permitió reducir la exigencia social y aprovechar recursos multimedia,

favoreciendo una adquisición del lenguaje distinta de la habitual y desvinculada de la socialización inmediata.

Como señalé al inicio, estos ensayos logran articular de manera notable distintas disciplinas. No solo aportan nuevas herramientas de estudio e investigación, sino que también evidencian la relevancia de esta convergencia para abordar desafíos contemporáneos como la salud mental, los aprendizajes no convencionales y los contextos políticos complejos.

REFERENCIAS

Hacker, P. M. (2002). Analytic philosophy: what, whence, and whither? In A. Biletzki, & A. Matar (Eds.), *The Story of Analytic Philosophy: Plot and Heroes* (pp. 3-34). Routledge.

Putnam, H. (1975). *Mind, language and reality: Philosophical papers* (Vol. 2). Cambridge University Press.