

ISBN 978-950-33-1584-2

Edición de  
**ANDRÉS A. ILCIC**  
**SOFÍA MONDACA**  
**PABLO TORRES**  
**A. NICOLÁS VENTURELLI**

# Jornadas de Epistemología e Historia de la Ciencia

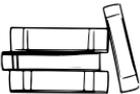
# 30°



# 30° Jornadas de Epistemología e Historia de la Ciencia

Edición de

Andrés A. Ilcic  
Sofía Mondaca  
Pablo Torres  
A. Nicolás Venturelli

Colecciones  
del CIFFyH 

30° Jornadas de Epistemología e Historia de la Ciencia / Sergio Aramburu... [et al.] ; editado por Andrés Ilcic ... [et al.]. - 1a ed. - Córdoba : Universidad Nacional de Córdoba. Facultad de Filosofía y Humanidades, 2020.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-950-33-1597-2

1. Filosofía de la Ciencia. 2. Historia de la Ciencia Argentina. I. Aramburu, Sergio. II. Ilcic, Andrés, ed.

CDD 306.4209

Ilcic, A. A., Mondaca, S., Torres, P., & Venturelli, A. N. (Eds.). (2020). *30° Jornadas de Epistemología e Historia de la Ciencia*. Córdoba, Argentina: Editorial de la Facultad de Filosofía y Humanidades, Universidad Nacional de Córdoba.

Publicado por

Editorial de la Facultad de Filosofía y Humanidades - UNC

Córdoba - Argentina

1° Edición



Área de

**Publicaciones**

Diseño de portadas: Manuel Coll

Diagramación: María Bella

Portada

Imagen superior: Isaac Newton, *Philosophiæ naturalis principia mathematica* (Amsterdam ed.), 1723, p. 267.

Imagen inferior: Oronce Fine, *Le sphere de monde: proprement dicte Cosmographie*, 1549, f. 8v.



Esta obra está bajo una Licencia Creative Commons  
Atribución-NoComercial-SinDerivadas 4.0 Internacional.

30° Jornadas de  
Epistemología e  
Historia de la Ciencia





# **Autoridades de la FFyH - UNC**

## **Decana**

---

Lic. Flavia Andrea Dezzutto

## **Área de Publicaciones**

---

Coordinadora: Dra. Candelaria De Olmos Vélez

## **Centro de Investigaciones de la FFyH María Saleme de Burnichon**

---

Dirección: Dr. Eduardo Mattio

Secretaría Académica: Dra. María Soledad Boero

Área Educación: Dr. Octavio Falconi Novillo

Área Feminismo Género y Sexualidades: Dra. Maite Rodigou Nocetti

Área Historia: Dra. Griselda Tarragó

Área Letras: Espec. Florencia Ortíz

Área Filosofía: Dra. Paula Hunziker

Área Ciencias Sociales: Dra. Gabriela Lugones



# Índice

<b>Presentación</b>	<b>15</b>
<hr/>	
<b>1. Filosofía general de la ciencia</b>	
Bruno Latour: Una nueva forma de idealismo metafísico por <i>Sergio Aramburu</i>	<b>21</b>
La relatividad ontológica del conocimiento científico y los límites de sus manifestaciones enunciativas por <i>Sergio Aramburu</i>	<b>32</b>
El compromiso racionalista y la historia epistemológica de las ciencias: Bachelard y Canguilhem por <i>Alejandra Gabriele</i>	<b>42</b>
La filosofía de los experimentos mentales por <i>Guadalupe Mettini</i>	<b>53</b>
¿La práctica hace a le expertx o le expertx hace a la práctica? por <i>Sofía Mondaca y Julián Reynoso</i>	<b>65</b>
Reproducibilidad y repetitividad experimental entre filósofos y científicos por <i>Marisa Velasco</i>	<b>76</b>
<hr/>	

## 2. Enseñanza de las ciencias

Análisis histórico-epistemológico preliminar sobre el modelo de endosimbiosis para enseñar biología celular y naturaleza de la ciencia al profesorado  
por *Joaquín Álvarez Soria y Eduardo Lozano* **89**

---

El modelo atómico de Lewis:  
Consideraciones epistemológicas y aplicaciones didácticas  
por *Mercedes Barquín, Guadalupe Quiñoa y Alfio Zambon* **98**

---

Los valores no epistémicos en la investigación del cambio conceptual  
por *José Antonio Castorina y Alicia Zamudio* **112**

---

El estudio de la química en clave histórica  
por *Sandra Sandoval Osorio* **124**

---

Los procesos de síntesis teórica y la estabilización de un campo de fenómenos  
por *Sandra Sandoval Osorio, José Francisco Malagón Sánchez, Marina Garzón Barrios y Liliana Tarazona Vargas* **135**

---

## 3. Historia y filosofía de las ciencias cognitivas

Una revisión de la actualidad de la atribución mental:  
Perspectiva de segunda persona  
por *Magalí Argañaraz* **149**

---

Mecanismos de herencia en la evolución cultural:  
Sperber y la naturalización de la cultura  
por *Malena León* **158**

---

Máquinas que simulan máquinas:  
El Homeostato como simulación de un sistema complejo  
por *Andrés Illić* **170**

---

Prácticas de vinculación disciplinar en las neurociencias  
cognitivas contemporáneas  
por *Agustín F. Mauro y A. Nicolás Venturelli* **182**

---

Las habilidades prácticas a la luz de la distinción  
personal-subpersonal  
por *Sofía Mondaca* **195**

---

Los parámetros normativos de corrección racional  
implícitos en las investigaciones en filosofía experimental:  
El caso del experimento que correlaciona la receptividad  
al bullshit con un déficit cognitivo  
por *María Natalia Zavadivker* **207**

---

#### **4. Historia y filosofía de las ciencias sociales y humanas**

El ángulo facial, una investigación cefalométrica publicada  
por Alda Mercante en 1912  
por *Luis M. Guimarey* **221**

---

Nuevas interpretaciones de lo humano ante el desafío  
tecno-científico  
por *Elizabeth Padilla* **229**

---

Características principales para una cartografía de las  
epistemologías del sur  
por *Gabriel Paravano* **239**

---

De la filosofía política a la economía política:  
La especificidad de la economía política como disciplina  
por *Ruth Pustilnik* **252**

---

Mediación e investigación histórica  
por *Edgar Rufinetti* **258**

---

Ciencias sociales: “imbricación” de valores  
epistémicos y sociales  
por *Eduardo Sota* **268**

---

## 5. Historia y filosofía de las ciencias físicas y naturales

- Secuela de la “Hipérbola de la química cuántica”:  
Una mirada sobre el presente de las simulaciones  
computacionales en este ámbito  
por *Penélope Lodeyro* y *Ma. Silvia Polzella* **281**
- 
- Cosas aprehendidas de la física de la gravitación  
por *Oswaldo M. Moreschi* **292**
- 
- Notas históricas sobre el puente Einstein-Rosen  
por *Victor Rodríguez* y *Pedro Walter Lamberti* **308**
- 
- Aportes al debate sobre la reducción de la química  
a la mecánica cuántica  
por *Esteban Guillermo Szigety* **319**
- 

## 6. Historia y filosofía de las ciencias biológicas

- El aporte de la ingeniería de ecosistemas y la construcción  
de nicho a las explicaciones paleontológicas  
por *Susana Gisela Lamas*, *Raúl Montero* y *Vicente Dressino* **331**
- 
- Transgénicos y agrotóxicos:  
el rol de la bioética en el sistema agroindustrial argentino  
por *Nahuel Pallitto* y *Guillermo Folguera* **338**
- 
- Redescubriendo a Claude Bernard  
por *Abigail Prchal* y *Jesús Alberto Zeballos* **346**
- 

## 7. Filosofía del lenguaje y teoría del conocimiento

- “Es como si nuestros conceptos estuvieran condicionados  
por un andamiaje de hechos”  
(Wittgenstein sobre la formación de conceptos)  
por *Gustavo Arroyo* **361**
- 
- De las metáforas de los gestos al lenguaje  
por *Irene Audisio* **367**
-

Alcances y límites epistemológicos de la metáfora, según Ricoeur por <i>Carlos Emilio Gende</i>	<b>376</b>
Tres virtudes del conservadurismo de la suposición por <i>Rodrigo Laera</i>	<b>386</b>
Atribución de creencias a animales: Un análisis de las posturas de Davidson y Wittgenstein a la luz de la evidencia empírica por <i>Juana Regues</i>	<b>393</b>
Razones para saber-cómo y saber-que por <i>Santiago A. Vrech</i>	<b>403</b>

## **8. Historia y filosofía de las ciencias formales**

¿Por qué no hay argumentos concluyentes en filosofía? El caso de los argumentos por regresión infinita por <i>Gustavo Arroyo</i>	<b>413</b>
Series temporales y regresión al infinito: el argumento de McTaggart por <i>Ulises Dávalos</i>	<b>420</b>
Lenguajes y arquitecturas por <i>Pío García y Javier Blanco</i>	<b>427</b>
Definiciones implícitas y estructuralismo matemático <i>ante rem</i> por <i>Eduardo N. Giovannini y Mayra Huespe</i>	<b>438</b>
Datos y semántica en computación: Análisis de la discusión de Fresco y Wolf con Floridi por <i>Xavier Huvelle</i>	<b>449</b>
El problema de significación de los símbolos por <i>Mariana Olezza</i>	<b>462</b>
Acerca de la definición de “número” Un abordaje wittgensteiniano por <i>Pablo Ruiz Lezcano</i>	<b>470</b>

La aplicación de los argumentos de regresión al infinito  
por *Omar Vásquez Dávila* **482**

---

## **9. Lógica y filosofía de la lógica**

Hacer lógica y saber historia de la lógica:  
Hacia una reconciliación metodológica  
por *Miguel Álvarez Lisboa* **493**

---

Una consideración del juicio reflexionante kantiano  
desde la lógica stit  
por *Inés Crespo, Sebastián Ferrando y Luis A. Urtubey* **504**

---

Acerca de estados de estados de cosas verdaderos en  
otros mundos, pero que nunca serán en este mundo  
por *Manuel Dahlquist* **516**

---

Pluralismo lógico y normatividad  
por *Alba Massolo* **527**

---

Aplicaciones de minería de argumentos y modelos  
de esquemas de argumentos  
por *Pablo Torres y Diego Letzen* **538**

---





## Presentación

Este libro recopila trabajos que fueron presentados originalmente en las 30<sup>o</sup> Jornadas de Epistemología e Historia de la Ciencia, que tuvieron lugar en la ciudad de Córdoba entre el 19 y el 21 de septiembre de 2019. Quienes participaron de las Jornadas en calidad de expositores fueron invitados a contribuir en esta recopilación con las versiones finales de sus ponencias. Los trabajos recibidos, habiendo ya sido evaluados y aceptados para su presentación en el evento original, fueron revisados por el comité editorial a fin de determinar que se cumplan las pautas formales establecidas para su publicación.

Aunque pequeña, creemos que la muestra es representativa de la diversidad de espacios que ocupa en la actualidad la reflexión epistemológica del continente en torno a las prácticas científicas en los distintos dominios. Este fue siempre el objetivo de los volúmenes, que desde el año 1993 albergaron los trabajos presentados en las Jornadas. La última edición de dicho formato fue en 2013.

Respondiendo al cambio en los estilos de circulación del conocimiento y a las nuevas plataformas, después de un arduo trabajo de preparación, en 2016 nació la revista *Epistemología e Historia de la Ciencia* con forma de publicación seriada y abierta a toda la comunidad. Si bien la revista es fruto de la labor y la experiencia de las publicaciones previas asociadas a las Jornadas, dada su independencia nunca pretendió reemplazar las

compilaciones que reflejaban la amplitud, la diversidad y la calidad de las comunicaciones presentadas y discutidas en el marco de las Jornadas, además de servir como archivo para la extensa comunidad latinoamericana que se encuentra trabajando en estas temáticas. Con este nuevo volumen pretendemos continuar con el espíritu de aquellas publicaciones, reflejando la calidad y amplitud temática de la reflexión epistemológica en torno de la ciencia dentro de nuestras comunidades académicas.

Los editores  
Córdoba, julio de 2020

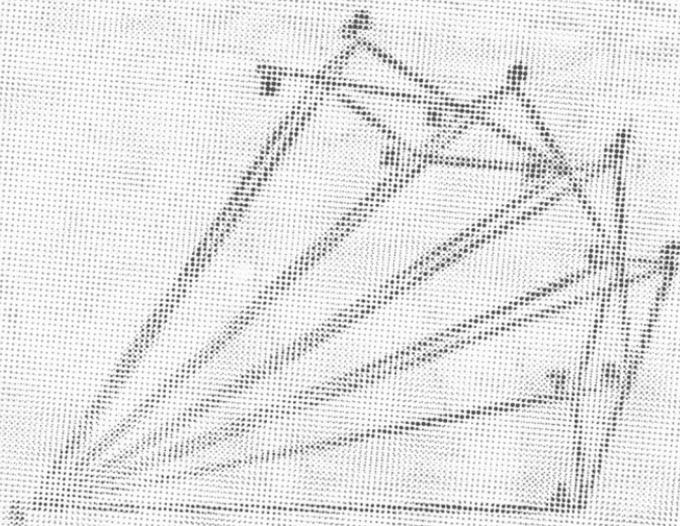






---

# 1. Filosofía general de la ciencia







## Bruno Latour

### Una nueva forma de idealismo metafísico

Sergio Aramburu\*

#### Las tesis de Latour

Alo largo de la mayor parte de sus escritos el filósofo francés Bruno Latour –autor más destacado del campo académico denominado “estudios de la ciencia” (*Science Studies*)– ha puesto en cuestión y pretendido reformular sobre nuevas bases no sólo los problemas, sino los supuestos mismos de la metafísica, la filosofía de la ciencia y la teoría del conocimiento mediante la presentación de una teoría metafísica que rechaza distinciones clásicas –de la ciencia, la filosofía y el realismo cotidiano–, como las establecidas entre un hecho y un enunciado que se refiere a él, “mente” y “mundo exterior”, sujeto y objeto, conocimiento y realidad y naturaleza y sociedad.

*La vida en el laboratorio: La construcción social de los hechos científicos* (Latour & Woolgar, 1979)<sup>1</sup> y *Ciencia en acción* (Latour, 1987) sostienen que los hechos y las entidades cuya existencia ha sido establecida por la ciencia son meras construcciones sociales que resultan de “prácticas” agonísticas en las que algún o algunos científicos imponen una “versión” o una “explicación ordenada” acerca de su existencia, y de esa manera los crean.

Así, asume un antirrealismo científico tan radical que considera que vivimos engañados acerca de las cosas que existen y cómo son, incluido el propio conocimiento científico (entendido como descripción de cosas independientes de él), por lo cual sostiene que “La idea de que los estudios de la ciencia pueden...contentarse con la filosofía de la ciencia o dejar de construir sus propias metafísica y ontología me es ajena” (Latour, 1992, *Post Scriptum*).

---

<sup>1</sup> Luego de la primera edición la palabra “social” fue eliminada del título, consecuentemente con la tesis de que no existen la sociedad humana ni la naturaleza como un ámbito diferente de ella.

\* Universidad de Buenos Aires  
seraramburu@yahoo.com.ar

Afirma que los “estudios de la ciencia” construyen su propia ontología porque la tesis de que no hay una distinción real entre los argumentos y aquello a lo que se refieren tiene que valer también para los propios escritos de Latour, que no pueden sostener consistentemente que hablan sobre un objeto de estudio que exista independientemente de ellos. Pero de ese modo se produce la contradicción de que cuando esos argumentos se refieren a afirmaciones y acuerdos científicos, tesis de la filosofía de la ciencia, la modernidad, Pasteur o el Instituto Salk de Estudios Biológicos tienen que referirse a lo que desde el realismo usual se consideran esas cosas para tener sentido: esas cosas no pueden ser “construcciones”.<sup>2</sup>

En *Nunca fuimos modernos* (Latour, 1991) extiende –desde lo que denomina “antropología simétrica”– el rechazo a otras dicotomías establecidas mediante una especie de crítica cultural en la que atribuye esas distinciones a “los modernos”. La colección de ensayos *La esperanza de Pandora* amplía la concepción metafísica afirmando, por ejemplo, que Pasteur ha creado una entidad en un párrafo mediante un “drama ontológico” (Latour, 1999, pp. 115-116), que no existen entidades individuales pues las cosas tienen “existencia parcial” y “existencia relativa”<sup>3</sup> y que, como los hechos son construidos socialmente mediante afirmaciones de existencia y atribución de características, se les puede “agregar realidad”: son “historias” que se refieren a cosas que no son metafísicamente diferentes de ellas (como lo narrado en una novela), pero que la gente considera reales (y actúa en consecuencia) porque ha aceptado la autoridad de la ciencia.

*Investigación sobre los modos de existencia* presenta una nueva taxonomía y un nuevo vocabulario metafísicos con el objetivo de “instaurar nuevos seres” para “contrarrestar” la ontología atribuida a la modernidad y crear un “mundo común”, pero ratificando que “hecho” significa “bien fabricado” (Latour, 2012, p. 10).

Así, la tesis central de la teoría metafísica formulada por Latour es que lo que llamamos hechos y entidades son inventos humanos realizados mediante el lenguaje en episodios históricos en los que la ciencia o la cultura establecen su existencia: son “construcciones sociales”, “historias” o hechos-fetiches (“*faitiches*”) mientras se lo afirme: tienen un modo de existencia social y lingüísticamente dependiente. Ello implica que el lenguaje humano no tiene función descriptiva: los enunciados y las “entidades externas” “son la misma cosa” (Latour & Woolgar, 1979, p. 177).

<sup>2</sup> Sobre la inconsistencia de los argumentos fundamentales de Latour puede consultarse Aramburu (2019, pp. 118, 119, 131-133, 144, 149).

<sup>3</sup> También en Latour (2000).

El rechazo de la existencia de naturaleza y realidad social humana como ámbitos metafísicos diferentes, y las correlativas de sujeto y objeto y conocimiento y realidad derivan así en la tesis de que la realidad toda consiste en una red de relaciones entre humanos “cuasisujetos” y no humanos “cuasiobjetos”, considerados tanto unos como otros “actantes” o “actores semióticos” que no existen independientemente de palabras y significados, teniendo un papel fundamental en ese proceso constitutivo el lenguaje natural, que no es considerado una entidad o un hecho individual (y, por tanto, ni mental, ni físico, ni abstracto), sino parte de un “fluido circulante” en el que lo real se autoconstruye colectivamente mediante afirmaciones de existencia.

### El idealismo metafísico

En la metafísica moderna el término “idealismo” ha sido empleado para designar un conjunto de posturas que, en términos generales, sostienen que no existen entidades independientemente de que se piense en ellas<sup>4</sup>. Se puede entender, por ejemplo, que cosas como los árboles son apariencias creadas por la conciencia, o que tienen un tipo de existencia metafísicamente derivada de ella. Sin embargo, se ha caracterizado de diferente modo lo asumido como existente primitivo, empleando para ello expresiones como “yo”, “conciencia”, “pienso (*cogito*)”, “pensamiento”, “idea”, “alma”, “espíritu”, “razón” “sujeto”, “mente”, etcétera, y otorgándoles sentidos diferentes a esas palabras en diversos argumentos.

Como postura metafísica moderna, es una de las respuestas a la pregunta sobre qué existe o qué existe primariamente y, en ese sentido, se opone al realismo metafísico, que sostiene que cosas como los árboles existen independientemente de que se piense en ellas. La pregunta, entonces, es específicamente acerca de la *existencia independiente de los objetos de pensamiento* y es típica de la filosofía moderna porque parte del supuesto – compartido por los argumentos que la responden de uno u otro modo– de que existe el pensamiento o conciencia que, como notó Franz Brentano, parece ser siempre conciencia de algo.

En la metafísica antigua, sin embargo, los argumentos no compartían supuestos acerca de la existencia de alguna entidad primitiva en particular, pero se asumía que cualquier teoría sobre la realidad debía dar cuenta

---

<sup>4</sup>Lo que no debe ser confundido con el llamado idealismo gnoseológico, según el cual en el acto de conocer lo determinante es el sujeto.

de lo que se denominó el “mundo sensible” (o materia, que la moderna llamó “sustancia extensa” y, siguiendo la ontología de la ciencia, hoy podríamos denominar “espaciotiempo”) y del “mundo inteligible”, y por eso se ha hablado de idealismo en contraposición al materialismo, aludiendo a las posturas que toman partido por una u otra alternativa como realidad primaria o como realidad toda.

Ese marco metafísico no contemplaba la existencia de la *res cogitans* como una entidad diferente tanto del dominio sensible (o material) como del inteligible de la que hubiese que dar cuenta. No es que no se haya advertido que los seres humanos tenemos estados o “contenidos” de conciencia, sino que “el alma” era considerada o bien una parte del mundo inteligible, o bien un producto metafísicamente derivado del mundo sensible, y los argumentos eran considerados actividad en el alma o de la razón, y no –por ejemplo– un producto cultural que pudiese existir objetivado en textos independientemente de cualquier mente (como sostuvo después Popper), porque el lenguaje humano era considerado algo de naturaleza psíquica que luego podía ser reflejado físicamente al hablar o escribir.

Pero al asumirse la existencia de una entidad primitiva como el *cogito*, se produce en la metafísica un cambio de programa de investigación filosófica (no sólo un cambio de pregunta, sino de los supuestos bajo los cuales se formula) pues, existiendo fuera de toda duda la sustancia pensante, la pregunta metafísica es si existe algo más y, en caso de que así sea, de qué naturaleza es.

Sin embargo, el nuevo programa metafísico no ha preguntado eso, sino si existe, además de la conciencia, algo de lo que experimenta como contenido, adoptando así una perspectiva antropocéntrica interna a ella o solipsista, y restringiendo así las posibilidades de lo que también puede existir: sólo se puede responder *haciendo referencia al sujeto o a la conciencia*, sea afirmando o negando la existencia independiente de lo percibido como sensible, de lo inteligible (por ejemplo, de números y entidades universales), o de ambas cosas. Las teorías idealistas modernas, empero, han negado fundamentalmente la existencia independiente de las entidades y hechos del mundo sensible, adoptando así una postura contraria al realismo metafísico de las personas, el lenguaje natural y el conocimiento científico.

Es decir, el problema metafísico moderno de la existencia independiente sólo admite respuestas *relativas*, y eso significa que no se indaga sobre si existen los árboles, sino si existe lo que la conciencia se representa

como árboles; no se pregunta directamente sobre lo que existe sino sobre una característica de los humanos: lo que se nos presenta en la conciencia como entidades externas o diferentes de ella.

La dificultad de este enfoque es que no está claro qué es lo que se quiere decir con “la conciencia” (o las expresiones alternativas) *en tanto entidad metafísica primitiva*. Si la expresión alude a la conciencia de un individuo humano (“pienso”), resulta problemático afirmar que lo único que existe fuera de toda duda y de modo no relativo es la conciencia de una persona (como Descartes), porque para que exista la conciencia de un ser humano tiene que existir el resto de él.

Y si se refiere a una especie de entidad universal (sustancia pensante), de la que forman parte todas las conciencias humanas individuales el problema es el mismo, porque hablar de una multiplicidad de conciencias *humanas* supone, por ejemplo, la existencia de cerebros, huesos y alimentos independientemente de ellas. Y, sean humanas o no, supone un punto de referencia externo, que ha de tener algún modo de existencia (como lenguaje quizás, pues si fuera otra conciencia las demás serían contenido de ella y no podrían ser primitivas). E implica que una es contenido de la otra, y si la conciencia A es contenido de la B y viceversa, ya no hay “mundo externo”, porque ninguna es la entidad metafísica primitiva.

Para la filosofía de la ciencia el problema es que, a pesar de que no parece existir una entidad universal como “la mente” ni se puede saber qué es lo metafísicamente primario ni lo “externo”, todavía en el siglo XXI se sigue argumentando en términos de “mundo externo” o “la mente y el mundo” (como si las “mentes” no formaran parte del mundo). Y los argumentos científicos frecuentemente son entendidos como algo “interno”, porque se interpreta que el lenguaje humano tiene lugar en “la conciencia” y, consecuentemente, que la existencia independientemente del lenguaje realista metafísico de la ciencia es existencia independientemente de la conciencia.

Es posible pensar que, de modo análogo a como en ciencia y filosofía política se puede hablar de un “contrato social”, sin que realmente haya tenido lugar tal contrato en ningún momento de la historia, en metafísica se puede considerar *el sujeto o la conciencia* como una especie de ficción filosófica útil,<sup>5</sup> pero en tal caso no cabe intentar interpretar los argumentos que aluden a ella literalmente, sino acaso como un recurso metodológica-

<sup>5</sup> Acaso un dios hecho a imagen y semejanza de las conciencias humanas: un fetiche que ha dominado la metafísica, diría Latour.

mente eficaz para abordar ciertos problemas, o que pudo haberlo sido en el pasado.

Pues el asunto es que, a lo largo del tiempo, muchos argumentos filosóficos –inevitablemente– asumen la ontología de la ciencia para referirse a cosas como hechos del pasado, genes, bombas atómicas o la inexistencia de un espacio físico absoluto, pero al hablar de “mundo externo” no lo hacen.

Y adoptar tal ontología, no sólo significa que no es posible ubicar esa conciencia en ningún lado, sino también que en el pasado las conciencias humanas no existieron pero otras cosas sí, y eso implica rechazar el supuesto básico de la metafísica moderna, pues cuando no existían humanos pero sí dinosaurios, ¿existían independientemente de que una conciencia humana –inexistente– pensara en ellos o sólo como contenidos de esa conciencia inexistente?

Porque la ciencia afirma de modo no relativo que El Universo existe,<sup>6</sup> no alude a lo que le ocurre a una presunta conciencia humana; pero ello desplaza el problema hacia el lenguaje, pues los argumentos científicos, que influyen en que muchas personas crean que existen virus y galaxias, son parte del lenguaje humano que asume una ontología realista metafísica (se refiere a cosas que considera independientes de él). Es decir, el problema de la existencia independiente parece subsistir, no como cuestión relativa a conciencias, sino semántica: cosas como hechos y entidades, ¿existen independientemente del lenguaje natural que puede referirse a ellas porque asume una ontología de hechos y entidades? Después de todo, ningún argumento parece haber explicado cuando empieza y cuando termina un hecho, ni por qué.

### **Lenguaje interno y lenguaje externo: un cisma en la semántica**

Reichenbach (1938, p. 17) señala que no se ha “enfaticado suficientemente” que el lenguaje, que es vehículo del conocimiento científico, consiste en objetos físicos como un “área de tinta” que relacionamos con hechos mediante una función llamada “significado”.

Pero algo más importante no ha sido adecuadamente remarcado: en la Viena del Círculo –en ese medio intelectual y en esos años– se ha pro-

---

<sup>6</sup> Y, en todo caso, de eso se sigue que existe independientemente de que alguien piense en él, pero también de cosas como que los perros ladren, que han de entenderse como aspectos de su existencia.

ducido un auténtico cisma en la semántica, una ruptura con una tradición de siglos, con la formulación de la tesis de que el lenguaje humano no es de naturaleza mental o “interna”.<sup>7</sup> Otto Neurath y el Carnap posterior al *Aufbau* también han sostenido que es de carácter físico, y Popper ha afirmado que no es mental sino abstracto y se representa físicamente mediante píxeles, tinta o vibraciones sonoras tal como un numeral representa un número.

Lo común a estas tesis tiene importantes consecuencias para el estudio del conocimiento científico (que consiste en o se expresa mediante enunciados), pero también para la metafísica, porque si el lenguaje humano existe y no es un estado “del alma”, toda afirmación sobre lo que existe no sólo tampoco lo es, sino que debe incluirlo con sus características (como la ontología que asume), lo que implica que también debe referirse a sí misma, y eso es lo que falta en el marco metafísico moderno (donde no ha llegado el “giro lingüístico”), porque “pienso” no es un estado psíquico sino una oración, como lo es “existe mi conciencia”.

Si el lenguaje es diferente de los estados de conciencia, el conocimiento humano no es pensamiento (ni los textos filosóficos ni los científicos están en mentes) y las preguntas epistemológicas no se refieren a un “mundo externo” sino al lenguaje que –sea físico o abstracto expresado físicamente–, a diferencia de los estados de conciencia, es algo público y contrastable: una expresión como “el conocimiento de una nueva especie de planta” no se refiere a un sujeto (ni a una actividad de conciencia ni al acto de tomar conocimiento), sino a un producto cultural como ciertas afirmaciones de una revista científica que, una vez publicadas, existen independientemente de cualquier conciencia.

De manera que lo que la ciencia afirma que existe y lo que las personas creen que existe son cosas metafísicamente diferentes; dicho de otro modo, y como ha señalado Popper: el conocimiento científico no es un sujeto con creencias sino un conjunto de afirmaciones descriptivas con significado; quienes tienen creencias son las personas. Esas afirmaciones están en la cultura, en la realidad social, y son modificadas históricamente, y ello influye en las creencias de las personas sobre lo que existe.

---

<sup>7</sup> El “giro lingüístico” ni es algo meramente filosófico ni ha sido superado, pues teorías lingüísticas como la de Roman Jakobson, que han considerado el lenguaje como un medio para la *comunicación* oral, lo han situado en un ámbito público y las lógicas actuales mayormente no consideran que el lenguaje natural sea pensamiento.

## Lo social como primitivo metafísico: la realidad como un sujeto colectivo que se autoconstruye

Así, parece que no es trivial la cuestión acerca de qué supuestos básicos son adoptados para responder preguntas metafísicas. Con el desarrollo de las ciencias sociales han surgido nuevas teorías sobre la índole de la realidad social y el papel de los contenidos de conciencia en ella. Por ejemplo, Berger y Luckmann (1966) en *La construcción social de la realidad* sostienen que la realidad social (no la realidad toda) posee dos dimensiones, una pública u “objetiva” y otra interna a las conciencias o “subjetiva”, porque como los humanos estamos necesariamente socializados (hemos incorporado, entre otras cosas, un lenguaje y significados que existían antes de que nacióramos), los estados de conciencia de las personas (por ejemplo, cuando observan o piensan en un árbol) son una parte de la realidad social, su dimensión “interna”.

Pero esta concepción obstaculiza concebir el lenguaje humano como algo metafísicamente diferente de las conciencias, pues se puede entender que lo social público tiene las características atribuidas a la sustancia pensante<sup>8</sup> o que hay una especie de entidad colectiva que las posee (“conciencia colectiva”, “nuestra mente”, “el pensamiento humano”), y que expresiones como “nuestras creencias” y “las afirmaciones científicas” son sinónimas o designan una misma sustancia que está repartida en ambos lados, pero esas cosas no son de la misma naturaleza ni siempre coincide su “contenido”.

Al negar la existencia individual no sólo de hechos, sino también del lenguaje, los argumentos de Latour no pueden distinguir metalingüísticamente entre argumentos científicos y entidades y hechos no lingüísticos (como una universidad o un árbol) y, por ello, entre lo que un conjunto de personas cree que existe y las propias afirmaciones científicas al respecto<sup>9</sup> ni, por tanto, advertir que, históricamente, las personas creen que existen ciertas cosas porque son afirmadas por la ciencia o en la sociedad.

Por lo que al identificar lo que en la cultura se afirma que existe con lo que existe (asimilando conocimiento con realidad y referencia con existencia) esta teoría, más que superar las bases de la metafísica moderna, ha desplazado su problema de la existencia independiente a otro terreno,

<sup>8</sup> Lo que explicaría que el conocimiento científico frecuentemente sea llamado “pensamiento”.

<sup>9</sup> Cfr. Aramburu (2019, p. 131).

donde ya no se indaga por contenidos de conciencia sino sobre referentes de afirmaciones de existencia establecidas y, al asumir como entidad metafísica primitiva –la realidad– una especie de sujeto colectivo (“la masa de las transformaciones”) respecto del cual nada existe independientemente porque se construye a sí mismo mediante afirmaciones de existencia, termina siendo una versión semántica y social del moderno idealismo metafísico.

## Conclusiones

Al negar la existencia independiente de aquello a lo que se refiere el lenguaje, el argumento de Latour no puede explicar cómo es que se refiere a algo, ni cómo es que aquello a lo que se refiere (“los modernos”, cierto laboratorio, investigadores, argumentos, acuerdos científicos, etc.) existe o existió independientemente de él. Pero además, si la hormona liberadora de tirotropina fue construida en un laboratorio en un contexto espaciotemporal, y el texto de Latour y Woolgar lo dice, está describiendo un hecho; si el bacilo de Koch es una historia que fue inventada en 1882 y si Pasteur creó una entidad en un párrafo, y textos de Latour lo dicen, entonces describen hechos. En definitiva, se describen “hechos” que no han ocurrido (la ontología inventada por los “estudios de la ciencia”); ese parece ser el precio por negar la existencia del lenguaje con función descriptiva (usando el lenguaje con función descriptiva): caer en inconsistencias. Quizás esto explique las más recientes afirmaciones de Latour (2004, p. 227) de que “el calentamiento global es un hecho, les guste o no” y que “no quisimos decir lo que dijimos”.

Si bien esta teoría expresa un cuestionamiento de los supuestos de la metafísica moderna, se halla limitada por la adopción de otros supuestos filosóficos tradicionales (como que la ciencia estudia la naturaleza), y por rechazar la ontología de la ciencia, lo que la lleva a omitir la existencia de hechos y entidades sociales, como el lenguaje en tanto producto cultural.

A pesar de que la ciencia tiene una ontología históricamente cambiante, sus preguntas y respuestas son diferentes de las del marco metafísico moderno, y eso genera dificultades para abordar algunos temas metafísicos como el de qué es lo que existe y si ello incluye lo que la ciencia afirma.

Pero tanto el problema metafísico de lo que existe como el de lo que existe independientemente de algo se enfrentan con una dificultad semántica quizás advertida en Carnap (1956): toda pregunta metafísica sobre

existencia y toda respuesta al respecto está formulada mediante el lenguaje, y cualquier lenguaje asume una ontología singular de la que no puede salir o prescindir. Cualquier argumento realista que afirme que existen árboles lo hace porque ha adoptado una ontología según la cual existen árboles independientemente de él: no puede afirmar que existe algo que no considere que existe asumiéndolo como referente, y eso significa que la afirmación no puede dejar de ser relativa a su ontología.

Consecuentemente, (como todo metaargumento metafísico es un argumento metafísico) ningún metaargumento puede establecer que lo que un argumento objeto dice que existe es lo que efectivamente existe, porque ninguno puede afirmar que existe algo que no considere que existe (asumiéndolo como su referente), lo que parece un problema semántico insoluble. Si se aduce que referentes (o, incluso, contenidos de conciencia) son representaciones o imágenes de los árboles, el problema subsiste, porque “los árboles” designa referentes de esta otra tesis: afirmar que algo existe independientemente del lenguaje es suponerlo.

El marco metafísico moderno podía afirmar que un contenido de conciencia es representación o imagen de un árbol porque sus afirmaciones metafísicas no tenían en cuenta el propio lenguaje que afirma ni, por tanto, la relación semántica de referencia, pero hacerlo revela que ninguna afirmación descriptiva puede adoptar un punto de referencia externo a ella, lo que significa que ningún argumento puede probar que algo existe independientemente del lenguaje.

Esto contrasta con los supuestos realistas tanto de la actitud natural como de la ciencia, aunque todo eso no deja de ser relativo, pues se trata de supuestos, contenidos de conciencia, conocimiento y lenguaje humanos, según la ontología actual del lenguaje natural que aquí se usa.

## Referencias

- Aramburu, S. (2019). Antirrealismo científico constructivista, lenguaje y realidad social. *Scientia in Verba Magazine*, 4(1), 118-151.
- Berger, P., & Luckmann, T. (1966). *The social construction of reality*. London: Penguin Books.
- Carnap, R. (1956) Empiricism, semantics and ontology. En *Meaning and necessity: A study in semantics and modal logic* (2<sup>nd</sup> enlarged ed., pp. 205-221). Chicago: The University of Chicago Press. (Obra original de 1950)



- Latour, B., & Woolgar, S. (1979). *Laboratory life. The construction of scientific facts*. Princeton: Princeton University Press.
- Latour, B. (1987). *Science in action. How to follow scientists and engineers through society*. Cambridge: Harvard University Press.
- Latour, B. (1991). *Nous n'avons jamais été modernes. Essai d'anthropologie symétrique*. Paris: La Découverte.
- Latour, B. (1992). *Ciencia en acción: Cómo seguir a los científicos e ingenieros a través de la sociedad*. (E. Aibar, R. Méndez & E. Ponisio, trads.). Barcelona: Labor. (Obra original de 1987)
- Latour, B. (1999). *Pandora's hope: Essays on the reality of science studies*. Cambridge, MA: Harvard University Press.
- Latour, B. (2000). On the partial existence of existing and nonexisting objects. En L. Daston (Ed.), *Biographies of scientific objects* (pp. 247-269). Chicago: The University of Chicago Press.
- Latour, B. (2004). Why has critique run out of steam? From matters of fact to matters of concern, *Critical Inquiry* 30(2) (Winter 2004), 225-248.
- Latour, B. (2012). *Enquête sur les modes d'existence: Une anthropologie des modernes*. Paris: La Découverte.
- Reichenbach, H. (1938). *Experience and prediction*. Chicago: The University of Chicago Press.



# La relatividad ontológica del conocimiento científico y los límites de sus manifestaciones enunciativas

Sergio Aramburu\*

## Introducción

**P**arece haber una inconsistencia en el hecho de que se acepta que en la historia de la ciencia hay cambios teóricos con cambio de ontologías, de maneras de explicar, de términos y significados científicos, se descubren cosas nuevas, se modifican teorías vigentes, surgen nuevas disciplinas científicas y se formulan nuevas teorías sobre nuevos aspectos de la realidad, y también que, según la propia ciencia, la realidad no permanece fija (pues el Universo se expande, antes no había seres vivientes en la Tierra, robots ni organismos genéticamente modificados, surgen nuevas especies, hay glaciaciones y cosas así) y, sin embargo, al mismo tiempo se suele considerar que lo que el conocimiento científico afirma que existe es lo que existe y se comporta como él lo describe.

Pero lo que la ciencia afirma que existe es su ontología, un conjunto de referentes de expresiones lingüísticas que las personas incorporamos en nuestra idea de realidad, de modo que luego de un cambio ontológico ésta resulta modificada y –por ejemplo– ya no creemos que exista el flogisto o un espacio físico absoluto, sino otras cosas. Esto significa que hablar de galaxias o de virus es usar un metalenguaje que se refiere a ciertas teorías científicas, y que decir que existen galaxias es asumir metalingüísticamente la ontología de la ciencia actual: las personas hoy no creen que exista el flogisto ni lo afirman porque no lo afirma la ciencia (no es un referente científico).

Consecuentemente, no es adecuado interpretar que una oración como “existen planetas” es verdadera en el sentido correspondentista (que es parte del realismo metafísico de la vida cotidiana o actitud natural y del lenguaje humano), porque todo argumento que afirme la existencia de algo independientemente de él, en realidad, lo supone al adoptar una ontología. Si no fuese así, cuando en el pasado se creía y se afirmaba la exis-

\* Universidad de Buenos Aires  
seraramburu@yahoo.com.ar



tencia del flogisto, miasmas o un espacio físico absoluto porque la ciencia lo afirmaba, esas oraciones hubiesen sido correspondentistamente verdaderas y habrían designado partes de la realidad: el flogisto, miasmas y el espacio físico absoluto.

## **Dos sentidos de “conocimiento científico”**

Afirmar o suponer que el conocimiento científico describe lo que existe entraña un problema: expresiones como “conocimiento científico” y “teoría científica” significan, en principio, que cierto argumento describe correctamente algo, que es conocimiento acerca de algo (primer sentido), pero eso implica que las teorías del pasado –no científicas, obsoletas, equivocadas, como la óptica corpuscular y la óptica ondulatoria– no son conocimiento científico pero, además, nunca lo fueron, porque nunca describieron bien su objeto de estudio (ya que la que lo describe bien es la teoría de los fotones). Sin embargo, en su momento fueron conocimiento científico, y si esta expresión significa “que describe correctamente”, entonces, tal como la teoría de los fotones lo hace hoy, cuando esas teorías estaban vigentes describían adecuadamente qué tipo de cosa es la luz, lo que es contradictorio. Porque lo que llamamos “luz”, no puede haber sido durante un tiempo corpúsculos, luego ondas y ahora fotones simplemente porque hayan sido aceptadas como científicas ciertas teorías que lo afirman.

En cambio, si se interpreta “conocimiento científico” como un conjunto de teorías aceptadas como tales (según criterios evaluativos y algo considerado como evidencia favorable a ellas) en un momento determinado (segundo sentido), se puede explicar que las del pasado –tal como está establecido– fueron científicas, a pesar de que no son ni eran descripciones correctas. El hecho de que en la historia de la ciencia haya cambio de teorías muestra (por *modus tollens*) que la decisión de aceptar una teoría como científica no garantiza que describa correctamente.

Sin embargo, generalizadamente se supone y se afirma que sí lo hacen, lo que permite explicar la falsedad de teorías científicas pasadas, por ejemplo, afirmando metafísicamente que la luz está constituida por fotones y no por corpúsculos, pero ello implica que ya no habrá cambio en la ontología de la ciencia. Y quizás las pasadas son erróneas porque, precisamente, son científicas (en el segundo sentido), es decir, puede que sea inherente al conocimiento científico estar equivocado. Si “teoría científica”

significa solamente “aceptada como tal según procedimientos y criterios epistemológicos”, entonces la historia de la ciencia nos enseña que la ciencia produce teorías erróneas. Y con la misma fuerza con la que afirmamos que científicos pasados creían que sus teorías eran correctas hemos de afirmar que nosotros creemos lo mismo de las vigentes, y eso significa que no es lícito afirmar que existen los quarks sino que creemos que existen.

Si se aduce que, no obstante, hay continuidad de la referencia luego de los cambios de teoría pues para la ciencia hay una misma cosa que es la luz, el problema es el mismo, porque al igual que en el caso de términos como “átomo” o “célula”, lo que llamamos “luz” ha sido caracterizado de diferente modo en la historia de la ciencia, y eso significa que sólo hay en común palabras, pero los significados y los referentes de las teorías no son los mismos.<sup>1</sup>

Esto muestra que identificar teorías vigentes con teorías correctas (o suponer que lo son) no está justificado (porque lleva a contradicción) e implica desestimar –no recordar– que antes se pensaba lo mismo, y obliga, luego de cada cambio teórico, a repetir sistemáticamente “antes estábamos equivocados, pero ahora tenemos la descripción adecuada de nuestro objeto de estudio”, y como eso siempre se hace después de que ocurrió el cambio, es sólo un criterio de infalibilidad ontológica, basado en una especie de mecanismo colectivo de olvido,<sup>2</sup> destinado a justificar el supuesto de que las teorías científicas vigentes son las que han dado con la ontología correcta.

Y es un criterio asimétrico y arbitrario porque –como ha sugerido Kuhn– toma las teorías actuales como patrón de referencia para juzgar las teorías obsoletas, como si las primeras poseyeran los términos y los significados definitivos de la ciencia y describiesen lo que existe y cómo es, lo que implica suponer que la realidad es fundamentalmente fija y desestimar que el proceso científico de conocer es históricamente cambiante, porque en él ocurren descubrimientos inesperados, cambios de teorías,

<sup>1</sup> Además, los argumentos (por ejemplo, del realismo convergente) en términos de “continuidad de la referencia” presentan un problema: un referente es parte de la ontología de un lenguaje, de modo que la continuidad de la referencia sólo significa que diferentes teorías asumen cierta misma ontología, pero ello no dice nada sobre lo que está más allá del lenguaje. Es decir, “referencia” es un término semántico, relacional, y lo que se trata de saber es si esos referentes, tal como son entendidos por la teoría de la que son referentes, existen independientemente de ella (lo que es un asunto metafísico).

<sup>2</sup> Quizás originado en el hecho de que las conciencias humanas siempre vivencian un presente con una idea de realidad; es decir: inevitablemente, siempre pensamos que la realidad es como la entendemos ahora y no como la entendíamos antes.

de taxonomías, de criterios clasificatorios, de términos y de significados, e implica privilegiar las teorías actuales respecto de las pasadas, y también de las que *–ceteris paribus–* habrá en el futuro, por un simple accidente histórico (que hoy son científicas).

## Lenguaje y realidad

Asimismo, como el lenguaje natural y el científico son realistas metafísicos, afirmar “existen los quarks” es asumir que esa oración es verdadera en el sentido correspondentista, lo que entraña, además de las epistemológicas, dificultades semánticas.

La afirmación frecuente de que los enunciados describen la realidad, así como la distinción de ciertos textos –por ejemplo, de lógica o el de Hempel que se cita a continuación– entre lenguaje y realidad, confunden, porque sugieren que, metafísicamente, la realidad está de un lado y el lenguaje del otro, fuera de ella, describiéndola, pero los enunciados son parte de la realidad. Lo que los caracteriza es lo que se ha llamado función descriptiva del lenguaje natural: la capacidad, de acuerdo con su ontología, de hacer referencia –mal o bien– a la realidad toda<sup>3</sup> o a partes de ella (que pueden ser incluso oraciones descriptivas).

Pero como los enunciados describen mediante palabras y significados, no pueden describir partes de la realidad, porque para describir cosas tal como son, ellas mismas tendrían, no que tener, sino que *ser* significados idénticos a los de las descripciones, pero las cosas que no son lenguaje no son en sí mismas semánticas, ya que los lenguajes son productos culturales de lo que los humanos llamamos ciertas especies animales creados, precisamente, para describir, designar, representar y cosas así.

Hempel (1935) señala que en los comienzos del Círculo de Viena, Otto Neurath y Rudolf Carnap coincidían –contra lo afirmado en el *Tractatus*– en que “Nadie entre quienes defienden la existencia de una escisión (*cleavage*) entre los enunciados y la realidad es capaz de explicar con exactitud cómo pueden compararse aquéllos y ésta, ni cómo podríamos averiguar la estructura de los hechos” (pp. 50, 51), por lo que consideraban que “los enunciados nunca se comparan con una ‘realidad’, con ‘hechos’”, sino con

---

<sup>3</sup> Las afirmaciones científicas no, pero el lenguaje natural humano (del que forman parte) sí puede hacer referencia a la realidad como un todo, y en todo caso la cuestión es si lo que la ciencia denomina “el Universo” se puede identificar con lo que el lenguaje natural llama “la realidad”.

otros enunciados (como también afirmó Popper). Sin embargo, afirmar que hay una escisión entre los enunciados y la realidad significa que los enunciados están separados metafísicamente de la realidad, es decir: que no son reales, lo que es contradictorio.

Esto quizás haya sido intuido por autores como Richard Rorty, Hilary Putnam y Bruno Latour, que rechazaron la tesis de que hay una división metafísica entre un enunciado que se refiere a un hecho y ese hecho, o por Nelson Goodman (1978, p. 3), quien sostuvo que “estamos confinados” dentro de los límites del lenguaje.<sup>4</sup> Pero de la afirmación de que los enunciados no están fuera de la realidad no se sigue la sustentada (descriptivamente) por argumentos pragmatistas y de autores como Rorty y Latour de que el lenguaje no tiene función descriptiva. No hay una división metafísica entre lenguaje y realidad, sino una asimetría semántica entre lenguaje descriptivo y aquello a lo que se refiere (que, además, en la mayoría de los casos, es metafísicamente diferente del lenguaje). La tesis autocontradictoria de la escisión entre enunciados y realidad proviene del supuesto de la metafísica moderna de que los enunciados son “internos” o mentales<sup>5</sup> y que el conocimiento es conocimiento del “mundo externo”, finalmente identificado con “el mundo”.<sup>6</sup>

### **“Correspondencia”**

Se suele entender que una oración como “Marte es un planeta del Sistema Solar” es verdadera, porque describe correctamente algo que existe y es como la oración lo afirma. El sentido de “verdadera” es el correspondentista, y en este caso se trata de una oración formulada en lo que Carnap llamó el modo material del discurso, lo que significa que se refiere a algo que es considerado no lingüístico (el planeta Marte). La postulada relación de correspondencia o verdad entre la oración y el planeta, ¿es un vínculo metafísico o semántico? Es decir, cuando se afirma –metalingüísticamente– que la oración es verdadera, ¿se alude a un vínculo que nunca se sale del lenguaje, debido a que dicha oración se refiere a algo que está caracterizado por el propio lenguaje que describe, o acaso la relación va

---

<sup>4</sup> Y también se refiere a lo que comenta Hempel: que respecto de la idea de verdad en el sentido correspondentista, “la naturaleza del acuerdo entre una versión y un mundo diferente de ella es notoriamente nebulosa” (Goodman, 1978, p. 17).

<sup>5</sup> Todavía en el siglo XXI muchos textos denominan al conocimiento científico “pensamiento científico”.

<sup>6</sup> Véase Aramburu (2017, 2019, n. 35).

más allá y sale del lenguaje para llegar a una parte de la realidad que es independiente de él? Es decir, "Marte" ¿designa una parte de la ontología del lenguaje científico o designa directamente una parte de la realidad?

El término "correspondencia" también mueve a confusión, pues en toda interpretación realista metafísica la relación de correspondencia<sup>7</sup> es considerada un vínculo adecuado entre oraciones descriptivas y aquello a lo que se refieren (entre "palabra y mundo", "lenguaje y realidad" o "proposiciones y hechos", por ejemplo). Pero al afirmar que un enunciado describe correctamente un hecho o una entidad no lingüística como, por ejemplo, un planeta, se está empleando un metalenguaje que asume que se refiere por un lado a algo lingüístico y por otro a algo que no lo es, describiendo una relación entre dos cosas de naturaleza diferente, y si hay una relación entre lo lingüístico y lo no lingüístico, esa relación no es lingüística sino metafísica (pues si fuera lingüística se trataría de la de referencia, y consecuentemente no se referiría a lo no lingüístico). Expresiones como "el ojo de dios" han sido usadas para expresar esta idea.

Sin embargo, para que el enunciado aluda a algo tiene que caracterizarlo de cierto modo mediante palabras y significados, y eso significa que sólo puede hacer referencia a algo que sea parte de una ontología que adopte. Por ejemplo, si afirma que algo se comporta ora como partícula, ora como onda, está asumiendo una ontología según la cual existen partículas y ondas, pero si se refiere a "ondículas", asignando un determinado significado a esta palabra, asume la existencia de ondículas (que son sus referentes): ninguna descripción puede afirmar que algo existe sin considerarlo existente; incluso afirmar "Pegaso no existe" es, en realidad, asumir la existencia de referentes de distinto tipo: caballos ficticiales y caballos no ficticiales, pues usar la palabra "Pegaso" es asumir la existencia de la mitología griega.

El término "verdad" es de carácter semántico y, en el sentido correspondentista, designa una relación postulada entre una entidad lingüística y aquello a lo que se refiere, pero como lo que describe es el lenguaje, no hay correspondencia conversa (porque no ocurre que aquello a lo que se refiere un enunciado esté también refiriéndose correctamente a él): la "flecha" de la "correspondencia" no va desde la conciencia hasta "la realidad", porque no es metafísica sino semántica, y tiene un solo sentido porque

---

<sup>7</sup> También confunde que muchos textos de lógica y de filosofía de la lógica se refieran a la relación de correspondencia como si fuera una propiedad de "portadores de verdad", denominando "verdades" a oraciones.

está hecha de significados puestos por la descripción y no por lo descrito, por lo que nunca se independiza del lenguaje que describe para llegar al destino metafísico: es una flecha que cree dar en el blanco deseado pero meramente atina en un referente del propio discurso.

Esta es la auténtica dificultad, y no es una escisión metafísica sino un problema semántico: la unilateralidad descriptiva, a través de significados que pone el propio lenguaje que describe, entre éste y aquello a lo que se refiere (sea lenguaje o no) produce la paradoja de que al describir se asume que se alude a algo que existe independientemente de la descripción, pero para ello hay que adoptar una ontología según la cual eso existe independientemente. De la distinción fregeana entre significado (“sentido”) y referencia, en tanto características al menos del lenguaje natural (que incluye el científico) se sigue que ningún discurso realista o metafísico carece de referentes, de manera que toda afirmación sobre lo que existe es, en realidad, sobre lo que se supone que existe al adoptar una ontología.

Es decir, no hay “correspondencia” en el sentido de equivalencia o simetría, o existencia de una relación converso o recíproca, sino que, por decirlo así, se envió una carta y nunca se obtuvo respuesta: un amor no correspondido entre el ámbito de la semántica y el de la metafísica (que no sabe escribir ni enviar cartas). Esta unilateralidad es una característica –o un problema– de la función descriptiva del lenguaje natural humano y tiene lugar independientemente de si se hace referencia a algo lingüístico o no, por lo que es más fundamental que la diferencia entre las oraciones descriptivas y las cosas no lingüísticas a las que se refieren.

Esto implica que toda descripción científica de alguna cosa considera o describe esa cosa con el lenguaje de la propia descripción y no con el de la cosa, que, además, mayormente no es ni tiene lenguaje (y si es lenguaje no describe recíprocamente la descripción), ni con otro lenguaje –un metalenguaje– que describa correctamente la descripción y lo descrito –como notaron, siguiendo a Carnap (1950/1956)– Thomas Kuhn (2000, p. 220) y Nelson Goodman (1978, pp. 4, 5), pues todo metalenguaje es descriptivo, es decir: sólo puede hacer referencia asumiendo una ontología singular, de modo que si se afirma que “Marte es un planeta del Sistema Solar” es una oración verdadera, en realidad se está asumiendo la ontología del lenguaje objeto científico, y por eso *hoy* no se afirma que “Plutón es un planeta del Sistema Solar” es una oración verdadera. Ninguna metaoración puede describir la “verdad como correspondencia” entre una oración

descriptiva y algo independiente de ella, porque ninguna puede salir del lenguaje y ser “el ojo” del dios.

## Ontología de la ciencia y metalenguaje

Esto significa que no puede haber una afirmación verdadera en sentido correspondentista que se refiera a, por ejemplo, planetas, porque toda afirmación acerca de planetas es, en realidad, acerca de lo que un lenguaje considera que son planetas como, siguiendo a Carnap, ha dicho Kuhn (2000, p. 220): es inevitable que aluda a referentes; es decir, o bien es una oración de una teoría científica que hace referencia a su ontología, o bien es una oración metalingüística (científica o no<sup>8</sup>) que hace referencia a esa teoría científica (y no directamente a lo que existe).

Por ejemplo, si se afirma al estilo Tarski “la oración ‘los planetas se desplazan recorriendo órbitas elípticas’ es verdadera si y sólo si los planetas se desplazan recorriendo órbitas elípticas”, puede parecer que la parte “material” de la oración, la que se refiere a los planetas, describe directamente una parte de la realidad, pero ello no es así, pues, como advirtió Carnap, ya es metalingüística: afirmar “los planetas se desplazan recorriendo órbitas elípticas” significa “lo que el conocimiento científico denomina ‘planetas’ se comporta como lo que ese mismo conocimiento considera que es recorrer órbitas elípticas”.

Es decir, no puede haber una oración científica verdadera acerca de una parte de la realidad, sino una oración científica que se refiera a lo que el propio conocimiento científico considere que es una parte de la realidad (empleando para ello palabras como “corpúsculos”, “entidades cuánticas”, “ondas lumínicas” o “fotones”, por ejemplo): el conocimiento científico puede modificar su ontología, pero no carecer de ella.

De acuerdo con el realismo asumido por el lenguaje humano, son cosas metafísicamente diferentes lo que existe y lo que según la ciencia existe (que son los referentes de un discurso), y eso vale tanto para lo que afirma ahora como para lo que afirmaba en el pasado y para lo que afirme en el futuro. Y el hecho es que la existencia misma del conocimiento científico (y del lenguaje descriptivo), que siempre ha de asumir una determinada ontología, produce esa distinción metafísica e impone la semántica: hablar

---

<sup>8</sup> No todas las oraciones científicas están formuladas en el modo material, pues un enunciado botánico que alude a la fotosíntesis empleando el término “luz” se refiere metalingüísticamente a lo que una teoría óptica física considera que es la luz.

de lo que existe no es hablar metalingüísticamente, hablar de lo que según la ciencia existe sí lo es. Afirmar que las ondas lumínicas y los corpúsculos no son componentes de la luz es tan metalingüístico como decir que los fotones sí lo son. Negar esto implica afirmar que en, en su momento la oración “existe el flogisto” describía directamente una parte de la realidad: el flogisto.

Fue Rudolf Carnap (1950/1956) quien advirtió que afirmar la existencia de fotones o corpúsculos es aludir a los referentes o parte de la ontología de un lenguaje (“valores de variables”), pero consideró, en virtud de los supuestos de su programa filosófico, que esas afirmaciones están mal enmarcadas o mal formuladas (porque ocultan que se refieren a un determinado marco lingüístico, haciendo creer que son descripciones directas de partes de la realidad “externa” a él), por lo que las consideró carentes de sentido cognitivo: pseudoenunciados. Sin embargo, se trata de oraciones metalingüísticas del lenguaje natural que simplemente describen usando el marco lingüístico del lenguaje objeto (asumiendo su ontología); es decir, existen muchas afirmaciones metalingüísticas dentro del propio conocimiento científico que asumen parte de la ontología de una teoría de la propia disciplina o de otra que se toma como conocimiento auxiliar.

## **Conclusiones**

Una postura falibilista sobre el conocimiento científico no solamente está apoyada por el hecho de que no se puede probar que una afirmación universal sea verdadera, sino que los cambios en la historia de la ciencia y muchos datos sobre cómo se produce el conocimiento (la publicación de artículos, el desarrollo de nuevas disciplinas científicas, teorías, cambios taxonómicos y de significados, etc.) van en la misma dirección. De la afirmación de Kuhn (1962) de que no resulta adecuado tomar las teorías científicas actuales para juzgar la ciencia del pasado se sigue, fundamentalmente, que hacerlo supone que no habrá cambio de ontología, instrumentos nuevos y más precisos, ni nuevas disciplinas y que los significados y taxonomías de la ciencia son definitivos, lo que contradice muchos datos empíricos sobre la naturaleza del conocimiento científico.

La tesis de que lo que existe es lo que la ciencia afirma que existe sólo puede ser mantenida por un tiempo al precio de olvidar que antes se pensaba y afirmaba lo mismo. Ello se debe a que hablar de virus o galaxias no es hacer referencia a entidades existentes sino a referentes de teorías cien-

tíficas que son considerados existentes. Afirmar que existen es incorporar esos referentes al propio discurso, porque tanto el lenguaje natural como las personas en la actitud natural asumen el realismo metafísico (que existe la realidad y es de cierto modo), de manera que al producirse un cambio en la ontología de la ciencia se modifica la idea de realidad de las personas y la ontología del lenguaje natural, porque cualquier creencia o discurso sobre la realidad necesariamente asume que es como se cree o se afirma ahora y no como se lo hacía antes.

Esto permite explicar por qué se cree y se afirma que existen ciertas cosas y luego de un cambio teórico eso es modificado, pero se vuelve a creer y a afirmar que existen ciertas otras cosas, a pesar de que muchos datos indican que eso ya ocurrió y sugieren que va a volver a ocurrir, lo que no puede ser explicado por la tesis de que lo que existe es lo que la ciencia afirma que existe, y apoya la de que el conocimiento científico es por naturaleza erróneo.

## Referencias

- Aramburu, S. (2017). Realismo estructurista y nominalismo científico, trabajo presentado en el *IV Workshop del Círculo de Buenos Aires*, Universidad de Buenos Aires. doi: 10.13140/RG.2.2.23387.21280
- Aramburu, S. (2019). Antirrealismo científico constructivista, lenguaje y realidad social. *Scientia in Verba Magazine*, 4(1), 118-151.
- Carnap, R. (1956). Empiricism, semantics and ontology. En *Meaning and necessity: A study in semantics and modal logic* (pp. 205-221). Chicago: The University of Chicago Press. (Obra original de 1950)
- Goodman, N. (1978). *Ways of worldmaking*. Indianapolis: Hackett Publishing Company.
- Hempel, C. (1935). On the logical positivists' theory of truth. *Analysis*, 2(4), 49-59.
- Kuhn, T. (1962). *The structure of scientific revolutions*. En O. Neurath (Ed.), *International encyclopedia of unified sciences. Foundations of the unity of science* (vol. I-II). Chicago: The University of Chicago Press.
- Kuhn, T., (2000). *The road since structure* (J. Conant & J. Haugeland, Eds.). Chicago: University of Chicago Press.



## El compromiso racionalista y la historia epistemológica de las ciencias: Bachelard y Canguilhem

Alejandra Gabriele\*

### El compromiso racionalista en el mapa de la historia de las ciencias francesa

Una entrada posible para abordar qué significa el compromiso racionalista en la historia de las ciencias francesa, es seguir un mapa de la filosofía francesa trazado por Foucault en “La vida: la experiencia y la ciencia” (1984/2009).<sup>1</sup> Allí se refiere a

una línea que separa una filosofía de la experiencia, el sentido y el sujeto, de una filosofía del saber, la racionalidad y el concepto. Por un lado, una filiación que es la de Sartre y Merleau-Ponty; por el otro, la de Cavaillès, Bachelard, Koyré y Canguilhem. (Foucault, 1984/2009, p. 42)

El punto que permite esta bifurcación en la filosofía francesa son las diferentes lecturas o anclajes respecto a las *Meditaciones cartesianas* de Edmund Husserl, publicadas en 1929. Una lectura desde la filosofía del sujeto aparece en 1935 en el artículo de Sartre sobre la “Trascendencia del ego”, buscando radicalizar la fenomenología de Husserl. La otra indaga sobre los problemas que fundan el pensamiento de Husserl: el formalismo y el intuicionismo. De esta última surgen las dos tesis de Jean Cavaillès en 1938 sobre el *Método axiomático y formalismo* y sobre *Notas sobre la formación de la teoría abstracta de conjuntos*.

---

<sup>1</sup> Este texto fue publicado en *La revue de métaphysique et de morale*, en 1984, número dedicado a George Canguilhem. Si bien Foucault pretendió escribir un artículo nuevo, sus condiciones de salud no se lo permitieron. En su lugar realizó algunas modificaciones a la introducción que había escrito para la edición estadounidense de *Lo normal y lo patológico* de Canguilhem. En español se publicó en 2009 en la compilación sobre escritos de biopolítica realizada por Giorgi y Rodríguez y el 2012 en un libro que reúne trabajos de Foucault bajo el título *El poder, esa bestia magnífica*. Las citas que aquí utilizamos corresponden a la primera publicación en español.

\* Universidad Nacional de Cuyo (UNCuyo).  
aalegabriele@gmail.com



La línea de trabajo desarrollada por Cavailleès puede interpretarse como más teórica y especulativa que las filosofías del sujeto, sobre todo porque no tematizaron sobre los problemas políticos inmediatos, “sin embargo –señala Foucault– fue la que tomó partido durante la guerra y combatió de manera directa, como si el fundamento de la racionalidad no pudiera disociarse de la interrogación sobre las condiciones actuales de la existencia” (Foucault, 1984/2009, p. 43). Tanto Cavailleès, como Bachelard, Koyré y Canguilhem, produjeron una filosofía en el marco de lo que Bachelard denominó compromiso racionalista. En el prólogo al libro que lleva este título, Canguilhem se refiere a otros compromisos para situar y distinguir las posiciones asumidas por Cavailleès y Bachelard:

Antes de Bachelard muchos racionalistas se creyeron comprometidos, aun cuando –a falta de una moda ideológica todavía inexistente– no se proclamaron como tales. Pero casi siempre se trataba de un compromiso de la razón contra la religión o contra el orden establecido de un poder tradicionalista, más bien que de un compromiso de la racionalidad de la razón contra su propia tradición. (Canguilhem, 1972/2001, p. 7)

El compromiso racionalista supone una dialéctica de la revocación, una revisión crítica de las construcciones de nuestra propia racionalidad, de radicalizar el compromiso filosófico como lo hizo Cavailleès,<sup>2</sup> “un filósofo matemático que fue muerto porque no creía que podía separar en su compromiso la razón y la existencia” (Canguilhem, 1972/2001, p. 7).

La recurrencia a Cavailleès en los trabajos de Bachelard y Canguilhem, sobre todo en este último, aparece como un estandarte filosófico y político que articula el compromiso racionalista como resistencia. Cavailleès sostenía que la filosofía de la conciencia debe ser reemplazada por la dialéctica de los conceptos. También aparece con frecuencia en las referencias a Cavailleès la desintegración de una filosofía de la conciencia: “Cavailleès quería de-subjetivar el conocimiento; en el mismo gesto consideró la Resistencia como una necesidad ineluctable, que ninguna referencia al Yo podía circunvenir” (Badiou, 2008/2009, p. 23). Se trata de una crítica de la propia racionalidad hasta descentrarse de toda identidad (personal, individual<sup>3</sup>)

---

<sup>2</sup> Jean Cavailleès fue filósofo y matemático, co-fundador del movimiento de resistencia *Libération-Sud* y fundador de la red de acción militar Cohors.

<sup>3</sup> Como una forma de “ironía objetiva”, tras ser torturado y fusilado en 1943, su cuerpo fue descubierto en una fosa común en la Ciudadela de Arras y bautizado en su momento como “Desconocido nº5” (cfr. Badiou, 2008/2009, p. 22).

que pueda interrumpir la lucha contra principios de orden que cercenen la vida.

Para continuar explorando este compromiso racionalista, retomamos el desarrollo argumental de Foucault en el momento en que se ocupa de la pregunta con la que Kant interroga a la filosofía de su tiempo: ¿qué es la Ilustración?, indagando sobre la forma que podía tomar en un momento particular y de los efectos que de ella podían esperarse:

De pronto, la cuestión del “presente” se vuelve una interrogación de la cual la filosofía no puede separarse: ¿en qué medida este “presente” depende de un proceso histórico general y en qué medida la filosofía es el punto en que la historia misma debe descifrarse a partir de sus condiciones? (Foucault, 1984/2009, p. 44)

A partir de este momento la historia se convierte en uno de los principales problemas de la filosofía, provocando diferentes destinos filosóficos. En Alemania toma la forma de una reflexión histórica y política sobre la sociedad (posthegelianos, Frankfurt, Luckács, Feuerbach, Marx, Nietzsche, Weber). En Francia, en cambio, fue la historia de las ciencias la que formuló la cuestión filosófica sobre lo que fue la Ilustración (cfr. Foucault, 1984/2009, p. 45). Es decir, la cuestión de la ilustración se desarrolló a través del positivismo desde las críticas de Saint-Simon al positivismo de Comte y sus sucesores. El punto de inflexión que cambiará el rumbo de la filosofía de la ciencia en Francia, será, como ya hemos dicho, el arribo de la fenomenología de Husserl en los años treinta.

Mientras que en Alemania el núcleo de las discusiones filosóficas se centra en la experiencia religiosa en su relación con la economía y el Estado, en Francia, las problemáticas filosóficas de la ilustración se desenvolverán entre

saber y creencia; forma científica del conocimiento y contenido religioso de la representación, o paso de lo precientífico a lo científico; constitución de un poder racional con una experiencia tradicional de fondo; surgimiento, en medio de una historia de las ideas y de las creencias, de un tipo de historia propia del conocimiento científico, origen y umbral de racionalidad. (Foucault, 1984/2009, p. 45)

En la tradición académica francesa, la historia de las ciencias ha estado asociada a la filosofía de las ciencias dado que fue introducida en las instituciones culturales por la escuela filosófica positivista (Canguilhem,

1952/1976). La primera cátedra de historia general de la ciencia fue creada en 1892 en el *Collège de France* y ocupada por el discípulo de Comte, Pierre Laffitte. Más adelante, se crea en la Sorbona una cátedra propiamente de “historia y filosofía de las ciencias” a la que Bachelard arribó en 1940.

El ingreso de la dimensión científica en el ámbito filosófico limitó de alguna manera ciertas prácticas filosóficas como “tratar supersticiosamente el conocimiento como una revelación, incluso largamente implorada, y la verdad como un dogma, incluso cualificado positivo” (Canguilhem, 1952/1976, p. 52). De esta manera se gestaron las condiciones filosóficas que hicieron posible el desarrollo de una epistemología histórica y de una historia epistemológica de las ciencias que, como señala Foucault interroga a la racionalidad que pretende ser universal negando su carácter contingente e histórico:

Tanto en la historia de las ciencias en Francia como en la teoría crítica alemana, se trata en el fondo de examinar una razón cuya autonomía de estructuras lleva inscrita la historia de dogmatismos y despotismos –una razón que, en consecuencia, tiene efectos de emancipación sólo a condición de que logre liberarse de sí misma. (Foucault, 1984/2009, p. 46)

Y para que la razón pueda liberarse de las supersticiones por ella misma creadas se necesita una “psicología de la conquista progresiva de las nociones en su contenido actual, como una puesta en forma de genealogías lógicas y, para emplear una expresión de Bachelard, ¡como un inventario de los ‘obstáculos epistemológicos’ superados!” (Canguilhem, 1952/1976, p. 52). Esto es, hacer una historia genealógica de los conceptos y sus filiaciones.

Se trata entonces de interrogar, desde la propia lógica occidental, sobre los límites de las racionalidades construidas y los poderes actuales de viejas y nuevas racionalidades:

al querer desvalorizar radicalmente, bajo el pretexto de una superación teórica, las intuiciones antiguas, uno pasa, insensible pero inevitablemente, a no poder comprender ya cómo una humanidad estúpida un buen día pasó a ser inteligente. . . . De suerte que la necesidad actual de una teoría más manejable y más comprensiva no puede sorprender a los espíritus incapaces de buscar en la historia de las ciencias el sentimiento de posibilidades teóricas de aquellos que la enseñanza de los solos y últimos resultados del saber las ha hecho familiares, sentimiento sin el cual no hay crítica científica, ni futuro de la ciencia. (Canguilhem, 1952/1976, p. 91)

## **Epistemología e historia de las ciencias**

Un rasgo distintivo de las propuestas teóricas de Bachelard y Canguilhem es la preocupación por pensar unidas la epistemología y la historia de las ciencias: “El reconocimiento de la historicidad del objeto de la epistemología impone una nueva concepción de la Historia de las ciencias. La epistemología de Gastón Bachelard era histórica; la historia de las ciencias de Canguilhem es epistemológica” (Lecourt, 1970/2005, p. XI).

El ingreso que elegimos a esta configuración histórico-epistemológica será la distinción que realiza Canguilhem entre la razón histórica, la razón académica y la razón filosófica, y las diferentes lógicas que cada una de estas implica en la realización de una historia de la ciencia. La razón histórica es la que en el contexto de la producción académica se ocupa de las conmemoraciones, el establecimiento de paternidades intelectuales, olvidando el marco intelectual y cultural en el que se produjo el conocimiento científico que supone ser objeto de estudio de esta historia de la ciencia. Para Canguilhem “tomar por objeto de estudio sólo asuntos de fuentes, invenciones o influencias, de anterioridad, simultaneidad o sucesión, equivaldría en el fondo a no distinguir entre las ciencias y otros aspectos de la cultura” (Canguilhem, 1966/2005, p. 17).

La razón científica es propia del orden de la investigación. Es la que llevan adelante los científicos en circunstancias en las que por falta de apoyo institucional necesitan validar sus logros investigando si su descubrimiento no ha sido ya pensado, construyendo así la imagen del predecesor. Estrategia fuertemente rechazada por Canguilhem dado que “la complacencia en buscar, encontrar y celebrar a precursores es el indicio más claro de la incapacidad para la crítica epistemológica” (Canguilhem, 1968/2009, p. 24). La búsqueda del predecesor conlleva la descontextualización histórica y cultural que se realiza sobre el pensador que se pretende precursor para lograr que sus conceptos, discursos y gestos sean coherentes con el contexto presente desde el que se pretende encontrar legitimidad. Deshistorizar al precursor es negar el valor de su pensamiento como respuesta a las problemáticas de su tiempo.

Finalmente, la razón filosófica es la que reconoce la implicancia mutua entre historia de la ciencia y epistemología. Permite distinguir dos tipos de historia de las ciencias, aquella que se ocupa de los conocimientos perimidos y aquella otra que hace la historia de los conocimientos estableci-

dos. Una historia de los conocimientos perimidos es una especie de museo de los errores de la razón humana:

Una tal actitud supone una concepción dogmática de la ciencia y, si se me permite, una concepción dogmática de la crítica científica, una concepción de los “progresos del espíritu humano” que es la *Aufklärung* [Ilustración], de Condorcet y de Comte. Esto que planea sobre dicha concepción es el espejismo de un “estado definitivo” del saber. En virtud de lo cual el prejuicio científico es el juicio de edades cumplidas. Es un error puesto que es de ayer. La anterioridad cronológica es una inferioridad lógica. (Canguilhem, 1952/1976, p. 48)

La razón filosófica en la historia de las ciencias es posible por la ruptura epistemológica que provoca Bachelard respecto de la concepción positivista al poner en tensión la dirección de la mirada de la historia de las ciencias con la de la epistemología. Canguilhem hace evidente esta tensión al distinguir la orientación temporal del historiador y del epistemólogo: “el historiador procede desde los orígenes hacia el presente, de suerte tal que, en alguna medida, la ciencia de nuestros días siempre se anuncia en el pasado.”, en cambio, “el epistemólogo procede desde lo actual hacia sus comienzos, de modo que el presente sólo funda en cierta medida una parte de lo que ayer se presentaba como ciencia” (Canguilhem, 1968/2009, p. 189).

A través de la mirada epistemológica, la razón filosófica permite des-ensmascarar una historia de continuidades progresivas del conocimiento científico, y la categoría que corre el velo no es otra que la de obstáculo epistemológico. A partir de Bachelard la historia de las ciencias ya no puede seguir siendo la acumulación de datos biográficos y doctrinas, sino que se trata ahora de una historia de las filiaciones conceptuales, precaria y destinada a la rectificación. Así, “De la historia de la ciencia, filosóficamente cuestionada, es decir, cuestionada en cuanto a la formación, la reforma y la formalización de los conceptos, surge una filosofía de la ciencia” (Canguilhem, 1968/2009, p. 185).

Una referencia frecuente respecto de la epistemología de Bachelard es su imposibilidad de ser definida en una doctrina sistemática y unívoca. Justamente esa imposibilidad de identificación canónica y de cierre doctrinal es lo que permite que la obra de Bachelard se convierta en una interpelación constante que moviliza nuestros supuestos teóricos y metodológicos.

El carácter histórico distintivo de esta epistemología es el perpetuo “recomienzo” y esencial estado de inacabado, consiste en producir un estado epistemológico que no deja de romper con su pasado liberándose de todo aquello que pueda producir inmovilidad (cfr. Dagognet, 2006). Para lograr ese estado Bachelard trabaja en dos dimensiones o dos epistemologías. Una que supone al “otro” como constituyente de una intersubjetividad que pretende ampliar la dimensión objetiva del conocimiento. Se trata de la dimensión pedagógica que valora de la misma manera la transmisión del saber como el saber en sí mismo. La otra dimensión de su epistemología es la que se organiza a partir de un autoanálisis del conocimiento objetivo: la “vigilancia intelectual de sí” que sólo es posible a partir de la figura de un Super Yo capaz de un control que fortalece la producción de conocimientos dinamizando la actividad racional (a diferencia del freudiano ligado a la neurosis, la culpabilidad y la censura). Entre estas dos epistemologías Bachelard supera tanto la lineal y unidireccional relación sujeto-objeto como su pretensión de un conocimiento plenamente objetivo.

Al mismo tiempo, al dialectizar la relación entre la organización racional del conocimiento y la experimentación objetiva creando el racionalismo aplicado, la epistemología histórica deja atrás cualquier pretensión de formalismo, rechazando todo racionalismo a priori y reivindicando ese nuevo territorio inaugurado por Bachelard que son los regionalismos epistemológicos. Sus preocupaciones por estos movimientos en el conocimiento científico, sobre todo en el de su época, están presentes en la gran cantidad de pruebas que aporta acerca del fenómeno de transracionalidad, como es la vinculación entre ciencia y técnica.

Concibe a la filosofía de la ciencia como una filosofía que estudia las aperturas del pensamiento, una filosofía *dispersada*, fundada en un pluralismo filosófico que es capaz de informar la diversidad de elementos empíricos y teóricos, teniendo en cuenta que *lo que es, deviene*. Por lo tanto, las tareas de la filosofía de las ciencias, lejos de la pretensión filosófica de universalizar, deben consistir en trabajar cada noción, hipótesis, problema y experiencia en su dimensión particular y diferente.

Debería fundarse una filosofía del detalle epistemológico, una filosofía científica *diferencial* que correspondiese simétricamente a la filosofía *integral* de los filósofos. Esta filosofía diferencial tendría a su cargo medir el devenir del pensamiento. . . . Meditando filosóficamente sobre cada noción, veríamos también con mayor claridad el carácter polémico de la

definición empleada, todo lo que esa definición distingue, suprime, niega. Las condiciones dialécticas de una definición científica diferente de la definición usual aparecerían entonces más claramente, y podría comprenderse, en el detalle de las nociones, lo que llamaremos la filosofía del no. (Bachelard, 1940/1993, p. 15)

En este sentido, una filosofía de las ciencias es una filosofía dispersada pero que supone una cohesión, la de su dialéctica: presume un progreso que se realiza en el sentido de un racionalismo creciente, eliminando respecto de todas las nociones, el racionalismo inicial. Esta dialéctica es diferente de las dialécticas filosóficas porque no se trata de una construcción *a priori* que traduce la marcha seguida por el espíritu en el conocimiento de la naturaleza como es el caso de la dialéctica hegeliana. Por el contrario, la dialéctica de la ciencia contemporánea, como la denomina Bachelard, es una dialéctica de la razón polémica que “nos sirve para cercar una organización superracional muy precisa. Nos sirve únicamente para virar de un sistema a otro” (Bachelard, 1940/1993, p. 113), provocando una ruptura con las nociones que se presentaron como obstáculos epistemológicos y creando un nuevo objeto, un superobjeto “resultado de una objetivación crítica, de una objetividad que retiene del objeto solamente lo que ella misma criticó. . . . Al destruir sus imágenes primeras, el pensamiento científico descubre sus leyes orgánicas” (Bachelard, 1940/1993, p. 115).

En este orden epistemológico, el proceso de producción de conocimientos científicos es racional (en el sentido de un racionalismo aplicado que no excluye la dimensión de la experiencia) y en su construcción elimina todo material de construcción irracional. Esos elementos irracionales (en tanto se diferencian de la racionalidad científica), si bien pretenden ser eliminados por la razón científica, dejan las huellas de los obstáculos epistemológicos que se pretendieron superar. Huellas que pueden observarse al realizar un perfil epistemológico de nuestras prácticas científicas, una especie de perfil psicológico de las diversas conceptualizaciones que están presentes en el sujeto que investiga. Trazando un perfil epistemológico “se podría medir la acción psicológica efectiva de las diversas filosofías en la obra del conocimiento” (Bachelard, 1940/1993, p. 37), poniendo en evidencia muchas zonas no tan claramente lógicas en lo que respecta a las nociones científicas.

Otro de los tópicos fundamentales de la apertura que ha significado la epistemología francesa es la propuesta epistemológica de Canguilhem, en la que la discontinuidad es tomada en el sentido de un combate per-

manente librado por la ciencia contra los “prejuicios”, las “resistencias” y los “obstáculos”. Si bien estos temas han sido abordados por Bachelard y Koyré, Canguilhem los trabajará como un procedimiento que forma parte de la historia de las ciencias. En palabras de Foucault:

Canguilhem hizo descender la historia de las ciencias desde la altura (matemática, astronomía, mecánica galileana, física de Newton, teoría de la relatividad) hasta regiones donde el conocimiento es mucho menos deductivo, regiones que han estado ligadas por mucho más tiempo al prestigio de la imaginación y que plantean una serie de cuestiones extrañas a los hábitos filosóficos. (Foucault, 1984/2009, p. 47)

La historia epistemológica de las ciencias canguilhemiana constituye su propia especificidad en tanto no se trata de una búsqueda de la verdad sino de la “historia de los ‘discursos de verdad’, es decir, de discursos que se rectifican, se corrigen, y que ejercen sobre sí mismos todo un trabajo de elaboración orientado por la tarea de ‘decir la verdad’” (Foucault, 1984/2009, p. 49). En tanto rectificación y corrección de los discursos, el error no es aquello que es eliminado por la evidencia de una verdad que venía desplegándose de manera silenciosa, sino que es aquello que es excluido de las sucesivas nuevas formas de decir la verdad. La discontinuidad es justamente la condición de posibilidad de la historia de las ciencias, en tanto esta se conforma en el siglo XVIII a partir de la conciencia de revoluciones científicas como las de la geometría algebraica, el cálculo infinitesimal y la cosmología copernicana y newtoniana.

Hacer la historia de los discursos de verdad, en tanto suponen correcciones, rectificaciones y cambios de rumbo en la dirección de la producción de conocimientos, supone también un ejercicio metodológico que debe retomarse cada vez para poder seguir garantizando una historia de las discontinuidades. En la construcción del discurso de verdad de la historia de la ciencia tradicional, hay una operación epistemológica que selecciona y elimina enunciados, teorías y objetos haciendo aparecer un discurso ordenado, en un espacio tiempo ideal. En este tipo de historia de las ciencias, el punto de vista epistemológico busca una normatividad interna en las diferentes actividades científicas tal como se han ido formando efectivamente.

Por otro lado, esta especificidad propia de la historia de las ciencias que tiene que ver con el punto de vista epistemológico, también implica una apertura a otras estrategias que no busquen ya la regularidad interna

de las teorías científicas en un momento dado, sino que posibiliten dar cuenta de las discontinuidades que forman parte en el proceso de producción y constitución de los discursos de verdad en la ciencia. Esta posibilidad epistemológica y metodológica es la que hacen posible las propuestas de Bachelard y Canguilhem. Específicamente, Canguilhem vuelve a ubicar a las ciencias de la vida en una perspectiva histórico-epistemológica.

En este breve ensayo que pretende puntear algunas claves de las aperturas que provocan los recorridos de Bachelard y Canguilhem en el campo de la historia epistemológica de las ciencias, buscamos garantizar el ejercicio epistemológico que desvela la historia de discontinuidades, contingencias y singularidades, que han hecho posible ese discurso científico y político que aparenta ser homogéneo, negando la heterogeneidad de elementos que lo han constituido en la historia de su formación. Garantizar el ejercicio crítico que tiene por objeto la construcción de nuestra propia racionalidad a través del compromiso racionalista permite conjurar el espíritu conservador que tiende a la construcción de uniformidades deshistorizadas, que aferrándose a las respuestas y dejan de preguntar. Como bien insiste Bachelard en *La formación del espíritu científico*,

precisar, rectificar, diversificar, he ahí los tipos del pensamiento dinámico que se alejan de la certidumbre y de la unidad, y que en los sistemas homogéneos encuentran más obstáculos que impulsos. En resumen, el hombre animado por el espíritu científico, sin duda desea saber, pero es por lo pronto para interrogar mejor. (Bachelard, 1938/1994, p. 19)

## Referencias

- Bachelard, G. (1993). *La filosofía del no* (N. de Labruno, trad.). Buenos Aires: Amorrortu. (Obra original de 1940)
- Bachelard, G. (1994). *La formación del espíritu científico: Contribución a un psicoanálisis del conocimiento objetivo* (J. Babibi, trad.). México: Siglo XXI. (Obra original de 1938)
- Badiou, A. (2009). *Pequeño panteón portátil* (M. Saúl, trad.). Buenos Aires, FCE. (Obra original de 2008)
- Canguilhem, G. (1976). *El conocimiento de la vida* (F. Cid, trad.). Barcelona, Anagrama. (Obra original de 1952)
- Canguilhem, G. (2001). Obertura (H. Beccacece, trad.). En G. Bachelard,

- El compromiso racionalista* (pp. 7-8). México, Siglo XXI. (Obra original de 1972)
- Canguilhem, G. (2005). *Lo normal y lo patológico* (R. Potschart, trad.). México, Siglo XXI. (Obra original de 1966)
- Canguilhem, G. (2009). *Estudios de historia y filosofía de las ciencias* (H. Pons, trad.). Buenos Aires, Amorrortu. (Obra original de 1968)
- Dagognet, F. (2006). Sobre una segunda apertura. En J. J. Wunenburger (coord.) & M. Serrado (trad.), *Bachelard y la epistemología francesa* (pp. 11-22). Buenos Aires: Nueva Visión. (Obra original de 2003)
- Foucault, M. (2009). La vida: la experiencia y la ciencia (F. Rodríguez, trad). En G. Giorgi. & F. Rodríguez (comps.), *Ensayos de biopolítica. Excesos de vida* (pp. 41-57). Buenos Aires, Paidós. (Obra original de 1984)
- Lecourt, D. (2005). La historia epistemológica de George Canguilhem. En G. Canguilhem, *Lo normal y lo patológico* (pp. VII-XXX). México, Siglo XXI. (Obra original de 1970)



## La filosofía de los experimentos mentales

Guadalupe Mettini\*

### Introducción

La manipulación de un fenómeno en una situación controlada ha sido considerada desde la Revolución Científica como la principal fuente de conocimiento del mundo natural. No obstante, los experimentos mentales han ocupado un lugar significativo en los episodios de cambio conceptual, en la formulación de conceptos y en el desarrollo de nuevas teorías. Esta clase de experimentos reportó un enorme beneficio en ocasiones en las cuales la experimentación ordinaria estuvo limitada por motivos técnicos y fue sumamente fértil para la exploración de hipótesis teóricas en campos en los que la experimentación real no era posible. Aunque existe una diversidad de ejemplares que pueden incluirse en esta categoría, los casos presentan algunos rasgos comunes. Consisten en la descripción de un escenario y/o de un aparato experimental y en la especificación de una serie de instrucciones sobre cómo introducir variaciones en la situación representada. La narrativa incluye además una interpretación de los resultados de estas acciones en el marco de algunos principios teóricos.

Desde el punto de vista filosófico los experimentos mentales presentan una propiedad asombrosa: proporcionan (o se proponen proporcionar) nuevo conocimiento acerca de fenómenos naturales sin la introducción de datos empíricos nuevos. Las primeras teorizaciones que repararon en esta cuestión coinciden con el desarrollo de la filosofía de la ciencia como disciplina independiente y se encuentran diseminadas en artículos y apéndices de textos canónicos. Estas contribuciones iniciales tendieron a expresar de manera no articulada, la paradoja central que motiva todavía hoy el tratamiento metateórico de estas prácticas: ¿cómo es posible obtener conocimiento nuevo del mundo natural solo razonando acerca de una situación imaginada? El estudio filosófico los experimentos mentales no ha logrado ofrecer todavía una explicación satisfactoria de su funcionamiento.

\* Universidad Nacional del Litoral.  
guadalupemettini@gmail.com

El objetivo de este trabajo es proporcionar un panorama general de la discusión filosófica acerca de los poderes epistémicos de los experimentos mentales, proveer un análisis crítico del estado actual de la discusión y formular los lineamientos de una perspectiva que integre la explicación del funcionamiento de estos dispositivos al examen de las prácticas experimentales. En vistas a estos fines, en primer lugar, se reconstruye sucintamente la discusión filosófica sobre el papel epistémico de los experimentos mentales en física. Se intenta mostrar que, aunque las posiciones inaugurales lograron formular el principal problema epistemológico que presentan estos dispositivos, ni estas perspectivas ni los desarrollos posteriores pertenecientes a la filosofía de la experimentación proporcionaron una explicación completa de su funcionamiento. En segundo lugar, se realiza una exposición crítica del estado actual de la discusión. Se establece allí que la falta de sistematicidad y la desvinculación entre los esquemas de análisis propuestos y algunas tesis ampliamente consensuadas acerca de las prácticas experimentales son limitaciones importantes de las recientes teorizaciones sobre los experimentos mentales. Finalmente se propone una línea de interpretación alternativa. Se sugiere que un marco más prolífico para examinar el funcionamiento de los experimentos mentales es el que proporciona la epistemología de los modelos científicos.

### **Los experimentos mentales en la ciencia física**

En los orígenes de la física moderna es posible identificar un número importante de experimentos mentales. Muchos de estos fueron empleados para plantear dudas acerca de la aplicabilidad de un principio teórico, establecer inconsistencias en una teoría o incluso dar sustento a una nueva hipótesis. Quizás el caso más famoso sea el propuesto pro Galileo (Galilei, 1638/1974). El experimento consiste en la representación de un escenario hipotético en el que se analiza el comportamiento de objetos de distinto peso en caída libre, lanzados individualmente y combinados. De la descripción del sistema experimental y de otras observaciones acerca del comportamiento de los cuerpos en medios de distinta densidad, Galileo deriva el principio de equivalencia. Contra la tesis aristotélica de acuerdo con la cual la velocidad (aceleración) de la caída de los cuerpos es proporcional a su peso, Galileo muestra que dos cuerpos de diferentes pesos, lanzados desde cierta altura, tocan el suelo casi al mismo tiempo en un medio sutil como el aire y caen con la misma velocidad en el vacío. Dado

que no se introducen nuevos datos, sino que los resultados se obtienen solo de la consideración de una serie de situaciones hipotéticas este caso ha sido considerado como el paradigma de los experimentos mentales en física. El empleo de idealizaciones como la eliminación de la resistencia o la desestimación de otras variables como el color de los cuerpos son rasgos característicos de estos dispositivos. Asimismo, la narrativa del experimento apela a actividades de visualización. El experimentador es instado a responder '¿qué sucedería?' en tales o cuales situaciones.

La noción de experimento mental fue introducida en la discusión filosófica por Mach (1897/1976), aunque el concepto aparece con anterioridad en el trabajo de Ørsted (1811/1998). La conceptualización de Mach apuntó algunos rasgos importantes de estos dispositivos: su parentesco con la experimentación ordinaria, su naturaleza mental o imaginaria y su importancia en el desarrollo de la ciencia moderna. A estas contribuciones iniciales, les fue seguido un largo silencio. Otros aportes importantes se encuentran recién hacia mediados del siglo XX. Una de las principales razones de la ausencia de los experimentos mentales en las reflexiones sobre razonamiento científico puede encontrarse en el influjo del empirismo del Círculo de Viena. Su vehemente rechazo a la filosofía especulativa motivó la idea de que estos dispositivos no constituyen herramientas metodológicas válidas para el conocimiento del mundo natural. Esta posición prevaleció entre los filósofos de la ciencia de la época. Popper (1952/2002) compartió la sospecha respecto de los usos de los experimentos mentales en el contexto de justificación. Aunque reconoció su fecundidad en la interpretación de la teoría cuántica, la termodinámica y la teoría de la relatividad, limitó sus usos legítimos a las funciones heurísticas y refutatorias. Consideró que los experimentos mentales son asimilables a métodos argumentativos y que el empleo de los mismos como evidencia a favor de una hipótesis es inadmisibile.

Otra de las razones de la ausencia de los experimentos mentales como tema de la reflexión filosófica es la centralidad que ocupó el estudio de las teorías en la filosofía de la ciencia hasta los años '60. La aparición de perspectivas historicistas, que pusieron de relevancia factores que había sido hasta el momento ignorados en el estudio del razonamiento científico, propició nuevas consideraciones acerca de la naturaleza y funciones de la experimentación. Koyré puso especial atención en la Revolución Científica y examinó los experimentos empleados por Galileo para impugnar la perspectiva aristotélica e introducir la propia. Kuhn (1964/1977), por

su parte, atribuyó a los experimentos mentales un importante papel en los procesos de reconceptualización que tienen lugar en las revoluciones científicas. En su opúsculo dedicado a esta cuestión se encuentra la primera formulación explícita la paradoja del conocimiento en los experimentos imaginarios:

Concediendo que todo experimento imaginario cuyos resultados han sido fructíferos incorpora en su diseño alguna información previa sobre el mundo, ésta no se halla en discusión dentro del experimento. [...] ¿Cómo es entonces que, confiando exclusivamente en datos familiares, se puede llegar con un experimento imaginario a una nueva comprensión de la naturaleza? (Kuhn, 1964/1977, p. 264)

Según Kuhn los filósofos que han considerado que los experimentos mentales ocupan un papel marginal en la justificación de teorías, fundan esta opinión en una mala comprensión del proceso de constitución de los conceptos. Los conceptos están imbricados en la matriz del paradigma, se adquieren en la experiencia y se vinculan con las expectativas empíricas a partir de las definiciones. La ejecución de un experimento mental que le permite al científico advertir que cierto concepto es, por ejemplo, inadecuado para dar cuenta de un fenómeno, le permite conocer simultáneamente algo acerca del mundo y acerca de su aparato conceptual. Al mostrar inadecuaciones entre las expectativas teóricas y la experiencia estos dispositivos le proporcionan al científico acceso a información que ya tiene a mano y que, a pesar de ello, le resulta de alguna manera inaccesible.

A pesar de los esfuerzos de los filósofos historicistas por mostrar el papel que los experimentos mentales desempeñaron en la historia de la física, a filosofía de la experimentación desarrollada durante 1980 interpretó a estos dispositivos en términos similares a los establecidos por Popper. Dentro esta línea de investigación los experimentos mentales fueron concebidos como tipos particulares de argumentos teóricos. Al depender de una serie de supuestos básicos tomados de un marco nomológico, sus poderes epistémicos se interpretaron como limitados a su capacidad retórica para exhibir contradicciones. A diferencia de los experimentos ordinarios, los experimentos mentales no tienen *vida propia* (Hacking, 1983, p. 150). Son fijos y, en gran parte, inmutables (Hacking, 1992, p. 307). Esto significa que no son capaces de madurar, evolucionar, adaptarse o reestructurarse, producir adelantos en la tecnología de los instrumentos ni desarrollar las habilidades de los experimentadores. Esta posición clausu-

ró el debate acerca de los poderes epistémicos de los experimentos mentales hasta los años 90 cuando algunos filósofos rehabilitaron el interés por estas prácticas. Estos autores introdujeron nuevas consideraciones acerca del acceso epistémico a las leyes de la naturaleza, el papel de las representaciones visuales y las estrategias argumentales empleadas en los experimentos mentales para ampliar nuestro conocimiento. Las nuevas reflexiones condujeron al desarrollo progresivo de un área de estudio robusta dentro de la filosofía de la ciencia.

### **Estado actual de la discusión**

La discusión actual en torno a de los experimentos mentales en ciencias fácticas se centra en el mencionado *problema epistemológico* de estos dispositivos más o menos en los mismos términos en los que fuera formulado por Kuhn (1964/1977, p. 264). En general, las propuestas intentan caracterizar a estos experimentos, determinar qué funciones pueden desempeñar respecto de hipótesis y teorías científicas, establecer cuáles son los alcances del conocimiento acerca del mundo natural que son capaces de proveer y cuáles son las fuentes de dicho conocimiento. Es posible agrupar las posiciones existentes respecto de este conjunto de problemáticas en cuatro perspectivas: la perspectiva empirista, la apriorista, la constructivista y las de los modelos mentales. Si bien las posiciones que pueden ubicarse bajo estos rótulos no constituyen grupos homogéneos, comparten un punto de vista general que habilita a presentarlas agrupadas de este modo.

Las perspectivas aprioristas sostienen que los experimentos mentales son capaces de proporcionar conocimiento nuevo del mundo natural prescindiendo de datos empíricos y coinciden en concebir a las matemáticas como paradigma de esta forma de conocimiento. Este punto de vista fue inicialmente sostenido por Ørsted (1811/1998), sugerido por Koyré (1973/1991) en base al estudio histórico de varios experimentos mentales de Galileo y elaborado extensamente por Brown (2011). Los experimentos mentales en física permiten, desde esta perspectiva, contemplar fenómenos naturales en la imaginación e intuir racionalmente las leyes de la naturaleza que los gobiernan, por lo que su funcionamiento es análogo al de los experimentos ordinarios y pueden desempeñar los mismos roles respecto de las teorías e hipótesis.

La perspectiva empirista representada principalmente por Norton, sostiene que los experimentos mentales científicos, a pesar de que en apariencia se parecen a los experimentos ordinarios, son argumentos encubiertos (Norton, 2004). Su alcance epistémico coincide con el de los argumentos teóricos y como tales no pueden apoyar hipótesis científicas del mismo modo que los experimentos reales. Se limitan a exhibir inconsistencias o mostrar las consecuencias indeseables de la asunción de algunos supuestos teóricos. Sus usos legítimos dependen por un lado de la validez de la inferencia empleada en los argumentos y por otro de la verdad de las premisas. Éstas se apoyan, a su vez, en información procedente del conocimiento empírico. Algunos experimentos mentales permiten organizar creencias empíricas que de otro modo se encuentran inarticuladas pero no poseen poderes epistémicos extraordinarios.

Las perspectivas constructivistas sugieren que los experimentos mentales pueden contribuir a una nueva comprensión del mundo natural porque son capaces de operar reformas en el aparato conceptual de los científicos. Esta posición, sustanciada inicialmente por Kuhn (1964/1977), fue posteriormente recuperada y ampliada por Gendler (2000). Desde este punto de vista los experimentos mentales eficaces son capaces de revelar lagunas en el sistema conceptual de los experimentadores, forzándolos a reconocer contradicciones en una perspectiva teórica o mostrando su incapacidad para dar sentido a un fenómeno que debería ser explicado por la teoría. El efecto de algunos experimentos mentales es comparable al que producen las observaciones y los experimentos reales especialmente en periodos de crisis paradigmática. Dentro de esta perspectiva Sorensen (1992) y Häggqvist (1996) propusieron esquemas para sistematizar el funcionamiento de los experimentos mentales empleando herramientas de lógica modal.

La perspectiva de los modelos mentales interpreta a los experimentos mentales como análogos estructurales de fenómenos empíricos. La descripción de un escenario o aparato permite a los experimentadores construir una representación espacio temporal de las relaciones y de la estructura causal de los eventos y las entidades involucrados. La interpretación de los experimentos mentales como modelos mentales fue introducida de manera independiente por Nersessian (1995) y Mišćević (1992). Ambas perspectivas comparten un trasfondo común de investigación en psicología cognitiva e intentan sustentar una propuesta de base empirista que permita ampliar la explicación proporcionada por Mach (1897/1976). En

general, sostienen que los modelos mentales representan situaciones de manera no arbitraria. Esta capacidad de representación es la extrapolación de habilidades cognitivas básicas al contexto de la investigación científica: representación visual, razonamiento contrafáctico y manipulación mental de una situación hipotética. Los modelos mentales que se forman a partir de la narrativa de un experimento mental permiten solucionar un problema planteado por la descripción de una situación hipotética filtrando alternativas a través del conocimiento empírico acumulado.

Si bien el panorama actual presenta diversidad en las perspectivas que se proponen explicar los poderes epistémicos de los experimentos mentales las propuestas existentes solo examinan parcialmente el problema. Solo los trabajos de Norton, Brown, Sorensen y Häggqvist desarrollan un estudio sistemático de los experimentos mentales, es decir, solo estas posiciones proveen una definición, una tipología y una fundamentación de sus funciones basada en el estudio de casos y solo la perspectiva representada por Norton proporciona un criterio explícito de confiabilidad epistémica. Asimismo, el estudio de casos es restringido. En general se halla limitado a experimentos procedentes del origen de la mecánica clásica y solo ocasionalmente al experimento EPR. Los experimentos mentales en mecánica cuántica y cosmología son escasamente examinados. Por otra parte, prevalece entre las perspectivas presentadas, una reticencia a definir apropiadamente la noción de experimento mental. En su lugar presentan ejemplificaciones, descripciones de rasgos perennes, comparaciones y analogías con otras prácticas. Esto dificulta el estudio sistemático ya que estas caracterizaciones resultan demasiado amplias o demasiado estrechas y no proporcionan condiciones suficientes para distinguir a los experimentos mentales de otras formas de razonamiento hipotético. Asimismo, aunque un número considerable de perspectivas sostiene que los experimentos mentales son esencialmente similares a los experimentos ordinarios, son escasas las referencias al análisis filosófico de estos últimos. Solo algunas posiciones emplean las conceptualizaciones elaboradas por la filosofía de la experimentación para iluminar el examen epistémico de los experimentos mentales. La tesis de la autonomía de la experimentación prácticamente no es discutida para el caso de los experimentos mentales, simplemente se asume que los diferentes propósitos con los que fueron empleados de hecho estos dispositivos son suficientes para atestiguar su independencia de la teoría. Dada la ausencia de la intervención del mundo natural, fundamento de la objetividad de los experimentos or-

dinarios, esta es una cuestión central para sostener que estos dispositivos pueden cumplir funciones en el contexto de justificación (más allá de las que puedan atribuírseles a los argumentos). En suma, de entre la variedad de posiciones presentadas no es posible obtener una interpretación del funcionamiento de los experimentos mentales que de sentido a la intuición de que se trata de prácticas experimentales y que sea consistente con la diversidad de casos existentes y con la multiplicidad de funciones que hasta el momento se asume que han desempeñado.

### **Hacia un estudio sistemático de los experimentos mentales**

Las conceptualizaciones recientes de los experimentos mentales no han logrado proporcionar una explicación satisfactoria de su funcionamiento. Una posible interpretación de la falta de consenso entre las posiciones y de las limitaciones en las interpretaciones vigentes es que las mismas se apoyan en falsas analogías. Algunas de las perspectivas que intentan dar cuenta del funcionamiento de estos dispositivos se basan en las similitudes entre los experimentos mentales y los argumentos que emplean suposiciones hipotéticas para derivar conclusiones acerca de los primeros. Pero la capacidad persuasiva de los experimentos mentales y a la rapidez con las que los resultados se obtienen a través de la ejecución de un experimento, son razones para argumentar que no se puede establecer una continuidad entre estas prácticas. Por otra parte, las posiciones que sostienen que estos dispositivos son experimentos usualmente refieren como motivos para ello las similitudes entre la narrativa de los experimentos mentales y la descripción de un experimento real y el rol activo que el experimentador asume en ambas prácticas. De acuerdo a este punto de vista, ambas prácticas se fundan en el método de la variación concomitante. No obstante, la falta de intervención en el mundo natural y la mayor dependencia que los experimentos mentales presentan respecto de los supuestos teóricos, parecen indicar que tampoco existe una continuidad entre éstos y los experimentos ordinarios.

Es posible argumentar que, los rasgos que aproximan a los experimentos mentales a las prácticas experimentales motivan su comparación con otro tipo de dispositivos. Los modelos mecánicos y las simulaciones computacionales, ampliamente utilizados en las prácticas científicas actuales, exhiben una semejanza más clara con los experimentos mentales que los argumentos y los experimentos ordinarios. Estas prácticas han sido incor-

poradas a la reflexión filosófica recientemente a partir de la noción de modelos científicos. Un modelo puede entenderse de manera general como:

[...] la descripción interpretativa de un fenómeno que facilita el acceso a ese fenómeno. Dicho acceso puede ser tanto perceptual como intelectual, cuando el acceso es no perceptual es en ocasiones facilitado por una visualización. Los modelos pueden descansar en idealizaciones, simplificaciones o analogías. Facilitar el acceso a un fenómeno implica en ocasiones, enfocarse en alguno de los aspectos específicos del fenómeno y desatender otros. A veces los modelos dependen de ideas o conceptos y frecuentemente emplean formalismo matemático pero siempre con la intención de proveer acceso a aspectos del fenómeno que se consideran esenciales. (Bailer-Jones, 2009, p. 1)

La interpretación de los experimentos mentales como modelos podría resultar prolífica para encausar su estudio sistemático. En principio la noción de modelo científico permitiría articular algunos rasgos salientes que los experimentos mentales exhiben: su carácter hipotético (no real), el empleo de simplificaciones, idealizaciones o abstracciones, el uso de representaciones visuales y su dependencia de una serie de principios teóricos. Asimismo, la noción de modelo empleada en estas teorizaciones tiene la acepción de modelo fenoménico, por lo que parece calzar correctamente con el objetivo cognitivo de estos dispositivos que en general es representar algún tipo de fenómeno. Por último, las elaboraciones recientes han desarrollado criterios de fiabilidad para el empleo de las simulaciones y los modelos, basándose en la idea de representación científica, sistema objetivo y usuario del modelo. Estas categorías podrían ser de utilidad para establecer criterios epistémicos similares para el caso de los experimentos mentales en los cuales en principio es posible trazar estas mismas distinciones. Estos criterios permitirían decidir además cuestiones relativas a la identidad de los experimentos, a sus repeticiones exitosas y a la validez de diferentes interpretaciones teóricas del mismo caso.

El aspecto más importante que estas perspectivas permitiría explorar es la legitimidad del uso de los experimentos mentales como evidencia a favor o en contra de teorías. Recientes desarrollos defienden el uso de modelos y simulaciones computacionales en el contexto de justificación. Esta posición descansa en una tesis empírica: en el contexto de la práctica científica las simulaciones y los modelos computacionales y mecánicos han sido consideradas como fuentes de conocimiento experimental. Asimismo, la proliferación del uso de simulaciones computacionales en el

contexto de justificación de teorías en las cuales la experimentación real no es posible, motiva su inclusión como potenciales fuentes de conocimiento acerca del mundo empírico. Por otra parte, la consistencia entre el conocimiento obtenido a través de estos dispositivos y el conocimiento experimental parece favorecer la tesis de acuerdo con la cual son fuentes confiables de conocimiento. En esta línea se ha argumentado que las simulaciones nos permiten crear el tipo de ambiente controlado en el que es posible variar las condiciones iniciales, los valores, los parámetros, etc., y en este sentido funcionan como el equipamiento de laboratorio usado para medir fenómenos físicos (Morrison, 2009, p. 34). Estos rasgos permiten postular que al menos cierta clase de simulaciones es capaz de proporcionar evidencia a favor de una teoría o incluso de proveer información en ámbitos en los cuales el desarrollo teórico es incipiente o nulo (Gelfert, 2016, p. 117). La interpretación de los experimentos imaginarios como modelos podría proporcionar argumentos análogos para considerar a estos dispositivos como fuentes de conocimiento experimental.

Una razón adicional para emplear las conceptualizaciones surgidas del examen de las simulaciones y los modelos al estudio sistemático de los experimentos mentales es que las interpretaciones filosóficas de estas estrategias no se comprometen en general con una epistemología exótica, ni requieren de la suscripción a compromisos metafísicos extraordinarios. Sería posible proveer, desde este punto de vista, una explicación del funcionamiento de los experimentos mentales que preserve el espíritu de la interpretación *machiana* de acuerdo con la cual la única fuente confiable de nuestro conocimiento es la experiencia apelando a diferentes estrategias de representación como medio usualmente empleado para ampliar el conocimiento.

## Referencias

- Bailer-Jones, D. (2009). *Scientific models in philosophy of science*. Pittsburgh: University of Pittsburgh Press.
- Brown, J. R. (2011). *The laboratory of the mind: Thought experiments in the natural sciences* (2<sup>nd</sup> ed.). New York: Routledge.
- Galilei, G. (1974) *Two new sciences* [*Discorsi e dimostrazione matematiche intorno a due nuove scienze*] (S. Drake, trad.). Madison: University of Wisconsin Press. (Obra original de 1638)

- Gelfert, A. (2016). *How to do science with models: A philosophical primer*. Cham: Springer.
- Gendler, T. (2000). *Thought experiment: On the power and limits of imaginary cases*. London: Routledge.
- Hacking, I. (1992). Do thought experiments have a life of their own? Comments on James Brown, Nancy Nersessian and David Golding. *Philosophy of Science*, 2, 302-308.
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.
- Hägqvist, S. (1996). *Thought experiments in philosophy*. Stockholm: Almqvist & Wiksell International.
- Koyré, A. (1991). *Estudios de historia del pensamiento científico* (E. Bustos & E. Perez Cerdeño, trads.). México: Siglo XXI. (Obra original de 1973)
- Kuhn, T. (1977). A Function for thought experiments. En *The essential tension* (pp. 240-265). Chicago: The University of Chicago Press. (Obra original de 1964)
- Mach, E. (1976) On thought experiments. En *Knowledge and error* (T. J. McCormack, trad.). Dordrecht: Reidel. (Obra original de 1897)
- Miščević, N. (1992), Mental models and thought experiments. *International Studies in the Philosophy of Science*, 6(3), 215-26.
- Morrison, M. (2009). Models, measurement and computer simulation: the changing face of experimentation. *Philosophical Studies*, 143, 33-57.
- Nersessian, N. (1995). Should physicists preach what they practice? *Science & Education*, 4, 203-226.
- Norton, J.D. (2004). Why thought experiments do not transcend empiricism. En C Hitchcock (Ed.), *Contemporary debates in the philosophy of science* (pp. 44-66). Malden, MA: Blackwell.
- Ørsted, H. C. (1998). First introduction to general physics. En K. Jelved, A. D. Jackson, & O. Knudsen (Eds.), *Selected scientific works of*

*Hans Christian Ørsted* (pp. 282-309). Princeton: Princeton University Press. (Obra original de 1811)

Popper, K. (2002). On the use and misuse of imaginary experiments, especially in quantum theory. En *The logic of scientific discovery* (pp. 464-480). New York: Routledge. (Obra original de 1952)

Sorensen, R. (1992). *Thought experiments*. New York: Oxford University Press.



## ¿La práctica hace a le expertx o le expertx hace a la práctica?

Sofía Mondaca\*.\*

Julián Reynoso\*

### 1. Introducción

La progresiva complejización de las prácticas científicas contemporáneas ha devenido en cúmulos de problemas que se manifiestan en distintos aspectos tales como la construcción de modelos, el diseño y la ejecución de experimentos, de simulaciones computacionales y de mediciones. La filosofía de la ciencia actual ha estudiado estas prácticas en mayor detalle a medida que se han convertido en estándares dentro de la ciencia. El conjunto de variables que puede afectar la modelización, la experimentación y la simulación de distintos fenómenos es cada vez mayor, no sólo por factores intrínsecos a la práctica, sino por costumbres, hábitos y tradiciones.

Y suele ser el caso que, a pesar de tener esas variables bajo control, parece ser necesario un “pase de magia”, un conocimiento adicional que no está explicitado en los protocolos de trabajo que suelen provenir de la mirada de profesionales con muchos años de experiencia bajo su manga. ¿Qué sucede, entonces, con le científicx expertx? ¿Cuál es su aporte en las prácticas científicas? Las nociones de experticia y saber práctico han sido muy discutidas en la filosofía de la mente desde mediados de siglo XX. Ryle (1949) inauguró este debate con la noción de saber-cómo [*know how*], señalando la importancia de reconocer el conocimiento práctico como un tipo de conocimiento disposicional relacionado con las capacidades prácticas de lxs sujetxs. No obstante, esta discusión no ha tenido, hasta el momento, gran repercusión en el ámbito de la filosofía de las prácticas científicas.

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH) Centro de Investigaciones María Saleme de Burnichon (CIFYH). Córdoba, Argentina.

\* CONICET, Instituto de Humanidades (IDH). Córdoba, Argentina.  
julianreynoso@unc.edu.ar

Intentando trazar una relación entre el gran campo de la filosofía de la mente y el gran campo de la filosofía de la ciencia, y en particular, las discusiones acerca de la práctica científica, el objetivo del presente trabajo será indagar de qué manera funciona la mirada práctica del expertx a la hora de tomar decisiones, a partir de dos casos particulares. Así, comenzaremos por revisar los puntos salientes del debate acerca del saber práctico señalado anteriormente. Luego, bajo esa lente, examinaremos de qué manera influye en el proceso teniendo en cuenta dos casos concretos de prácticas científicas: uno llevado a cabo por Hines et al. (2014) y el segundo de Reside et al. (2019). El propósito será mostrar la importancia del llamado saber práctico experto a la hora de tomar ciertas decisiones en contextos concretos de investigación. Si bien tal afirmación parece ser compartida por la mayoría de lxs científicxs, consideramos que se encuentra poco explorada filosófica y científicamente. De tal modo, pretendemos ofrecer unos primeros lineamientos para tomar tal debate con la importancia que merece.

## 2. Perspectivas y miradas sobre lxs expertxs y su experticia

La distinción entre un tipo de conocimiento teórico y un tipo de conocimiento práctico ha estado presente en las discusiones desde los orígenes de la filosofía. Ya Aristóteles trazaba la distinción entre *epistêmê*, entendida como teoría pura, y *technê*, entendida como práctica, y se esforzaba por discutir acerca de qué tipo de conocimiento implicaba la renombrada sabiduría. Sin embargo, recién a mediados del siglo XX, con la publicación de “Knowing how and knowing that” (1946) y posteriormente con *El concepto de lo mental* (1949), una de las obras más reconocidas de G. Ryle, el problema sobre el saber-qué [*know that*] vs. el saber-cómo [*know how*] ha tomado una importancia trascendental.

El saber-qué ha sido entendido como un tipo de conocimiento teórico y proposicional y se lo ha identificado con el hecho de tener cierta información acerca de algo: saber qué sostiene la teoría de la relatividad, saber qué función tienen ciertos instrumentos en un laboratorio, etc. Por otro lado, el saber-cómo ha sido entendido como un tipo de conocimiento práctico que se relaciona con las habilidades o la experticia que tiene unx sujetx para saber cómo realizar ciertas acciones: saber cómo usar un instrumento, saber cómo tomar muestras para un preparado de microscopio, saber cómo realizar una medición, etc. Basándose en las tesis ryleanas, muchos autores se han esforzado por mostrar que todo conocimiento, en

el fondo, consiste en poseer una habilidad para realizar determinado tipo de comportamiento.

Así, en la actualidad, podemos encontrar diversas interpretaciones de lo que es tener un saber-cómo o un saber práctico –que es tener experticia sobre algo– y, en consecuencia, qué significa ser unx agente expertx.<sup>1</sup> B. Montero nos plantea ciertas preguntas que resultan de gran interés para profundizar acerca de lo que significa ser unx expertx en cierta materia:

¿Deberíamos definir la pericia en referencia a algún tipo de estándar social, quizás diciendo que los expertos son aquellos que se han convertido en “profesionales” en su campo? ¿O deberíamos confiar en una prueba de habilidad, diciendo quizás que, independientemente de si son reconocidos como tales, los expertos actúan de una manera relativamente superior? ¿O podría haber una norma objetiva con respecto a la cual deberíamos medir la experiencia? ¿Podemos definir la experiencia como el logro de un cierto calibre de habilidad, como, por ejemplo, lo que en el modelo de adquisición de habilidades de Fitts y Posner (1967) cuenta como el nivel más alto de destreza? ¿O debería pensarse en la pericia como la acumulación de conocimientos? ¿Haber formado durante un cierto número de años es un componente necesario para ser un experto? ¿O podría un experto nacer en lugar de ser hecho? (2016, p. 52)

Podemos encontrar una gran cantidad de respuestas a tales preguntas. Pero todas parecen acordar con la idea de que, incluso en las actividades propiamente cognitivas, como las actividades académicas, no sólo es necesario poseer cierta información relevante para el campo determinado de especialización, sino también tener ciertas habilidades o técnicas (Goldman, 2001).

Tales preguntas habilitan la discusión acerca de cómo debemos entender nuestro conocimiento y, en particular, la discusión acerca de si existe de hecho una independencia entre tipos de conocimiento, cuál es la relación que hay entre ambos, qué clase de objetos o entidades decimos que se conocen con el *know that* o con el *know how*, cómo debemos entender los procesos de aprendizaje de ambos tipos de conocimiento, etc. No obstante, hasta el día de hoy, estas preguntas no han sido resueltas. Es razonable

---

<sup>1</sup> Dentro de la filosofía de la mente y la teoría del conocimiento, ha habido un amplio desarrollo de la distinción inaugurada por Ryle (1946, 1949). Principalmente podemos distinguir dos grandes enfoques, el intelectualismo y el anti-intelectualismo, que discuten hasta qué punto se puede sostener la distinción y qué tipo de características tienen dichos tipos de conocimiento. No obstante, entrar en tales inquisiciones excede los propósitos del presente artículo.

pensar que los debates sobre en qué consiste exactamente el conocimiento continuarán mientras exista algo a lo que llamemos “filosofía”.

Si bien Ryle ha influido principalmente en el campo de la filosofía de la mente y de la teoría del conocimiento, sus aportes pueden ser discutidos en otros campos de análisis filosófico. En particular, consideramos aquí que la noción de “saber cómo” resulta relevante para el análisis de la práctica científica. Con frecuencia, nos encontramos con situaciones que nos señalan que el aprender cierto conjunto de reglas generales no parece ser suficiente para saber cómo aplicarlas correctamente dependiendo la situación particular. Esto último requiere, al parecer, de una sabiduría especial e irreductible al cúmulo teórico. Cartwright afirma al respecto:

Nuestras mejores y más poderosas ciencias deductivas parecen apoyar sólo un tipo muy limitado de cierre: siempre y cuando los únicos factores relevantes en el trabajo sean aquellos que puedan ser modelados apropiadamente por la teoría, la teoría puede producir predicciones exactas y precisas. Pero esto está muy lejos de la esperanza de que todas las situaciones se presten a predicciones exactas y precisas. (1999, p. 188)

Luego, Barrotta y Montuschi señalan:

Incluso cuando las teorías están “personalizadas” para satisfacer las necesidades especiales del caso en cuestión (por ejemplo, elaborando modelos que incluyen mejoras en el ámbito de la aplicabilidad de la teoría general), no contienen de antemano todos los recursos necesarios para representar todos los casos posibles fuera del ámbito de competencia de los ámbitos en los que funcionan con éxito. (Barrotta & Montuschi, 2018, p. 3)

Ante la pregunta ¿cómo podemos aplicar el conocimiento general a circunstancias prácticas específicas?, algunos autores han ofrecido pensar en algunos tipos particulares de conocimiento. En ocasiones se dice que debemos tener en cuenta el llamado conocimiento local<sup>2</sup> a la hora de ejercer nuestras prácticas científicas (Barrotta & Montuschi, 2018), a veces también se ha hecho alusión a “factores de apoyo” para señalar aquellas condiciones locales que nos ayudan a aplicar el conocimiento general en circunstancias específicas (Cartwright & Hardie, 2012, p. 25).

---

<sup>2</sup> Dicho conocimiento local hace referencia a un conocimiento particular que proviene de fuentes que no son las científicas teóricas. Comúnmente se hace referencia a un conocimiento que proviene de la sociedad que es a menudo necesario para llenar el vacío entre el conocimiento de lxs expertxs y las decisiones correctas.

Lo que es cierto es que por más variadas que sean las respuestas parecen coincidir en que lxs expertxs científicxs necesitan ser aristotélicamente sabios, esto es, tienen que saber cómo aplicar sus conocimientos teóricos a las circunstancias locales y específicas (Barrotta & Montuschi, 2018). Los conocimientos teóricos pueden servir como una guía para la acción, como un cúmulo de reglas que conducen cómo debemos dirigir nuestra práctica científica. Sin embargo, siempre conservan un grado de generalidad tal que no pueden señalar cómo tomar muchas decisiones particulares y dependientes del contexto que, en reiteradas ocasiones, resultan de gran importancia tanto en el desarrollo de cierta práctica científica como también en la capacidad de predicción de ciertos resultados. “El papel del conocimiento local es crucial para asegurar un uso correcto –más específicamente, ‘relevante’– del conocimiento experto en circunstancias prácticas” (Barrotta & Montuschi, 2018, p. 7).

A continuación, presentaremos dos casos que muestran la relevancia del llamado conocimiento práctico, en distinción al conocimiento teórico en contextos de prácticas científicas.

### **3. Cómo influye la experticia sobre la construcción de modelos**

Los casos que hemos tomado para el presente trabajo pretenden rescatar distintos aspectos del impacto que la mirada experta tiene en las prácticas científicas. El primero involucra el análisis de muestras de tejidos en estudios de cáncer de mama, mientras que el segundo apunta al ajuste de modelos para predecir el impacto en hábitats de especies protegidas del cambio climático en regiones de Australia.

#### **1) El diablo está en los detalles (Hines et al., 2014)**

En un artículo publicado en 2014, Hines et al. describen con detalle las dificultades que encontraron al llevar a cabo una investigación conjunta entre dos laboratorios dedicados al estudio del cáncer de mama:

A pesar de utilizar métodos, reactivos y muestras aparentemente idénticos, nuestros dos laboratorios no pudieron reproducir los perfiles de clasificación celular activada por fluorescencia (FACS, por sus siglas en inglés) de las células mamarias primarias. (Hines et al., 2014, p. 779)

El comienzo de la colaboración estipulaba dos meses para conciliar los datos que se iban a utilizar para las investigaciones sucesivas. Sin embargo, y para sorpresa de ambos grupos involucrados, tomó más de dos años entender cuáles eran las diferencias en la metodología que arrojaban resultados completamente diferentes.

A pesar de que ambos grupos comenzaron con tejidos mamarios primarios de mamoplastia de reducción y el conjunto de perfiles FACS obtenidos en cada laboratorio fue consistentemente reproducible, los perfiles obtenidos en Boston y Berkeley no fueron similares. La pregunta era por qué. (Hines et al., 2014, p. 779)

El primer sospechoso de causar la diferencia fue el equipamiento que utilizaban. Si bien eran similares, eran modelos diferentes de equipos parecidos. No obstante, el mismo fue exonerado tras una prueba que involucró una calibración de microesferas que reveló patrones de línea celular de manera correcta.

Otros sospechosos fueron los métodos de obtención de las muestras, la posibilidad de que los tejidos de las muestras tuvieran vastas diferencias entre sujetxs de las costas este y oeste de Estados Unidos, la fuente de los sueros y los aditivos, protocolos empleados, pero nada parecía ser la causa. Tras un año de idas y vueltas, buscando la raíz de la diferencia que les impedía comenzar la investigación, se encontró que cuando los tejidos se dirigieron en organoides en un lugar y se enviaron a otro lugar para ser separados y analizados, los perfiles FACS resultantes coincidieron con la institución original donde se realizó la digestión de colagenasa del tejido. En palabras de Hines et al. “cada uno de los tejidos llevaba la firma del laboratorio que lo había procesado” (p. 780).

Un tejido fresco, nuevo, fue dividido en partes iguales y distribuido a ambos laboratorios con particular atención al método y se procesó según la técnica empleada en cada uno de ellos. Siguiendo los procedimientos, notaron que la manera en la que las muestras fueron centrifugadas en cada uno de los laboratorios diferían radicalmente.

## **2) Más allá del modelo: sapos, pájaros y expertxs**

Uno de los casos más interesantes para ilustrar la intrincada relación entre modelización y la mirada del experto es el estudio que Reside et al. (2019)

llevaron adelante sobre la región conocida como Humedales Tropicales Australianos, al noreste del país.

A partir de la necesidad de comenzar preparativos para adaptar las áreas protegidas ante las modificaciones que suponen distintos escenarios de cambio climático, Reside et al. argumentan que el enfoque habitual basado en modelos correlativos<sup>3</sup> falla al no capturar la complejidad del movimiento migratorio de las especies amenazadas.

Reside et al. tomaron un enfoque particularmente interesante. A partir de resultados de modelos correlativos y los sometieron a un proceso de escrutinio por parte de expertxs, especializados en hábitats y la geografía de la zona (potencialmente) afectada. Tras el escrutinio, formularon siete “principios generales” que dan cuenta cómo los modelos correlativos sin ajuste por parte de los expertxs sobre (o sub) estiman la aptitud de los hábitats en cuestión. Más allá de estos principios, intentaremos reconstruir de qué manera fue llevado a cabo este relevamiento y el modo en el que los modelos fueron refinados siguiendo las pautas propuestas por los expertxs.

Dada la urgencia y la magnitud de los cambios climáticos que se esperan, aun con las proyecciones más optimistas, Reside et al. sostienen que es necesario incorporar otros enfoques para evaluar los resultados de los modelos correlativos: “El conocimiento experto es utilizado cada vez más para aconsejar políticas de manejo e informar el estado del arte para situaciones en las que han de tomarse decisiones, pero los datos empíricos no son abundantes” (Reside et al., 2019, p. 2).

### 3) Modelos y expertxs

Reside et al. tomaron un enfoque algo más interesante. Su objetivo era encontrar aquellos casos en los que los modelos tenían un peor desempeño y desarrollar guías que tuvieran en cuenta principio ecológicos para mejorar los resultados de los modelos. Para ello tomaron tres elementos:

---

<sup>3</sup> Los modelos de distribución de especies correlativos modelan la distribución observada de una especie en función de variables predictoras climáticas referenciadas geográficamente utilizando múltiples enfoques de regresión. Dado un conjunto de presencias observadas referidas geográficamente de una especie y un conjunto de mapas climáticos, un algoritmo encuentra los rangos ambientales más probables dentro de los cuales vive una especie. Se asume que las especies están en equilibrio con su medio ambiente y que las variables ambientales relevantes han sido adecuadamente muestreadas. Estos modelos permiten la interpolación entre un número limitado de especies.

datos de predictores con grillas de 1 km de resolución; los modelos “Accuclim” de 250 m de resolución; y por último estos mismos modelos tras haber sido escrutados por expertxs. Según Reside et al.:

De estas proyecciones, cuantificamos la extensión y ubicación de las diferencias espaciales entre los *outputs* de los tres modelos y examinamos los factores que fueron omitidos por los primeros dos modelos. Utilizamos el *output* refinado por los expertos para predecir áreas que potencialmente podrían albergar micro-refugios. (Reside et al. 2019, p. 2)

Para llevar a cabo el estudio, Reside et al. tomaron 17 especies de vertebrados endémicos de la zona, de manera tal que la muestra fuera lo más representativa posible. Ocho de las especies eran ranas y las nueve restantes de pájaros. Los datos sobre la ocurrencia de las especies fueron obtenidos por relevamiento llevado adelante por el Centro de Biodiversidad Tropical. Por último, los datos climáticos fueron tomados del Proyecto Australiano de Disponibilidad de Agua, cuya resolución fue ajustada de 5 km a 250 m utilizando métodos propuestos en estudios ya realizados.<sup>4</sup> Con todas estas herramientas en la mano, se definió un umbral para evaluar la adecuación de un hábitat potencial, teniendo en cuenta la ubicación actual de las especies.

#### 4) Taller de expertxs

Para el proceso de examinación de los modelos, fueron convocados ocho expertxs en biología, ecología, en distribución de las especies estudiadas y en los Humedales Tropicales Occidentales. Dicen los autores: “Lxs expertxs comenzaron evaluando la distribución actual de las especies para verificar los modelos en cuestión, o para identificar área que sobre (o sub) estimaran hábitats potenciales guiados por preguntas predefinidas” (Reside et al., 2019, p. 4).

Es importante destacar que el proceso de consulta con lxs expertxs no siguió ninguna pauta estructurada, de las múltiples que se han desarrollado en distintos protocolos, dado el número relativamente pequeño de participantes. Reside et al. sostienen que para próximos estudios de este tipo se considerarán pautas más formales.

---

<sup>4</sup> Aquí Reside et al. (2019) echan mano al trabajo realizado por Williams et al. (2010).

## 5) Resultados

Utilizando herramientas interactivas de visualización científica, lxs expertxs formularon los ya mencionados siete principios que describen de qué manera la modelización de la distribución de las especies falla al estimar la idoneidad de los hábitats. Fue llamativa la incapacidad de capturar con precisión en una resolución suficientemente fina, las condiciones de la selva tropical húmeda y fresca. Estos factores son claves para la determinación de la idoneidad del hábitat para las especies estudiadas.

## 4. Comentarios finales (para empezar)

Los casos que hemos tomado para el presente trabajo pretenden rescatar distintos aspectos del impacto que la mirada experta tiene en las prácticas científicas. Ryle mismo nos dice que de hecho cuando hablamos de conocimiento, no nos interesamos tanto en el aprendizaje teórico adquirido sino más bien en las capacidades cognitivas que se aprenden, las que se ponen en juego, las que se muestran en las prácticas. Esta afirmación, creemos, cobra importancia en el análisis de los casos aquí presentados.

En el primer caso presentado (FCAS), nos interesa rescatar la dificultad que encontraron dos grupos de científicxs con un alto grado de experticia en investigación de punta sobre cáncer de mama al lograr replicar las muestras para una investigación conjunta. Para un profano en el área puede resultar absolutamente implausible que algo así ocurra, dado que se trata de profesionales altamente capacitados y especializados. Aun así, fue necesario un profundo análisis de las prácticas y costumbres en cada uno de los laboratorios para dar con la fuente de la discrepancia en el preparado de las muestras. Esto no necesariamente tiene que ver con aspectos “teóricos”, en el sentido de conocimiento de fondo específico de la disciplina, sino con aspectos más bien prácticos acerca de cómo se preparan las muestras. Algo que, da la impresión, daban por sentado que la manera hacerlo de cada uno de los laboratorios era “la manera” en la que debía hacerse. O, aún más, que el protocolo sobre cómo realizar el experimento era suficiente para determinar sólo una manera de aplicar los conocimientos teóricos que poseían.

Por otro lado, el caso que presentan Reside et al. (2019) sobre el trabajo conjunto entre modeladores y especialistas en disciplinas del *target* ilus-

tra de qué manera es posible integrar la mirada de expertxs para corregir y ajustar resultados del modelo para toma de decisiones en áreas protegidas. Resulta particularmente interesante que se dé este proceso de *feedback* entre expertxs en modelización y en la disciplina específica del sistema en cuestión. Dada la urgencia con la que ciertas decisiones deben ser tomadas, atacar el problema con ambas miradas parece haber producido resultados que sin dicha integración no hubieran sido del todo satisfactorios.

Para responder a la pregunta que lleva el título del presente trabajo: ¿la práctica hace al expertx o le expertx hace a la práctica?, podríamos decir que no es ni una ni otra. Lxs expertxs científicxs sin duda necesitan tener un conocimiento teórico acabado que le permita guiar su práctica, pero, al parecer, esto no es suficiente. El conocimiento teórico general no es suficiente por sí solo para explicar dicha práctica. Los casos mencionados pretenden ser muestra de esto. Tanto los protocolos armados para llevar a cabo un experimento específico como la interacción entre la comunidad de modeladores y expertxs en el sistema target no logran capturar toda la complejidad de los factores que están involucrados en las prácticas científicas. Esto parece sugerirnos que la forma de protocolizar mejor no tiene que ver con elaborar reglas y pasos a seguir teóricamente más específicos, esto es, con teorizar aún más las prácticas, sino, de lo contrario, con comprender mejor el fenómeno de las prácticas científicas, atendiendo a las complejidades que nos presenta. En tal sentido, el conocimiento práctico –el llamado saber-cómo– resulta relevante ya que es justamente lo que permite que lxs científicxs puedan aplicar los conocimientos generales a situaciones particulares. A pesar de ser un tema que ha despertado muchísimo interés en distintas disciplinas, hemos encontrado cierta dificultad a la hora de recabar conclusiones que tenga un grado de generalidad mayor a casos específicos.

## Referencias

- Barrotta, P., & Montuschi, E. (2018). Expertise, relevance and types of knowledge. *Social Epistemology*, 32(6), 387-396. <https://doi.org/10.1080/02691728.2018.1546345>
- Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Cambridge University Press.

- Cartwright, N., & Hardie, J. (2012). *Evidence-based policy: A practical guide to doing it better*. Oxford University Press.
- Goldman, A. I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research*, 63(1), 85-110
- Hines, W. C., Su, Y., Kuhn, I., Polyak, K., & Bissell, M. J. (2014). Sorting out the FACS: A devil in the details. *Cell Reports*, 6(5), 779-781.
- Montero, B. (2016). *Thought in action: Expertise and the conscious mind*. Oxford: Oxford University Press.
- Reside, A. E., Critchell, K., Crayn, D. M., Goosem, M., Goosem, S., Hoskin, C. J., Sydes, T., Vanderduys, E. P., & Pressey, R. L. (2019). Beyond the model: Expert knowledge improves predictions of species' fates under climate change. *Ecological Applications*, 29(1), e01824. <https://doi.org/10.1002/eap.1824>
- Ryle, G. (1946). Knowing how and knowing that: The presidential address. *Proceedings of the Aristotelian Society*, 46(1), 1-16. <https://doi.org/10.1093/aristotelian/46.1.1>
- Ryle, G. (1949). *The concept of mind*. University of Chicago Press.
- Wieten, S. (2018). Expertise in evidence-based medicine: A tale of three models. *Philosophy, Ethics, and Humanities in Medicine*, 13(1), 2. <https://doi.org/10.1186/s13010-018-0055-2>
- Williams, S. E., VanDerWal, J., Isaac, J., Shoo, L. P., Storlie, C., Fox, S., Bolitho, E. E., Moritz, C., Hoskin, C. J., & Williams, Y. M. (2010). Distributions, life-history specialization, and phylogeny of the rain forest vertebrates in the Australian Wet Tropics: *Ecological Archives* E091-181. *Ecology*, 91(8), 2493-2493. <https://doi.org/10.1890/09-1069.1>



## Reproducibilidad y repetitividad experimental entre filósofos y científicos

Marisa Velasco\*

### Introducción

Hace casi una década que escuchamos hablar cada vez más a menudo de “crisis de reproducibilidad” en las ciencias experimentales. Las más prestigiosas revistas científicas dedican casi un número al año, muchos editoriales, cartas de lectores y encuestas e investigaciones entre sus lectores sobre el problema de la crisis de reproducibilidad. Muchas publicaciones han cambiado sus políticas editoriales y los mecanismos de aceptación y evaluación por estas razones. Desde aproximadamente 2010 la literatura en revistas científicas sobre el problema es muy abundante, y a su vez en ciertas áreas disciplinares se han creado organizaciones para enfrentar el problema.

Por otra parte, los filósofos de la ciencia han discutido profusamente el problema de la replicabilidad y la reproducibilidad experimental, especialmente desde la década del 80 del siglo pasado.

Sin embargo, es posible notar que las reflexiones de científicos y filósofos van en direcciones en parte divergentes y que, al menos, existe una tensión en el modo en que caracterizan y el rol que dan a la reproducibilidad y replicabilidad experimental.

En la primera parte de este trabajo reconstruiré de un modo muy breve los conceptos filosóficos de replicabilidad, repetitividad y reproducibilidad experimental, así como los problemas asociados con esos conceptos. La literatura es abundante y las problemáticas discutidas, así como las posiciones asumidas, por los filósofos son muy diversas. Por razones de espacio, y porque esta literatura es más conocida en el ámbito filosófico, sólo presentaré algunas distinciones y problemas asociados.

En una segunda parte, intentaré caracterizar y sistematizar los múltiples problemas que se esconden detrás del concepto de “crisis de reproducibilidad” en la literatura científica. Esta presentación intentará mostrar

\* Universidad Nacional de Córdoba, Facultad de Filosofía y Humanidades.  
marisa@ffyh.unc.edu.ar



que algunos de esos problemas están muy lejos de las reflexiones filosóficas y otros, aunque están más cerca, han recibido un tratamiento completamente diferente al que recibieron en ámbitos filosóficos.

## El problema entre filósofos de la ciencia

En gran parte de los textos clásicos de filosofía de la ciencia se puede encontrar reflexiones sobre el tema, generalmente tratando de elucidar por qué la replicabilidad es la marca del método científico. Por ejemplo,

Sólo cuando se da la recurrencia de ciertos acontecimientos de acuerdo con reglas o regularidades —y así sucede con los experimentos repetibles— pueden ser contrastadas nuestras observaciones por cualquiera (en principio). Ni siquiera tomamos muy en serio nuestras observaciones, ni las aceptamos como científicas, hasta que las hemos repetido y contrastado. Sólo merced a tales repeticiones podemos convencernos de que no nos encontramos con una mera «coincidencia» aislada, sino con acontecimientos que, debido a su regularidad y reproductibilidad, son, en principio, contrastables intersubjetivamente. . . . Puede definirse el efecto físico científicamente significativo como aquél que cualquiera puede reproducir con regularidad sin más que llevar a cabo el experimento apropiado del modo prescrito. (Popper, 1962, p. 44)

En Collins (1975) podemos encontrar la primera versión de su tesis del “regreso del experimentador”. Esta idea fue desarrollada luego con detalle en *Changing order* (1985) y posteriormente en diversos artículos. En forma breve, y en las propias palabras del autor, la tesis sostiene que:

. . . el *regreso del experimentador* . . . es una paradoja que se produce para aquellos que usan la replicación como un test de la verdad de las afirmaciones del conocimiento científico. El problema es que, puesto que la experimentación es un problema de habilidades prácticas, nunca puede estar claro si un segundo experimento que ha sido realizado suficientemente bien cuenta para chequear los resultados del primero. Algún test adicional es necesario para testear la calidad del experimento —y así sucesivamente. (Collins, 1985, p. 2)

¡Pero no sabremos si hemos construido un buen detector hasta que lo usemos y obtengamos el resultado correcto! (Collins, 1985, p. 84)

La primera es la versión experimental, la segunda es la versión instrumental de la tesis. Las repercusiones de las ideas de Collins han sido muy

variadas, así como las interpretaciones que se han dado de esta tesis. Hasta hace unos pocos años la más común era que Collins no estaba sosteniendo un argumento epistemológico, sino uno sociológico, porque sólo puede romperse el regreso infinito a través de factores no epistémicos. La idea de conocimiento implícito, a la que recurrió el propio Collins, no jugaría en estas interpretaciones un papel epistémico. Estas discrepancias, en el marco de las disputas con las perspectivas sociales de la ciencia, llevaron, por ejemplo, a encendidas discusiones entre Allan Franklin y Harry Collins. Sin embargo, hoy es posible encontrar un escrito común entre Franklin y Collins (2016) hablando exactamente de este problema en el que reconocen más acuerdos que discrepancias.

A partir de las ideas iniciales de Collins, y especialmente después del nacimiento de lo que dio en llamarse epistemología de la experimentación, es posible encontrar muchas reflexiones filosóficas sobre replicabilidad y reproducibilidad experimental, aunque no necesariamente en continuidad con la tesis de Collins.<sup>1</sup>

La literatura filosófica no ha sido demasiado uniforme en el uso terminológico, aunque sí en las distinciones conceptuales por lo que utilizaré los términos con el siguiente significado:

- **Repetibilidad** del experimento en sus procedimientos y condiciones materiales.
- **Reproducibilidad** del resultado de un experimento a través de diferentes procedimientos materiales para obtener el mismo resultado experimental.
- **Replicabilidad** experimental incluye, como término genérico, la repetibilidad y la reproducibilidad.<sup>2</sup>

Esta literatura ha abordado como tópicos de interés filosófico los siguientes problemas:

---

<sup>1</sup> Cf., por ejemplo, entre muchos otros trabajos, Radder (1992, 2003), Steinle (2002, 2016), Bogen (2001), Franklin (1994, 1997), Leonelli (2018), Norton (2015), Schickore (2010, 2011), Feest (2016, 2019).

<sup>2</sup> El vocabulario no sólo no es uniforme, sino que además se incorporan otras distinciones como, por ejemplo, reproducción directa y reproducción conceptual, pero que no aportan a lo que este trabajo quiere mostrar. La terminología adoptada aquí sigue casi textualmente la adoptada por Radder, aunque no necesariamente se siguen las ideas de este autor.

- a) La posibilidad misma de que un experimento sea repetible. Es a partir de tratar de elucidar qué significa repetir un experimento que los conceptos anteriores (repetibilidad, reproducibilidad y replicabilidad) aparecen como distinciones necesarias.
- b) Sólo los experimentos repetibles/reproducibles (repetidos/reproducidos) constituyen evidencia científica. Poner en cuestión esto implica, por un lado, distinguir los experimentos repetibles de los experimentos repetidos, pero, por otro lado, también obliga a repensar el propio concepto de evidencia experimental y su rol dentro del conocimiento científico.
- c) La replicabilidad experimental es la marca del método científico, al menos desde el siglo XVII. Diversos trabajos han mostrado el carácter histórico de la noción de replicabilidad, esto es, han mostrado diferentes modos de entender la replicabilidad en distintas disciplinas durante los siglos XVII, XVIII y XIX. Pero estos trabajos también han mostrado que la no-replicabilidad experimental ha tenido diferentes roles en diferentes períodos históricos.
- d) El problema de la confiabilidad de los experimentos únicos. Esto es, las prácticas científicas muestran experimentos que han sido diseñados para no ser repetidos (por diversas razones prácticas y teóricas) y sin embargo sus resultados gozan de una gran credibilidad.
- e) El control instrumental es un modo alternativo de replicabilidad. Esto es, gran parte de la confianza en los resultados experimentales no proviene de su replicabilidad, sino de que podamos tener un control y calibración instrumental como base de esa confianza.

Muchos de los tópicos de esta lista, que está muy lejos de ser exhaustiva, están profundamente relacionados y en algunas ocasiones han sido la base del desarrollo de programas de investigación en filosofía de la ciencia. Mi intención con esta enumeración es hacer patente que estos son problemas acerca de los cuales han pensado muchos filósofos de la ciencia de modo explícito desde por lo menos la década del 1980.

Aunque los modos de abordar estos problemas y las soluciones propuestas por los filósofos de la ciencia no son uniformes, como es habitual en esta disciplina, creo que casi todos ellos acordarían con la siguiente cita:

Distintos casos históricos muestran una sorprendente variedad de actitudes hacia la replicabilidad. Esto deja en claro que la replicabilidad no fue ni es el único medio para aceptar los resultados experimentales. . . . La replicabilidad puede ser una condición suficiente pero no es una condición necesaria para la aceptación de los resultados experimentales. (Steinle, 2016, pp. 57-58)

## Reproducibilidad entre científicos: “crisis de reproducibilidad”

Entre el 2010 y 2015 se realizaron una serie de estudios sobre problemas de reproducibilidad de resultados publicados en prestigiosas revistas científicas<sup>3</sup>. Los primeros estudios afectan de modo especial a investigaciones en psicología e investigaciones pre-clínicas de cáncer y muestran tasas muy altas de problemas de replicabilidad experimental. Esta discusión en el ámbito científico recibe comúnmente el nombre de “crisis de reproducibilidad”. Aunque el vocabulario también es variable en este ámbito en general se refieren alternativamente a lo que hemos denominado más arriba como repetibilidad, reproducibilidad y replicabilidad, alternativamente. Los problemas, perspectivas y disciplinas son muy amplios y variados en la literatura de científicos e investigadores de lo que se ha denominado *metascience*. Sin embargo, el marco general que comparten podría resumirse en la siguiente cita que introduce la presentación de un simposio de científicos sobre la temática:

La *regla de oro* de la ciencia es la reproducibilidad. Idealmente, las investigaciones son dignas de atención, publicación y cita si investigadores independientes pueden reproducirlas. . . . Pero en gran parte de la literatura los resultados no son reproducibles. (Grant, 2012, el énfasis es mío)

Son muy pocas las ocasiones donde la “regla de oro” se pone en tela de juicio y es por esta razón que la llamo el marco de la discusión y ese marco explica por qué las dificultades en la replicabilidad dan lugar a una crisis.

---

<sup>3</sup> Podrían citarse muchos estudios en esta dirección, así como mucha discusión, en el ámbito de las revistas científicas más prestigiosas, sirvan a modo de ejemplo (D. Baker, Lidster, Sottomayor, & Amor, 2012; M. Baker, 2015; Begley, 2013; Begley & Ellis, 2012; Open Science Collaboration, 2015; Schooler, 2014). Tanto es así que casi todas ellas tienen anualmente a partir de esa fecha algún editorial y alguna sección especial dedicada al tema. *Nature* realizó una encuesta entre sus lectores y autores en 2017 cuyos resultados y análisis pueden encontrarse en <https://media.nature.com/original/magazine-assets/d41586-018-04590-7/15675426>. Las publicaciones en esa misma revista sobre problemas de reproducibilidad entre 2012-2018 puede encontrarse en <https://www.nature.com/collections/prbfkwmvzw>

Aunque a lo largo de estos años la literatura científica ha abordado diversos temas en torno a lo que denominaron crisis de reproducibilidad, la siguiente lista muestra los tópicos más recurrentes en la discusión:

- a) Se da por sentado que habitualmente los resultados publicados no son reproducidos, ni repetidos por la comunidad científica. Esto recibe tratamiento de diverso tipo: las revistas importantes no publican normalmente reproducciones ni repeticiones experimentales, sean estas exitosas o no, los científicos jóvenes pierden tiempo de carrera académica si se dedican a este tipo de actividad. Buena parte de la discusión tomó esto como un hecho que lleva a la discusión del punto siguiente.
- b) Las publicaciones de resultados experimentales rara vez incluyen datos suficientes para llevar adelante repeticiones experimentales, y hasta puede resultar complejo llevar adelante reproducciones. Aunque en algunas disciplinas resulte complejo pensar qué significaría repetir un experimento, en general la regla de oro no es puesta en tela de juicio. Solo en trabajos más recientes esta variable ha empezado a tenerse más en cuenta.
- c) El tratamiento estadístico de los datos publicados no siempre resulta transparente en una publicación, por lo que se sugiere que el problema no es el diseño y la ejecución experimental sino la inferencia estadística involucrada en los datos publicados.
- d) Casi como una consecuencia de los datos anteriores (y porque además es un problema para las disciplinas que más impacto recibieron por el análisis de reproducibilidad) a veces parece que el problema dejaría de ser tal si las muestras experimentales tuvieran el tamaño adecuado para las inferencias que se llevan adelante con esos resultados.
- e) Esta literatura parece suponer, y a veces lo enuncia explícitamente, que la crisis de reproducibilidad no está uniformemente repartida a través de las disciplinas, las más afectadas parecen ser la psicología, los tratamientos clínicos (especialmente los de cáncer) y las investigaciones ligadas a neurociencias.<sup>4</sup>

---

<sup>4</sup> Resulta llamativo que en la discusión al interior de algunas de estas disciplinas parece suponerse, aunque no se diga de modo explícito, que hay disciplinas o áreas disciplinares que resultan casi inmunes a los problemas de replicabilidad. Más adelante se listan algunas estrategias para salir de la crisis de reproducibilidad, muchas de las estrategias hacen patente que habría disciplinas que constituyen modelos a seguir.

- f) Finalmente, aunque en términos epistemológicos puede resultar el aspecto menos relevante, se discute si la crisis de reproducibilidad es una forma de fraude dentro de la ciencia.<sup>5</sup>

Esta discusión, que aquí ha sido fuertemente resumida y esquematizada, tuvo diversos impactos en los modos de hacer y publicar resultados científicos. La siguiente lista, también de modo resumido y esquemático, muestra las respuestas a los problemas planteados por parte de las institucionales y a través de la modificación de prácticas de investigación:

- a) Cambios en políticas editoriales. Las revistas científicas exigen más datos, generalmente en forma de anexos disponibles al menos para evaluadores, antes de la evaluación para publicación, la mayoría de las revistas han elaborado *check lists* para el envío de trabajos empíricos basadas en problemas detectados por esta discusión, se han incorporado evaluadores especialistas en estadística, que no son necesariamente especialistas disciplinares, entre otras modificaciones.
- b) Movimiento *Open Science*. En buena medida el movimiento de Ciencia Abierta es una respuesta a esta discusión. Se espera que la disponibilidad “completa”<sup>6</sup> y abierta de los resultados experimentales subsanen en gran parte los problemas que la crisis de reproducibilidad ha dejado al descubierto.
- c) Trazabilidad experimental. Como consecuencia de los puntos anteriores (de ambos, o al menos de uno de ellos) se espera que la trazabilidad experimental sea una fuente de transparencia y confiabilidad en los resultados que la crisis puso en tela de juicio.
- d) “Elaboración y estandarización metodológica” y reproducción parcial. Los menos optimistas respecto a que los puntos anteriores logren evitar los problemas puestos en evidencia por la crisis de reproducibilidad, aunque sin renunciar a la regla de oro, han puesto sus esperanzas en estos dos aspectos. El supuesto es que disciplinas altamente estan-

---

<sup>5</sup> En este punto hay un acuerdo bastante unánime que la incidencia del fraude intencional tiene un impacto poco significativo. Lo que hace que la crisis de reproducibilidad tenga consecuencias principalmente epistemológicas y metodológicas.

<sup>6</sup> No he podido evitar las comillas porque la propia idea de completa, así como la idea de estandarización ideal del punto d) de esta lista, resulta cuanto menos ingenua desde una perspectiva filosófica.

darizadas en sus protocolos experimentales no tienen los problemas antes mencionados, y que si no es posible la replicabilidad al menos se debe aspirar a una reproducción parcial. Estas ideas son producto de discusiones que intentan interpretar o poner en tela de juicio los resultados que presenta esta disciplina llamada *meta-science*, por lo que bien podríamos llamarla *meta-meta-science*.

## Consideraciones finales

Los filósofos y los historiadores de la ciencia parecen haber puesto en cuestión lo que hemos llamado la “regla de oro” de la ciencia, al menos en una versión irrestricta. De esta forma la replicabilidad experimental resulta en uno de los múltiples modos en que la evidencia experimental gana en confiabilidad. Los científicos parecen no seguir tampoco esta regla en sus prácticas. Esto genera, al menos en la versión pública de la ciencia, desconfianza en los resultados publicados, así como un conjunto de reflexiones y acciones para salir de lo que han llamado “crisis de reproducibilidad”. En síntesis, la tensión parece profunda entre científicos y filósofos en lo que respecta al rol de la reproducibilidad. Sin embargo, las estrategias sugeridas, y en algunos casos llevadas adelante, por los científicos tienden a acercar posiciones.

## Referencias

- Baker, D., Lidster, K., Sottomayor, A., & Amor, S. (2012). Reproducibility: Research-reporting standards fall short. *Nature*, *492*(7427), 41.
- Baker, M. (2015). First results from psychology’s largest reproducibility test. *Nature*. doi: 10.1038/nature.2015.17433
- Begley, C. G. (2013). Six red flags for suspect work. *Nature*, *497*(7450), 433-434. doi: 10.1038/497433a
- Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531.
- Bogen, J. (2001). ‘Two as good as a hundred’: Poorly replicated evidence in some nineteenth-century neuroscientific research. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *32*(3), 491-533. doi: 10.1016/S1369-8486(01)00013-9

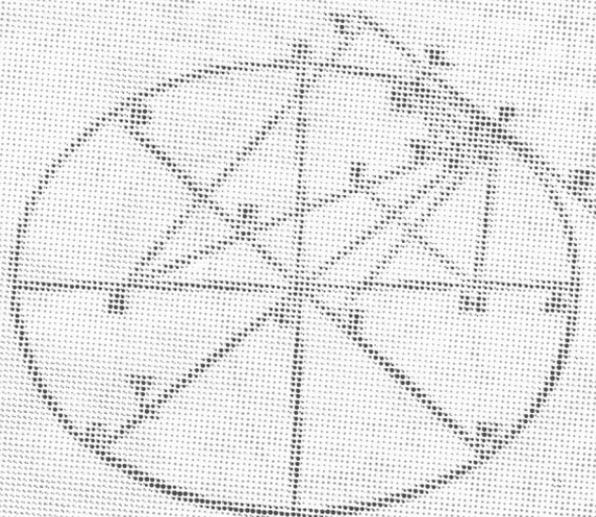
- Collins, H. M. (1975). The seven sexes: A study in the sociology of a phenomenon, or the replication of experiments in physics. *Sociology*, 9(2), 205-224. doi: 10.1177/003803857500900202
- Collins, H. M. (1985). *Changing order: Replication and induction in scientific practice*. London; Beverly Hills: Sage Publications.
- Feest, U. (2016). The experimenters' regress reconsidered: Replication, tacit knowledge, and the dynamics of knowledge generation. *Studies in History and Philosophy of Science Part A*, 58, 34-45. doi: 10.1016/j.shpsa.2016.04.003
- Feest, U. (2019). Why replication is overrated. *Philosophy of Science*, 86(5), 895-905. doi: 10.1086/705451
- Franklin, A. (1994). How to avoid the experimenters' regress. *Studies in History and Philosophy of Science Part A*, 25(3), 463-491. doi: 10.1016/0039-3681(94)90062-0
- Franklin, A. (1997). Recycling expertise and instrumental loyalty. *Philosophy of Science*, S42-S52.
- Franklin, A., & Collins, H. (2016). Two kinds of case study and a new agreement. En T. Sauer & R. Scholl (Eds.), *The philosophy of historical case studies* (pp. 95-121). doi: 10.1007/978-3-319-30229-4\_6
- Grant, B. (2012). Science's reproducibility problem. *The Scientist Magazine*. Recuperado de <https://www.the-scientist.com/news-opinion/sciences-reproducibility-problem-40031>
- Leonelli, S. (2018). Rethinking reproducibility as a criterion for research quality. En L. Fiorito, S. Scheall, & C. E. Suprinyak (Eds.), *Research in the history of economic thought and methodology* (Vol. 36, pp. 129-146). doi: 10.1108/S0743-41542018000036B009
- Norton, J. D. (2015). Replicability of experiment. *THEORIA. An International Journal for Theory, History and Foundations of Science*, 30(2), 229-248. doi: 10.1387/theoria.12691
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251). doi: 10.1126/science.aac4716

- Popper, K. R. (1962). *La lógica de la investigación científica* (V. Sánchez de Zavala, Trad.). Madrid: Tecnos.
- Radder, H. (1992). Experimental reproducibility and the experimenters' regress. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 63-73. Recuperado de <http://www.jstor.org/stable/192744>
- Radder, H. (2003). *The philosophy of scientific experimentation*. Pittsburgh, Pa.: University of Pittsburgh Press.
- Schickore, J. (2010). Trying again and again: Multiple repetitions in early modern reports of experiments on snake bites. *Early Science and Medicine*, 15(6), 567-617. doi: 10.1163/157338210X526629
- Schickore, J. (2011). The significance of re-doing experiments: A contribution to historically informed methodology. *Erkenntnis*, 75(3), 325-347. doi: 10.1007/s10670-011-9332-9
- Schooler, J. W. (2014). Metascience could rescue the 'replication crisis'. *Nature News*, 515(7525), 9. doi: 10.1038/515009a
- Steinle, F. (2002). Experiments in history and philosophy of science. *Perspectives on Science*, 10(4), 408-432.
- Steinle, F. (2016). Stability and replication of experimental results: A historical perspective. En H. Atmanspacher & S. Maasen (Eds.), *Reproducibility: Principles, problems, practices, and prospects* (pp. 39-63). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118865064.ch3>



---

## 2. Enseñanza de las ciencias







# Análisis histórico-epistemológico preliminar sobre el modelo de endosimbiosis para enseñar biología celular y naturaleza de la ciencia al profesorado

Joaquín Álvarez Soria\*

Eduardo Lozano\*

## Los estudios con enfoque histórico epistemológico (EHE) en el campo de la enseñanza de las ciencias

En las investigaciones que llevamos a cabo, y para colaborar con el diseño de propuestas de enseñanza integradoras de aspectos disciplinares de la biología y de contenidos metacientíficos (Lozano et al., 2016, 2018), realizamos estudios desde una perspectiva histórico-epistemológica (Matthews, 1994; Adúriz Bravo, 2010), que dan lugar a la adopción de una postura metacientífica explícita en el proceso de enseñanza y también articulada con la propia naturaleza de los modelos a enseñar.

Por una parte, estos estudios se enfocan en las variables internas de la dinámica de la producción de los modelos científicos abordados, de aquí que pueden considerarse estudios “internalistas”, y los análisis producidos, de orden epistemológico, permiten describir y comprender aspectos del desarrollo de los modelos, la estructura que sostiene las vinculaciones entre el modelo teórico y las parcelas de realidad que explica, las tensiones, la dinámica argumentativa de los ajustes teórico-empíricos realizados en el tiempo, entre otros aspectos (Giere, 1988/1992, Adúriz Bravo, 2012). Además, los EHE proporcionan información sobre las circunstancias históricas y también sobre los contextos de pensamiento en que surgieron los modelos teóricos y permiten describir y analizar el devenir que los modelos han tenido hasta su estado de actual consolidación (Kuhn, 1977/1996; Orozco, 1992).

En el marco de este enfoque, y en función de diseñar una unidad didáctica para la enseñanza del modelo de endosimbiosis para la materia: Introducción a la Biología, de una carrera de Profesorado de Nivel Medio

\* Centro de Estudios e Investigación en Educación (CEIE), Universidad Nacional de Río Negro (UNRN).  
jasoria@unrn.edu.ar

y Superior en Biología, desarrollamos un estudio que permitió abordar las siguientes preguntas: ¿Cómo fue la génesis, evolución y desarrollo del modelo de endosimbiosis? ¿Qué dificultades, obstáculos, y discusiones aparecieron en su producción y en qué contextos se dieron? Dichas preguntas se formulan con el objetivo de identificar episodios históricos que ofrezcan ideas, materiales, recursos, analogías, enfoques y también textos que contextualicen, que sirvan como escenario para ambientar la discusión sobre determinados temas metacientíficos a enseñar, pertinentes a esos episodios y consideradas de interés para la formación de los futuros profesores en biología.

La metodología utilizada para estos estudios implica un relevamiento exhaustivo de fuentes primarias y secundarias, y también de otros estudios calificados, que ofrezcan una versión consistente y robusta de la historia y evolución del modelo científico abordado.

## **Periodización preliminar**

El modelo de endosimbiosis, desde finales del Siglo XIX hasta la actualidad, se ha estructurado alrededor de la siguiente cuestión: tratar de determinar qué tipo de implicancias tuvieron los antepasados bacterianos en el origen y evolución de la célula eucariota. Las investigaciones y las controversias en torno a esta problemática se dieron en diferentes momentos y contextos teóricos y tecnológicos, y esto ha posibilitado caracterizar tres períodos, en los casi 150 años de desarrollo.

### **a) Primeras aproximaciones teóricas sobre endosimbiosis**

Las primeras investigaciones sobre endosimbiosis fueron formuladas por científicos que trabajaron en diferentes disciplinas, tanto en Europa como en Estados Unidos (Margulis, 2002; Sapp, 2014). El botánico francés Andreas Schimper (1856-1901) y el ruso Andrei Sergeivich Famintsyn (1835-1918), investigaban en el campo de la fisiología vegetal y se encontraban interesados particularmente en las células y sus cloroplastos. En un contexto en el cual se afianzaba la teoría celular (Baker, 1952; Harris, 2000) y se discutía la recientemente formulada teoría darwiniana de la evolución, propusieron que los cloroplastos tendrían un origen simbiótico relacionado con un organismo fotosintetizador llamado cianobacteria, y se llevaron a cabo intentos experimentales de laboratorio con el fin de cultivar y estudiar la naturaleza de dichos cloroplastos.

El zoólogo francés Paul Portier (1866-1962), propuso el origen bacteriano de las mitocondrias en el libro *Les symbiotes* (1918) y el biólogo celular Ivan Wallin (1883-1969), planteaba también que las mitocondrias podrían ser poblaciones de bacterias y que la diferenciación de estas originaba los demás orgánulos celulares. Por otro lado, el ruso Kostantine Merezhkovsky (1855-1921) jugó un papel central en la formulación teórica de la endosimbiosis, acuñando el término “simbiogénesis” (1909), definido como “el origen de organismos por la combinación o asociación de dos o más seres que entran en simbiosis” (Merezhkovsky, 1910, en Sapp, 2014, p. 92). Su compatriota, Boris Michailovich Kozo-Polyansky (1890-1957), un botánico evolucionista, extendió estas ideas en su libro de 1924: *Simbiogénesis: Un nuevo principio evolutivo*.

Así, en este período, diversos científicos plantearon hipótesis robustas sobre el origen bacteriano de algunos orgánulos que componen estructuralmente las células eucariotas (Margulis, 2002) pero, como propone Sapp (2014), tuvieron un lugar marginal dentro de la biología durante la mayor parte del siglo XIX y principios del XX.

Esta cuestión puede atribuirse a la fuerte influencia de una teoría emergente en esa época, la que establecía el origen microbiano de las enfermedades (Sapp, 2014). Bajo esta concepción del mundo, era determinante en las discusiones el rol de los microorganismos como agentes infecciosos y no la historia de ese grupo biológico de gran diversidad y antigüedad (Lazcano Araujo, 2002). De forma sintética, Ivan Wallin dejó plasmada esta idea en su obra *Symbiogenicism and the origin of species*:

La propuesta de que las bacterias, [...] pueden representar el factor causal fundamental en el origen de las especies es bastante chocante. La evidencia de las actividades constructivas de las bacterias ha estado ahí desde hace muchos años, pero las concepciones populares de las bacterias han quedado teñidas principalmente por sus actividades destructivas, cuya representación es la enfermedad. (Wallin, 1921, p. 8)

Para dar cuenta de otros obstáculos de conocimiento que se debieron sortear, Margulis (2002) describe que, “[...] hasta la llegada del microscopio electrónico, un instrumento de enorme resolución, había cerca de veinte términos para esos pequeños cuerpos que hay dentro de las células, más tarde conocidos como mitocondrias” (p. 97). Específicamente, el citólogo y anatomista canadiense Cowdry (1918), en un recorrido histórico sobre el estudio de estas organelas, plantea que, desde su primera clasifica-

ción en el año 1874, fueron reclasificadas setenta y una veces en un lapso de treinta años (pp. 44-45).

Así, la calidad de las observaciones al microscopio óptico, colaboró quizás en el desarrollo de un lenguaje poco preciso sobre la naturaleza misma de los orgánulos, que a su vez, pudo haber influido en la forma de concebir el papel evolutivo que podrían haber representado las bacterias en el origen de la célula eucariota.

## **b) Período de latencia**

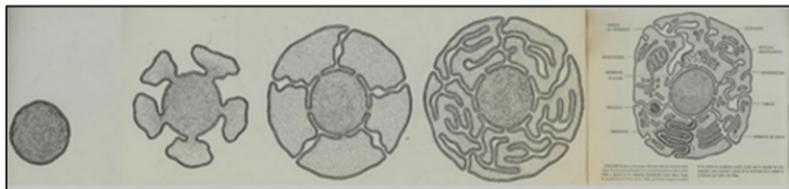
Las décadas del treinta al cincuenta del siglo XX, fueron una época de gran desarrollo de la teoría sintética de la evolución (Salgado y Arcucci, 2016). A partir de ella “[...] los arquitectos clásicos de la llamada síntesis evolucionista, habían convenido desde hace tiempo que la mutación y la recombinación genética dentro de las especies eran las únicas fuentes de innovación evolutiva por selección natural” (Sapp, 2014, p. 94). En este contexto, poco propicio para el desarrollo de visiones alternativas sobre la evolución, no se han encontrado publicaciones relacionadas con la endosimbiosis.

A medida que la teoría sintética de la evolución continuaba desarrollándose, los avances tecnológicos en microscopía electrónica y las investigaciones en genética molecular, abrieron nuevos caminos que dejaron en relieve nuevas estructuras subcelulares que interpretar y nuevos problemas que la teoría debía resolver. En este contexto, uno de los ejemplos más claros que representa la influencia del programa darwiniano (adaptacionista/gradualista) en la explicación evolutiva del origen de la célula eucariota, es la hipótesis teórica formulada por Robertson (1965/1970), autor del modelo de “unidad de membrana” que se observa en la figura 1, la que explica el origen de las mitocondrias por invaginaciones sucesivas de la membrana.

## **c) Período de resurgimiento**

Durante los años sesenta y setenta del siglo XX se discutieron ideas sobre endosimbiosis en más de cincuenta artículos científicos (Sapp, 2014).

El resurgimiento se vio impulsado, básicamente, por una corriente crítica de las posturas evolutivas “nucleocéntricas” predominantes, las



**Fig. 1:** Origen de organelas por invaginación gradual de membranas.

*Nota:* Imagen de extraída de “La membrana de la célula viva” de D. Robertson. En J.R. Villanueva, (Ed.), *La célula viva. Selecciones de Scientific American*. Madrid: Blume. 1970, p. 82.

efectuaban jóvenes biólogos estadounidenses y eran afines a la revalorización de las ideas formuladas en las publicaciones del primer período (Sampedro, 2007). Ahora bien, el contexto se modificó, y el nuevo terreno en el cual se buscarían pruebas y también en el cual se dirimirían controversias fue el de la biología y la genética molecular.

el punto de inflexión para la aceptación de la simbiosis en la evolución celular fue a principios de los sesenta, cuando se hallaron moléculas de ADN y ribosomas en las mitocondrias y los cloroplastos, y se examinaron los orgánulos al microscopio electrónico. (Sapp, 2014, p. 84)

La presencia de ADN en mitocondrias y cloroplastos, implicó a biólogos del campo de la teoría sintética en la necesidad de profundizar los modelos de evolución de la célula eucariota, discutir el supuesto origen bacteriano de esas organelas y formular explicaciones que asimilaran ese hecho al marco teórico de referencia. En esta línea, Max Taylor de la Universidad Británica de Columbia, propuso un modelo que denominó “filiación directa” (Taylor, 1974, 1976), por el cual, la presencia de ADN de mitocondrias y cloroplastos, “[...] evolucionaron sin mediación de la simbiosis, [...] todos ellos surgieron como pellizcos del ADN del núcleo” (Margulis, 2002, p. 96).

La controversia sobre el ADN presente en las organelas, se dirimió en favor del modelo endosimbiótico, primero, porque a inicios de la década del sesenta se determinó su parecido con el ADN bacteriano, y luego, cuando el microbiólogo estadounidense Carl Woese (1928-2012) logró establecer que el ARN ribosómico contenido en cloroplastos y mitocondrias tenía un origen alfa-proteobacteriano y cianobacteriano respectivamente (Margulis, 2002; Sapp, 2014).

En el año 1967, la estructuración del modelo encuentra un punto de inflexión con la publicación de Margulis de *On the origin of mitosing cells* y, a partir de allí, del desarrollo de un programa de trabajo que unificó conocimientos provenientes de diferentes campos, los articuló y apoyó con una serie de datos morfológicos, bioquímicos, genéticos e incluso geológicos que permitieron consolidar otro enfoque explicativo del origen de la célula eucariota. En la actualidad, y luego de transcurridos casi 50 años, el análisis de los manuales para la formación de biólogos, depositarios de los modelos de ciencia normal, da como resultado una coexistencia de modelos para explicar el origen de la célula eucariota. Mitocondrias y cloroplastos son reconocidas como organelas endosimbióticas y otras estructuras celulares como retículo endoplásmico y aparato de Golgi, se proponen como el resultado de invaginaciones graduales de membrana.

### **Implicancias para la formación en el eje “Naturaleza de las ciencias”**

La perspectiva Naturaleza de la ciencia constituye una línea de investigación y un área de enseñanza, integra aspectos de filosofía, historia y sociología de la ciencia y tiene como objetivo desarrollar la educación metacientífica de los y las estudiantes y lograr que las imágenes que sobre la ciencia circulan en la sociedad sean más actualizadas y críticas (Lederman, 1992; McComas, 1998). En este campo existe un amplio debate de perspectivas respecto de cuáles son las temáticas metacientíficas que deberían enseñarse y el modo de organizarlas (Gilbert & Justi, 2016). Para nuestra investigación adoptamos una propuesta que se orienta a la construcción de entidades metateóricas específicas (Adúriz-Bravo & Ariza, 2012), necesarias para estructurar un campo metacientífico que sea fecundo a los estudiantes del profesorado y que potencie y haga más significativa la formación disciplinar y didáctica que lleven a cabo. El modo de especificar los aspectos a enseñar es mediante la formulación de ideas clave metacientíficas, afirmaciones sencillas pero muy significativas para los y las estudiantes ya que constituyen el aspecto específico de la ciencia que se pretende abordar en cada campo. Desde esta perspectiva y de manera preliminar, al analizar los resultados obtenidos en el estudio consideramos que:

- La influencia que tuvo la teoría microbiana de la enfermedad en el desarrollo de la teoría endosimbiótica y la formulación de nuevos problemas para su desarrollo bajo el paradigma molecular de la biología

en el tercer período, son episodios que pueden implicarse en el desarrollo de una idea clave sobre “paradigma”.

- Las discusiones sobre el origen del ADN de mitocondrias y cloroplastos y la formulación del modelo alternativo propuesto por Taylor, son episodios que pueden implicarse en el desarrollo de ideas claves sobre: la “carga teórica de las observaciones” y “la construcción de los hechos científicos”.

Un posterior análisis más profundo permitirá identificar nuevos episodios e implicar otras entidades metateóricas a enseñar en el desarrollo de unidades didácticas.

## Referencias

- Adúriz-Bravo, A. (2013). A ‘semantic’ view of scientific models for science education. *Science & Education*, 22(7), 1593-1611.
- Adúriz-Bravo, A. (2012). Aproximaciones histórico-epistemológicas para la enseñanza de conceptos disciplinares. *Revista Virtual EDUCyT*, 1(1), 107-126.
- Adúriz-Bravo, A. & Ariza, Y. (2012). Importancia de la filosofía y la historia de la ciencia en la enseñanza y el aprendizaje de las ciencias. En Z. Monroy Nasr, R. León-Sánchez & G. Álvarez Díaz de León (comps.), *Enseñanza de la ciencia* (pp. 79-92). México: DGAPA.
- Arcucci, A. & Salgado, L. (2016). *Teorías de la evolución: Notas desde el sur*. Argentina: Editorial UNRN.
- Baker, J. R. (1952). The cell-theory: a restatement, history, and critique. *Quarterly Journal of Microscopical Science*, 93 (2): 157-190.
- Cowdry, E. V. (1918). *The mitochondrial constituents of protoplasm*. Carnegie Institution of Washington.
- Giere, R. (1992). *La explicación de la ciencia: Un acercamiento cognoscitivo* (C. Gidi, trad.). México: Consejo Nacional de Ciencia y Tecnología. (Obra original de 1988)
- Harris, H. (2000). *The birth of the cell*. New Haven, CT: Yale University Press.

- Kuhn, T. S. (1996). *La tensión esencial* (R. Helier, trad.). México: Fondo de Cultura Económica. (Obra original de 1977)
- Lazcano A. (2002). El origen del nucleocitoplasma. Breve historia de una hipótesis cambiante. En Margulis, L. (Ed.), *Una revolución en la evolución* (pp. 182-183). Valencia: Universitat de Valencia.
- Lozano, E., Bahamonde, N. & Adúriz-Bravo, A. (2016). Análisis histórico-epistemológico sobre los modelos de membrana celular para enseñar biología celular y naturaleza de la ciencia al profesorado. *Filosofía e Historia da Biología, São Paulo*, 11(1), 49-68.
- Lozano, E., Bahamonde, N., Cremer C. & Mut, P. (2018). El desarrollo de una línea metacientífica para la enseñanza del modelo de presión arterial en la formación del profesorado en biología. *Revista Electrónica de Enseñanza de las Ciencias*, 17 (3), 564-580
- Margulis, L. (2002). *Una revolución en la evolución* (Vol. 20). Valencia: Universitat de València.
- Margulis, L. & Olendzenski, L. (1996). *Evolución ambiental: efectos del origen y evolución de la vida sobre el planeta Tierra*. España: Alianza editorial.
- Matthews, M. (1994). Historia, filosofía y enseñanza de las ciencias: la aproximación actual. *Enseñanza de las ciencias: revista de investigación y experiencias didácticas*, 12 (2), 255-278.
- Orozco, J. C. (1992). Consideraciones para un enfoque histórico-epistemológico de la enseñanza de las ciencias. *Tecné Episteme y Didaxis TED*, 7, 22-35.
- Robertson, D. (1970). La membrana de la célula viva. En Villanueva, J. R (Ed.), I. García Acha, M. Losada & C. Fernández Heredia (trads.), *La célula viva* (pp. 75-83). *Selecciones de Scientific American*. Madrid: Blume. (Obra original de 1965)
- Saap, J. (2014). Demasiado fantástica para la sociedad formal. Una breve historia de la teoría simbiótica. En D Sagan. (Ed.), *Lynn Margulis. Vida y legado de una científica rebelde* (pp. 79-96). Barcelona: Tusquets Editores.

- Sampedro, J. (2002). *Deconstruyendo a Darwin*. Argentina: Ed. Drakontos bolsillo.
- Taylor, F. J. R. (1974). II. Implications and extensions of the serial endosymbiosis theory of the origin of eukaryotes. *Taxon*, 23(2-3), 229-258.
- Taylor, F. J. R. (1976). Autogenous theories for the origin of eukaryotes. *Taxon*, 377-390.
- Wallin, I. E. (1927). *Symbiogenesis and the origin of species*. Baltimore: Waverly Press.



# El modelo atómico de Lewis

## Consideraciones epistemológicas y aplicaciones didácticas

Mercedes Barquín<sup>\*,\*\*</sup>

Guadalupe Quiñoa<sup>\*,\*</sup>

Alfio Zambon<sup>\*</sup>

### Introducción

Uno de los problemas principales que enfrentamos en la práctica docente de la química, tanto en el nivel secundario como universitario, es la falta de cohesión entre los temas que forman parte de la currícula. Temas como: estructura atómica, configuración electrónica, enlace químico, tabla y propiedades periódicas; aparecen dispersos y con pocos nexos de unión.

En ese contexto, un caso histórico como el modelo atómico desarrollado por Gilbert N. Lewis en las primeras décadas del siglo pasado, puede resultar una herramienta valiosa para relacionar los diferentes temas, y además posibilita un espacio para la reflexión acerca de la naturaleza del conocimiento científico en química, dado las interesantes aristas epistemológicas que presenta.

En este trabajo, presentaremos el modelo atómico de Lewis y resaltaremos las consideraciones epistemológicas que muestran aplicaciones didácticas. Para ello, en primer lugar, expondremos algunas características salientes de los modelos en química, presentaremos brevemente algunos modelos atómicos que suelen destacarse en los cursos. Luego revisaremos el desarrollo del concepto histórico de enlace y nos focalizamos en el modelo atómico de Lewis, para finalmente discutir las ventajas y limitaciones de la propuesta.

\* Universidad Nacional de la Patagonia San Juan Bosco, Comodoro Rivadavia, Chubut.

\*\* Colegio Deán Funes, Comodoro Rivadavia, Chubut.

\* Colegio Abraham Lincoln, Comodoro Rivadavia, Chubut.  
azambon@infovia.com.ar

## Los modelos en química

Tradicionalmente los modelos en ciencia permiten articular el conocimiento científico con la representación de ese conocimiento. En las últimas décadas los modelos han ganado preeminencia sobre las teorías en filosofía de la ciencia (Lombardi, Accorinti, & Martínez, 2016). El objetivo de este trabajo no es reflexionar acerca de la noción de modelo en general, sino emplear la noción de una manera utilitaria para reflexionar acerca de problemas puntuales en la educación en química, pero creemos importante presentar de manera sucinta algunas consideraciones desarrolladas en ese fructífero campo de investigación.

Adúriz Bravo (2012) realiza una revisión de las distintas posiciones en el campo de la didáctica de las ciencias con aplicación a la química de la noción de modelo. Presenta una discusión detallada de aspectos epistemológicos y didácticos de diferentes posturas, reconoce tres puntos de partida sobre los que hay un cierto consenso en la comunidad de la didáctica de la química: 1) si bien las nociones de modelo y modelización han estado presentes de manera implícita en los currículos de todo los niveles educativos, en estos últimos años la presencia de los mismo se trata de manera más explícita, 2) la producción didáctica en torno a la noción ha alcanzado niveles de detalle y sofisticación importantes, pero aún quedan discusiones acerca de aspectos básicos, y 3) parece que está emergiendo una nueva forma de entender los modelos desde la didáctica de las ciencias. Posteriormente analiza las relaciones entre los modelos, las analogías, las teorías y la mediación.

En tanto, Chamizo (2006) expone las caracterizaciones generales de la naturaleza de los modelos científicos, y resalta su importancia en todos los niveles educativos. Este autor identifica las ocho características “menos controvertidas” de los modelos científicos: 1) Los modelos son representaciones de objetos, sistemas, fenómenos o procesos, de manera que un modelo siempre es un modelo de algo. En tanto, las representaciones simplifican lo representado y facilitan su comprensión. 2) Son instrumentos para intentar responder las preguntas científicas. 3) Presentan alguna analogía con los fenómenos que representan, y permiten derivar hipótesis susceptibles de ser puestas a prueba. 4) Son más simples que la realidad que buscan representar. 5) Los modelos se pueden ampliar y corregir. 6) Se desarrollan a lo largo de la historia, en un proceso de continua revisión.

sión para ordenar la nueva evidencia empírica. 7) Deben ser aceptados y consensuados por la comunidad científica. Y, finalmente, 8) se pueden clasificar en tipos de modelos: icónicos y conceptuales.

## **Los primeros modelos atómicos**

A partir del descubrimiento de las partículas subatómicas, se intentó dilucidar cómo esas partículas interactuaban entre sí, lo cual motivó a muchos científicos a proponer “modelos atómicos”. Entre los primeros modelos están los propuestos por Philip Lenard y Hantaro Nagaoka (Solís & Selles, 2004). Joseph J. Thomson también planteó un modelo atómico muy destacado. Propuso imaginar al átomo como una esfera de materia sólida que tenía carga positiva, en cuyo interior estaban los electrones dispuestos de manera de formar una serie de puntos de carga negativa. Este modelo fue conocido como “pastel de pasas” (Sánchez Ron, 1992).

Entre los primeros modelos atómicos, el primero en ofrecer evidencia empírica en su favor fue el formulado por Ernest Rutherford en 1911. A partir de los resultados de su investigación, infirió que el átomo consistía en un núcleo de carga positiva. Habría igual número de electrones que protones, los que estarían girando alrededor del núcleo de manera análoga a los planetas alrededor del Sol (Sánchez Ron, 1992).

Las dificultades teóricas planteadas por el modelo atómico de Rutherford encontraron una vía de resolución mediante la aplicación, propuesta por Niels Bohr, de la nueva teoría cuántica al sistema atómico. El principal problema que lo preocupaba era el de la inestabilidad del modelo de Rutherford. El punto de partida de su trabajo no fue en absoluto la idea de que un átomo fue una especie de sistema planetario a pequeña escala, imagen que Bohr nunca se tomó textualmente (Pullman, 1995/2010), sino el problema de la estabilidad de la materia.

En 1913 Bohr postuló un nuevo modelo de la estructura atómica mediante tres artículos. En el primero propuso un modelo atómico para el átomo de hidrógeno, tomando como base: el modelo atómico “planetario” de Rutherford y las líneas espectrales descritas por la ecuación de Balmer. Consideró que para poder mantener la estabilidad del sistema, se debían cumplir tres postulados: 1) el electrón en los átomos monoeléctricos podrá solamente “girar” en algunas órbitas circulares de energía y radio determinadas, 2) No todas las órbitas serán posibles; sólo estarán permitidas aquellas que cumplan con un conjunto de valores discretos, 3)

El electrón sólo podrá emitir energía en los pasajes de una órbita de mayor a una de menor nivel energético, y absorberá energía cuando el pasaje se verifique en sentido contrario. Se puede decir que el modelo de Bohr es un modelo “híbrido”, ya que se basaba en una alianza muy extraña entre las concepciones y las fórmulas de la dinámica clásica y los métodos de la teoría cuántica. Esta conjugación teórica un tanto forzada fue objeto de críticas en algunos casos rigurosas, como las brindadas por de Broglie o Schrödinger, entre otros (Pullman, 1995/2010).

## El concepto de enlace durante el siglo XIX

Las primeras nociones de que el enlace químico era de naturaleza eléctrica, fueron resultado de los experimentos de William Nicholson y Antony Carlisle quienes lograron descomponer el agua en hidrógeno y oxígeno mediante el paso de corriente eléctrica (electrólisis). En 1807 Humphry Davy emplea la electrólisis para diferenciar los elementos de los compuestos, descubrió dos nuevos elementos, sodio y potasio.

En 1802, Jöns Berzelius descubrió que las sales alcalinas se descomponían en ácidos y bases durante la electrólisis. En 1819 propone su teoría electroquímica y su sistema dualístico (materia-electricidad). Según Berzelius la electricidad era el *primun movens* de todos los procesos químicos: los átomos de los diversos elementos eran considerados por él como dipolos eléctricos con una carga predominantemente positiva o negativa, salvo el hidrógeno que era neutro. La combinación de los elementos se daría debido a su diferente carga eléctrica y distribución de carga, pero esta unión no necesariamente produciría la neutralización de las cargas. Previamente, el sueco Torben O. Bergman y el francés Claude Bertholet, atribuyeron la estabilidad del enlace químico a la fuerza gravitacional que actúa entre las partículas que lo forman. Sin embargo, resultó que la afinidad química no correspondía a las masas de los átomos que se unen para formar moléculas (Cruz-Garriz, Chamizo, & Garriz, 1991).

En 1852, Edward Frankland desarrolló el concepto de atomicidad, que equivale a lo que hoy se conoce como valencia, investigando sobre las propiedades de compuestos organometálicos. Sugirió que cada elemento formaba compuestos con cantidades definidas de otros elementos, ya que cada uno tiene una “atomicidad” fija, considerando esta como la capacidad de combinación de los átomos de cada elemento, en comparación con la del hidrógeno que siempre es una. Es importante anotar que Frankland

nombró como elementos electro-negativos al cloro, bromo y yodo, así como al etilo y otros radicales (a cuyas partículas nombraba como átomos) y como electro-positivos a los elementos metálicos, aunque no esbozó ninguna definición de estos términos (Gallego Badillo et al., 2004). En 1868 C.W. Wichelhaus introduce el término valencia o capacidad de combinación, pero el “remate” respecto a la veracidad de la relación entre la valencia y la estructura molecular se alcanza con Jacobus van 't Hoff y Joseph Le Bel, quienes en 1874 resuelven el enigma de la isomería óptica descubierta por Louis Pasteur, al proponer que la tetravalencia del carbono tiene que ver con su geometría tetraédrica (Garritz, 1997).

El descubrimiento del electrón por Thomson en 1897 proporcionó la base de la teoría electrónica del enlace, en tanto el trabajo magistral de Alfred Werner sobre la naturaleza de los compuestos de coordinación, gracias también al estudio de su conductividad eléctrica en disolución, lo llevó a considerar dos tipos de valencia para ellos: la *Hauptvalenz* o valencia principal, considerada idéntica con lo que hoy conocemos como número o estado de oxidación del metal central, y la *Nebervalenz*, o valencia auxiliar, hoy reconocida como el enlace coordinado entre el metal y sus ligantes.

## El concepto de enlace durante el siglo XX

En 1904 Richard W. H. Abegg describe la “regla de ocho” donde propone que los elementos mostrarían una valencia electropositiva máxima (valencia normal) y una valencia electronegativa máxima (contravalencia) en el que la suma de las dos valencias siempre es igual a ocho. En ese mismo año Thomson propuso la primera teoría electrónica de valencia, en la que el enlace químico sería una atracción electrostática (Scerri, 2016), y en 1907 publica *The corpuscular theory of matter* en la que propone que la unión entre átomos en un compuesto es causada por la transferencia de electrones. Esta teoría es respaldada por George Falk y John Nelson quienes en 1910 publican “The electron conception of valence” e introducen el uso de flechas en los enlaces, sugeridas por Thomson, para indicar la dirección de transferencia (concepto de enlace vectorial) (Brock, 1993/1998).

La aplicación de la teoría a los compuestos de química orgánica o compuestos no polares condujo a serias dificultades. En 1913 William C. Bray y Gerald E. K. Branch publican “Valence and tautomerism” donde introducen los términos “números de valencia total” y “números polares”



(Brock, 1993/1998). El primer número se encontraría solo en la química orgánica y el segundo en la química inorgánica elemental. Esta publicación le permitió a Lewis realizar un comentario acerca de la valencia y señaló una nueva distinción, entre los enlaces no polares y polares:

En el primer tipo, los electrones ocupan posiciones fijas dentro del átomo. En el segundo tipo, los electrones se mueven libremente de átomo a átomo dentro de la molécula. En el tercer tipo o metálico, el electrón es libre de moverse incluso fuera de la molécula. (Shaik, 2007, p. 53)

Thomson, en 1914, publica “Forces between atoms and chemical affinity” donde concluye que hay dos tipos enlaces: polares y no polares (Shaik, 2007). Mediante cálculos matemáticos se determinó que en el enlace participaban dos líneas de fuerzas, o dos electrones. Ese mismo año, Alfred Parson sugirió que el enlace podía ser magnético y no eléctrico; de ese modo el electrón se convertía en electroimán. En la teoría del “magnetón”, los magnetones no se disponían en esferas sino en las esquinas de un cubo que rodeaba una carga positiva central (Brock, 1993/1998).

En 1916, Walter Kossel publicó una teoría del enlace polar, la cual era una teoría electrostática del enlace químico, actualizada con el concepto del electrón. Observó la “regla de ocho” al organizar los electrones en grupos concéntricos, cada anillo de electrones tiene un máximo estable de ocho, excepto en el primer caso cuando el máximo es dos (Nye, 1994). Ese mismo año Lewis publica “The atom and the molecule” donde introduce la teoría del par de electrones compartido (Brock, 1993/1998) (véase más adelante).

En 1919 Irving Langmuir en el *Journal of the American Chemical Society* publica un artículo donde dio los nombres “covalente” y “electrovalente” a los enlaces químicos que se distinguen por la partición del par de electrones entre dos átomos, extendiendo conscientemente la teoría de Lewis (Nye, 1994).

En 1923 Lewis publica *Valence and the structure of atoms and molecules* (Brock, 1993/1998) (véase más adelante).

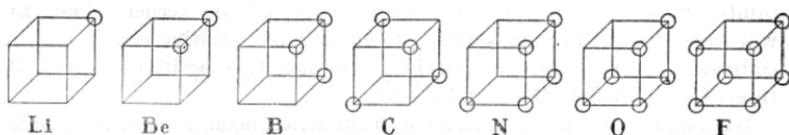
En 1925 Werner Heisenberg presenta la mecánica matricial y en 1926 Erwin Schrödinger, la mecánica ondulatoria. En 1927, Walter Heitler y Fritz London publican la teoría del enlace de valencia y Friedrich Hund propone el método de orbitales moleculares, en simultaneo que Robert S. Mulliken. En 1932 Linus Pauling publicó una serie de artículos “The nature of chemical bond”, en los que introduce conceptos como re-

sonancia, electronegatividad relativa, enlace iónico y enlace covalente, sus aportes quedan plasmados en un libro publicado en 1939 (Garritz, 2014). El modelo de Repulsión de los pares electrónicos de la capa de valencia fue descrito por R. J. Gillespie y R. S. Nyholm en 1957. A lo largo del siglo XX se siguieron sumando aportes a la noción de enlace químico y la noción del mismo continúa en estudio y discusión como lo muestra Ball (2011).

## El modelo atómico de Lewis

En 1916 se publicaron dos trabajos sumamente importantes acerca de la teoría electrónica de valencia. El primero, de Walter Kossel, fue un intento explícito de desarrollar y extender la teoría de Abegg. Kossel desarrolló una gráfica del total de electrones por átomo o ion, versus su número atómico, para los 57 primeros elementos. El segundo artículo fue el de Lewis, quien retomó los aspectos fundamentales de su trabajo de 1902, y llegó a un modelo dualista de transferencia de electrones (Jensen, 1984). De manera general, la propuesta de Lewis consistió en un modelo de átomo cúbico, de naturaleza estática, con los electrones ubicados en los vértices (Partington, 1937/1945), de manera de poder formar pares de unión con otro átomo (Fig. 1).

En la propuesta de Lewis, a medida que se avanzaba en cada periodo de la tabla periódica, se agregaba un electrón a uno de los vértices, observa Scerri (2007) que la elección de un cubo puede parecer extraña desde la perspectiva actual, donde la idea que los electrones circulan en derredor del átomo está muy extendida, pero el modelo de Lewis tenía un aspecto muy importante en el desarrollo de la tabla periódica, ya que sugería que la periodicidad química y las propiedades de los distintos elementos se rigen por el número de electrones que se ubicaban en cada cubo (Scerri, 2007).



**Fig. 1:** El átomo cubico de Lewis.

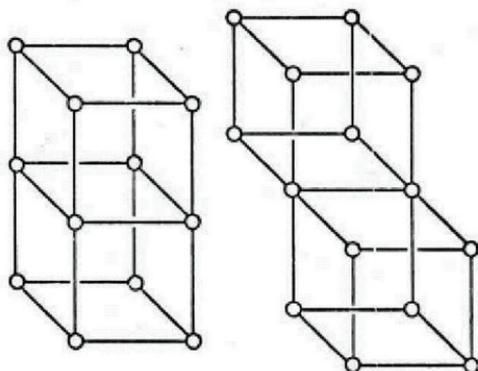
*Nota:* Imagen extraída de “The atom and the molecule” de G.N. Lewis, *Journal of the American Chemical Society*, 38(4), 1916, p. 767.

Para Lewis, las capas atómicas exteriores son mutuamente interpenetrables, de modo que un electrón puede formar parte de una misma capa exterior de dos átomos diferentes y no podría afirmarse su pertenencia a ninguno de los dos de forma exclusiva. Sin la concepción de Lewis del vínculo de par de electrones compartido, que luego fue utilizado con éxito por los químicos orgánicos, la interpretación de los mecanismos de reacción no habría podido desarrollarse. Del mismo modo, sin la idea del enlace de par compartido, la aplicación de la mecánica cuántica al enlace químico habría comenzado desde un terreno mucho menos seguro (Kohler, 1971).

Lewis, de manera similar a Thomson, desarrolló un modelo estático del átomo y a partir de ese modelo justificó el concepto de repetición periódica de la valencia en términos electrónicos y las particulares estabildades de los gases nobles. El concepto del enlace del par de electrones compartido era el aspecto saliente del trabajo de Lewis. La suposición que para mantener unidos los átomos en las moléculas, era fundamental la presencia de un par de electrones conjuntamente retenido por dos átomos, proporcionó el enlace faltante para darle consistencia al comportamiento experimental, y refutó la creencia de que todos los tipos de unión química son esencialmente el mismo (Brock, 1993/1998).

Al igual que sus predecesores, Lewis había reconocido tanto la importancia del número ocho en las relaciones de valencia y su correlación con las estructuras de los gases nobles. Sin embargo, a diferencia de otras propuestas como el modelo iónico, formulado por Kossel (donde esta correlación sólo era posible en el caso de estados de máxima oxidación), Lewis presentó un modelo de mayor amplitud, donde se resaltaba la importancia del par de electrones compartido, lo que permitía la extensión de la regla del ocho a especies en estado de oxidación inferior, así como a especies homonucleares como el  $\text{Cl}_2$  y el  $\text{N}_2$  (Jensen, 1984). En el modelo de Lewis tenían cabida la representación de los dobles y triples enlaces, mediante la unión de una arista o una cara respectivamente (Fig. 2), en tanto el modelo se limitaba a los elementos de los bloques *s* y *p* de la tabla periódica (Scerri, 2007).

Langmuir, en su primer y más ambicioso artículo, intentó remediar este defecto extendiendo el modelo para los elementos del bloque *d*. Mediante el uso de 11 postulados básicos que justificaban la disposición de los electrones en átomos aislados y su participación preferencial en compuestos (Fig. 3). Langmuir buscó deducir toda la estructura teórica y des-



**Fig. 2:** Representación de los dobles y triples enlaces en el modelo de Lewis.

*Nota:* Imagen extraída de “Abegg, Lewis, Langmuir and the octet rule” de W. B. Jensen, *Journal of Chemical Education*, 61, 1984, p. 194.

criptiva de la química a partir del modelo del átomo cúbico (Jensen, 1984).

A pesar de su entusiasmo, la extensión de la propuesta del átomo estático de Lewis-Langmuir demostró ser de corta duración. Para 1921, tanto Bohr sobre la base de la evidencia espectral, como Charles Bury, sobre la base de evidencia química, habían demostrado que el postulado de Langmuir sobre la secuencia de llenado de las capas electrónicas tenía incorrecciones, en especial el postulado de que sólo podría iniciarse una nueva capa de electrones si las anteriores estaban completamente llenas. Además, la mayoría de los físicos teóricos para este momento estaban firmemente comprometidos con un modelo de átomo dinámico, y rechazaban el modelo estático extendido de Langmuir (Jensen, 1984).

Por lo tanto, en la práctica real, el desarrollo del modelo de Lewis realizado por Langmuir, estaba restringido a los mismos átomos que Lewis había trabajado originalmente. Tenía varias limitaciones intrínsecas, sin embargo, varios conceptos valiosos pudieron evolucionar a partir de los refinamientos de Langmuir, como los principios de electroneutralidad e isoelectrónico (isostérico), que han sido caracterizados como “la última mecánica importante no cuántica respecto del enlace químico” (Jensen, 1984, p. 196). También fue importante para la discusión actual en química, la propuesta de Langmuir del uso de postulados para deducir de manera matemática la expresión de la regla del “octeto” (Fig. 4).

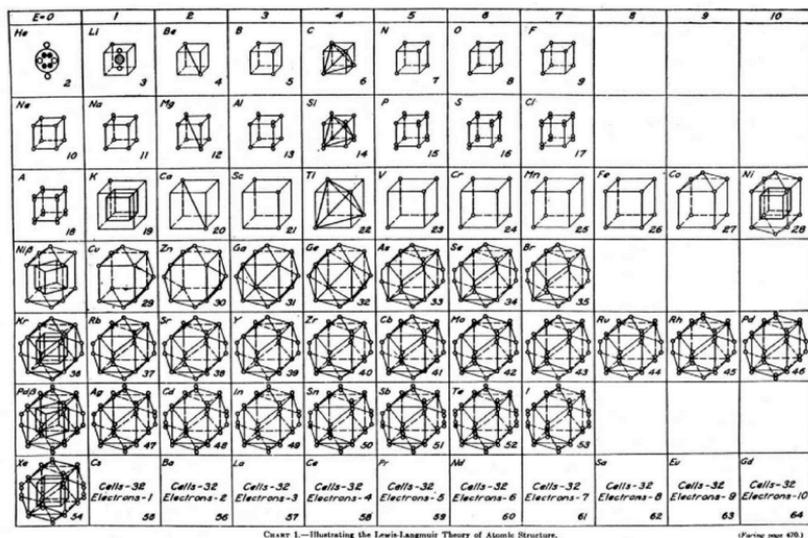


Fig. 3 : Extensión del modelo cúbico propuesta por Langmuir.

Nota: Imagen extraída de “Abegg, Lewis, Langmuir and the Octet Rule” de W. B. Jensen, *Journal of Chemical Education*, 61, 1984, p. 196.

Table 4. Examples of Langmuir’s Octet Equation

Species	$\sum e$	$n$	$H$	$\nu = \frac{1}{2}(8n - \sum e) + H$	Structure
NH <sub>3</sub>	8	1	3	3	$\begin{array}{c} \text{H} \\   \\ \text{N} - \text{H} \\   \\ \text{H} \end{array}$
CO	10	2	0	3	$ \text{C} \equiv \text{O} $
H <sub>2</sub> CO	12	2	2	4	$\begin{array}{c} \text{H} \\   \\ \text{C} = \text{O} \\   \\ \text{H} \end{array}$
NaCl	8	1	0	0	$\text{Na}^+ [ \bar{\text{C}}\bar{\text{I}} ]^-$
CS <sub>2</sub>	16	3	0	4	$(\text{S} = \text{C} = \text{S})$

Fig. 4: Estructuras obtenidas mediante la ecuación de Langmuir.

Nota: Imagen extraída de “Abegg, Lewis, Langmuir and the octet rule” de W. B. Jensen, 1984, *Journal of Chemical Education*, 61, p. 196.

## **Discusión y reflexiones finales**

El aprendizaje es un proceso en constante cambio, construcción y reelaboración conceptual, donde los nuevos conocimientos interactúan y se integran a los conocimientos previos. El aprendizaje sucede cuando estos nuevos conocimientos adquieren sentido y significado. El modelo atómico de Lewis, que integra naturalmente el enlace químico en la concepción del átomo, permite integrar conceptos tales como: estructura atómica, configuración electrónica y tabla periódica, entre sí y con algunas propiedades macroscópicas de la materia. De esa manera, lo cotidiano para el alumno, se puede relacionar con las elaboraciones científicas, conformando un nexo significativo entre lo abstracto y lo concreto, lo que representa un estímulo para el análisis crítico y el desarrollo del pensamiento científico.

En el nivel secundario, el eje temático enlaces y uniones químicas se estudia en 4to o 5to año, dependiendo de la orientación de la escuela. Entre los contenidos a enseñar, se encuentran: parámetros de la estructura molecular, enlaces intramoleculares e intermoleculares, aspectos energéticos y estructura cristalina, enlaces covalentes e iónicos, polaridad molecular, disociación iónica, geometría molecular, teoría de los orbitales moleculares y fuerzas intermoleculares. Estos son temas de difícil comprensión por parte de los alumnos, y en el proceso muchos recurren al aprendizaje memorístico, por ejemplo, el concepto de enlace covalente “dativo” o coordinado es complejo de comprender y en muchas ocasiones se recurre al empleo de metáforas antropocéntricas, como quitar, dar, sacar, prestar, entre otras. El empleo de aristas o caras de cubos que interactúan, representa un aporte clarificador en ese sentido.

En la práctica docente el empleo del concepto de modelo es muy útil y particularmente fértil, ya que permite trabajar con representaciones que pueden hacer posible transformar el contenido científico en un saber enseñable (transposición didáctica). El modelo de Lewis es un modelo icónico ya que es una representación de un concepto abstracto no observable. Permite, a través de la utilización de pictogramas, construir y comprender lo que sucede en el enlace químico. También llegar al concepto de la regla del octeto. Es así que el modelo atómico de Lewis muestra una idea intuitiva del concepto de enlace, que ayuda a entender y profundizar el mismo.

El modelo de Lewis, abordado desde un punto de vista histórico, permite presentar el conocimiento químico como un saber en construcción, y resaltar que los conceptos que hoy se enseñan, fueron motivo de discusión, de estudio, de intentos de refutación científica y de una construcción



colectiva, y conforma un espacio donde se resalta la importancia de la historia de la ciencia en la enseñanza de la ciencia.

## Referencias

- Adúriz Bravo, A. (2012). Algunas características clave de los modelos científicos relevantes para la educación química. *Educación Química*, 23(2), 248-256.
- Ball, P. (2011). Beyond the bond. *Nature*, 469(7328), 26-28.
- Bray, W. C., & Branch, G. E. (1913). Valence and tautomerism. *Journal of the American Chemical Society*, 35(10), 1440-1447.
- Brock, W. (1998). *Historia de la química* (I. Medina, P. Burgos, A. del Valle & E. García Hernández, trads). Alianza Editorial. (Obra original de 1993)
- Chamizo, J. A. (2006). Los modelos de la química. *Educación Química*, 17(4), 476-282.
- Cruz-Garriz, D., Chamizo, J. A., & Garriz, A. E. A. (1991). *Estructura atómica: Un enfoque químico*. Editorial Addison-Wesley Iberoamericana.
- Falk, K. G., & Nelson, J. M. (1910). The electron conception of valence. *Journal of the American Chemical Society*, 32(12), 1637-1654.
- Gallego Badillo, R., Pérez Miranda, R., Uribe Beltrán, M. V., Cuéllar Fernández, L., & Amador Rodríguez, R. Y. (2004). El concepto de valencia: su construcción histórica y epistemológica y la importancia de su inclusión en la enseñanza. *Ciência & Educação (Bauru)*, 10(3), 571-583.
- Garriz, A., & Rincón, C. (1997). Valencia y número de oxidación. Corolario para docentes. *Revista Educación Química*, 8(3).
- Garriz, A. (2014). Historia de la química cuántica. *Educación en química*, 25(1), 170-175.
- Jensen, W. B. (1984). Abegg, Lewis, Langmuir and the octet rule. *Journal of Chemical Education*, 61, 191-200.
- Kohler, R. E. (1971). The origin of G. N. Lewis's theory of the shared pair bond, *Historical Studies in the Physical Sciences*, 3, 343-376.

- Kossel, W. (1916). Über molekülbildung als Frage des Atombaus. *Annalen der Physik*, 354(3), 229-362.
- Lewis, G. N. (1916). The atom and the molecule. *Journal of the American Chemical Society*, 38(4), 762-785.
- Lewis, G. N. (1923). *Valence and the structure of atoms and molecules* (No. 14). Chemical Catalog Company, Incorporated.
- Lombardi, O., Acorinti, H., & Martínez González, J. C. (2016). Modelos científicos: el problema de la representación. *Scientiae Studia*, 14(1), 151-174.
- Nye, M. J. (1994). *From chemical philosophy to theoretical chemistry: dynamics of matter and Dynamics of disciplines, 1800-1950*. University of California Press.
- Partington, J. R. (1945). *Historia de la Química* (C.E. Prelat, trad.). Editorial Espasa Calpe. (Obra original de 1937)
- Pauling, L. (1932). The nature of the chemical bond. III. The transition from one extreme bond type to another. *Journal of the American Chemical Society*, 54(3), 988-1003.
- Pauling, L. (1932). The nature of the chemical bond. IV. The energy of single bonds and the relative electronegativity of atoms. *Journal of the American Chemical Society*, 54(9), 3570-3582.
- Pullman, B. (2010). *El átomo en la historia de la humanidad* (J. Sarret, trad.). Editorial Biblioteca Buridán. (Obra original de 1995)
- Ron, J. M. S. (1993). *Espacio, tiempo y átomos: Relatividad y mecánica cuántica* (Vol. 51). Ediciones AKAL.
- Shaik, S. (2007). The Lewis legacy: the chemical bond—a territory and heartland of chemistry. *Journal of computational chemistry*, 28(1), 51-61.
- Scerri, E. R. (2016). *A tale of seven scientists and a new philosophy of science*. New York: Oxford University Press.
- Scerri, E. R. (2007). *The periodic table: Its story and its significance*. New York: Oxford University Press.
- Solís, C., & Selles, M. (2003). *Historia de la ciencia*. Editorial Espasa Calpe.

Thomson, J. J. (1907). *The corpuscular theory of matter*. A. Constable & Company, Limited.

Thomson, J. J. (1914). LXXXVIII. The forces between atoms and chemical affinity. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 27(161), 757-789.



## Los valores no epistémicos en la investigación del cambio conceptual

José Antonio Castorina\*  
Alicia Zamudio\*\*

### Ciencia y valores: objetividad y racionalidad sin neutralidad valorativa

Como es sabido, la filosofía clásica de la ciencia ha rechazado la idea según la cual los valores pueden intervenir positivamente en la elaboración de los conocimientos científicos, ya que atentarían contra su objetividad al impregnar la producción, y sobre todo la justificación de las teorías, de componentes subjetivos. Tal como lo señalara Carnap (1963/1992), los enunciados de valor absoluto carecerían de significado cognitivo. Se ha argumentado con frecuencia que la neutralidad valorativa es indisoluble de la objetividad. Según la tradición empirista la racionalidad científica requiere imprescindiblemente de la distinción entre hechos y valores preservando así a la ciencia, y especialmente al contexto de justificación -sustentado en la exclusiva objetividad de los hechos-, de la influencia de valor alguno. Cabe sin embargo señalar que aún dentro del Círculo de Viena, se plantearon discusiones en torno a este problema, en particular a través de Neurath (Cartwright et al., 2007; Carnap, 1963/1992).

Pero, desde mediados del siglo XX, se profundizaron los debates entre los defensores y los críticos de la dicotomía entre hechos y valores y, por lo tanto, las discusiones acerca de la tesis de la “neutralidad valorativa” de la ciencia. Las perspectivas críticas consideraron que los valores son inherentes al trabajo científico, entendiendo a la ciencia como una práctica social. Como ha argumentado convincentemente Putnam (2002), la práctica de cualquier ciencia está atravesada por valores –sean o no epis-

\* Universidad Pedagógica Nacional (UNPE). CABA, Buenos Aires, Argentina; Universidad Nacional de Tres de Febrero (UNTREF). Sáenz Peña, Buenos Aires, Argentina; CONICET. Buenos Aires, Argentina.

\*\* Universidad Nacional de Lanús (UNLa). Remedios de Escalada, Lanús, Buenos Aires, Argentina; Universidad de Buenos Aires (UBA). CABA, Buenos Aires, Argentina.

aliciazamudio62@yahoo.com.ar



témicos– de modo que los juicios sobre hechos presuponen juicios de valor y recíprocamente. Así, no se han podido excluir justificadamente los enunciados científicos que son descriptivos y valorativos simultáneamente. Putnam y Gómez han argumentado para oponerse a la dicotomía de hechos y valores, justificando su indisociabilidad, al tiempo que les atribuyen un papel positivo fundamental como componentes de la práctica de la ciencia. Los valores, y específicamente los no epistémicos, considerados por los defensores de la neutralidad valorativa contrarios al carácter racional de la ciencia, son entendidos en esta perspectiva como componentes fundamentales e ineludibles de la racionalidad científica. En efecto, Gómez (2014) señala que, contrariamente a la perspectiva clásica, una auténtica (no utópica) objetividad científica solo es posible y alcanzable si se reconocen y explicitan los valores no epistémicos que intervienen en la investigación.

Se puede afirmar asimismo que las filosofías que subyacen a la investigación científica en tanto presupuestos ontológicos y epistemológicos, son asimiladas por los científicos a su propia práctica, desde la especificidad de su elaboración de conocimientos. Estas se asocian con su trasfondo socio cultural incluyendo los valores no epistémicos, referidos a aspectos normativos morales y políticos, debido a que esas concepciones del mundo son expresión de las relaciones y disputas sociales.

Un problema relevante, para este trabajo, es si al introducir tan fuertemente los valores, particularmente los morales o políticos –los más cuestionados por los positivistas– en todos los contextos de la producción científica, se abandona la objetividad del conocimiento. Sostendremos una crítica de la objetividad basada en “los hechos” escindidos de las prácticas mediante las cuales los construimos y, por lo tanto, de los valores relativos a dichas prácticas. Las epistemologías feministas, cuestionaron la tesis de la “neutralidad” valorativa según la cual los investigadores pueden responder a exigencias de objetividad, sin examinar sus propios compromisos históricos con lo que estudian, sin asumir el significado de los orígenes de sus problemáticas de investigación (Anderson, 2004; Harding, 2004; Longino, 2015). Es preciso reconsiderar pues el problema de la objetividad desde la negación de la tesis de la neutralidad valorativa, y adoptar una tesis diferente acerca de la objetividad, a la que interpretamos como un resultado alcanzado progresivamente por un proceso social e histórico.

Por su parte, Habermas (1963/1974) ha sostenido que la neutralidad valorativa no solo es imposible, sino falsa, y que el juicio es una condición

para poder explicar, especialmente en las ciencias sociales. Es decir: para explicar una acción es necesario comprender su sentido, y esto supone comprender las pretensiones de validez inherentes a las acciones, las que se apoyan a su vez argumentativamente en razones que son susceptibles de juicio. Por eso comprender implica juzgar, tomar partido a favor o en contra a través del juicio de esas acciones. Para él, la idea de racionalidad no se restringe a su carácter instrumental, sino que debe ser tan amplia como para incluir aspectos normativos, expresivos y comunicativos, por lo tanto, ha admitido la posibilidad de discutir racionalmente acerca de los mejores valores, “los más racionales.”

Interpretamos pues la objetividad como un resultado histórico, una marcha hacia adelante, “un proyecto”, en el sentido de una conquista a alcanzar por la investigación y no un dato que es anterior al conocimiento mismo, como en la versión positivista. No se trata del acuerdo o correspondencia de alguna representación cognitiva con un mundo ya dado, sin intervención de las creencias del investigador, sino de actividades que finalmente son limitadas por el mundo al que se dirigen, ya que de algún modo éste resiste a las interpretaciones. La objetividad se predica de una actividad social, una clase de intersubjetividad en la validación de hipótesis, de logro de consensos en un proceso de análisis crítico y debate por los grupos de investigadores.

Nos proponemos utilizar estas tesis para examinar cómo intervienen los valores no epistémicos en algunas de las teorías del cambio conceptual (CC), las que pretenden dar cuenta del cambio cognitivo que tiene lugar en procesos de aprendizaje de conceptos científicos, relacionando estos conocimientos con los saberes previos de los estudiantes, con sus implicaciones para los procesos de enseñanza de la ciencia. No se trata de descalificar a las teorías o al resultado de las investigaciones por la intervención de valores no epistémicos, sino por el contrario, de analizar su intervención por ser inherentes a cualquier práctica científica. Todo investigador asume alguna perspectiva valorativa que se puede discutir apelando a buenas razones y se la puede comparar con otras que intervienen en la investigación. En nuestro caso, analizaremos la presencia de valores en la elaboración y justificación de las teorías del cambio conceptual, atendiendo a sus investigaciones, así como a las condiciones sociales o académicas en que se realizan y al rigor conceptual de los análisis reconstructivos.

## **Análisis de la intervención de valores no epistémicos en la investigación sobre cambio conceptual**

En términos generales, se pueden distinguir zonas o niveles del proceso de investigación en los que intervienen los valores no epistémicos: el modo de formular los problemas, el recorte del objeto de investigación, las unidades de análisis, incluso, la elección de los modelos explicativos. Aquí solo presentaremos brevemente algunos problemas, que a veces forman parte de los debates entre los investigadores del CC, y que pueden ser analizados y discutidos en términos de valores. Pretendemos apuntar a una revisión crítica de las propias investigaciones atendiendo a la perspectiva que hemos presentado como marco general de este trabajo.

Las teorías del cambio conceptual (TCC) –surgidas en el seno de las psicologías del desarrollo y de la enseñanza de las ciencias, desde hace cuatro décadas–, han abordado empíricamente el problema de cómo los conceptos de los que disponen los sujetos en diferentes áreas temáticas, cambian con el aprendizaje y el desarrollo. Se trata de explicar las dificultades de los “estudiantes” en aprender los conceptos más avanzados y contra intuitivos de las ciencias particulares; y sobre todo, dar cuenta de la reorganización de sus saberes cotidianos. La categoría de cambio conceptual adoptada, supone –en muchos autores– una analogía con el “cambio teórico” en la historia de la ciencia, siguiendo las filosofías del giro historicista originado en Kuhn. Strike y Posner (1992) formularon inicialmente la pregunta ¿cómo pasan los estudiantes de una concepción C1 a una concepción C2?, entendiendo por “concepción” a una organización cognitiva compleja. Para ellos, este proceso podía asimilarse a las ideas de Kuhn sobre el cambio conceptual en la ciencia de un paradigma a otro en un proceso de revolución científica; asemejaron, asimismo, la resistencia al cambio de un paradigma en la historia de la ciencia con la persistencia de las denominadas “concepciones erróneas”. De este modo, el cambio conceptual sería radical y el conflicto cognitivo tendría un papel central en el reemplazo de ideas anteriores por nuevas ideas. Desde esta versión pionera a la actualidad, se ha configurado un campo heterogéneo de investigaciones muy detalladas de carácter descriptivo en torno al “conocimiento disponible”, sea cotidiano o ingenuo, relativo a dominios específicos de conocimientos, y respecto a la explicación del proceso de cambio, destacándose, entre otras, las teorías de Carey y Vosniadou. En

este marco, es posible identificar controversias entre los investigadores que incluyen tesis ontológicas y epistemológicas que, como veremos, involucran valores no epistémicos.

Para ilustrar alguna disputa, evocamos una cuestión clásica en la literatura del campo, relacionada con las concepciones infantiles sobre la forma y posición de la tierra, y su reorganización, para alcanzar una débil aproximación al conocimiento científico. Vosniadou (2006, 2013), ha identificado modelos infantiles que consisten en la representación de la tierra plana apoyada en el agua, pasando por el modelo dual, en el cual la tierra es hueca y en cuyo interior hay una plataforma donde “vive la gente”, que representa un intento de sintetizar las ideas espontáneas y la instrucción escolar; por último, algunos sujetos llegan a representarse a la tierra esférica, cuyos habitantes viven en toda su superficie. Según esta perspectiva, estos modelos se inscriben en teorías marco, conformadas por ciertos presupuestos ontológicos y epistemológicos. La revisión de los modelos, –de la tierra plana y el sintético o dual–, proviene de una reorganización trabajosa de estos presupuestos. (Vosniadou 2006). En las antípodas de esta interpretación, la teoría contextualista (Ivarsson, Schoultz, & Saljo, 2002) cuestiona la idea misma de cambio conceptual, basándose en que las concepciones de la tierra como plana o como esfera hueca no aparecen cuando los niños han sido expuestos a un globo terráqueo, en tanto instrumento cultural. Se considera así que no se pueden estudiar las ideas de los niños al margen de los instrumentos culturales asociados. Más aún, lo único que interesa son las relaciones con estos últimos a través del diálogo discursivo entre alumnos y docentes. En términos estrictos, no resultan relevantes los saberes previos como tales en términos del sujeto individual, sino los producidos en intercambios discursivos (Castorina & Zamudio, 2019). En estas disputas, el modo de plantear los problemas, o los procedimientos metodológicos están asociadas a los supuestos ontológicos y epistemológicos de los investigadores y, como veremos, a valores no epistémicos implicados.

La indagación empírica de carácter descriptivo sobre “saberes cotidianos” o ideas previas, tal como se ejemplifica en las concepciones infantiles sobre la forma de la tierra, referidas en el párrafo anterior, constituye una de las líneas de investigación más desarrolladas por las TCC. Sin embargo, la cuestión de la naturaleza epistémica de esos saberes y su relación con el conocimiento científico constituye uno de sus puntos más críticos. Gran parte de la literatura sobre cambio conceptual asimila “sabe-

res cotidianos”, o conocimiento pre-instruccional a “conceptos erróneos” [*misconceptions*]. Se pueden identificar tres posiciones: la perspectiva TT (Teoría-Teoría), para la cual los conocimientos pre y post enseñanza son equiparables en su forma o estructura, excepto que el primero es científicamente incorrecto; la denominada KIP (*Knowledge in pieces*), sostiene que el conocimiento pre-instruccional es menos coherente, consta de conjuntos de ideas que pueden resultar fuentes positivas para el aprendizaje, y permitir a los estudiantes tomar conciencia de lo que es incorrecto; la perspectiva denominada OV (*Ontological View*) que identifica diferencias entre el conocimiento pre y post enseñanza en referencia, más que a su contenido específico, a las ontologías supuestas relativas al tipo de entidades que se conciben como existentes. En las perspectivas TT y OV el conocimiento pre-instruccional no solo es erróneo sino “mal concebido” (Di Sessa, 2017). La versión del KIP conlleva la idea de la riqueza, complejidad y diversidad del conocimiento ingenuo, y pone en cuestión la pertinencia categorial de la noción de teoría aplicada a estos conocimientos por parte de las otras dos posiciones, que conllevan la consideración del carácter erróneo del conocimiento disponible.

Halldén, Scheja y Haglund (2008) caracterizan al acento puesto en la falibilidad y la dificultad, como un “racionalismo negativo” en el que el saber disponible es identificado como obstáculo respecto de la posibilidad del cambio conceptual. Sin duda, hay instancias del proceso de CC en las que ciertos conocimientos cotidianos hacen obstáculo epistemológico para la reconstrucción del saber a enseñar, pero lo que se discute es si esta es la única caracterización posible, lo que quita significado propiamente cognitivo y entidad epistémica propia a aquellos conocimientos, y al esfuerzo intelectual que supone. Las diferencias conciernen entonces al tipo de relaciones que pueden establecerse entre viejas y nuevas concepciones en los procesos de aprendizaje. La discusión no parece resolverse recurriendo solo a investigaciones empíricas, porque incluye tesis epistemológicas y compromisos valorativos en los investigadores del CC. Así, la mayoría de las formulaciones plantean una asimetría entre conocimiento científico y conocimiento de la vida cotidiana. Si solo se miran los saberes cotidianos a la luz del conocimiento que se enseña, es irremediable considerarlos como un conocimiento de menor valor que debe ser eliminado, completamente reemplazado o sustituido, y al que se juzga como erróneo. No se reconoce al tratar el problema del CC que el conocimiento cotidiano puede ser descrito como una entidad epistémica que adquiere valor

cognoscitivo en los contextos en los que opera y también en las situaciones didácticas en las que interviene. Al mismo tiempo es importante analizar si al afirmar el carácter erróneo del conocimiento de la vida cotidiana, se legitima una descalificación de los saberes preexistentes. La idea de error en el mundo escolar es una idea primordialmente valorativa ya que se asocia al juicio y la evaluación negativa del aprendizaje.

En otra discusión, se contraponen la concepción de un sujeto individual, que hace una elaboración cognoscitiva interna, con escasas o nulas referencias a la intersubjetividad, con el carácter social de los procesos de aprendizaje de conocimientos científicos en el espacio escolar. En un análisis crítico a las versiones clásicas del CC, Miyake (2008) destaca la necesidad de considerar el cambio de conceptos con eje en la colaboración y reflexión compartida entre alumnos y docentes y entre pares. La reducción del problema a investigar a la dimensión cognitiva individual o, por el contrario, la consideración de la colaboración y la intersubjetividad, involucran diferentes orientaciones a seguir en todas las fases de investigación. Así, Inagaki y Hatano hacen referencia explícita a la dimensión valorativa, cuando dicen que las indagaciones del CC han sido “demasiado cognitivas y demasiado individualistas” (Inagaki & Hatano, 2008, p. 259). Claramente, se valora positivamente el conocimiento individual, con exclusión de la actividad intersubjetiva, y una desvalorización del saber docente, al desconocer su protagonismo en la promoción del CC.

Asimismo, muchas de las investigaciones tienen consecuencias para la enseñanza de las ciencias, pero las situaciones didácticamente configuradas en contextos escolares no constituyen variables relevantes en la mayoría de los estudios empíricos. Las condiciones específicas de producción de conocimientos en el espacio escolar no forman parte del problema tal como es presentado y analizado por las teorías del CC. Quizás sea este uno de los principales problemas que enfrenta esta orientación para la enseñanza de las ciencias: no considerar a la acción del docente como constitutiva de las situaciones que se crean para el aprendizaje. El desconocimiento del rol activo de docentes y alumnos en la situación didáctica define un problema u objeto distinto a investigar y restringe el cambio conceptual –por afuera de dicha situación– a una actividad mental individual, descontextualizada. Aquí hay una carga valorativa respecto del lugar que se le otorga al docente y a los alumnos. La no consideración de los contextos, pone en evidencia la intervención de valores en la investigación, sea el individualismo como aspiración a que cada uno produzca por

sí mismo las ideas, o aspirar a que el curso de acción de los alumnos siga un procedimiento constructivo espontáneo y universal, no asociado a las condiciones en que se formulan las tareas. Revisar y discutir esa carga valorativa puede contribuir a modificar la interpretación del conocimiento que se configura en la situación didáctica y por ende el que se produce en la investigación.

Según Habermas (1981/1989), todo acto enunciativo contiene un enunciado y una enunciación. Esta última remite a la pretensión de validez de lo que se enuncia de modo que el concepto que conforma el enunciado y la enunciación en tanto pretensión de validez conforman una unidad inescindible en el acto enunciativo. Así el carácter erróneo atribuido al conocimiento de la vida cotidiana, el carácter individual del proceso cognoscitivo o la “omisión” de la actividad del docente en la configuración de la situación didáctica son algo más que una mera descripción, sino que involucran indudablemente valores.

Finalmente, cabe examinar los aspectos valorativos que involucra la supuesta homogeneidad cultural del sujeto del cambio conceptual, así como las relaciones entre cambio conceptual y heterogeneidad cultural. En la mayoría de los estudios de CC se considera un sujeto con independencia de su pertenencia a un mundo cultural con sus propias características, una expresión de la preeminencia de un cierto patrón cultural. Es decir, el de la cultura occidental y, en particular ciertos “patrones de racionalidad universal” relativos al conocimiento científico, que deben ser alcanzados como meta de los procesos de cambio y que desconocen las diferencias sociales y culturales, incluso dentro de una misma sociedad. Por el contrario, una posición favorable a la interculturalidad y diversidad de identidades en los estudios de CC, puede reorientar profundamente las indagaciones. Se trata de identificar la riqueza y variedad de los conocimientos previos de los estudiantes, entendiendo su inserción o pertenencia a diversas culturas, cuyas interacciones se basen en el reconocimiento de su valor, igualdad y dignidad, asumiendo que se influyen enriquecedoramente en el intercambio. Justamente, suponer la homogeneidad hace invisible la diversidad de las sociedades y contextos escolares, de los sujetos y grupos culturales que las integran; en tanto que reconocer la heteorgeneidad hace visible el conflicto o las asimetrías de poder que siempre han permeado la diversidad cultural, y muchas veces la desigualdad en la distribución social del conocimiento en los sistemas escolares. Por ello, los estudios del CC en su pretensión de influir el aprendizaje de

conocimientos científicos, deben discutir los valores morales y políticos involucrados en la interpretación de los sujetos de aprendizaje, sus identidades y la distribución del conocimiento en la diversidad cultural. Ya que los intercambios son activos, se habilita la construcción de espacios de transformación, convirtiéndose en una práctica política como respuesta a la geopolítica hegemónica de conocimiento.

### **Conclusión: ¿El reconocimiento de los valores en el estudio del cambio conceptual elimina su objetividad?**

Finalmente, la cuestión central: ¿el reconocimiento de los valores en el estudio del cambio conceptual, elimina la objetividad de su estudio? Al respecto sostendremos la tesis según la cual la consideración de la intervención de esos valores no supone que los problemas se deben formular de modo tal que conduzcan a una conclusión predeterminada; por el contrario, una hipótesis sugerida por valores morales o políticos, pueda ser falsada por la experiencia. De lo contrario, la intervención de los valores en la investigación se volvería inaceptable. Y lo que es fundamental, es preciso reconsiderar el concepto mismo de objetividad: se abandona la tesis de que el conocimiento es una apropiación completa de una realidad existente, de modo impersonal y desinteresado. Por el contrario, la objetividad se concibe como interactiva, producida intersubjetivamente luego de un proceso social de las investigaciones, consensuando el modo de arbitrar las discusiones. Se puede afirmar que el cuestionamiento de ciertos valores, en base a buenas razones, puede ayudar al logro de la objetividad, cuando aquellos obstaculizan el planteo de ciertos problemas o de consecuencias que son debatibles desde otros valores no epistémicos. Estos análisis conducen a promover la reflexión crítica por parte de los investigadores.

Las controversias abiertas en el campo de investigación que hemos referenciado respecto de los distintos ámbitos de intervención de valores, dan cuenta de la propia revisión crítica de los investigadores, que se aproximan a la identificación de preeminencias valorativas diferentes. La explicitación de los presupuestos ontológicos y epistemológicos, así como las distintas valoraciones implicadas, puede promover el avance de las propias investigaciones empíricas y el análisis crítico de las condiciones sociales y de los valores que intervienen, en particular morales y políticos. Se contribuye así al incremento de los consensos que constituyen la obje-

tividad de los conocimientos sobre el cambio conceptual; una objetividad sin neutralidad valorativa en los términos que hemos sostenido en este breve trabajo.

De este modo, es posible hacer una reconstrucción crítica de los valores que obstruyen el avance de los conocimientos, o de los que la facilitan. Por ejemplo, la evaluación de los alumnos como “sujetos asociales y aculturales” o la invisibilización del papel de los docentes en los procesos de CC son obstáculos epistemológicos para la investigación. El cuestionamiento argumentado a esos valores morales y políticos se convierte en un componente del proceso de conquista de la objetividad; en lugar de ser ajena a éste, es imprescindible. Así, los valores no epistémicos condicionan y forman parte de la investigación, pero lo hacen de un modo por lo general implícito. Por ser parte del “sentido común académico”, es preciso explicitarlos. Se trata de una reflexión que los tematiza para establecer cuando dificultan la ampliación de las evidencias o si, por el contrario, orientan la obtención de un más amplio rango de evidencias.

## Referencias

- Anderson, E. (2004). Uses of value judgment in science: A general argument, with lessons from a case study of feminist research of divorce. *Hypatia*, (19), 1–24.
- Carnap, R. (1992). *Autobiografía intelectual* (C. Castells, trad.). Barcelona: Paidós. (Obra original de 1963)
- Cartwright, N., Cat, J., Fleck, L., & Uebel, T. (1996). *Otto Neurath: Philosophy between science and politics*. Cambridge: Cambridge University Press.
- Castorina, J.A. & Zamudio, A. (2019). Los supuestos ontológicos y epistemológicos en las teorías del campo conceptual. *Revista de Epistemología e Historia de la Ciencia*, 3(2), 49–68
- Di Sessa, A. (2017). Conceptual change in a microcosm: Comparative learning analysis of a learning event. *Human Development*, 60, 1–53.
- Gómez, R. (2014). *La dimensión valorativa de las ciencias*. Bernal: Universidad Nacional de Quilmes.

- Habermas, J. (1974). *Theory and practice* (J. Viertel, trad.). Boston: Beacon Press. (Obra original de 1963)
- Habermas, J. (1989). *Teoría de la acción comunicativa* (vol. 2, M. Jiménez Redondo, trad.). Buenos Aires: Taurus. (Obra original de 1981)
- Hallden, O., Scheja, M. & Haglund, L. (2008). The contextuality of knowledge: An intentional approach to meaning making and conceptual change. En S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 509–532). Nueva York: Routledge.
- Harding, S. (2004). Introduction: Standpoint theory as a site of political, philosophic and scientific debate. En S. Harding (Ed.), *The feminist standpoint theory reader* (pp. 1–15). Nueva York: Routledge.
- Inagaki, K. & Hatano, G. (2008). Conceptual change in naïve biology. En S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 240–262). Londres: Routledge.
- Ivarsson, J., Schoultz, J. & Saljo, R. (2002). Map reading versus mind reading: Revisiting children's understanding of the shape of the earth. En M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 77–100). Dordrecht: Kluwer Academic Publishers.
- Longino, H. (2015). The social dimensions of scientific knowledge. En E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition). <https://plato.stanford.edu/archives/spr2015/entries/scientific-knowledge-social/>
- Miyake, N. (2008). Conceptual change through collaboration. En S. Vosniadou (Ed.), *Handbook of research on conceptual change* (pp. 240–262). Londres: Routledge.
- Putnam, H. (2002). *The collapse of fact-values dichotomy and other essays*. Londres: Cambridge University Press.
- Strike, K. & Posner, G. (1992). A revisionist theory of conceptual change. En R. Duschl & R. Hamilton (Comps.), *Philosophy of science, cognitive psychology, and educational practices* (pp. 147–176). Nueva York: State University of New York Press.

- Vosniadou, S. (2006). Investigaciones sobre el cambio conceptual: direcciones futuras y de vanguardia. En M. Carretero & S. Vosniadou (Comps.), *Cambio conceptual y educación* (pp. 35–52). Buenos Aires: Aiqué.
- Vosniadou, S. (2013). Conceptual change in learning and instruction. En S. Vosniadou (Ed.), *International handbook of research on conceptual change* (2da ed., pp. 11–30). Nueva York: Routledge.



## El estudio de la química en clave histórica

Sandra Sandoval Osorio\*

### ¿Cómo se asume historia y la epistemología para la enseñanza?

Cuando se acude a la historia de las ideas científicas y a las reflexiones sobre cómo cambian estas ideas, el historiador y el educador se preguntan de manera distinta: cómo se aproxima y cuáles son sus intereses no es una simple cuestión de metodología, como tampoco se resuelve por el tránsito de las metodologías del historiador o del filósofo al campo de la enseñanza de las ciencias, porque aunque se han estandarizado bastante bien los intereses y métodos del historiador, e incluso en muchas ocasiones acogemos sus productos como herramientas adecuadas para la enseñanza, no se puede asumir que los intereses de comprensión son idénticos.

A pesar de la distinción de intereses o intencionalidades, lo que sí es importante acentuar es que comprender el cambio de las ideas científicas en su historicidad no es una lectura desde el pasado hasta el ahora. Ya los historiadores afirman que para su comprensión de la historia se construyen categorías para analizar los hechos del pasado. Por lo tanto, hemos comprendido que la ciencia no es una acumulación o sumatoria de ideas que se van imponiendo para superar los errores que los anteriores habían cometido. En ese sentido, se acoge la idea de Bachelard (1953) de un espíritu científico que se rectifica.

Sin embargo, en los textos de enseñanza de las ciencias, muchas veces se sigue insistiendo en una versión de historia anecdótica de ideas triunfadoras que superan y corrigen errores del pasado. Procurar una presentación de una historia crítica de las ciencias es más deseable, pues desde allí se ven dinámicas de las ciencias en las que se estudian las preguntas y las formas como los sujetos las han abordado, las influencias que han tenido las formas como se ve el fenómeno estudiado, las dinámicas experimentales que se le imponen a las situaciones de estudio. Pero su presentación no

\* Estudiante del Doctorado Interinstitucional en Educación. Énfasis en Filosofía y Enseñanza de la Filosofía. Universidad Pedagógica Nacional de Colombia.  
ssandoval@pedagogica.edu.co

asegura una modificación en la postura de los sujetos frente al saber, ésta no se transforma sencillamente por la aparición de nuevas perspectivas históricas en la enseñanza.

Entonces, en este orden de ideas se hace énfasis en qué significa tener una relación con el saber: estar involucrado en los procesos de comprensión, complacerse en el camino recorrido, ser protagonista en la formulación de hipótesis y en la rectificación de las propuestas. Además de procurar una versión crítica de la historia de la ciencia, también se procura que el sujeto acuda a los textos de los científicos principalmente porque hay coincidencias en las preguntas y preocupaciones cognitivas, y perspectivas de estudio que discutir.

### **El caso de la energía química**

Propongo una perspectiva histórica que haga sus interpretaciones desde un contexto específico contemporáneo, pero reconociendo los cambios de las teorías y de sus dinámicas. En este sentido, retomamos la forma en la que Mach recurre a los estudios históricos:

No tiene por objetivo revivir en nuestros ojos las ideas de los primeros investigadores, para restaurar los primeros bosquejos de las doctrinas que han adoptado sus sucesores, de seguir en todos sus detalles la evolución de los datos que son poco a poco organizados, distinguidos, terminados, para convertirse en teorías más extensas y detalladas. Si se apela a la historia, es solamente para comprender el significado de las fórmulas económicas<sup>1</sup> que constituyen hoy la ciencia. (Duhem, 1903/1996, p. 118)

Ahora bien, comprender los desarrollos de las ideas sobre energía química que promueve reemplazar las ideas metafísicas del ímpetu a las ideas contemporáneas, da elementos para cuestionar, por una parte, nuestras acepciones en las clases de ciencias; y, por otra, nuestras concepciones de conocimiento. Este caso lo esbozaremos más adelante. Al mostrar los fe-

---

<sup>1</sup> Cuando se hace referencia a fórmulas económicas, en Duhem y Mach, podemos entender que la ciencia establece conceptos y teorías que una vez comprendidas ya no es necesario volver a cada caso particular que explica, sino que su validez se extiende a todos los casos planteados hasta el momento, así como a los posibles que surjan en las mismas condiciones y consideraciones. Duhem dice que una teoría es un conjunto de formulaciones matemáticas, que constituyen la forma más sencilla, clara y completa de representar un conjunto de regularidades experimentales contempladas en el proceso de estudio de los fenómenos físicos; y una vez establecida la teoría no se requiere cada vez volver a las experiencias a las que hace referencia, sino que se comprende que ya están y otras posibles de aparecer se encuentran cobijadas en la formulación teórica.

nómenos a los que nos referimos cuando hablamos de energía química, se vinculan elementos a las reflexiones sobre la construcción de fenómenos en la enseñanza de las ciencias.

En la construcción de la idea de energía química podemos, de un lado, narrar los diferentes momentos que identifican la transformación de estas ideas, las cuales traen como consecuencia nuevas formas de percibir; y, de otro lado, destacar maneras de comprensión científica que repercuten en las teorías científicas que tomamos en consideración. Con esto queremos mostrar que el docente de ciencias asume una lectura histórica que está en consonancia con su relación con el saber científico: si el conocimiento es absoluto y se renueva por superación de errores que se cometieron en el pasado, la perspectiva histórica que mejor entenderá es aquella que le proporciona ejemplos para defender su argumento de grandes hombres con mejores ideas que triunfan sobre otros, ahora bien, si el docente está más cercano a un saber en donde la pregunta es constitutiva del mismo, una perspectiva crítica es una en la cual el docente podrá comprender que las preguntas, los argumentos, los despliegues experimentales son formas de diálogo con los científicos de diferentes épocas. De ahí que consideramos que lo que podemos hacer, en el proceso de formación de docentes, es cuestionar las distintas perspectivas históricas poniéndolas en diálogo con otras lecturas históricas que involucran al docente como parte del ejercicio de comprender el objeto de su disciplina, es decir, no hay una relación prevista entre historia, epistemología y enseñanza. Son campos que se pueden interceptar y se interceptan por la acción del docente y sus preguntas e intenciones.

En ese ámbito, hay una preocupación legítima por el acceso y clasificación de las fuentes que se utilizan para hacer historia de las ciencias, así como por el rasgo de rigurosidad que le da valor al producto final, los historiadores presentan reconstrucciones de una realidad pasada basados en investigaciones de hechos precisos y discuten esas reconstrucciones en términos de la adecuación de tales hechos, pero muy pocas veces estas discusiones conducen a un consenso, sin embargo, en las divulgaciones que se hacen para la enseñanza de las ciencias se destacan los consensos o los aciertos y son precisamente los disensos los que le dan más valor a los criterios de cambio en las tareas científicas.

Con lo anterior, se destaca que para la clase de ciencias tomamos los términos, metodologías, protocolos que conforman los desarrollos actuales. Hay enfoques para la enseñanza de las ciencias que toman esta pers-

pectiva de los protocolos, o en otras palabras los resultados estandarizados. Pero epistemológicamente, la pregunta es más la causa dinámica de la ciencia y por tanto puede haber una enseñanza de las ciencias que trabaje centrada en tales preguntas. Además, con ellas –las preguntas– acudimos a conversar con los autores, en otras palabras, el análisis histórico reporta una ruta en la cual el docente examina esta relación con el saber que planteamos antes como central en esta discusión, principalmente en el terreno de la enseñanza.

Pasando ya, al caso que queremos presentar, teniendo el preámbulo como perspectiva para la lectura, vamos a destacar algunos elementos en los que inicialmente se ha podido examinar cómo se aborda el estudio de las transformaciones químicas, donde la afinidad química aparece como una fuerza que determina la clase de cambios que se dan. Luego para medirla se asocian formas y aparatos de medida que llevan al vínculo con una clase de energía que hemos llamado energía química.

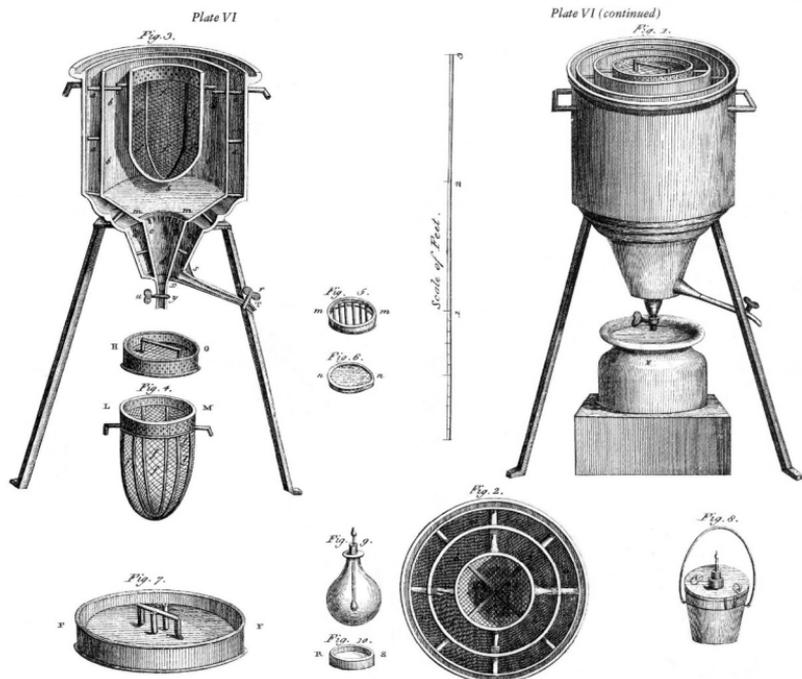
Al inicio del siglo XIX, recién constituido el campo de los fenómenos de transformación de las sustancias, ordenar las relaciones entre las reacciones químicas y la cantidad de calor generado o absorbido, es una de las problemáticas planteadas. Así, se perfeccionaron los métodos para tomar medidas de calores de reacción, que estaba asociado a la cantidad que se requería o que se liberaba en un proceso químico. Para esto se construyen aparatos especiales como las bombas calorimétricas de Lavoisier y se ponen en consideración maneras modernas de concebir el calor<sup>2</sup>.

La pregunta por la relación química / calor de reacción orienta y hace necesaria la construcción de instrumentos y experiencias para avanzar en el abordaje de estas preguntas.

Se destaca que Lavoisier, por ejemplo, relacionaba las diferentes clases de reacciones con su afinidad con el calórico: “el hidrógeno y el oxígeno, tienen una gran tendencia a unirse con el calórico y convertirse en gas; mientras que el carbono, por el contrario, es un principio fijo, y que tiene muy poca afinidad con el calórico” (Lavoisier, 1789, p. 97).

---

<sup>2</sup> La afinidad química, antes de estos estudios es registrado como una posible fuerza de atracción que une a unas sustancias, que hace una sustancia prefiera atraer a una y no a otra. La medida de esta fuerza que hace que las sustancias se transformen en otras (que un metal se deje corroer por un ácido y más por un ácido nítrico o clorhídrico que por un ácido acético, o que un ácido reaccione con una base para formar una sal), es la que la vincula por una parte a las medidas de calor generado en las reacciones químicas y por otra parte a la rapidez con la que una reacción se extendía.



**Fig. 1:** El calorímetro de Lavoisier y Laplace.

*Nota:* Reproducción de la placa VI de Lavoisier, A. L. (1780). *Elements of Chemistry, In a New Systematic Order, Containing all the Modern Discoveries* (R. Kerr, Trad.). Recuperado de <http://www.gutenberg.org/ebooks/30775> [Traducción al inglés de Lavoisier (1789)].

Se plantea en la memoria sobre el calor de Lavoisier y Laplace (1780) que la cantidad de calor que se libera en un proceso de combustión puede derretir cantidades medibles de hielo. Laplace sugiere que la cantidad de hielo derretido por un objeto caliente al enfriarse a cero (el punto de fusión del agua) es proporcional a la cantidad de calor que emitió, siempre que el hielo estuviera inicialmente a cero. Si el recipiente de hielo tiene la forma de una esfera hueca de hielo, la temperatura del hielo luego subiría lentamente y comenzaría a derretirse, pero sólo en la superficie exterior y el centro hueco estaría protegido de la entrada de calor de tal manera que un objeto caliente colocado en el centro se enfriaría entonces hasta el cero, todo el calor emitido sería absorbido por el hielo formando agua a cero, y ningún calor escapa.

Las mediciones realizadas por Lavoisier y Laplace con respecto a los calores de reacción, o las de Bertholet, constituyen una propuesta para la medida calorimétrica de la afinidad química. A continuación, se describen algunos aspectos del calorímetro de hielo construido para estas determinaciones.

En la memoria de Lavoisier y Laplace sobre el calor se presenta el principio de funcionamiento del calorímetro del hielo (Figura 1) y numerosas experiencias realizadas como la determinación de calor de reacción del aceite de vitriolo con el agua, de la cal viva con el agua y con el ácido nítrico y calores de combustión de diferentes sustancias. La última parte de la memoria está destinada a consolidar la teoría del oxígeno (Laplace & Lavoisier, 1780). La siguiente imagen es el calorímetro, que tenía aproximadamente tres pies de altura y estaba hecho de metal con tres compartimentos concéntricos.

Las discusiones sobre si el calor es un fluido material no es el que más ocupa la atención de estos científicos. Cuando se está tratando con la química se habla del calor como algo que permite establecer una relación entre los estados iniciales y finales de los sistemas químicos. Estos estados son cuidadosamente definidos, no son solamente puestos en un recipiente de reacción y colocados los termómetros para que se midan las variaciones en el calor, también se requiere que este en las mismas circunstancias, la misma cantidad de sustancia, a las mismas condiciones de temperatura y presión, esto más que controlar el fenómeno, es construirlo.

Se dan otros métodos de medida como el uso del tiempo de disolución de pequeños cilindros metálicos sumergidos en soluciones ácidas de Wenzel en 1777 o la medida de la capacidad de saturación realizada por Kirwan en 1781, es decir, la cantidad de ácido utilizado para saturar una cantidad de base dada (Ostwald, 1902).

Más adelante, los estudios como los de Waage y Gulberg (1864/1986) establecen una relación entre la reacción química y su velocidad de reacción, cuando abordan esta cuestión de la afinidad, el ímpetu con la que la reacción se sucedía era un indicio sin duda de la facilidad o dificultad con la que se realizaba la transformación de las sustancias que intervenían. La afinidad química se concebía como una fuerza que interviene en la actividad de las sustancias para su mayor o menor probabilidad de reacción química. El asunto es que, para este entonces, abundaban las tablas de afinidad que organizaban, por ejemplo, la afinidad de los metales por el oxígeno, o por los ácidos, o la afinidad entre un ácido y diferentes álcalis

o viceversa. Todas estas tablas dan un orden entre clases de sustancias que producen clases de reacciones distintas, sin embargo, no responden la pregunta por cuánto más o menos es afín una sustancia por otra. La cuantificación no solo va a responder la anterior pregunta, sino que también va a unificar todas estas ordenaciones.

De la relación entre la reacción química con el calor de reacción y de la relación con la fuerza química que se presenta en las reacciones con mayor rapidez de reacción se abren dos caminos para el estudio de la actividad química: uno que se convierte en un proyecto de desplegar la manera de entender la mecánica, y en particular la cinética, aplicada a las reacciones químicas, pues se hace énfasis en cuantificar las fuerzas de acción química y se empiezan a relacionar los procesos físicos con las transformaciones químicas, haciendo analogías entre los comportamientos de las sustancias con las leyes de la mecánica. El otro camino es la relación de las medidas de calor con la mayor o menor afinidad que se desarrolla en la termoquímica y que incluso relaciona la actividad química con otras clases de actividades como la eléctrica.

La medición de afinidad química en virtud de su velocidad es abordada por los científicos Berthollet, Wenzel, Wilhelmy, Guldberg y Waage. El ejercicio de esclarecer la relación entre la afinidad química y las fuerzas químicas o entre ella y los calores de reacción, repercute en formalizaciones matemáticas y en desarrollos experimentales que a la vez esclarece el significado de la afinidad química.

Wenzel propone maneras de medir las fuerzas de acción química, empleando el tiempo como una cualidad para organizar de manera creciente o decreciente las reacciones químicas; después Wilhelmy formula una expresión matemática para describir la pérdida de reactivo en el tiempo y Guldberg y Waage reafirman la velocidad de reacción como una medida de afinidad química, postulando las leyes de acción de masas y volúmenes; y expresando la ecuación de afinidad química en términos de velocidad de reacción, integrándose conceptos de la mecánica física a las transformaciones químicas (Ostwald, 1908/1909).<sup>3</sup>

Como se ha dicho antes, paralelamente se destaca la relación de la actividad química con otras clases de actividades como la eléctrica o la calórica. Entonces se establece que la mayor o menor actividad química

---

<sup>3</sup> Un particular desarrollo de estos aspectos se encuentra en Pulido (2016) en donde en la revisión histórico crítica establece que la velocidad de reacción química es una medida derivada de la afinidad o reactividad química.

puede estar en relación directa con la mayor o menor energía eléctrica o calórica que se pueda medir durante un proceso químico. Se ha pasado de hablar de afinidad a acuñar el término “actividad”. En el caso de los fenómenos químicos con los fenómenos eléctricos, Ritter establece la analogía entre las escalas de tensión eléctrica de los metales que había construido Volta<sup>4</sup> y la escala de la afinidad de los metales por el oxígeno (Ostwald, 1910/1912).

Vincular al estudio de las reacciones químicas, y a la actividad química de las sustancias, medidas de energía, abre una perspectiva conceptual, desde donde se enfrenta el hecho de que en las reacciones químicas no solamente se produce o requiere calor, o mejor dicho, no toda la energía química que se transforma en calor se puede transformar en trabajo, esto quiere decir que es necesario distinguir que proporción de energía química es la que se debe a asociar a la medida de la afinidad química.

Es fundamental esta distinción, pero además también la distinción de magnitudes asociadas a clases de fenómenos:

Una diferencia de temperatura determina la existencia de un fenómeno de calor; una diferencia de Tensión, la de un fenómeno eléctrico. Y una diferencia de presión, la de un fenómeno mecánico correspondiente. Un estudio más profundo, muestra que, para cada tipo de energía, es posible determinar una magnitud. . . . Para la energía química un potencial químico es la expresión exacta de lo que buscábamos, sin saberlo bien, bajo el nombre de afinidad química. (Ostwald, 1908/1909, p. 224)

Con los anteriores desarrollos se muestra que el estudio de la energía química como una cantidad propia de las reacciones químicas, enlaza no solo el problema de cómo se transforman unas sustancias con otras, en qué condiciones, con qué especificidad, etc., sino que estas preguntas también se relacionan con la aparición del dominio de la electroquímica y de la termoquímica, como ámbitos que abordan cómo suceden las reacciones químicas y que terminan siendo la relación de un campo específico de la química con otro también específico y propio de los fenómenos eléctricos o de los fenómenos calóricos.

La afinidad química en el tránsito del siglo XIX dejó de ser una cualidad metafísica de la tendencia de un elemento a unirse con otro para pasar a ser una cantidad física que daba cuenta de cuánto más o menos trabajo estaba involucrado una reacción química.

---

<sup>4</sup> Volta construye estas escalas antes de la construcción de la batería, en 1799.

Estos elementos aportan de manera fundamental a la comprensión de cómo se transforman las sustancias químicas y dan una perspectiva diferente para la enseñanza de la misma, con la que no siempre los profesores nos familiarizamos.

### **Primeras consideraciones para la enseñanza de las ciencias**

En un énfasis por los contenidos a trabajar en la clase de ciencias, rápidamente nos daremos cuenta que el estudio de las transformaciones químicas en relación con los criterios desde los cuales se puede medir la afinidad química se puede poner de presente tanto aspectos cinéticos como termodinámicos. Pero además se hace muy visible esto que hemos llamado la relación con el objeto, el sujeto exhibe sus maneras de comprensión en las que por ejemplo se encuentran situaciones en las cuales el animismo y sustancialismo particularmente se hacen presentes en sus formas de razonar.

Organizar, distinguir y delimitar los efectos sensibles producidos por las disposiciones experimentales, y de acuerdo con ello, establecer el aumento o disminución de la afinidad química nos muestra que los efectos producidos no son cotidianos, naturales o inmediatos. Como tampoco lo son las preguntas ni los montajes experimentales involucrados.

En la enseñanza de las ciencias el conocimiento se construye sobre la base de contradecir las ideas seguras e inmediatas (Bachelard, 1972/1973, p. 12). Con ello se pone en juego nuestra relación con los otros. Vamos a defender que la ciencia construye un mundo, como fruto de las relaciones entre las formas como son vistos los objetos por el sujeto(s) y las posturas que se asumen frente a ellos. Esta relación –basada en autores como Bachelard– fundamenta nuestra discusión.

También se pone de presente que es posible conectar las preguntas actuales con las de los científicos de hace 2 o 3 siglos, por ejemplo. Las preguntas que se hacen en distintos tiempos tienen sentidos compartidos, e incluso pueden responder a una estructura lógica que hace que se pueda establecer una comunicación entre ideas separadas en el tiempo.

Esto nos lleva a evaluar que las preguntas sobre una situación de estudio no son ocurrencias espontáneas y que llevarlas al ámbito de la enseñanza implica construir un contexto preciso donde tengan un sentido que los lleve a la indagación, esto significaría muchas veces reconocer sus actuales explicaciones a la vez que provocar desajustes deliberados con

tales explicaciones. Es por ello que afirmamos que lo que llamamos en la escuela experiencia no es más que un preámbulo.

Abundan las pedagogías del elogio a la experiencia inmediata o en el otro extremo las estrategias para la corrección de ideas erróneas. Pero no solo ello, sino que esta pedagogía se ha construido contra una que ha hecho de lo dicho en los libros de textos, que es una ciencia bastante socializada, además una ciencia con respuestas a preguntas que no nos hacemos, una ciencia inmovilizada, que por su repetición año tras año, logra parecer natural.

Esta pedagogía donde lo natural, lo cotidiano, lo útil en el aula, olvida que el sujeto conquista su postura en relación con los otros, pero además el sujeto es reversible por cuanto esa relación con los otros lo hace volver una o varias veces a asumir posturas anteriores, y que los docentes creemos ya superadas, porque confiamos en una didáctica que nos asegura el cambio como permanente.

Para finalizar cabe destacar que el docente enfrenta un obstáculo que es removible pedagógicamente si deliberadamente se construyen escenarios para el desajuste con las ideas actuales, esto no sucede igual en la ciencia. Además, acudir a los textos de los científicos involucrados en la construcción de las ideas que estamos estudiando son oportunidades para generar espacios de comunicación con aquellos que han aportado a nuestras formas de pensar, pero sobre todo espacios de diálogo con nosotros mismos.

## Referencias

- Bachelard, G. (1953). *Le matérialisme rationnel*. Paris: Les Presses universitaires de France.
- Bachelard, G. (1973). *El compromiso racionalista* (H. Beccacece, trad.). Buenos Aires, Argentina: Siglo XXI editores. (Obra original de 1972)
- Berthollet, C. L. (1803). *Essai de statique chimique*. Vol. 2. Paris, Badouin.
- Duhem, P. (1996). Analysis of Mach's 'The science of mechanics: A critical and historical account of its developments'. En R. Ariew & P. Barker (Eds., trads.), *Pierre Duhem. Essays in the history and philosophy of science* (pp. 112-130). Indianapolis; Cambridge: Hackett Publishing Company. (Obra original de 1903)

- Lavoisier, A. L. (1789). *Traité élémentaire de Chimie*. Paris, Francia: Gauthier Villars, Éditeur.
- Lavoisier, A. L., & Laplace, P. S. (1780). *Mémoire sur la Chaleur*. Paris, Francia: Mémoires de l'Académie des sciences.
- Ostwald, W. (1909). *L'évolution d'une science: La chimie* (M. Dufour, trad.). Paris, Francia: Ernest Flammarion éditeur. (Obra original de 1908)
- Ostwald, W. (1912). *L'évolution de l'électrochimie* (E. Philippi, trad.). Paris, Francia: Librairie Felix Alcan. (Obra original de 1910)
- Pulido, D. (2016). *Estudio de la velocidad de reacción química como medida de la reactividad química* [tesis de Maestría en docencia de la química no publicada]. Universidad Pedagógica Nacional de Colombia.
- Waage, P. & Gulberg, C. M., (1866). Studies concerning affinity (H.I. Abrash, trad.). *Journal of chemical education*, 63(12), 1044-1047. (Obra original de 1864)



# Los procesos de síntesis teórica y la estabilización de un campo de fenómenos

Sandra Sandoval Osorio\*

José Francisco Malagón Sánchez\*

Marina Garzón Barrios\*

Liliana Tarazona Vargas\*

## Introducción

En esta comunicación presentamos el estudio que hemos adelantado de la constitución de la electroquímica como un campo en el cual se establecen relaciones entre los fenómenos químicos y los fenómenos eléctricos. Este estudio, en el cual se acude a la revisión histórico crítica de los textos de científicos, entre los cuales destacamos a Volta, Davy, Faraday y Ostwald, se sustenta en la perspectiva de contribuir a una enseñanza de las ciencias centrada en la constitución de fenomenologías, esto es como docentes de ciencias hacemos un análisis de textos de científicos (fuentes primarias) que abordan problemáticas que son objeto de nuestro estudio y cuya lectura esta direccionada por nuestras preguntas.<sup>1</sup>

En el estudio adelantado describimos las relaciones de equivalencia que se concretan en el caso de la electroquímica. A partir de esto, se han generado propuestas experimentales para la enseñanza, que hasta este momento nos permiten describir cómo se establece la equivalencia entre magnitudes como la corriente eléctrica y cantidad de sustancia, o entre tensión eléctrica y oxidación de las sustancias. Evidenciamos, una vez más, que la constitución de magnitudes, la construcción de formas de medida y aparatos de medida, son imprescindibles tanto en las operaciones

---

<sup>1</sup> Ayala, Malagón y Sandoval (2013) presentan claramente que esta es una lectura intencionada con fines pedagógicos a la que hemos denominado estudios histórico críticos y que no corresponde al ejercicio de un historiador de la ciencia o a un epistemólogo de las ciencias que centra su objeto de estudio en la reconstrucción de preguntas, metodologías y concepciones de los científicos en una época y contexto específico.

\* Grupo de Estudios Histórico Críticos y Enseñanza de las Ciencias EHC\*EC. Universidad Pedagógica Nacional de Colombia.  
ssandoval@pedagogica.edu.co

de producción de efectos sensibles, a la vez que en la producción de teorización sobre los campos de estudio. En estos procesos, la determinación de una cualidad (o cualidades) que da cuenta del fenómeno que se estudia, el análisis de efectos producidos y las relaciones de orden permiten estabilizar el fenómeno (Malagón, Ayala, & Sandoval, 2013, p. 98).

En el caso del fenómeno electroquímico, es necesario definir cuáles son las cualidades que lo caracterizan, qué efectos se relacionan, cuál es la ordenación que se logra. La pregunta por cuánto más o menos está un punto con respecto a otro en una ordenación, lleva al problema de la construcción de la magnitud y del aparato con el que se mide. Con el trabajo de Faraday, hemos encontrado una primera organización cuantitativa de los efectos electroquímicos (la cantidad de sustancia transformada es proporcional a la cantidad de electricidad que pasa por el electrolito), además del establecimiento de equivalentes al comparar efectos (las cantidades de sustancias que viajan con cantidades iguales de electricidad son químicamente equivalentes). Ostwald (1910/1912) afirma que para establecer esta relación fue necesario identificar los “efectos susceptibles de ser medidos”, por un lado, Faraday reconoce “la desviación de la aguja imantada del galvanómetro (recientemente inventado), y, por otra parte, la descomposición química” (p. 96). Cada efecto vinculado a un dominio diferente, el galvanómetro: indicador del paso de corriente, y la cantidad de sustancia descompuesta: indicador de transformación química, ambos efectos visibles en la electrólisis.

En la serie V de los *Experimental reasearchs in electricity* Faraday (1849) señala varios aspectos importantes:

- La descomposición electroquímica depende esencialmente de la corriente de electricidad (p. 146).
- Algunos científicos asumen que hay dos clases de fluido eléctrico, pero para Faraday es un eje de poder que tiene fuerzas contrarias, exactamente iguales en cantidad, contrarias en dirección, (p. 148).
- El efecto depende esencialmente de las afinidades químicas entre las partículas de clases opuestas (p. 149).

Así la producción de metales o sales en los polos se debe a la presencia de otras sustancias con las que son afines, igualmente pasa en el caso de los gases que se liberan, los ácidos y álcalis que quedan en el fluido que rodea los polos. (la explicación de la descomposición por la atracción de

los polos no da cuenta de estos casos en los que las sustancias no quedan unidas a los polos).

Yo supongo que los efectos son debidos a la modificación, por la corriente eléctrica, de la afinidad química de las partículas que atraviesan o por las cuales la corriente pasa, dando el poder de activar con más fuerza en una dirección que en otra y en consecuencia haciéndolos pasar por una serie de sucesivas descomposiciones y recomposiciones en direcciones opuesta, y finalmente causando su expulsión o exclusión en los límites del cuerpo bajo descomposición, en la dirección de la corriente, y en cantidades pequeñas o grandes, de acuerdo con que la corriente sea más o menos poderosa. (Faraday, 1849, p. 152)

Faraday encuentra que la descomposición electroquímica de sustancias no dependía de la distancia entre los polos por los que entraba y salía la corriente. Para Faraday la atracción y repulsión entre los polos y las sustancias en contacto con éstos no da respuesta a la descomposición electroquímica. Con los experimentos en los que uno solo de los polos genera la descomposición, muestra que los electrolitos se descomponen y sus compuestos no necesariamente se acumulan sobre un polo. Lo que sí es necesario para que se dé la descomposición es la corriente (como lo describen los otros científicos).

Para Faraday será la corriente la causa de la descomposición de las sustancias que ésta atraviesa, más que la acción de los polos. Los polos serán el límite entre el electrolito y el camino por donde entra y sale la corriente eléctrica.

Otro aspecto que se deriva de esta explicación es que si la afinidad de los componentes de la sustancia es muy fuerte la corriente producida por la pila no podrá descomponerlo. En caso contrario, la capacidad del electrolito de sufrir descomposición está vinculada con la capacidad de conducir electricidad. Faraday lo expresa de la siguiente manera:

La conclusión es casi irresistible, que en los electrolitos el poder de transmitir la electricidad a través de la sustancia depende de la capacidad de sufrir descomposición; teniendo lugar solo mientras se están descomponiendo y siendo proporcional a la cantidad de elementos separados. (Faraday, 1849, p. 201)

La relación que establece Faraday es: a deflexiones iguales de la aguja, corresponden cantidades iguales de sustancias descompuestas, en tiempos iguales, sin que la fuente de electricidad tenga incidencia.

La relación de proporcionalidad, expuesta por Faraday, entre los efectos de transformación de sustancias dada una cantidad de electricidad, como forma de encontrar sustancias químicamente equivalentes, no es la única relación que encontramos al estudiar la constitución del fenómeno electroquímico. También destacamos procesos en los que la conversión de efectos químicos en efectos eléctricos son fundamentos para dar cuenta de la conducción eléctrica en soluciones (Arrhenius, van 't Hoff), de la relación de las fuerzas electromotrices en las células galvánicas y el trabajo generado en los procesos químicos surgidos en las pilas (Gibbs, Helmholtz). En esta línea, el estudio de la electrólisis y la construcción de pilas nos han permitido encontrar la equivalencia como una categoría epistemológica necesaria para dar cuenta de este nuevo dominio fenomenológico.

Destacamos dos aspectos relevantes dentro de la constitución de la electroquímica: la definición de la electricidad como síntesis teórica y la relación de la electricidad con los cambios químicos.

### **La electricidad: convergencia de tres campos fenomenológicos**

Durante el siglo XVIII coexisten tres clases de fenómenos a los cuales se les conoce como electricidad: electrostática en la cual se han configurado magnitudes como la carga, el potencial y el campo eléctrico para hablar de la cantidad de electricidad. El otro la electricidad galvánica o animal en la cual se estudian los espasmos y conmociones como características de una electricidad que era considerada como propia de los seres vivos. Y, por último, el caso de la electricidad voltaica producida por el contacto de metales diferentes y que da origen a las corrientes que implica tener circuitos cerrados.

En este apartado destacamos dos aspectos sobre el experimento en la conformación de un dominio fenomenológico constituido por la convergencia de tres fenomenologías: la electrostática, la electricidad animal y la electricidad voltaica; etapa en la cual se generan las condiciones para explicar el fenómeno de la electricidad animal en términos análogos a cómo se explica la electricidad voltaica, ejemplificar cómo los aparatos de medida terminan siendo un factor de unificación, y medir electricidades débiles.

En esta síntesis que unifica son varias las cuestiones que juegan un papel importante, pero destaquemos los instrumentos y la producción de efectos que minuciosamente son operados una y otra vez hasta que se

establecen las magnitudes y formas de medida como la botella de Leyden y el galvanómetro de Cumming (1821).

Ya Volta había hecho el esfuerzo de utilizar el condensador multiplicador de Nicholson para medir los efectos que producen por el contacto de dos metales, pero se da cuenta que no puede distinguir entre la medida debida a las mismas piezas del aparato y los debidos a los fenómenos que ocurren en las disposiciones de metales en las pilas construidas

El galvanómetro de Cumming se introduce más tarde, por Davy (1826) para evidenciar los efectos que provocan las placas de metal de una pila y el cambio en la dirección de la corriente y con ello la polaridad (positiva o negativa) de las placas.

En el caso del multiplicador de Schweigger (fig. 1) se coloca una aguja magnetizada para contrarrestar el efecto del campo magnético terrestre. El sistema de Schweigger logra amplificar o multiplicar el efecto de una corriente eléctrica sobre una aguja magnetizada. En la medida que se hicieran más giros del cable conductor alrededor del marco de madera se lograba un efecto magnético mayor sobre la aguja magnetizada y permite medir y trabajar con corrientes débiles como las que ya habíamos descrito en los sistemas electroquímicos. La deflexión de la aguja es considera-



**Fig. 1:** Representación del multiplicador de Johann Schweigger.

*Nota:* Imagen extraída de “Schweigger Multiplier – 1820—MagLab”. <https://nationalmaglab.org/education/magnet-academy/history-of-electricity-magnetism/museum/schweigger-multiplier>

da proporcional a la intensidad de la corriente que circula por la bobina cuando se producen cambios químicos. Esto es, en síntesis, que la magnitud del cambio químico es medida por el efecto magnético que produce la circulación de corrientes eléctricas generadas en tal proceso químico.

## **El electrolito como generador de nuevas preguntas**

El estudio de los efectos y reacciones que suceden en el electrolito genera un campo amplio de investigación. Volta y Galvani se habían centrado en establecer las causas de la fuerza eléctrica que se exhibía en los músculos de la rana y en caracterizar los efectos del contacto de los metales.

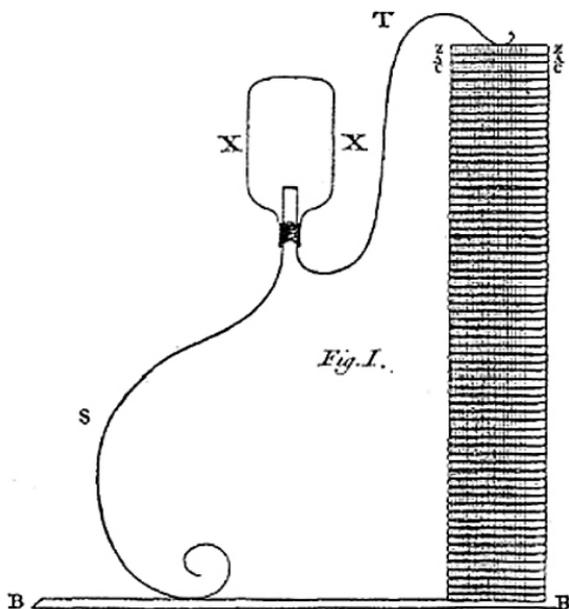
Nicholson y Carlisle (1788) habían observado la descomposición del agua que se utilizaba en una pila cuando el circuito se cerraba. Lo sorprendente fue que los dos componentes del agua se separaron, el oxígeno en uno de los hilos y el hidrógeno en el otro. Carlisle y Nicholson también observaron que en el electrodo donde se recogía oxígeno, el agua era ácida, y que era alcalina en el electrodo donde el hidrógeno se liberaba (Ostwald, 1910/1912, p. 76).

Pero ya Volta había observado este fenómeno. Él afirma:

El recipiente X-X [Figura 2] está lleno de agua destilada. A medida que el extremo del cable TT toca la superficie de la última lámina, las pequeñas burbujas de aire se forman sobre la superficie del cable SS en la parte sumergida en el agua, que crece gradualmente y se forman un chorro continuo de aire, que se recoge en la parte superior del tazón XX. Examinado este aire se encuentra que es aire inflamable. La superficie de los alambres metálicos está a la cabeza de unos pocos minutos calcinada. En los experimentos que realizó el profesor JACQUIN esta mañana, se utilizaron cables de oro, plata, hierro, plomo y latón. Todos estos hilos tienen agua descompuesta y han proporcionado una gran cantidad de gas inflamable. (Volta, 1800/1923, p. 4)

Se plantea dos interrogantes, si es posible acelerar o hacer más lenta la descomposición del agua, utilizando diversos metales interpuestos en la disposición de columnas. Y si es posible determinar la cantidad de agua descompuesta a partir de la cantidad de gas inflamable producido en un determinado tiempo. Seguramente, va a ser precisamente Faraday quien mejor aborde estas cuestiones, como ya lo adelantamos en la primera parte de este texto.

Otros investigadores como Haldane, Cruikshank, Davy, mostraron que, al trabajar con soluciones de sales de plata, zinc, cobre, plomo y otros



**Fig. 2:** Imagen incluida en la carta de Alessandro Volta al profesor Brugnatelli, “Sobre algunos fenómenos químicos obtenidos con el nuevo aparato eléctrico”. Otoño de 1800.

*Nota:* Imagen reproducida de “Lettera del prof. Alessandro volta al prof. Brugnatelli sopra alcuni fenomeni chimici ottenuti col nuovo apparecchio elettrico”. *Le opere di Alessandro Volta*, Vol. 2. Milano: Hoepli, 1923, p. 4.

metales, estas sales se podían precipitar o depositar en el electrodo donde se liberaba hidrógeno, mientras que en el otro electrodo se daban otros procesos de transformación química diferentes.

Estos hechos, que hemos estudiado en *Bakerian Lecture* (Davy, 1826), son los que permiten indicar que la actividad eléctrica que se da en las pilas o baterías está muy relacionada con las reacciones químicas producidas, en las cuales la clase de electrolito o conductor húmedo merece una investigación a profundidad.

Estas revisiones han sido de capital importancia para describir los cambios químicos que se dan en estos procesos en términos de cantidades de electricidad y viceversa, así como también la afinidad química puede ser medida en relación con las tensiones generadas en estos montajes.

La producción de ciertas sustancias en uno u otro electrodo de la pila, permitiría entonces dar cuenta de la polaridad de las placas tanto como

de las sustancias. Esto generó la pregunta si la electricidad descomponía sustancias y componía o generaba otras sustancias. Davy avanza en este asunto:

Yo concluí que las combinaciones y descomposiciones por electricidad eran atribuibles a la ley de atracciones y repulsiones eléctricas, y adelanté la hipótesis, “esas atracciones química y eléctrica fueron producidas por la misma causa, actuando en un caso sobre partículas, en el otro sobre las masas”; y que la misma propiedad, bajo diferentes modificaciones, fue la causa de todo el fenómeno exhibido por diferentes combinaciones voltaicas. (Davy, 1826, pp. 389-390)

Con esta conclusión e hipótesis que expone Davy se puede decir que los efectos de la electrodescomposición implicaron que los científicos del momento explicaran cómo se producían éstos (por atracciones eléctricas), si era posible concebir el movimiento de esas sustancias mientras eran afectadas por la electricidad, y cuántos de esos efectos eran producidos por las acciones químicas entre sustancias.

En estos trabajos se muestran dos relaciones muy importantes:

- La comparación de las escalas de tensión con las escalas de afinidad de los metales por el oxígeno, en la que se termina estableciendo una equivalencia entre las afinidades eléctricas con las afinidades químicas.
- A medida que se produce una tensión, por el contacto de dos metales, se producen cambios en las sustancias en solución (electrolito) o en las placas del metal (electrodos), cambios químicos que se explican por el paso de la corriente.

Estas relaciones nos ayudan a argumentar que cuando se logra hablar de un fenómeno en términos de otro, a través del análisis de las magnitudes que caracterizan a cada uno de los fenómenos, se establecen procesos de conversión de las magnitudes medibles de un fenómeno en términos de las magnitudes que organizan el otro fenómeno. Estudiar este tipo de convertibilidad entre magnitudes y fenómenos requiere construir formas de expresión que indiquen cuánto equivale el cambio del fenómeno A en términos del fenómeno B. La equivalencia es entonces una categoría de síntesis o de formalización de este cambio que se puede realizar o concretar en los aparatos de medida y la construcción de magnitudes.

## **Estabilización del dominio fenomenológico**

De acuerdo con los trabajos relacionados hemos allegado criterios de estabilización del fenómeno que están ligados al hecho de que:

- Volta construye y explica el funcionamiento del aparato eléctrico artificial y observa, los cambios químicos que se suceden en estos montajes tanto en el aparato de columna como en la disposición de tazas. En esta etapa, sobresale la estabilización de la producción de fuerzas electromotrices por la organización de los pares de metales según los efectos que generan. Cuando se tiene estandarizado el proceso de construcción de la batería, se logran producir corrientes de diferentes intensidades y este pasa a hacer parte de los instrumentos eléctricos que son usados para la descomposición de sustancias, como la descomposición del agua o de las bases, lo cual da condiciones de posibilidad para estudiar los diferentes procesos que se dan, lo que llamaremos procesos electroquímicos.
- Para Davy es importante explicar lo que sucede a nivel de las reacciones químicas en las pilas y para ello establece condiciones de diferenciación entre clases de sustancias álcalis, ácidos, metales y sustancias en estado gaseoso. Estas diferencias se establecen en relación con los cambios eléctricos y químicos que se evidencian en las pilas.
- Y Faraday hace una revisión de las condiciones experimentales, de las teorías existentes y de las nuevas leyes que establece para este campo.

La estabilización cobra importancia en términos de la utilización de las baterías en los procesos electroquímicos, en asociar aparatos de medida para la medición de los efectos químicos y eléctricos, elaboración de leyes y principios que estructuran las teorías en el campo electroquímico.

Después de esta dinámica de discusión, repetición de experimentos en los que se encuentra que cambiando las clases de metales se obtienen las mismas fuerzas electromotrices, esto va acompañado de una operación en la cual se establece el orden de estas junturas de metales o disposición de orden de los metales en los montajes de tazas; igualmente esta posibilidad lleva a hacer comparación de resultados las cuales producen las escalas de tensión de los metales, o el establecimiento de las leyes de Volta. Estas ob-

servaciones y afirmaciones entran a hacer parte de la base de experiencias que conforman el dominio fenomenológico que hemos extendido. En este caso hemos estabilizado la construcción del instrumento.

## Trabajo en el aula

En la enseñanza de las ciencias hemos llevado lo estudiado para trabajar con nuestros estudiantes de las carreras de ciencias de la Facultad de Ciencia y Tecnología de la Universidad Pedagógica Nacional y hemos desarrollado algunas actividades que tienen que ver con la construcción de montajes que permiten establecer relaciones que inicialmente no existían para el estudiante: la posibilidad de construcción de columnas o de tazas con diferentes disposiciones se busca establecer la generación de estos fenómenos que hoy conocemos como eléctricos y además establecer si los fenómenos dependen de los metales utilizados o de las sustancias que actúan como electrolitos.

Luego la pregunta por el electrolito y el papel que juega en las soluciones nos permiten también establecer cómo se pueden cuantificar los cambios que se observan sobre las placas de metales.

Finalmente se cuestionan los modelos que han sido aprendidos acríticamente de cargas + y - moviéndose en las soluciones, o peor aún de electrones que saltan de una sustancia a otra y que por ello producen cambios en las soluciones, a establecer sistemáticamente como se dan los cambios y de qué factores depende.

## Agradecimientos

Este trabajo recoge algunos de los desarrollos en el marco del proyecto DQU509-19 La actividad experimental en el aula. Estabilización de un campo fenomenológico y construcción de instrumentos y formas de medida como síntesis teóricas, auspiciado desde la Subdivisión de Gestión de Proyectos CIUP – Universidad Pedagógica Nacional de Colombia.

## Referencias

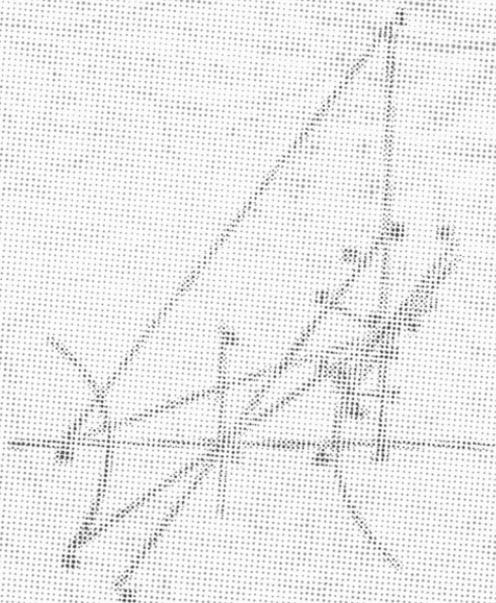
Ayala, M. M., Malagón, J. F., & Sandoval, S. (2013). La historia en la enseñanza de las ciencias. Una relación polémica. En M. M. Ayala, J. F. Malagón, & S. Sandoval (Eds.), *Construcción de fenomenologías*

- y procesos de formalización (pp. 21-37). Bogotá: Editorial Universidad Pedagógica Nacional.
- Cumming, J. (1821). On the connection of galvanism and magnetism. *Transactions of the Cambridge Philosophical Society*, Vol I, Part II, 269-280.
- Davy, H. (1826). The Bakerian lecture. On the relations of electrical and chemical changes. *Philosophical Transactions of the Royal Society of London*, 116, 383-422. <https://doi.org/10.1098/rstl.1826.0031>
- Faraday, M. (1849). *Experimental researches in electricity*. Londres: R. & J. F. Taylor.
- Malagón, J. F., Ayala, M. M., Sandoval, S. (2013). La actividad experimental: construcción de fenomenologías y procesos de formalización. En M. M. Ayala, J. F. Malagón, & S. Sandoval (Eds.), *Construcción de fenomenologías y procesos de formalización* (pp. 87-104). Bogotá: Editorial Universidad Pedagógica Nacional.
- Nicholson, W. (1788). A description of an instrument which, by the turning of a winch, produces the two states of electricity without friction or communication with the earth. In a letter from Mr. William Nicholson to Sir Joseph Banks, Bart. P. R. S. *Philosophical Transactions of the Royal Society of London*, 78, 403-407. <https://doi.org/10.1098/rstl.1788.0028>
- Ostwald, W. (1912) *L'évolution de l'électrochimie* (E. Philippi, trad.). Paris, Francia: Librairie Felix Alcan. (Obra original de 1910)
- Volta, A. (1923). Lettera del prof. Alessandro volta al prof. Brugnatelli sopra alcuni fenomeni chimici ottenuti col nuovo apparecchio elettrico. Autunno 1800. En *Le opere di Alessandro Volta. Edizione Nazionale*, Vol. 2 (pp. 2-5). Milán: Ulrico Hoepli Editore. (Obra original de 1800)



---

### **3. Historia y filosofía de las ciencias cognitivas**







# Una revisión de la actualidad de la atribución mental

## Perspectiva de segunda persona

Magalí Argañaraz\*

### Cartografía de la atribución mental

Aunque las reflexiones en torno a la comprensión del otro, de sus pensamientos y acciones, podrían rastrearse desde la antigüedad, por cuestiones materiales y temporales, la narración historiográfica que aquí planteo integra los desarrollos en filosofía de la mente respecto a los debates en atribución mental desde los años '80 hasta la actualidad. Antes de continuar, creo necesario aclarar que esta presentación no busca ser exhaustiva ni excluyente, sino más bien introductoria.<sup>1</sup>

Hasta los '70 la teoría predominante y mejor consolidada en la explicación del modo en que atribuimos estados mentales a otros y a nosotros mismos fue la teoría de la teoría proveniente del funcionalismo. La teoría de la teoría, y sin querer incurrir en especificaciones, sostiene que la posibilidad de que alguien pueda decir de otro que tiene creencias, deseos, emociones, intenciones (estados mentales), supone que tal posea un repertorio y competencia conceptual respecto de tales términos a partir de los cuales poder inferir, esto es, poder describir, explicar y/o predecir los estados mentales del otro y sus acciones correspondientes. Aquellos autores que sostienen esta teoría, comparten las siguientes tesis (Brunsteins, 2010):

- 1) La teoría del sentido común (esto es, la que poseen los sujetos para atribuir estados mentales) es ampliamente tácita, el acceso a la mayoría de la información que se utiliza en la explicación y predicción no es consciente, no es accesible de un modo introspectivo.
- 2) Los estados mentales postulados en la teoría, tienen contenido, lo que significa que versan sobre lo que alguien piensa, cree, desea o imagina acerca de algún rasgo del mundo.

<sup>1</sup> Para un estudio más completo véase Brunsteins (2010).

\* Universidad Nacional de Córdoba  
arganaraz.mgl@gmail.com

- 3) Los estados mentales se conciben como actitudes proposicionales, se dirigen al contenido de un enunciado proposicionales. El modo de dirigirse puede variar, con lo cual podemos tener distintitas actitudes ante el mismo contenido.
- 4) Los estados mentales desempeñan un rol explicativo.
- 5) Los estados mentales son funcionalmente discretos.

Ahora, estas tesis que fueron establecidas por diferentes razones, si bien en la actualidad poseen cierta fortaleza, tienen el problema de que explican las atribuciones mentales en un adulto normal, quedando excluidas las niñas, la naturaleza y dinámica de las enfermedades mentales, el sueño, las operaciones psicológicas automáticas, las ilusiones perceptuales, entre otras.

En los '80 suceden algunos acontecimientos que moverán los debates de la atribución mental. Por un lado, pensadores como Alvin Goldman (1992) y Robert Gordon (1996) comienzan a pensar otra estrategia de la atribución mental, a saber, la simulación, la cual buscaba resolver ciertos problemas que se le presentaban a la teoría de la teoría. Por el otro, en esos años se formuló el experimento "Lo que Mary ignora" (Jackson, 1986) que puso en cuestión la explicación funcionalista y trajo lo que actualmente se conoce como el problema de los *qualia*: si uno cree que Mary conoce efectivamente algo nuevo respecto del rojo, entonces la cuestión es ¿cómo describir lo que siente/sabe Mary?; y si uno cree que Mary no conoce nada nuevo, sin embargo debe asumir que hay ciertas propiedades fenoménicas que no son funcionales y que son, además, privadas. Finalmente, otros pensadores, siguiendo lo que dice Broncano (2013), recuperan lo que Wittgenstein decía respecto de que es un hecho gramatical que el sujeto es sujeto en primera persona: nadie puede afirmar "tengo su dolor de muela" porque nadie puede sentir los dolores del otro, aunque sí pueda entender y saber que el otro tiene dolor de muela. Así, en el contexto de aparición de la teoría de la simulación, aparecen concomitantemente dos dimensiones que reivindican la autoridad de primera persona, por un lado, la irreducibilidad de lo fenoménico y, por el otro, la irreducibilidad gramatical de la primera persona.

La teoría de la simulación, tiene como gran aporte a la explicación de la atribución mental el hecho de que la simulación es un mecanismo de atribución mental espontáneo y automático, a diferencia de la teoría del

sentido común. Ello permite sostener la ingenuidad teórica del simulador respecto de las teorías mentales o de las generalizaciones legaliformes, pudiendo así desprenderse de compromisos teóricos muy fuertes que asumía la teoría de la teoría y sin caer en reducciones al *infinitum*. La simulación sería un proceso o rutina cognitiva en el cual el sujeto análoga su propia psicología para comprender al otro. Ahora, aunque desde el punto de vista empírico puede rastrearse su punto de partida en los mecanismos imitativos, pareciera ser que desde la empiria también se presenta el hecho de que nuestras predicciones acerca de las conductas y estados mentales son penetrables cognitivamente, lo cual entra en contradicción con la consideración de la simulación (Brunsteins, 2010).

Por último, siguiendo nuevamente a Broncano (2013), vale mencionar las reflexiones de W.V.O. Quine y D. Davidson a partir del experimento mental conocido como “Contexto de interpretación radical”, por los aportes que ellas significaron en torno al debate de la comprensión del otro. Si Quine a partir de tal experimento diagnosticó la esencial inestabilidad del significado basado en la observación de la conducta, Davidson fue más allá, extrayendo consecuencias a nivel de todas las relaciones humanas en las que estén implicadas las interacciones mutuas de las acciones: si queremos comprender al otro (ya sean sus acciones, sus deseos o lo que sabe), debemos suponer que compartimos un mundo (“mi mundo es a grandes rasgos similar al suyo”, “mi mente es a grandes rasgos similar a la suya”, “mi conocimiento es a grandes rasgos similar al suyo”, etc.), esto es, debemos socializarnos, incluir al otro en nosotros. Así,

La tercera persona no es tan tercera, ni la primera tan primera. No puede existir una absoluta alteridad sin pagar el precio de la comprensión: el Otro absoluto es ininteligible. . . . En adelante, el antropólogo desencarnado estará condenado al peor de los destinos: la incompreensión radical. (Broncano, 2013, p. 98)

Ahora, aunque esta línea de pensamiento concuerda con ciertas ideas de la perspectiva de segunda persona, es preciso destacar que el enfoque davidsoniano preserva una visión teoricista de la atribución mental y una concepción de interdependencia entre la intencionalidad mental y la intencionalidad lingüística, de modo que solo de los sujetos lingüísticos puede decirse que tengan pensamientos y creencias (Scotto, 2002, p. 139).

## **Segunda persona**

En este contexto, a principios del S. XXI, la perspectiva de segunda persona aparece como una estrategia de atribución y comprensión mental que busca dar cuenta del modo en que conocemos a otros y a nosotros mismos afirmando: 1) que reconocemos –implícitamente– estados mentales en los demás de manera automática en las situaciones comunicativas cara a cara; 2) y este mecanismo cognitivo intersubjetivo es la perspectiva básica que le niño es capaz (en relación primordial con el adulto responsable) para acceder a la realidad mental (Gomila, 2001). A partir de aquí, y sin presentarse como una alternativa excluyente de las otras estrategias de atribución mental, se puede argumentar que la perspectiva de primera y tercera persona se desarrollan posteriormente y a partir de la de segunda.

Al ser una perspectiva nueva en el estudio de la atribución mental, al igual que en sus inicios la teoría de la simulación se formaba partiendo de sus diferencias con la teoría de la teoría, así también lo primero para comprender lo propio de la perspectiva de segunda persona es su diferencia con las otras dos perspectivas: la atribución que realizamos a otro no la hacemos como espectadores/observadores de las circunstancias que le afectan sino que es una atribución que realizamos en un contexto particular en que somos agentes y en donde nuestra actitud depende de las del otro y, recíprocamente. Así, los propósitos de la atribución mental (en la segunda persona) no son metas epistémicas como predecir la conducta del otro (teoría de la teoría) o comprender lo que hizo (teoría simulacionista), sino que resultan de la necesidad de la interacción misma, saber cómo actuar, coordinar, corregir las propias expectativas, persuadir, ser aceptado o perdonado, etc., lo cual puede ser visto como conocimiento práctico (y no proposicional). En estas circunstancias la atribución no se guía ni por la inferencia de una teoría de la mente ni por la proyección de mí en las circunstancias del otro, antes bien lo hace por la percepción significativa de las señales (el rostro, los gestos, los movimientos corporales, la mirada, la atención, además de las señales lingüísticas) que expresan la actitud del otro en tales circunstancias. La perspectiva de segunda persona, luego, apunta a ser la capacidad de la que depende la intersubjetividad, esto es, la interacción significativa con otros. Por su parte, la estructura de estas interacciones es reflexiva (en estas atribuciones cada uno se da cuenta de que el otro se da cuenta de la reciprocidad de la atribución), práctica (la fun-

ción de las atribuciones es mediar la interacción) y comunicativa (la capacidad de atribución es la que puede activar un receptor en una situación comunicativa, lo cual presupone un emisor semejante) (Scotto, 2002).

Lo segundo es que, en la perspectiva de segunda persona, la atribución en la interacción es automática o espontánea, lo cual vuelve difícil considerarla como una atribución intencional principalmente porque tanto las teorías de la teoría como las teorías simulacionistas consideran que la atribución se realiza de manera consiente y deliberada. El punto aquí es que tradicionalmente se ha supuesto que el sujeto capaz de atribución es el adulto –cuando no este adulto ha sido interpretado como científico o filósofo–, mientras que de lo que la perspectiva de segunda persona da cuenta es que ella es la perspectiva básica que se encuentra en niños humanos como en animales no humanos a partir de la cual ellos pueden ir adquiriendo las demás perspectivas.

Lo tercero es que el ámbito de aplicación de esta perspectiva son las intenciones comunicativas que se producen en las interacciones comunicativas, para lo cual es necesario reconocer las funciones comunicativas del paralenguaje y de la gestualidad, porque son las primeras adquisiciones de los niños. La interacción emocional (cuya estructura es semejante a la de la segunda persona) es el primer modo de relación que le niño conoce y que es una fuerte motivación a la interacción diádica y del desarrollo de la capacidad para distinguir configuraciones faciales y responder en consecuencia. En conjunción con estos modos comienzan a desarrollarse los gestos comunicativos como la ritualización y, posteriormente, los gestos simbólicos (icónicos) y los deícticos. Ahora, las expresiones emocionales, así como el resto del repertorio gestual y facial es diverso, situacional y social, esto es, corresponden a un contexto particular y son expresiones que no son simples reflejos irrefrenables sino hábitos adquiridos socialmente (Gomila, 2001). Y, entonces, la apuesta es que la comprensión de las expresiones y los gestos dependen de la activación de la perspectiva de segunda persona.

Ahora, principalmente en torno a la emocionalidad presente en la comprensión del otro y de sí, es que la teoría de la simulación y la perspectiva de segunda persona entran en un campo de disputa: ambas afirman esa emocionalidad, ambas afirman ser procesos espontáneos y ambas buscan ligarse al razonamiento práctico. En este sentido, sería enriquecedor en el futuro poder establecer diálogos entre la perspectiva de segunda persona y los desarrollos contemporáneos de la empatía, la cual se presenta

como una estrategia de atribución mental que daría cuenta de manera más acabada de lo que en sus comienzos se planteó con la simulación, pues integraría tanto aspectos cognitivos como emotivos (Brunsteins, 2017).

Un caso a destacar, y que mostraría cómo la perspectiva de primera surge a partir de la de segunda, es el hecho de que muchos teóricos de la evolución afirman que no nacemos siendo conscientes de nosotros mismos y que el fenómeno de la atención conjunta resulta fundamental para el surgimiento de esta (Ferrer de Luna, 2014). Los bebés, ya a muy temprana edad, se involucran en formas de relación interpersonal con la madre (padre o adulto responsable) que divergen radicalmente con la forma en que se relacionan con objetos inanimados; y antes del primer año de vida, el bebé es capaz de coordinar acciones y actitudes interpersonales con su cuidador. Gomila sostiene que no es para nada descabellado pensar esta interdependencia entre el niño y quienes le cuidan, pues desde nuestro nacimiento dependemos de otros para sobrevivir (Gomila, 2001, 2008). El problema principal es la “brecha ontogenética” que separa al niño del adulto, la cual puede ser superada al pensar que son las respuestas intencionales mediadas por una discriminación perceptiva la que precede a los conceptos intencionales y a la discriminación lingüística, y que mediante esta discriminación perceptiva sensible al contexto se da un tipo de interacción comunicativa intencional no lingüística. La posibilidad de tal afirmación es que percibir la atención no es otra cosa que percibir un estado que apunta a algo diferente del que atiende, sin importar en primera instancia, el contenido y que, posteriormente, el niño capta el objeto de atención mediante el seguimiento con la mirada de la mirada del adulto hacia el objeto de atención. Esto da cuenta de una cierta predisposición del niño a percibir la atención de su madre/padre y esta predisposición supone una captación de perspectiva de segunda persona. Luego, antes de que el niño perciba un objeto externo a la madre/padre y a sí mismo, lo que percibe es que la atención de la madre/padre se dirige a sí mismo, siendo así el objeto de atención. De modo que, algunos teóricos sostienen, que la conciencia de sí mismo no deriva del concepto de sí mismo o de Yo sino de una percepción temprana de la atención y las actitudes de otros dirigidas hacia uno, lo que en otras palabras significa que la perspectiva de segunda persona es la que hace experimentar el sí mismo como objeto para los demás (Ferrer de Luna, 2014).

## Para seguir pensando...

A modo de cierre, quisiera retomar las críticas y observaciones que Rabossi (2003) hace a Gomila para distinguir su propuesta o bien como una complementación a las otras dos perspectivas o bien como una alternativa a las mismas en tanto reconsideración de la concepción del sujeto (de las distinciones subjetivo/objetivo, interno/externo). Y al respecto Gomila afirma que su intención es ser una alternativa y, cito:

Este estímulo me llevó posteriormente a tratar de mostrar que los teóricos de la teoría y de la simulación se mueven en los mismos términos dialécticos en que se plantea el problema clásico del conocimiento de las otras mentes. (Gomila, 2008, p. 164)

Esto es, el problema del conocimiento de las otras mentes, lo enfrentan de manera directa cualquier perspectiva de primera persona (racionalista o emotivista) que conceda al sujeto la posición privilegiada de fundamento ontológico y epistémico, y donde se suele caer en los argumentos por analogía (le otre es definido como algo que es como Yo pero que no soy Yo); pero, las filosofías que parten de perspectivas de tercera persona, si bien dejan de tener el problema (ontológico) de cómo conocemos a otras mentes, pues explican las relaciones de manera causa-efecto, el precio que pagan es la pérdida del sujeto agente.

Luego, la idea de que la perspectiva de segunda persona, desde el punto de vista ontogenético (sin caer en la ontología), busca ser la perspectiva básica en tanto que sería la primera adquirida por los niños a través de la cual ellos entran al mundo de la vida mental, desarrollando sus capacidades cognitivas (poder reconocerse a uno mismo, poder reconocer a otros, poder atribuirse intenciones, emociones y atribuirle intenciones a otros, etc.). En esta línea, la perspectiva de segunda persona ofrece una alternativa al esquema tradicional, desplazándose del interés de explicar o comprender al otro –suponiendo una distinción sujeto/objeto, observador/observado– hacia la interacción entre dos sujetos que buscan relacionarse (comunicativa y emocionalmente). Ahora, también puede pensarse que modificar el esquema en el que se piensa la atribución mental no implica presentar como alternativa excluyente la perspectiva de segunda persona frente a las otras perspectivas, pues los mismos autores la presentan como básica de las otras dos, no como excluyente de ellas.

Pensar la atribución mental en un esquema gnoseológico no cartesiano, donde le individue no es un Yo que se encuentra con el mundo y con otros a los cuales debe justificar y con los cuales debe aprender a relacionarse, en primer lugar, significa salir del singular “El Hombre” y dejar de pensar al humano como un Yo siempre adulto que no nace. El cambio de paradigma invita a pensar a que los humanos nacemos y permanecemos inmersos en un mundo de cosas y de humanos, sin los cuales no sobreviviríamos, ni podríamos desarrollarnos: requerimos de otros que nos enseñen a caminar, a comer, a hablar y comunicarnos, a cuidar nuestro cuerpo, a escribir, a pensar, a jugar, a compartir, a habitar el mundo. Así, la intersubjetividad no sería una relación exterior de dos sujetos, sino que sería constitutiva de nuestra animalidad humana y la interacción intencional, intersubjetiva –que posibilita la creación de significados públicos, compartidos– desentendería de un mecanismo básico de comprensión social, que sería la de segunda persona.

## Referencias

- Broncano, F. (2013). *Sujetos en la niebla*. Barcelona: Editorial Herder.
- Brunsteins, P. (2003). Discusión. En A. D. Duarte & E. A. Rabossi (Eds.), *Psicología cognitiva y filosofía de la mente* (pp. 219-229). Madrid: Editorial Alianza.
- Brunsteins, P. (2010). *La psicología folk*. Buenos Aires: Ediciones del signo.
- Brunsteins, P. (2017). La empatía y su contribución en el ámbito de los derechos humanos. En G. Lariguet (Ed.), *La urdimbre de la razón: ensayos de filosofía teórica y práctica contemporáneos* (pp. 236-254). Mar del Plata: Kazak Ediciones.
- Goldman, A. (1992). *Liaisons: philosophy meets the cognitive and social sciences*. MIT Press.
- Gomila, A. (2001). La perspectiva de segunda persona: mecanismos mentales de la intersubjetividad. *Contrastes Revista Internacional de Filosofía. Suplemento VI: “Filosofía actual de la mente”*, 65-86.
- Gomila, A. (2008). La dimensión moral de la perspectiva de segunda persona. En D. Pérez & L. Fernández (Eds.), *Cuestiones Filosóficas: Ensayos en honor de Eduardo Rabossi* (pp. 155-173). Buenos Aires: Editorial Catálogos.

- Gordon, R. (1996). Radical simulation. En P. Carruthers & P. Smith (Eds.), *Theories of mind* (pp. 11-21). Cambridge University Press.
- Ferrer de Luna, J. (2014). El papel de la segunda persona en la constitución del auto-conocimiento. *Daimon Revista Internacional de Filosofía*, (62), 71-86.
- Escajadillo, C. (2018). La importancia filosófica de la perspectiva de la segunda persona: agencia y explicación intencional. En C. Monteagudo & P. Quintanilla (Eds.), *Los caminos de la filosofía, diálogo y método* (pp. 261-273). Lima: Fondo Editorial PUCP.
- Jackson, F. (1986). What Mary didn't know. *The Journal of Philosophy*, 83(5), 291-295.
- Rabossi, E. (2003). Discusión. En A. D. Duarte & E. A. Rabossi (Eds.), *Psicología cognitiva y filosofía de la mente* (pp. 230-237). Madrid: Editorial Alianza.
- Scotto, C. (2002). Interacción y atribución mental: la perspectiva de la segunda persona. *Análisis Filosófico*, XXII(2), 135-151.



# Mecanismos de herencia en la evolución cultural

## Sperber y la naturalización de la cultura

Malena León\*,\*

### Introducción

Los estudios sobre la evolución cultural se basan en la idea de que la cultura evoluciona de acuerdo con principios darwinianos. Sin embargo, no hay acuerdo sobre cuán lejos pueda llevarse la analogía entre la evolución cultural y la biológica, por lo que sigue abierta la cuestión acerca de cuáles serían los mecanismos específicos y las “unidades” de selección en este dominio (Acerbi & Mesoudi, 2015). Para algunos, la transmisión cultural sería un mecanismo de preservación que permite la selección entre diferentes rasgos culturales (Cavalli-Sforza & Feldman, 1981; Boyd & Richerson, 1985), para otros, un proceso de transformación en el que los individuos recrean esos rasgos cuando los transmiten. Esto último es lo que propone la teoría de Sperber (1996/2005) acerca de los *atractores culturales*; que hacen referencia a un cierto patrón de transformaciones en la transmisión de las representaciones culturales. En el primer grupo podríamos ubicar también a la controvertida teoría de la memética, que sostiene que no se trata de meramente una analogía con la selección natural de la teoría de Darwin para la evolución de las especies, sino que el mismo mecanismo opera en la cultura.

Respecto de esta variedad de enfoques, que difieren acerca de cuánto de la teoría darwiniana sobre la evolución de las especies debe ser extrapolado al cambio cultural, según Dennett (2010), hay un sentido en el que necesariamente toda teoría sobre la evolución cultural debe ser darwiniana. Sostiene que toda teoría de la evolución cultural que merezca más de un segundo de atención tiene que ser darwiniana en el sentido mínimo de ser consistente con la teoría de la evolución por selección natural del *Homo sapiens*:

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH) Centro de Investigaciones María Saleme de Burnichon (CIFYH). Córdoba, Argentina.

\* CONICET, Instituto de Humanidades (IDH). Córdoba, Argentina.  
malena.leon.m@gmail.com



El nombre de nuestra especie está bien escogido y es la cultura la que nos hace homínidos conocedores, por lo que una teoría mínimamente darwinista de la cultura debe sostener que los rasgos fenotípicos que hacen a la cultura acumulativa posible –a saber, el lenguaje y los hábitos de sociabilidad– han evolucionado por selección natural. (Dennett, 2006, p. 133)

De aquí en adelante llamaremos *darwinismo mínimo* al criterio que establece que las teorías en cuestión deben entender que los rasgos fenotípicos que hacen posible a la cultura acumulativa son el resultado del proceso de selección natural.

Después de establecer el “piso de acuerdo” para poder mantener una discusión significativa, queda abierta la cuestión de si las teorías en cuestión deben estar basadas en el hecho de que la transmisión de la información en las interacciones humanas implica siempre una transformación (como sostiene Sperber, junto con otros antropólogos cognitivos), o bien, deben apoyarse en un modelo de replicación y selección. La alternativa de la replicación y selección es defendida, entre otros, por Dennett (1990). Dennett considera que las teorías que se ubican en este grupo son más darwinistas que las otras, dado que sostienen que el cambio cultural se puede explicar por el mismo mecanismo de selección natural. Llamaremos a este requisito *darwinismo máximo*.

Ahora bien, se podría pensar que siendo la teoría de Dennett un buen exponente del *darwinismo máximo*, sería esperable que sea también una muy buena representante del *darwinismo mínimo*. Sin embargo, lo que me propongo mostrar es que, si comparamos las teorías de la memética dennettiana y la teoría de los *atractores culturales* de Sperber, la segunda parece dar cuenta de una mejor explicación en términos de *darwinismo mínimo*. Si bien Dennett sí cumple con el requisito en su versión básica, dado que su teoría admite que los rasgos fenotípicos que hacen posible a la cultura acumulativa son el resultado del proceso de selección natural, el único vínculo que guarda esto con las tesis centrales respecto de cómo se produce el cambio cultural es de compatibilidad. Para mostrar esto, en la primera parte de esta presentación voy a exponer sintéticamente la teoría dennettiana acerca de la memética, esto es, lo que el mismo Dennett entiende como una forma de *darwinismo máximo*. Luego, presentaré la teoría de Sperber, para, a continuación, prestar especial atención al modo en el que su teoría extrae implicancias concretas de la idea de que los seres humanos son productos de la evolución por selección natural. Con esto pretendo

concluir que, si bien Sperber cree que es inadecuada la extrapolación del mecanismo de selección natural al ámbito de la cultura, como propone el *darwinismo máximo*, su teoría constituye un ejemplo mucho más acabado y radical de *darwinismo mínimo*.

## Memética

El mismo Darwin (1877) señaló que el proceso de selección natural es neutral respecto del dominio. Se trata de un proceso que puede ocurrir siempre que se den ciertas condiciones abstractas, con independencia de que se trate de organismos biológicos. Algunas formas de nombrar estas condiciones son:

(1) variación, es decir, continua abundancia de elementos diferentes; (2) herencia o replicación, los elementos tienen la capacidad de crear copias de sí mismos; y (3) diferencias en aptitud, es decir, que el número de copias de un elemento creadas en un tiempo dado varía dependiendo de las interacciones entre las características de este elemento y las del medio ambiente en el cual persiste. (Dennett, 1995, p. 563)

Sobre estas bases, se puede sostener que en la cultura opera un proceso de selección natural literalmente, no es una mera analogía.

El primero en explotar esta vía de análisis, que luego sería retomada por Dennett, entre otros, es Dawkins (1976). Dawkins parte de una caracterización abstracta de la selección como un proceso que requiere de entidades que se reproducen, como los progenitores, que dejan descendencia que se les asemeja. La teoría memética adopta la visión de Dawkins de que las entidades que tienen la habilidad de hacer copias fieles de ellos mismos, llamados “replicadores”, son necesarios para explicar esta semejanza trans-generacional. En los modelos biológicos estándar se asume que los genes son los replicadores relevantes: hacen copias de ellos mismos y explican por qué los organismos de la descendencia son semejantes a los progenitores.

Ahora bien, para aplicar el mismo principio a la cultura, habría que encontrar un replicador cultural que haga lo mismo: para eso se postula la existencia del meme. Un meme sería la unidad de la selección cultural, el elemento más pequeño que se replica a sí mismo con fiabilidad y fecundidad (cfr. Dennett, 1995, p. 564). Ejemplos de memes son: canciones, ideas, frases pegajosas, modas de ropa. También pueden ser técnicas. La idea

central de la memética es que las ideas (o técnicas) pueden ser conceptualizadas como replicadores que van de mente en mente (o, más en general, de huésped en huésped), realizando copias de ellas mismas. Si bien la mayoría de los que predominan serían beneficiosos para sus huéspedes, podría haber algunos que no lo sean y, aun así, sean buenos replicadores. ¿Cómo sucedería esto? Según Dennett, algunos memes tendrían la propiedad de hacer su propia replicación más verosímil, con independencia de si son beneficiosos para sus huéspedes. Lo que comparten es que producen el efecto de desactivar las fuerzas selectivas que irían en contra de ellos. Algunos ejemplos son: las teorías conspirativas, que interpretan la evidencia en su contra como evidencia de que la conspiración es muy poderosa; los mensajes al final de una carta que solicitan que la carta sea copiada, o, de lo contrario, caerá una maldición en la persona que no lo haga; la idea de fe, que implica que la irracionalidad de alguna de sus tesis no es un problema, dado que justamente en eso consiste la fe.

La teoría memética ha sido criticada mucho y por distintos motivos. Las principales críticas se podrían reunir en torno a dos cuestiones: la de si en realidad existe algo así como los memes (algo que sea análogo al gen para el caso de la cultura) y la de si la teoría sirve para explicar algo. Para nuestros propósitos, basta con lo expuesto para ilustrar qué sería para Dennett el *darwinismo máximo* –con el que contrapone el *darwinismo mínimo*–: una teoría que sostiene que la selección natural es el principal mecanismo evolutivo en todos los dominios, por lo que en cada uno de ellos debe existir algo que funcione como replicador.

Ahora bien, ¿cómo satisface esta teoría el requisito del *darwinismo mínimo*? En términos estrictos, lo satisface, dado que es “compatible” con el mismo. Sin embargo, las consecuencias teóricas que se siguen del principio y que son parte de la teoría memética son muy generales: la selección natural dio lugar a organismos con mentes, en donde pueden ser alojados distintos tipos de memes. Ahora bien, las características que tienen esas mentes por ser adaptaciones sólo están incluidas en las consideraciones de por qué algunos memes son más exitosos que otros, a partir de una única dimensión. Hay un gran supuesto básico: la racionalidad. Podemos suponer que algunos memes les va bien porque demuestran ser beneficiosos para nosotros y, comportándonos de modo racional, los seguimos propagando. Además, uno de los ejemplos que brinda Dennett de “buenos replicadores” es el de la fe, porque desactiva las presiones selectivas que podría tener en su contra. En síntesis: cumple con ser compatible con

el requisito. Las consecuencias teóricas que se siguen de él son mínimas: simplemente están vinculadas a la racionalidad que exhiben los agentes.

## **La teoría de Sperber sobre la evolución cultural**

Su propuesta consiste en que la cultura sea estudiada a partir de una epidemiología de las representaciones, dado que explicar la cultura es explicar por qué y cómo hay ideas que son contagiosas. La epidemiología hace un uso ecléctico de modelos explicativos: genética de poblaciones, ecología, psicología social y también puede elaborar modelos nuevos cuando sea necesario. Lo que los modelos epidemiológicos tendrían en común, que los distingue de las explicaciones holísticas, es que explican macro-fenómenos a escala de las poblaciones como un efecto acumulativo de micro-procesos que provocan acontecimientos individuales.

El objeto de estudio de la epidemiología de las representaciones son las cadenas causales constituidas por las representaciones mentales y públicas y el mecanismo según el cual los estados mentales de los organismos causan la modificación del medio (producción de signos) y estas modificaciones del medio causan la modificación de estados mentales de otros organismos humanos. Explicar una representación cultural es identificar los factores psicológicos y ecológicos que mantienen la cadena causal.

Sperber sostiene que los modelos que lo han precedido le han concedido un lugar muy limitado a la psicología, es decir, a las mentes humanas. Ahora bien, si su explicación va a realizarse a partir de los micro-procesos, es menester señalar que los micro-mecanismos que provocan la propagación de las ideas son psicológicos; más específicamente cognitivos. En ese sentido, señala que

la lección más obvia de los trabajos cognitivos recientes es que el recuerdo no es el almacenamiento a la inversa y la comprensión no es la expresión a la inversa. La memoria y la comunicación transforman la información. Por tanto, tratar las representaciones, mentales o públicas, como causas materiales entre otras causas materiales implica enraizar el estudio del pensamiento y de la comunicación en la psicología cognitiva. (Sperber, 1996/2005, p. 37)

En línea con lo anterior, una de las diferencias que tiene una epidemiología de las representaciones respecto de la epidemiología de enfermedades es que en la segunda la replicación es la regla y la mutación es la

excepción, mientras que, en la primera, la transformación es la regla y rara vez permanecen estables. Por eso, en el caso de la cultura lo que hay que explicar es la permanencia.

Su concepción acerca de las capacidades cognitivas humanas como transformadoras de la información no niega que existan permanencias ni estabilidad en la cultura humana. Antes bien, Sperber señala que hay una gran cantidad de regularidades transculturales. Esas regularidades son justamente lo que pretende explicar el concepto de *atractor cultural*: se trata de un constructo estadístico abstracto, que hace referencia a los patrones de transformaciones en la transmisión de las representaciones culturales.

Ahora bien, ¿en qué consisten las explicaciones que brinda Sperber? Cabe señalar que Sperber sienta las bases teóricas de una epidemiología de las representaciones, pero no se dedica a desarrollarla. Tampoco se trata de una explicación exhaustiva sobre cómo se deben buscar los modelos que operan en cada caso. En lo que queda voy a brindar algunos ejemplos que él da y que, a mi parecer, ilustran bien el tipo de explicaciones en las que está pensando.

Antes de hacerlo es necesario señalar algo respecto del objeto con el que trataría la epidemiología de las representaciones. Cuando Sperber dice que existen representaciones mentales y representaciones públicas, las está considerando según su concreción material. Sin embargo, las representaciones también pueden abordarse como objetos abstractos, esto es, en virtud de sus propiedades formales. Las propiedades formales pueden considerarse de dos maneras: como propiedades de objetos abstractos, considerados en sí mismos (“enfoque platónico”) o como propiedades que puede atribuirles y explotar un dispositivo procesador como lo es la mente humana (“enfoque psicológico”). En otros términos, las propiedades formales de las representaciones pueden considerarse como propiedades psicológicas potenciales. Tomar a una representación según sus propiedades formales psicológicas sería, por ejemplo, preguntarse por qué determinada representación (un relato de lo que pasó hoy en la bolsa) es más difícil de recordar y transmitir que otra (el cuento de Caperucita). Parece evidente que hay representaciones más difíciles de interiorizar, recordar y exteriorizar que otras, ¿a qué se debe? No es a la complejidad: “Caperucita” no es menos compleja que un número de 20 dígitos. Lo que hace que algunas lo sean no tiene que ver con un carácter intrínseco de sus contenidos, sino con las capacidades cognitivas y comunicativas humanas. Ejemplificaré con una de las explicaciones brindadas por Sperber.

Hay muchos ritos asociados a la mortalidad perinatal. Según Sperber (1996/2005), desde el punto de vista epidemiológico debemos entender a la recurrencia de un peligro de determinado tipo como un factor psicológico, que es capaz de estabilizar la práctica ritual. En ese sentido, “cualquier práctica concebida como defensa contra un tipo recurrente de peligro se reactualiza con regularidad merced a ese mismo peligro. Sin embargo, sólo es así si las personas interesadas tienen fe en la eficacia de la práctica.” (p. 53). En muchos casos podemos asumir que esa fe en la práctica se basa en la confianza en los ancianos. Sin embargo, sería de esperar que la aparición reiterada de desgracias que se están intentando evitar, tenga efectos sobre la fe en la práctica. Así, si la práctica es muy ineficaz, debería ir desgastándose progresivamente la fe en ella.

Sperber señala 4 tipos de casos que, puestos en consideración, deberían decir a las personas si su práctica es eficaz o no (p. 53):

- 1a. La práctica se ha seguido estrictamente y no se ha producido la desgracia;
- 1b. La práctica se ha seguido estrictamente y se ha producido la desgracia;
- 2a. La práctica no se ha seguido estrictamente y no se ha producido la desgracia;
- 2b. La práctica no se ha seguido estrictamente y se ha producido la desgracia.

En principio, la misma experiencia debería, después de un tiempo, llevar a las comunidades a notar que ciertas prácticas son ineficaces. Sin embargo, las mismas se mantienen. Hay dos alternativas: pensar que se mantienen de forma irracional o pensar que se sostienen por inferencias injustificadas. Sperber va por la segunda. Dice que, si bien está demostrado que muchos animales somos capaces de realizar inferencias espontáneas, muchas veces los cálculos de probabilidades se distorsionan de modo sistemático. El análisis que realiza Sperber señala que hay 3 razones para pensar que puede otorgarse una importancia excesiva a los casos de tipo 2b. (1) La desgracia siempre requiere una explicación, (2) cuando el incumplimiento estricto de una práctica viene seguido de la desgracia, da la sensación de haberla causado y (3) explicar una desgracia diciendo que fue causada por la conducta de ciertas personas hace posible señalar responsabilidades y dar, al menos, una respuesta social ante una situación crítica. En condiciones de ese tipo, seguir estrictamente la práctica en cuestión protege del riesgo social de ser culpado por provocar una eventual desgracia (sería una forma de eficacia).

Este sería un ejemplo de una disposición cognitiva, es decir de un factor psicológico. A su vez, el mismo interactúa con un factor ecológico: la frecuencia con la que se produzcan cada uno de los casos mencionados. Así, por ejemplo, si es una sociedad en la que la desgracia ocurre muy pocas veces, la práctica se puede haber ido diluyendo. Si se trata de una sociedad que en la que ocurre muchas veces, habría muchos casos de 1b y 2b. Dado que el 2b sería más tomado en cuenta y el mismo funciona mostrando la eficacia de la práctica, en esa comunidad la práctica seguiría estando arraigada. Todo esto podría probarse de modo experimental.

Sperber también predice que las prácticas ineficaces destinadas a impedir desgracias inevitables están condenadas a desaparecer (podríamos pensar que, si son todos casos 1b y 2b, la inferencia indica al final que es ineficaz). Según esta consideración, predice que no deben existir culturas que mantengan prácticas destinadas a evitar que la gente muy mayor muera.

Todo esto buscaría explicar por qué una práctica que es ineficaz, y que pretende proteger contra diversos tipos de desgracias, puede estabilizarse, aunque carezca de eficacia. Eso es sólo un factor de los múltiples que intervienen para explicar la permanencia de una representación en una cultura. Un paso siguiente sería explicar el contenido específico de esa práctica. Toma de los estudios interpretativos y estructuralistas antropológicos la consideración de que de hecho hay cierta coherencia entre los conjuntos culturales, al menos al interior de cada cultura. Para Sperber esa coherencia no explica nada, sino que es algo que hay que explicar.

Como dijimos, según su teoría, en el proceso de transmisión las representaciones se transforman, pero no al azar, sino que lo hacen con ciertas direcciones. Lo hacen en la dirección de los contenidos que requieren menos trabajo mental y producen mayores efectos cognitivos. Las representaciones se van transformando progresivamente, en la dirección de optimizar la razón entre efecto y trabajo. Dicha razón recibe el nombre de “relevancia” en la teoría de Sperber y Wilson (1986): “Sostenemos que los procesos cognitivos humanos se orientan a la maximización de la relevancia” (p. 109). Lo que ellos dicen, en su teoría sobre la comunicación humana, es que esta solo es posible porque hay cierto grado de semejanza entre los pensamientos del comunicador y los del auditorio. Desde este enfoque, en el que los procesos inferenciales adquieren un rol central, la replicación sin modificaciones es considerada un caso límite.

La teoría de los *atractores culturales* de Sperber está vinculada con la de la multimodularidad de la mente. Si bien no será desarrollada en esta instancia, consiste en la idea de que la mente está compuesta por una multitud de mecanismos procesadores de la información, que operan en dominios específicos de forma autónoma o casi autónoma (Tooby & Cosmides, 1992). En esto Sperber se apoya en determinados desarrollos de la psicología evolucionista, que sostienen que dichos módulos son el resultado de la acción de la selección natural.

¿Cómo se forman los conceptos en las mentes humanas? Apoya la hipótesis de que la ostensión puede dar origen a los conceptos básicos si antes hay una estructura lógica previa: se trataría de un esquema innato y una disposición a aplicar ese esquema cada vez que se presente oportunidad. Si los humanos tenemos una disposición a elaborar cierto tipo de conceptos, esos son los conceptos “contagiosos”.

La propuesta es que tenemos una disposición innata para elaborar conceptos de acuerdo a determinados esquemas. Tenemos diferentes esquemas para distintos campos (taxonomía para los seres vivos, caracterización por la función para los artefactos, etc.). Se trata de los conceptos básicos. Están presentes en todas las lenguas y son más o menos parecidos.

A diferencia de lo que la antropología socio cultural estándar ha estado haciendo en el último tiempo, Sperber y sus colaboradores han enfatizado la existencia de regularidades transculturales en dominios como la religión (Boyer, 2001), la clasificación social (Hirschfeld, 1998) y la biología folk (Atran, 1998).

En función de lo expuesto, es claro que para Sperber el hecho de que en el proceso de transmisión de información están involucrados seres humanos tiene implicancias teóricas de gran peso. Las mismas son no sólo compatibles con la teoría de la evolución darwiniana por selección natural, sino que están explicadas a través de ella: se trata de rasgos que podrían ser comprendidos como adaptaciones:

Las capacidades cognitivas humanas, genéticamente determinadas, son el resultado de un proceso de selección natural. Tenemos razones para suponer que son adaptativas, es decir, que han ayudado a la especie a sobrevivir y extenderse. Esto no quiere decir que todos sus efectos sean adaptativos. (Sperber, 1996/2005, p. 67)

Apoyado en esta concepción, propone una distinción, aunque a veces resulte confusa, entre disposiciones y susceptibilidades. Mientras que

las disposiciones son aquellas capacidades que se han seleccionado positivamente en el proceso de evolución biológica, las susceptibilidades son efectos colaterales de las disposiciones. Las primeras hallan las condiciones adecuadas para desarrollarse en el ambiente en el que se desarrollaron filogenéticamente; en cambio, las susceptibilidades sólo pueden desarrollarse como resultado de un cambio en las condiciones ambientales.

Valiéndose de esta distinción, Sperber utiliza el caso de las meta-representaciones, con el objeto de explicar algunos casos que resultan paradójicos si entendemos que las habilidades cognitivas de los seres humanos son, en general, adaptativas. Por ejemplo: la creencia de que Dios está en todas partes. La respuesta de Sperber consiste en distinguir entre el conocimiento empírico cotidiano (o creencias intuitivas) y las meta-representaciones (o creencias reflexivas). Las capacidades meta-representativas son muy importantes porque permiten, por ejemplo, dudar, pensar que algo es falso, etc. También nos permiten procesar información que no comprendemos por completo.

Las disposiciones meta-representacionales son muy útiles. Es difícil pensar cómo ocurriría el desarrollo ontogenético si un niño no pudiera retener información antes de entenderla del todo. Sin embargo, esta capacidad también puede generar susceptibilidades: que las mentes humanas sean invadidas por misterios conceptuales imposibles de resolver. Ambas son racionales, pero de distinto modo. Las creencias intuitivas deben su racionalidad a mecanismos innatos, perceptivos e inferenciales. Las reflexivas son racionales en virtud de su origen. Así, las creencias religiosas no se desarrollan a causa de una disposición, sino a causa de una susceptibilidad. No tenemos cerebros diseñados para creer cosas irracionales, pero sí para creer cosas que no entendemos del todo y para confiar en lo que nos dicen aquellos a quienes consideramos confiables.

Podríamos pensar que este modo de extraer consecuencias del *darwinismo mínimo* que tiene Sperber lo vuelve un determinista genético. Si bien algunas críticas a su teoría no lo han tenido en cuenta, es importante señalar que, para Sperber, los *atractores culturales* no son exclusivamente cognitivos. Si bien es cierto que Sperber dedica la mayor parte de su libro a intentar identificar estos factores cognitivos siempre señala que en estas cadenas causales intervienen factores psicológicos y ecológicos.

Ahora bien, ¿de qué modo responde Sperber al requisito del *darwinismo mínimo*? La teoría de Sperber entiende que las capacidades cognitivas humanas que participan de la evolución cultural son el resultado de la evolución por selección natural de formas muy específicas:

- (1) tendencia de la mente humana a transformar la información maximizando la relevancia. Heredados genéticamente según la teoría de la multimodularidad de lo mental.
- (2) herencia de conceptos básicos, muchos de los cuales, podemos suponer, son adaptativos (ej.: clasificar los artefactos técnicos por su utilidad)
- (3) capacidad de tener creencias racionales de diferentes formas. Una por mecanismos perceptivos e inferenciales. La otra en virtud de su origen. Esto tiene un fuerte sentido adaptativo, dado que mientras la primera es importante para la supervivencia por motivos obvios (conforma lo que solemos entender como “racionalidad”), la segunda sería muy importante para la adquisición de conocimientos más complejos en el desarrollo ontogenético

Por lo que podríamos decir que Sperber no sólo cumple con el requisito de ser compatible con el *darwinismo mínimo*, sino que es una versión muy radical del mismo: las consecuencias de entender a la mente como un producto evolutivo son muy específicas.

## Conclusión

Sin tomar partido por una u otra, en este trabajo he argumentado que la teoría de los *atractores culturales* de Sperber, que no cumple los requisitos impuestos por un *darwinismo máximo*, según Dennett, constituye un ejemplo mucho más acabado y radical de *darwinismo mínimo* que la teoría memética.

## Referencias

- Acerbi, A., & Mesoudi, A. (2015). If we are all cultural Darwinians what's the fuss about? Clarifying recent disagreements in the field of cultural evolution. *Biology & philosophy*, 30(4), 481-503.
- Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and brain sciences*, 21(4), 547-569.
- Boyer, P. (2001). *Religion explained*. Random House.

- Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.
- Dennett, D. C. (1990). Memes and the Exploitation of Imagination. *Journal of Aesthetics and Art Criticism*, 48(2), 127-135.
- Dennett, D. (1995). *Darwin's dangerous idea*. Londres: Penguin Books.
- Dennett, D. (2006). From typo to thinko: When evolution graduated to semantic norms. En P. Jaisson, & S. Levinson (Eds.), *Evolution and culture* (pp. 133-145). Cambridge, MA: MIT Press.
- Hirschfeld, L. A. (1998). *Race in the making: Cognition, culture, and the child's construction of human kinds*. MIT Press.
- Sperber, D. (2005). *Explicar la cultura [Explaining culture: A naturalistic approach]* (P. Manzano, trad.). Madrid: Morata. (Obra original publicada en 1996)
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.
- Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. En J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163-228). Oxford University Press.



## Máquinas que simulan máquinas El Homeostato como simulación de un sistema complejo

Andrés A. Ilcic<sup>\*,\*</sup>

A better demonstration can be given by a machine, built so that we know its nature exactly and on which we can observe just what will happen in various conditions. (We can describe it either as “a machine to do our thinking for us” or, more respectfully, as “an analogue computer”). One was built and called the “Homeostat”.

– W. Ross Ashby, *Design for a brain*, p. 97.

El objetivo de este trabajo es mostrar cómo el uso de modelos por parte de la cibernética estuvo mayoritariamente acompañado por una noción de máquina cuya propiedad fundamental es la de poder simular el comportamiento de un sistema. Ross Ashby es el personaje central. Se trata de una figura un tanto descuidada en la mayoría de las reconstrucciones de la cibernética, pero sin duda se trata de una de las mentes más profundas detrás del movimiento cibernético. En una nota autobiográfica de 1962, Ashby escribió sobre Ashby, contando que “desde 1928 Ashby ha prestado la mayor parte de su atención al problema de ¿cómo puede el cerebro ser al mismo tiempo mecánico y adaptativo? Obtuvo la solución en 1941 pero no fue sino hasta 1948 que el Homeostato fue construido para encarnar el proceso especial. . . . Desde entonces ha trabajado para hacer más clara la teoría de los mecanismos semejantes al cerebro [*brainlike mechanisms*]” (Ashby, 1962, p. 452, citado por Pickering, 2010, p. 98). La pregunta fundamental de Ashby se vuelve así la pregunta por los límites de las máquinas y, en el proceso, por su capacidad de servir como herramienta de estudio para comprender sistemas excesivamente complejos como el cerebro humano.

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH)  
Centro de Investigaciones María Saleme de Burnichon (CIFYH). Córdoba, Argentina.

\* CONICET. Córdoba, Argentina.  
ailcic@ffyh.unc.edu.ar

Como lo anticipa el título del primer libro de Ashby, el concepto de diseño va a ser central.<sup>1</sup> Esto es porque va a entender que aplicando el método cibernético al fenómeno que se pretende estudiar, se debería poder construir, material o abstractamente, una máquina que sea capaz de duplicar o simular el comportamiento de dicho fenómeno, sin que deban existir entre ellos semejanzas estructurales. Una verdadera solución al problema sólo puede, por lo tanto, aceptarse si es que en última instancia permite la construcción de un sistema artificial semejante:

Para ser consistentes con los supuestos hechos, hemos de suponer (cosa que el autor acepta) que una solución real a nuestro problema permitiría la construcción de un sistema artificial que, como el cerebro vivo, sería capaz de adaptar su comportamiento. Así, si este trabajo es exitoso, incluirá (al menos por implicación) especificaciones para la construcción de un cerebro artificial que, de manera similar, se autocordine. (Ashby, 1952b, p. 10)

*Design for a brain* (Ashby, 1952b) es un libro destinado enteramente a explicar con detalles un modelo mecánico particular del cerebro –o, quizás mejor aún, de alguna clase posible de cerebro– pero también es un libro que elabora las conexiones entre los modelos mecánicos y los modelos matemáticos ya que, para justificar su construcción, Ashby debe introducir al lector en numerosos detalles de la teoría de los sistemas dinámicos. De hecho, si no fuera por el título uno creería que el libro poco tiene de cerebros. Esto se debe a que primero que nada es un tratado de por qué el comportamiento a gran escala del cerebro en su entorno puede entenderse como un sistema dinámico, esto es, un sistema que evoluciona en el tiempo cambiando de estados. Es esta concepción de alto nivel, en el que los detalles físicos o estructurales del fenómeno quedan excluidos –su interpretación como una caja negra con entradas y salidas– la que le va a permitir diseñar una máquina que tenga, desde este nivel, la misma clase de comportamiento. (Ashby, 1952) es un libro en el que se pone en práctica el método cibernético para poder responder la pregunta que lo guía. La reflexión sobre el método y su justificación tienen su punto máximo en la *Introducción a la cibernética* (Ashby, 1956), un texto que es complementario a *Design for a brain* y con reflexiones mucho más profundas. Lo que me interesa rescatar de esta serie de reflexiones es la actitud frente a las máqui-

---

<sup>1</sup> La segunda edición del libro de 1960 se tradujo al español como *Proyecto para un cerebro: El origen del comportamiento adaptativo* (Ashby, 1960/1965).

nas como modelos de los fenómenos de la naturaleza que encarna Ashby y que reflejan el ideal de la cibernética, detrás de los cuales se encuentra la necesidad de simular un comportamiento mediante una máquina, entendida ésta última en sentido amplio. Son las máquinas, entonces, las que le van a permitir a la cibernética *poner a prueba* sus modelos. Esto es, van a ser el recurso empleado para determinar que un modelo de un sistema complejo no está eliminando mediante sus idealizaciones el conjunto de elementos cuya interacción se puede decir compleja y que es la causa del fenómeno que se pretende estudiar. Aquí yace una tensión del modelado ya que precisamente hay que *modelar* la complejidad (*i.e.* simplificarla de alguna manera) sin perderla del todo. Esto lleva a que la cibernética reconozca a la complejidad como un fenómeno en sí mismo, que debe ser estudiado científicamente y a postularse ella misma como uno de los métodos para poder estudiar dicha complejidad:

En el estudio de algunos sistemas, sin embargo, la complejidad no puede evadirse completamente. . . . Pero la ciencia actual está dando los primeros pasos en el estudio de la “complejidad” encarada como tema autónomo. La cibernética se destaca entre los métodos para lidiar con la complejidad. Rechaza las ideas vagamente intuitivas que adquirimos al emplear máquinas tan simples como el reloj despertado y la bicicleta, y se presta a construir una disciplina rigurosa sobre lo complejo. . . . La cibernética ofrece la esperanza de proporcionar métodos efectivos para el estudio y el control de sistemas que son intrínseca y extremadamente complejos. (Ashby, 1956, p. 5-6)

Aquí Ashby señala dos aspectos muy importantes. El primero, y sobre el que elaboro más adelante, es que la complejidad es siempre relativa a un sistema (en oposición a un objeto). El segundo es la interpretación clásica de las máquinas y sus limitaciones que recién con la aparición de la cibernética puede ponerse completamente en cuestión dado que se puede comenzar a pensar una clase diferente de máquina, cuyo funcionamiento no esté atado a un propósito “de bajo nivel” y, por lo tanto, que dicho funcionamiento no esté limitado por las intenciones del diseñador y los órganos que le dotó a la misma durante su fabricación. Uso el término “órgano” en el sentido cartesiano ya que debemos a Descartes una de las mejores formulaciones sobre la interpretación clásica de las máquinas. Esta concepción cartesiana de máquina es la que opera detrás de lo que Turing (1950) llamó la “objeción de Ada Lovelace” a la posibilidad de “inteligencia maquina”, ya que cuando Lovelace describía el motor analítico de Babbage dejaba en claro que el motor no podía inventar ni descubrir nada “nuevo”:

El Motor Analítico no tiene pretensión alguna de *originar* algo. Puede realizar cualquier cosa que *sepamos cómo ordenarle* que la realice. Puede seguir el análisis; mas no tiene poder alguno de *anticipar* ninguna relación o verdad analíticas. Su lugar es el de asistirnos en hacernos *disponible* aquello con lo que ya estamos familiarizados. (Lovelace, 1843, p. 689. Énfasis en el original)

Como suele ocurrir algunas veces, la “solución” a este problema o límite de las máquinas apareció poco tiempo después de esta afirmación, de la mano de la biología. La teoría de la selección natural, de hecho, puede leerse como una solución al problema de la originalidad por medios meramente mecánicos. Ahora bien, incorporar el método de la selección natural al diseño de las máquinas es algo que debió esperar casi cien años más. Ashby y Turing llegan a conclusiones similares aproximadamente al mismo tiempo y los dos ven a las máquinas como el recurso ideal para comprender los procesos naturales. Para Ashby, el cerebro humano debe realizar una tarea análoga a la de la selección natural para producir comportamiento adaptativo y no hay razón lógica alguna por la que dicho método no pueda implementarse mecánicamente:

espero mostrar que un sistema puede ser de naturaleza mecanística y pese a eso puede dar origen a un comportamiento que sea adaptativo. Pretendo mostrar también que la diferencia esencial entre el cerebro y cualquiera de las máquinas hasta ahora construidas no es otra que en que el cerebro hace considerable uso de un método que hasta el momento se utilizando muy poco en las máquinas. Espero mostrar que mediante dicho método podemos hacer que el comportamiento de una máquina sea tan adaptativo como queramos, y que es posible que con el mismo método pueda explicarse incluso la adaptabilidad humana. (Ashby, 1952b, p. 1)

La carga de la prueba estará, claro, en nuestra capacidad de crear máquinas que utilicen un proceso semejante al de la evolución, pudiendo utilizar una regla simple para generar “más diseño y adaptación que la regla usada para generarla” (Ashby, 1952a, p. 50). En el proceso de crear dicha máquina, Ashby da una interpretación de la teoría de la evolución en términos de sistemas dinámicos, la que le permite ver a los dos procesos como análogos. La máquina que logra fabricar Ashby en 1948 para implementar físicamente lo que ya había imaginado a principio de esa década fue bautizada con el nombre de “Homeostato” ya que su comportamiento se puede describir, parafraseando a Proust, como el de una “búsqueda de la homeostasis perdida”.

Desde el punto de vista técnico, el Homeostato es una máquina bastante simple cuya novedad radica en su diseño y, fundamentalmente, en el comportamiento que puede lograr pese a su simplicidad.<sup>2</sup> Estrictamente el Homeostato no es una máquina sino cuatro máquinas idénticas (llamadas unidades) conectadas entre sí, de manera tal que el estado de cada una de ellas afecta al estado de las otras. Cada una de las unidades es una pequeña torre metálica, arriba de la cual hay un imán suspendido, que es el foco de atención del dispositivo. El comportamiento del imán está afectado por el campo magnético que generan cuatro bobinas que se encuentran a su alrededor. La corriente de las bobinas proviene de cada una de las unidades, incluyendo una bobina conectada a la misma unidad que se puede entender como de auto-alimentación. Al frente de cada imán hay una pequeña cuba de agua en cuyos extremos hay un electrodo, uno a -2 V y en el otro extremo a -15 V, creando un diferencial. El imán está suspendido de una pequeña aguja, sobre la que pivotea un alambre. Uno de sus extremos se hunde en la cuba de agua, tomando el potencial de acuerdo al lugar en la cuba en el que se encuentre, mientras que el otro extremo del alambre está conectado a un triodo, de donde cada unidad emite su corriente de salida; así la intensidad de dicha corriente depende de la posición del alambre que depende, a su vez, del movimiento del imán. Una resistencia conectada al triodo asegura que cuando el imán esté justo en la posición central la corriente de salida sea nula. Por la forma en la que las unidades están conectadas, esta corriente de salida funciona como una de las entradas de cada una de las otras tres unidades. Esto hace que cada vez que la máquina es encendida los imanes se muevan por las corrientes de las unidades, pero el movimiento de los imanes causa diferencias en las corrientes, por lo que constantemente se generan nuevos movimientos.

Las unidades tienen distintos elementos con los que se pueden modificar las “reacciones” del aparato. Por ejemplo, la polaridad de la corriente de entrada a cada una de las bobinas puede ser invertida mediante un interruptor, mientras que, a su vez, cada una de las bobinas cuenta con un potenciómetro por el que se puede regular qué fracción de la corriente de entrada total llega finalmente a la bobina correspondiente. Algunas de las

---

<sup>2</sup> Este es otro *insight* de la cibernética, que está atado a corregir el sesgo detrás de la concepción cartesiana de que sólo lo infinitamente complejo puede crear algo infinitamente complejo. En realidad, un sistema extremadamente simple puede producir un comportamiento increíblemente complejo a tal punto que no se pueda reconstruir (esto es, modelar) el sistema original a partir de la observación del comportamiento.

configuraciones posibles llevan al sistema a un estado de equilibrio, en el que todos los imanes quedan en la posición central, lo que representa un equilibrio activo o dinámico ya que el sistema está en constante funcionamiento para permanecer en dicho estado y si se lo perturba de alguna manera busca volver a dicho estado. Otras de las configuraciones iniciales son completamente inestables, por lo que los imanes nunca quedan fijos en su lugar central.

La facultad de retornar al estado de equilibrio (que fue fijado “a mano” mediante la configuración inicial) tras recibir una perturbación es una propiedad interesante del aparato, pero no es la central. El comportamiento más interesante del Homeostato se da cuando su configuración no está fijada por los controles manuales, sino que está sujeta a la posición de un uniselector disponible en cada una de las unidades. Un uniselector es una clase de telerruptor o interruptor de paso (también conocido como relé paso a paso). Estos dispositivos electromecánicos permiten que una corriente de entrada sea realizada o direccionada a una de muchas posibles conexiones de salida, cuyo control está regulado por una serie de pulsos eléctricos, aunque también puede ser controlado por un motor rotario.<sup>3</sup> El uniselector de Ashby tiene 25 posiciones que fueron seleccionadas al azar, con los valores numéricos procedentes de un libro de números al azar, como los que se podían comprar en esa época. El patrón de la retroalimentación de las unidades está determinado por el valor de salida de los uniselectores de cada una de las unidades, lo que deja al Homeostato con un total de 390.626 combinaciones posibles. La bobina que acciona el uniselector sólo se activa cuando la corriente de salida de la unidad excede el nivel de umbral del relé al que está conectado. Cuando el Homeostato está regido solamente por los uniselectores, emerge un nuevo comportamiento del sistema ya que cada vez que el comportamiento en una configuración dada es inestable (lo que equivale a decir que el imán está lejos de la posición central), el uniselector varía su estado. Se puede decir que el sistema empieza a buscar su estado de equilibrio probando las distintas configuraciones posibles. Esta es la manera en la que Ashby describe el funcionamiento del Homeostato bajo este modo automático de operación:

---

<sup>3</sup> Este era el caso de los uniselectores utilizados por la “Bombe”, la computadora electromecánica de uso específico diseñada por Alan Turing para descifrar los mensajes que los alemanes codificaban mediante la máquina Enigma. Quizás el uso más popular de esta clase de relés fue en las viejas centrales telefónicas.

En otras palabras, la máquina empieza a cazar una combinación de configuraciones de los uniselectores que dan un sistema estable, esto es, que dan las retroalimentaciones internas correctas. Cuando encuentra una combinación con las retroalimentaciones correctas, la mantiene y demuestra que ha armado el sistema de retroalimentación que resulta en un mantenimiento coordinado de sus variables en un valor óptimo, como un ser vivo. El punto importante es que encuentra su propia disposición [*arrangement*] de retroalimentación, el diseñador meramente se limitó a proveerle de mucha variedad. (Ashby, 1949, p. 78)

Dicho mantenimiento de las variables del sistema dentro de un rango determinado es la forma en la que Ashby entiende a la homeostasis y, la noción propia de variable es esencial para la cibernética, ya que va a ser el recurso usado para determinar el sistema que está bajo estudio. Algo interesante que puede simularse en el Homeostato es la adaptación a un determinado ambiente, ya que se puede establecer la configuración de una o dos unidades manualmente y dejar que los otros dos sean controlados por el uniselector. Una vez encontrado un equilibrio, cualquier modificación manual de los valores puede ser visto como un cambio en el ambiente, al cual el sistema nuevamente debe adaptarse. Las desestabilizaciones también pueden ser no diseñadas o previstas, ya que se puede intervenir el aparato, por ejemplo, moviendo el imán manualmente, conectando mecánicamente dos imanes para que se muevan en conjunto, cambiarle el recorrido dentro de la cuba, etc. Bajo cualquier intervención no destructiva, el sistema vuelve a explorar el espacio de configuraciones disponibles hasta lograr un estado de equilibrio.

Ashby se pregunta si el Homeostato es un cerebro y responde que no lo es, que claramente está muy lejos de *ser* un cerebro. El principal punto en contra en la comparación con el cerebro de un mamífero es que el Homeostato no tiene forma de guardar una configuración a la que volver después frente a un estímulo; su falta de memoria le obliga a descartar toda adaptación para poder adaptarse a las nuevas condiciones que el medio ambiente le impone. Ahora bien, ya para Ashby a finales de los años 1940, la dificultad es menor ya que se puede solucionar con más unidades y pequeñas alteraciones en las conexiones. De todas formas, el Homeostato nunca pretendió *ser* un cerebro sino simplemente servir como un prototipo de una clase diferente de máquinas, para mostrar que uno de los comportamientos más indicativos de un cerebro –su adaptabilidad– es posible de realizarse mediante procesos meramente mecánicos. Es pre-

cisamente en esta adaptabilidad en la que Ashby ve el medio para lograr cualquier clase de *imitación* de un comportamiento inteligente:

La creación de un cerebro sintético requiere ahora poco más que tiempo y trabajo. Pero hay un punto que debe quedar muy en claro: un cerebro sintético adecuado debe desarrollar *su propia* inteligencia [*cleverness*] –no debe ser un mero loro. Sin importar qué tan sorprendente sea el comportamiento, siempre debemos preguntar qué tanto de dicho comportamiento ha sido forzado por el diseñador y qué tanto es contribuido por la máquina misma. (Ashby, 1949, p. 79. Énfasis en el original.)

Si bien el homeostato no es un cerebro, sí es una “prueba de concepto” de que una de las características principales del cerebro puede ser entendida maquinicamente, siempre y cuando se extienda la categoría de máquina para integrar esta nueva *clase* de máquinas, que van más allá de las máquinas clásicas o cartesianas porque no se limitan a llevar a cabo un procedimiento específico que heredan del diseñador, sino que pueden aprender y cambiar su configuración. Si no fuera por la mención del Homeostato, la frase siguiente pareciera estar hablando de un modelo de aprendizaje maquínico o automatizado [*machine learning*] contemporáneo, describiendo a grandes rasgos lo que se conoce como aprendizaje supervisado:

La otra clase de máquina, el Homeostato, está basado en principios diferentes. No necesita instrucciones detalladas, sólo necesita algún método por el que se le informe de la ocurrencia de movimientos no legales y de los jaques mates. Cómo debe hacer la máquina para evitar estas informaciones (retroalimentación) se deja que ella misma lo averigüe. (Las adaptaciones ya demostradas por el homeostato fomentan la confianza de que con sólo desarrollos menores la máquina va a tener éxito). (Ashby, 1949, p. 79)

Esto permite ver al Homeostato y al cerebro como instancias de una misma *clase de máquina* dentro del conjunto máquinas posibles, conjunto que no se agota en las dos categorías descritas hasta aquí, aunque sí lo hace para la cibernética temprana. Una máquina está en la misma clase si puede usarse para *simular* o *imitar* el comportamiento de una máquina que ya esté en esa clase, lo que vuelve a poner en escena al concepto de comportamiento y el de caja negra, ya que no se presume los detalles específicos de la estructura interna, sino que se entiende a la máquina abstraída de tales detalles (cf. Rosenblueth & Wiener, 1945). Esto es importante porque no sólo señala que debe usarse un método para estudiar al comportamiento

de las máquinas abstractas, el núcleo de la cibernética, sino que también deben considerarse la forma mínima en la que una máquina abstracta debe estar configurada, para lograr su comportamiento, lo que hace a su ubicación en la clase de máquina. Ahí yace el balance a lograr entre las explicaciones de alto nivel o comportamentales/funcionales y las de bajo nivel o estructurales, no en sentido específico, sino más bien amplio. Aquí también se puede observar parte de la razón por la que cibernética es la primera ciencia de la computación en sentido amplio, ya que, si bien al comienzo no trató específicamente de sistemas computacionales directamente, los recursos que desarrolló para el estudio de las clases de máquinas sirvieron de base para el estudio teórico de la computación. Esta manera de entender la relación entre una máquina y un fenómeno a estudiar también puede ser usada para iluminar algunos aspectos del debate filosófico contemporáneo sobre el uso de las simulaciones computacionales y de los modelos en ciencia, tema que escapa a las limitaciones de este trabajo.

## **Conclusiones**

Creo que no es exagerado decir que el Homeostato encarna la contribución filosófica más importante de la cibernética en tanto logra mostrar cómo la teoría de la selección natural, el gran avance de la biología del siglo XIX, puede tener lugar en el mundo de las máquinas, no sólo en el sentido de que los seres humanos inventan distintas máquinas entre las cuales los mejores diseños son seleccionados para ser usados y mejorados, sino también en el sentido en que la misma operación básica de la máquina, a su misma escala, puede hacer uso de un análogo de la selección natural para inventar su propio propósito –y método– de acción. Así, efectivamente, se pueden crear máquinas que simulan haber sido diseñadas y tener un propósito particular pese a que, estrictamente, fueron diseñadas para *no* tener un propósito, sino para buscarlo. La dinámica interna, abstractamente concebida, le da un rol fundamental al tiempo, otra característica de los procesos computacionales, que heredan de nuestra concepción de los sistemas dinámicos. Norbert Wiener utiliza el término “transitorio” para referirse a la falta de propósito de estas “máquinas darwinianas”, otro término que acuñó y que recomendó a Ashby para etiquetar a la clase de máquinas como el cerebro y el homeostato:

Así resulta que en la máquina de Ashby, como en la naturaleza de Darwin, aparece un propósito en un sistema que no fue construido para que lo tuviera, simplemente por ser transitoria la carencia de finalidad de acuerdo con su misma naturaleza. A la larga, el gran propósito trivial de la entropía máxima parecerá ser el más duradero de todos. Pero en las etapas intermedias un organismo o una sociedad de ellos tenderá a permanecer más tiempo en aquellos modos de actividad en los que las diferentes partes funcionan conjuntamente, según una estructura [*pattern*] que tendrá más o menos sentido. (Wiener 1954/1969, p. 36)

La clave está en que dicho “patrón significativo” de acuerdo al cual las partes se comportan para mantenerse alejado del propósito final de cualquier sistema, el estado de máxima entropía al que todo está condenado, no tiene que estar incorporado en el diseño, sino que puede ser creado por la misma operación de la máquina, pudiendo encontrar así otras formas de configuración para mantenerse en ese “equilibrio dinámico” lejos del estado de equilibrio termodinámico. Wiener nota, que la clase de máquinas que sugiere el Homeostato de Ashby es una contribución filosófica fundamental, no sólo porque cambia totalmente la concepción limitada de las máquinas que era propia de la época sino porque al hacerlo abre la puerta a una infinidad de nuevas posibilidades para la automatización de las tareas de la humanidad:

Creo que la brillante idea de Ashby del mecanismo arbitrario sin propósito que busca uno propio mediante un proceso de aprendizaje es no sólo una de las más valiosas contribuciones de nuestra época a la filosofía, sino además algo que conducirá a progresos sumamente útiles en la automatización. No sólo podemos introducir un propósito en las máquinas, sino que, en la inmensa mayoría de los casos, la máquina diseñada para evitar determinadas fallas de funcionamiento buscará por sí misma un propósito que puede llevar a cabo. (Wiener 1954/1969, p. 36)

El mismo Ashby ya había considerado la posibilidad de que estas máquinas puedan ser utilizadas para controlar aquellos sistemas que son extremadamente difíciles de predecir y controlar debido a que su naturaleza escapa a la comprensión de la humanidad en un determinado momento, lo que implica que dentro de las capacidades de la máquina estaría la de encontrar leyes de la naturaleza que todavía no conocemos y que quizás sólo podamos conocer mediante dicha clase de máquinas:

La construcción de una máquina que reaccione exitosamente a situaciones más complejas que las que pueden ser manejadas actualmente por el cerebro humano transformaría mucha de nuestras dificultades y perplejidades. Una máquina semejante podría ser utilizada, en el futuro distante, no meramente para conseguir una respuesta rápida a una pregunta difícil sino también para explorar regiones de sutileza y complejidad intelectual que de momento están más allá de los poderes humanos. Por ejemplo, los problemas políticos y económicos del mundo a veces parecen involucrar complejidades que van incluso más allá de las capacidades de los expertos. Quizás a una máquina semejante se le pueda suministrar enormes tablas de estadística, cantidades de hechos científicos y otros datos, de manera tal que después de un tiempo puede llegar a emitir como resultado una amplia serie de instrucciones complicadas, un tanto sin sentido para quienes tengan que obedecerlas pero que, sin embargo, lleven de hecho a una resolución gradual de las dificultades políticas y económicas gracias a su entendimiento y uso de principios y leyes naturales que a nosotros todavía nos son oscuros. (Ashby, 1949, p. 79)

Si bien puede parecer una solución un tanto drástica a algunos problemas, soluciones semejantes se están aplicando actualmente en dominios mucho más acotados mediante el uso de modelos que sugieren decisiones, siendo sus resultados producto de la aplicación de redes neuronales artificiales cuyo funcionamiento interno no nos es transparente, ni siquiera para quienes la implementaron. Quizás lo más importante a remarcar aquí es como ya Ashby señalaba que existen ciertos dominios en los que no se puede pretender soluciones simples y elegantes, sino que dada la naturaleza intrínsecamente compleja del fenómeno que se pretende estudiar o controlar, cualquier modelo que se considere un buen modelo para tales efectos va a ser un modelo a su vez intrínsecamente complejo. Si bien es posible crear modelos más sencillos y entendibles que puedan ser usados para comprender algunos aspectos del fenómeno, es necesario reconocer cuáles son los límites del modelado. El Homeostato representa un caso prototípico de modelado de sistemas complejos, mediante el cual se puede reconstruir el comportamiento de un sistema complejo, haciendo que sus descripciones sean equivalentes sólo en determinado nivel de abstracción. Reconocer dicho nivel de abstracción y lo que se gana y se pierde al fijar el mismo es parte esencial de cualquier práctica de modelado y es de vital importancia para justificar el conocimiento que se extrae de los modelos. Esto implica también poder acomodar dicho modelo en la jerarquía de modelos y cómo los mismos están conectados entre sí (cf. Rosenblueth & Wiener, 1945). Quizás la principal lección a extraer es que la relación *mo-*

*delo-modelo* es al menos tan importante como la relación *modelo-mundo*, y probablemente sea aún más importante que esta última. Después de todo es establecer relaciones entre modelos es precisamente lo que permite una simulación.

## Referencias

- Ashby, W. R. (1949). The electronic brain. *Radio Electronics, Special Television Number*(March), 77–80.
- Ashby, W. R. (1952a). Can a mechanical chess-player outplay its designer? *The British Journal for the Philosophy of Science*, 3(9), 44–57
- Ashby, W. R. (1952b) *Design for a brain*. London: Chapman & Hall.
- Ashby, W. R. (1956). *An introduction to cybernetics*. New York: J. Wiley.
- Ashby, W. R., (1965). *Proyecto para un cerebro: El origen del comportamiento adaptivo* (V. Sánchez de Zavala, Trad.) Madrid: Tecnos. (Obra original publicada en 1960)
- Lovelace, A. K. (1843). Notes on ‘Sketch of the analytical engine invented by Charles Babbage, by LF Menabrea, officer of the military engineers, with notes upon the memoir by the translator’. En R, Taylor (Ed.). *Scientific memoirs*, 3 (pp. 666–731). London: Richard and John E. Taylor.
- Pickering, A. (2010). *The cybernetic brain: Sketches of another future*. Chicago; London: University of Chicago Press.
- Rosenblueth, A., & Wiener, N. (1945). The role of models in science. *Philosophy of Science*, 12(4), 316–321. <https://doi.org/10.1086/286874>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(October), 433–60.
- Wiener, N. (1961). *Cybernetics or control and communication in the animal and the machine* (2<sup>nd</sup> ed.). MIT press.
- Wiener, N. (1969) *Cybernetics y sociedad* [*The human use of human beings: Cybernetics and society* (2<sup>nd</sup> ed.)] (J. Novo Cerro, trad.). Buenos Aires: Sudamericana. (Obra original publicada en 1954)



# Prácticas de vinculación disciplinar en las neurociencias cognitivas contemporáneas

Agustín F. Mauro\*

A. Nicolás Venturelli\*,‡

## Introducción

Las neurociencias se han convertido en una de las disciplinas científicas de mayor crecimiento en las últimas décadas. En diversos niveles de análisis y con una gran variedad de técnicas, miles de científicos intentan desentrañar las capacidades de los cerebros humanos y no humanos. Consecuentemente en los últimos años se ha dado una proliferación de neuro-disciplinas. Desde las clásicas neurobiología y neuropsicología hasta disciplinas nuevas como la neuroantropología, neurociencia educacional, neuroeconomía o *neuro-law* (neuro-criminología). Las neurociencias cognitivas (NCC) han jugado un rol crucial en ese proceso dada la poderosa combinación de los modelos cognitivos del procesamiento de la información y las técnicas no invasivas de neuroimagen. En la mayor parte de los casos estas nuevas disciplinas suponen la confluencia entre las NCC y otra disciplina ya establecida. Por lo tanto, este avance de las neurociencias supone la creación y transformación de vínculos disciplinares tanto dentro de las NCC como entre otras disciplinas y las NCC. Este proceso supone una negociación constante de ontologías, fronteras disciplinares, competencias, proyectos de investigación, criterios epistémicos y no epistémicos, etc., convirtiendo los vínculos disciplinares en un problema central para los científicos.

Como es esperable, los vínculos disciplinares también son relevantes para la epistemología y la filosofía de la ciencia. Estos procesos de colaboración científica dan lugar a la creación de nuevos temas de investigación, nuevas metodologías, determinan la división del trabajo cognitivo, tienen consecuencias sobre las características del conocimiento producido (porque el proceso determina el producto), y porque todos desean o afirman

\* CONICET, Instituto de Humanidades (IDH). Córdoba, Argentina.

‡ Facultad de Psicología, Universidad Nacional de Córdoba. Córdoba, Argentina.  
agustinfmauro@gmail.com



realizar investigación interdisciplinaria, pero ha sido muy difícil caracterizarla de modo satisfactorio (Frodeman, 2010; MacLeod, 2018).

En este trabajo nos interesa hacer una propuesta para caracterizar algunos elementos de las prácticas de vinculación disciplinar en las NCC contemporáneas. De modo que podamos aportar a los debates en filosofía de la ciencia sobre la interdisciplina y que ayudemos a repensar los espacios de colaboración científica con las NCC. A este fin, vamos a historizar y caracterizar algunas prácticas interdisciplinarias en ciencias cognitivas (CC) y en NCC, y posteriormente analizar críticamente las prácticas interdisciplinarias actuales en NCC.

### **Precisiones conceptuales sobre “vinculación disciplinar”**

Previo a desarrollar el trabajo, consideramos importante hacer una breve aclaración sobre qué entendemos por “vinculación disciplinar” y cuál es nuestro abordaje del tema. Existe una vasta literatura que discute los conceptos de “interdisciplina”, “multidisciplina” y “transdisciplina”. Si bien hay múltiples definiciones en juego y hay un reconocimiento generalizado de la necesidad de realizar proyectos interdisciplinarios, no se comprende adecuadamente en qué consistirían o cómo deberían hacerse. La literatura sobre el tema no es de gran ayuda, en la mayoría de los casos se establecen taxonomías que diferencian multi-, inter-, o trans- disciplina, u otros prefijos adicionales. Julie Thompson Klein se destaca en el ámbito de los estudios sobre interdisciplinaria y realiza una propuesta estándar con relación a las taxonomías de la interdisciplinaria.

Klein considera que un programa es multidisciplinario si yuxtapone disciplinas “pero, las disciplinas permanecen separadas, los elementos disciplinares mantienen sus identidades originales y las estructuras de conocimiento existentes no son cuestionadas”<sup>1</sup> (Klein, 2010, p. 17). Por otra parte, un programa se convierte en interdisciplinario cuando la integración y la interacción se realizan de modo proactivo, y explícitamente se reestructuran las perspectivas existentes. Por último, un programa transdisciplinario “surge cuando existe colaboración entre los expertos académicos y otros actores sociales que proveen de conocimiento local e intereses contextuales para buscar soluciones democráticas a problemas controversiales” (Klein, 2010, p. 25), por ejemplo, el caso de los problemas am-

---

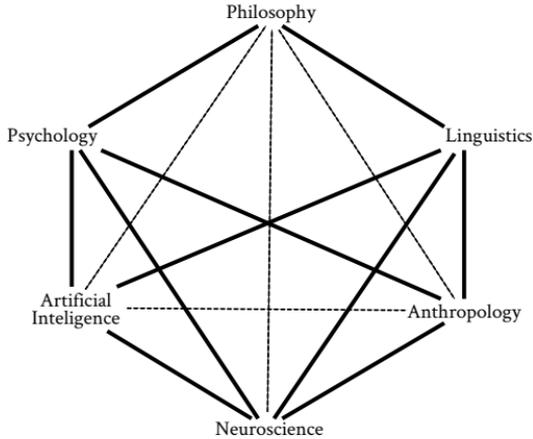
<sup>1</sup> Todas las traducciones son propias.

bientales. Más aun, su taxonomía propone las siguientes categorías para dar cuenta de la variedad de abordajes interdisciplinares: *Narrow ID*, *Wide ID*, *Shared ID*, *Cooperative ID*, *Methodological ID*, *Theoretical ID*, *Generalizing ID*, *Integrated ID*, *Conceptual ID*, *Bridge building*, *Restructuring*, *Unifying ID*, *Instrumental ID*, *Critical ID*, *Strategic* u *Opportunistic ID*. Si bien no podemos desarrollar estas categorías en detalle, consideramos que como tales no permiten dar cuenta de los procesos de vinculación disciplinar. La propuesta de Klein, es decir, las categorías “multi-, inter-, trans- disciplina” y su taxonomía carecen de dinámica, muestran la estructura de un vínculo disciplinar en un momento dado, pero no muestran el modo en que se transforma ese vínculo en el tiempo. Sin embargo, siguiendo la tradición de la filosofía de las prácticas científicas, consideramos que hay que entender la producción de conocimiento científico en clave histórica, más aún tratándose de un campo sumamente dinámico como el de interés aquí.

Por lo tanto, preferimos utilizar el término “vínculos disciplinares” como una categoría más general que sólo presupone la creación de algún vínculo entre disciplinas, pero sin caracterizar en qué consiste. Tomamos esta decisión por dos razones. En primer lugar, porque de este modo podemos utilizar una estrategia *bottom-up* para la caracterización de los regímenes de producción de conocimiento que están en juego. Es decir, podemos primero caracterizar el conjunto de fenómenos relacionados a los vínculos disciplinares y posteriormente realizar un trabajo conceptual para diferenciar exactamente el tipo de vínculos disciplinares que hay en juego. En segundo lugar, esta estrategia nos permite capturar los aspectos dinámicos de los vínculos disciplinares, carentes en la propuesta de Klein. En la conclusión se podrá observar que mediante esta estrategia pudimos recuperar la importancia de los instrumentos científicos en el ordenamiento de los vínculos disciplinares, y encontramos además dos dimensiones de análisis de la disciplinariedad: vínculos entre programas de investigación y vínculos entre diferentes formas de experticia. Estos aspectos de los vínculos disciplinares permanecen ocultos con las categorías de Klein.

## Interdisciplina en ciencias cognitivas

El planteamiento programático de los años inaugurales de las CC fue muy ambicioso en diversos aspectos, pero muy especialmente en lo que hace a los ideales proyectados de cohesión e integración disciplinar. La expresión



**Figura 1:** Hexágono de la ciencia cognitiva.

*Nota:* Imagen modificada de la cubierta del *Reporte de la Fundación Sloan* de E. Walker, 1978.

original, ciencia cognitiva, explícitamente en singular, pretendía remarcar estos ideales. El famoso hexágono de la ciencia cognitiva, presentado en varias ocasiones y originalmente en el conocido reporte de la *Sloan Foundation* de 1978, pretendía, entre otras cosas, expresar la idea de un campo unificado erigido sobre sólidos fundamentos teóricos que hacían posible el trabajo conjunto de sub-disciplinas dispares como la inteligencia artificial, la antropología o la lingüística. Esta es una visión muy enraizada, defendida tanto “desde adentro” por científicos activos como por historiadores y filósofos de la ciencia. Como un ejemplo de los primeros, podemos citar a Walker (1978):

la trama fuertemente articulada de interconexiones entre estos sub-dominios es la que sustenta la tesis de que ha surgido una ciencia autónoma de la cognición en la última década. [L]a red de disciplinas en interacción que aquí se muestra debiera concebirse en su conjunto bajo el nombre de ciencia cognitiva. (p. 75)

La historiadora M. Boden, por su parte, refiere explícitamente a su carácter interdisciplinario, haciendo hincapié sobre la relevancia de un plano teórico de acuerdo:

La historia de la ciencia cognitiva está marcada por una profunda y continua interdisciplinariedad. Esta es una relación intelectual más íntima que

la mera multidisciplinariedad. Una y otra vez, los investigadores en un área han tomado *ideas teóricas*, no sólo *datos*, prestadas de otra área. (Boden, 2006, p. 12)

El filósofo P. Thagard también sigue la misma línea, arriesgando una versión aun más explícita sobre el modo en que el carácter interdisciplinario de las CC ha tomado forma:

Sólo es útil para investigadores de diferentes disciplinas tratar de colaborar si hay poderosas ideas teóricas que cruzan los límites disciplinares. Para la ciencia cognitiva, las ideas integradoras principales han sido representación y computación, que pueden iluminar la naturaleza del pensamiento de modos que son útiles para todos los campos de la ciencia cognitiva –psicología, neurociencias, inteligencia artificial, filosofía, lingüística y antropología. (Thagard, 2010, p. 241)

Desde una perspectiva epistemológica, posiblemente el aspecto que más distanció las CC del conductismo es la prioridad otorgada al modelado teórico, atravesado por la computadora en su doble rol como fuente de heurísticas y como herramienta de modelado. Casi todas las líneas de investigación hasta fines de los años 70 fueron dominadas por asunciones teóricas relacionadas con la computadora digital como marco para abordar la cognición. En la mayoría de los modelos, el procesamiento consistía de operaciones discretas ejecutadas en serie, el componente de memoria era distinto del procesador y las operaciones de este podían ser descritas en términos de reglas del tipo de las de los lenguajes de programación. Además de las consabidas repercusiones en torno de la autonomía explicativa respecto del sustrato sobre el que los algoritmos se ejecutan, la computadora ofrecía un lenguaje –por ejemplo, la información en tanto que recibida, almacenada, recuperada, transformada, transmitida, etcétera– y una herramienta concreta para la elaboración de modelos. De forma indirecta, alimentó también un conjunto de métodos experimentales dirigidos a estudiar la estructura de etapas y sub-procesos computacionales del sistema interno de procesamiento: cabe mencionar técnicas de uso difundido como la medición de los tiempos de reacción (Sternberg, 1969) y el análisis de protocolo (de reportes verbales introspectivos), tal como fue desarrollado, entre otros, por H. Simon.

## Viraje hacia las neurociencias cognitivas

Si pasamos ahora a considerar el viraje que las CC dieron a partir de la década del 90, debemos pensarlo más bien en términos del nacimiento de las NCC como proyecto heredero de aquellas, pero en definitiva como un nuevo proyecto, con otros objetivos y con otras herramientas para alcanzarlos. Lo que dio gran impulso al campo fue, sin lugar a dudas, el desarrollo y la creciente accesibilidad de técnicas de neuroimagen funcional (Bub, 2000; Horwitz et al., 2000), que posibilitan la exploración de cerebros humanos, vivos y en actividad. Más allá del proyecto específico cuyo portavoz más visible ha sido M. Gazzaniga –ver, por ejemplo, Gazzaniga (1995) y la colección *Cognitive Neuroscience*– la noción de “neurociencias cognitivas” es utilizada a menudo para denotar un área vasta y diversa que comprende cualquier sector de la investigación neurocientífica que aborda fenómenos y categorías tradicionalmente asociados con la investigación psicológica. Las NCC de este modo combinan el uso de paradigmas experimentales y el marco teórico-conceptual de la psicología cognitiva con las técnicas de neuroimagen (resonancia magnética funcional, tomografía por emisión de positrones) y de registro electrofisiológico (microelectrodos, electroencefalograma). Tanto los abordajes experimentales como de modelado computacional han crecido y se han consolidado bajo el objetivo general de estudiar las bases neuronales de los distintos fenómenos descriptos por las distintas ramas de la psicología. En términos del mismo Gazzaniga (1995), las NCC comprenden el estudio del comportamiento y la cognición humana utilizando una combinación de métodos de las neurociencias y las CC, uniendo los niveles de explicación cognitivo y neuronal.

Ahora, en lo que hace a los posicionamientos en torno del tipo de vínculos interdisciplinarios que subyacen al (o que articulan el) campo, Boone y Piccinini (2016) recientemente han hecho explícita una postura clara y ponderada al respecto. Según estos autores, en el viraje antes descrito se ha dado lo que denominan una revolución neurocognitiva, expresión que usan para subrayar la idea de un cambio radical en el abordaje predominante de los fenómenos cognitivos. Identifican además dos principios sobre los cuales se asentaría esta nueva revolución: por un lado, el rechazo del supuesto de autonomía entre los niveles de descripción mencionados arriba por Gazzaniga y, por otro, un modelo de explicación alternativo,

esto es, la explicación por mecanismos en múltiples niveles (consistente en la integración de funciones y estructuras representacionales y computacionales a lo largo de múltiples niveles de organización en el cerebro). Por más que ambos principios sean de por sí criticables, queremos llamar la atención sobre un punto más general pero que a la vez entendemos como crítico.

Se trata de una idea que mancomuna tanto este posicionamiento por parte de Boone y Piccinini como la posición más o menos unificada que presentamos previamente en torno de cómo opera –u operó, dependiendo de la medida en que alguien acepte o no el relato revolucionario– la interdisciplina en CC. En ambos casos, de hecho, el posicionamiento sobre el modo en que se establecen los vínculos entre disciplinas dirigidas a un mismo objeto de estudio está atado a la idea de que existe un marco o un conjunto de principios teóricos de base que sostiene los intentos por explicar la cognición. Tanto el relato revolucionario de Boone y Piccinini (2016) como las ideas defendidas por los historiadores y filósofos mencionados comparten así el supuesto de la necesidad de un armazón filosófico/teórico para llevar adelante un programa de investigación interdisciplinario: dado lo que unos y otros ven como el exitoso establecimiento de un proyecto interdisciplinario, lo que se está afirmando es la existencia de un conjunto compartido de supuestos epistemológicos y metodológicos generales –como puede ser la idea de que existe un formato paradigmático de explicación en las NCC– que viabiliza las prácticas científicas. Es sobre esta plataforma común que, de acuerdo con esto, se puede articular la investigación dirigida a un conjunto de fenómenos tan complejo como lo es la cognición.

## **División del trabajo disciplinar**

Uno de los principales debates que se dieron ante la difusión de técnicas de neuroimagen como la tomografía por emisión de positrones y, especialmente, la resonancia magnética funcional (fMRI) fue el de la necesidad de modelos cognitivos para guiar la búsqueda experimental en las NCC. La idea aquí era que, por más poderosas o innovadoras estas técnicas pudieran ser, sin una manera de guiar su correcta utilización para explorar la actividad del cerebro, no iban a ofrecer más que cataratas de datos difícilmente aprovechables para responder preguntas en torno de cómo opera el cerebro humano. Un conocido artículo de Kosslyn (1999) se pre-

guntaba, en tono irónico: si las neuroimágenes son la respuesta, ¿cuál es la pregunta? Su intención era subrayar la necesidad de dirigir la búsqueda experimental mediante hipótesis específicas en términos de procesamiento de la información, volviendo a la vez a poner el foco sobre los paradigmas experimentales psicológicos en los que las tareas son diseñadas para aislar y controlar procesos cognitivos específicos, esto es, aquellos procesos blanco de la exploración mediada por neuroimágenes. Más recientemente, y en esta misma línea, Cacioppo y Decety (2009) afirman que los científicos cognitivos deberían dedicarse a desarrollar modelos teóricos rigurosos que especifican la estructura y los procesos de la mente, mientras que los neurocientíficos se deberían dedicar a la exploración del cerebro mediada por técnicas de neuroimagen.

Esta particular división del trabajo fue clave para cimentar la tradición localizacionista en las NCC, dominante en estudios con neuroimágenes funcionales. Un ejemplo representativo de esto ha sido, hasta aproximadamente 2005, el dominio del paradigma de sustracción en el campo, por el cual el foco está puesto en determinar las bases neuronales de un particular proceso cognitivo a través de comparaciones entre tareas para aislar ese proceso de interés. Está claro que un abordaje experimental de este tipo depende de una descomposición de un fenómeno cognitivo complejo en procesos más simples. El del paradigma de sustracción es presentado aquí solo como un caso ilustrativo, pero que cristaliza una determinada (y aparentemente nítida) división del trabajo disciplinar que, como es esperable en un campo dinámico como el de interés, fue modificándose en los años siguientes: hoy, la sustracción simple ha sido prácticamente abandonada.

La aludida reivindicación del rol específico de psicólogos y lingüistas de descomponer un fenómeno cognitivo complejo en operaciones que pudieran localizarse en el cerebro puede contrastarse con estudios actuales, caracterizados por el uso de técnicas variadas de análisis de datos, el abandono del paradigma de sustracción y en general una mayor autonomía respecto de los modelos cognitivos. Algunos cambios en las prácticas experimentales que vale la pena mencionar para ilustrar, de modo muy esquemático, los desarrollos ocurridos en los últimos 15 años son los siguientes: el crecimiento de la fMRI en estado de reposo (*resting state fMRI*); un creciente interés en fenómenos en los márgenes de la psicología cognitiva (*mind wandering, daydreaming*); un gran crecimiento de los estudios de conectividad funcional (dado especialmente el advenimiento

de técnicas como las neuroimágenes por tensores de difusión); el uso de diferentes técnicas para el análisis estadístico de los mismos *data sets*; la preponderancia del uso de técnicas multivariadas (donde el foco es puesto sobre la interacción entre áreas) por sobre de las univariadas (donde el foco es puesto sobre la activación promedio en una región, favoreciendo así los esfuerzos dirigidos a la localización de un determinado proceso cognitivo) y, finalmente, el crecimiento de estudios exploratorios (a través de técnicas diversas de *data mining*). Este proceso supuso un cambio en la división del trabajo disciplinar, especialmente en lo que respecta el rol de los paradigmas experimentales y modelos cognitivos provenientes de la psicología experimental.

### La preponderancia de la fMRI

Como se pudo observar, con el giro de las CC a las NCC la investigación y la organización de los vínculos disciplinares comienzan a girar alrededor de la fMRI, por lejos la técnica experimental más difundida. Esto es lo que decidimos llamar la preponderancia de la fMRI. Callard y Fitzgerald, dos investigadores en programas interdisciplinares en neurociencias, presentan dos situaciones que nos permiten ilustrar esta situación. En primer lugar, relatan que en un workshop para investigadores interesados en la intersección entre las neurociencias, las humanidades y las ciencias sociales, al hablar de interdisciplina en realidad se buscaba un tipo de interdisciplina particular: debía conseguirse un grupo que integre científicos sociales y neurocientíficos, y que realicen una investigación de tipo experimental. Esto ya muestra cómo se estructura la interdisciplina sobre la base de la metodología neurocientífica. Pero, en segundo lugar, comentan cómo el plan de trabajo del proyecto de investigación queda estructurado por la fMRI (o la técnica de neuroimagen elegida). Dicen los autores:

Tan pronto como el estudio con fMRI en neurociencias cognitivas se instala en los preparativos de la colaboración, los investigadores están consensuados por dos cronogramas particulares. El primero: ¿Cuándo estará disponible el escáner? (ésta es siempre la pregunta controvertida). Segundo: ¿Cuánto tiempo hay para encontrar sujetos experimentales para situar en el escáner? (Callard & Fitzgerald, 2015, p. 84)

Lo que señalan los autores es que la estructura de la colaboración interdisciplinar con las NCC siempre supone la realización de un estudio

de neuroimagen, que no existe la posibilidad para que la investigación de neuroimagen deba esperar los resultados de un método como la etnografía, o que un proyecto pensado para refinar los constructos de investigación en el estudio de neuroimagen termine determinando que un estudio de neuroimágenes no es el tipo de procedimiento experimental más apropiado para el tema de investigación (Callard & Fitzgerald, 2015). En otras palabras, la fMRI estructura qué se entiende por interdisciplina en las NCC y con las NCC. Se estabiliza, “cajanegriza”, una estructura de colaboración interdisciplinar que gira alrededor de situar sujetos experimentales en un escáner y obtener datos de la actividad cerebral.

Ahora bien, ¿qué hacemos con esta forma de interdisciplinariedad? El problema con esta forma de llevar a cabo los proyectos interdisciplinarios es que limita las posibilidades de la colaboración disciplinar en las NCC y con las NCC. Uno de los objetivos de la colaboración disciplinar es justamente innovar en los modos de realizar investigaciones, pensar nuevos problemas, crear nuevas estrategias de intervención. Sin embargo, la preponderancia de la fMRI predetermina los vínculos disciplinares, la división del trabajo disciplinar, el plan de trabajo, las preguntas viables, el tipo de resultado a obtener, etc.

Esta preponderancia ilumina dos dimensiones de análisis respecto de los vínculos disciplinares: la diferencia entre los vínculos entre programas de investigación y los vínculos entre diversas experticias. La dimensión de análisis de los programas de investigación surge cuando se evalúa la preponderancia de la fMRI en los vínculos disciplinares para pensar, por ejemplo, las múltiples neuro-disciplinas que han surgido en las últimas décadas. Estas neuro-disciplinas (neuro-educación, neuro-economía, neuro-estética, neuro-antropología, etc.) en la mayor parte de los casos suponen la vinculación entre las NCC y alguna disciplina establecida. Cabe preguntarse entonces ¿qué significa esta preponderancia de la fMRI en los vínculos disciplinares para pensar las múltiples neuro-disciplinas que están surgiendo?

Podemos hipotetizar que tiende a perderse variedad y riqueza metodológica si en todas las disciplinas involucradas se introducen sujetos experimentales en un escáner y se trata de inferir características del aprendizaje, la toma de decisiones, los juicios estéticos, etc. Por ejemplo, ¿qué proyecto es deseable para la neurociencia educacional y cuál es el lugar del fMRI en ese proyecto? La neurociencia educacional se propone vincular la investigación neurocientífica con la investigación educacional, pero to-

davía es una disciplina metodológicamente indefinida. Existen programas que enfatizan el uso de las metodologías e instrumentos de las NCC para estudiar fenómenos de interés en educación (Gabrieli, 2016) y existen programas que enfatizan la importancia de articular la experimentación en el aula, capaz de capturar fenómenos y problemas relevantes para los docentes, con la experimentación en el laboratorio (Coch & Ansari, 2012). En este sentido, la preponderancia de la fMRI estructura los proyectos que la neurociencia educacional puede perseguir y por lo tanto el tipo de conocimiento producido. Son vínculos disciplinares de este tipo los que mejor permiten problematizar la forma que adquiere la interdisciplina con las NCC y llamar la atención sobre dicha preponderancia.

Por otro lado, la preponderancia de la fMRI nos permite observar otra dimensión de los vínculos disciplinares. Es de esperar que un instrumento tan complejo requiera la articulación de múltiples experticias. Por ejemplo, investigadores en neurociencias, psicólogos u expertos en paradigmas experimentales y expertos en el tema de investigación, matemáticos, físicos o computólogos, entre otros. En ese sentido el instrumento exige equipos interdisciplinares. Podemos pensar en este tipo de vínculos como inherentes al uso mismo de la técnica, pero nada tiene que ver con la vinculación entre programas de investigación, sobre la que hemos hecho hincapié aquí.

En definitiva, el vínculo entre diversas disciplinas puede ser un vínculo entre programas de investigación o un vínculo entre experticias. Mientras que la primera dimensión de análisis puede ser puesta en tela de juicio y problematizada, la segunda nos permite sugerir que la producción de conocimiento neurocientífico en la actualidad requiere ensamblajes complejos (de disciplinas, programas de investigación, metodologías, experticias, instrumentos, modelos, etc.) y que no hay una única experte o disciplina capaz de manejar las diferentes variables en juego.

## **Conclusión**

Pudimos observar que la interdisciplina en las NCC y con las NCC actualmente gira en torno a establecer diferentes tipos de vínculos alrededor del uso de técnicas de neuroimagen. Es decir, las técnicas de neuroimagen estructuran gran parte de los proyectos interdisciplinares. Se generan novedosos vínculos disciplinares, pero la investigación termina gravitando en torno de la neuroimagen.



En función de este diagnóstico consideramos que se puede problematizar la práctica de los científicos y de los filósofos de la ciencia. Proponemos a los científicos atender a la articulación entre el programa de investigación, el diseño y planificación de la investigación, y los instrumentos, y cómo se adecúan mutuamente. Por ejemplo, si queremos vincular neurociencias y educación, ¿es necesaria la investigación de neuroimagen o tal vez sería necesario realizar primero una investigación en el aula? Si nos interesa el vínculo entre neurociencias y estética, ¿solo podemos preguntarnos en qué lugar del cerebro hay mayor activación cuando ocurre una experiencia estética?

Proponemos a filósofos de la ciencia comprender el problema de la interdisciplina o los vínculos disciplinares en términos dinámicos. Aproximaciones como las de Klein (2010) asumen un programa de investigación estático donde los vínculos ya están establecidos y no cambian con el tiempo. Nuestra propuesta muestra más claramente que de hecho cambian, por ejemplo, cuando surgen nuevos instrumentos o técnicas de investigación. Por otro lado, se propone a los filósofos de la ciencia que centrarse en los instrumentos y las metodologías asociadas puede iluminar cómo se construyen los vínculos disciplinares y puede contrarrestar las perspectivas teoricitas como las de Thagard (2010) o Boone y Piccinini (2016).

## Referencias

- Boden, M. (2006). *Mind as machine: A history of cognitive science*. Oxford: Oxford University Press.
- Boone, W., & Piccinini, G. (2016). The cognitive neuroscience revolution. *Synthese*, 193, 1509–1534.
- Bub, D. (2000). Methodological issues confronting PET and fMRI studies of cognitive function. *Cognitive Neuropsychology*, 17, 467–484.
- Cacioppo, J., & Decety, J. (2009). What are the brain mechanisms on which psychological processes are based? *Perspectives on Psychological Science*, 4(1), 10–18.
- Coch, D., & Ansari, D. (2012). Constructing connection: The evolving field of mind, brain and education. En S. Della Sala & M. Anderson (Eds.), *Neuroscience in education: The good, the bad and the ugly* (pp. 33–46). Oxford: Oxford University Press.

- Fitzgerald, D., & Callard, F. (2015). *Rethinking interdisciplinarity across the social sciences and neurosciences*. Hampshire: Palgrave Macmillan.
- Frodeman, R. (2010). Introduction. En R. Frodeman, J. Thompson-Klein, & C. Mitcham (Eds.), *The Oxford handbook of interdisciplinarity* (pp. xxix–xxxix). Oxford: Oxford University Press.
- Gabrieli, J. D. E. (2016). The promise of educational neuroscience: Comment on Bowers (2016). *Psychological Review*, 123(5), 613–619. <https://doi.org/10.1037/rev0000034>
- Gazzaniga, M. (Ed.). (1995). *The cognitive neurosciences*. Cambridge: MIT Press
- Horwitz, R., Friston, K., & Taylor, J. (2000). Neural modeling and functional brain imaging. *Neural Networks*, 13, 829–846.
- Klein, J. T. (2010). A taxonomy of interdisciplinarity. En R. Frodeman, J. Thompson-Klein & C. Mitcham (Eds.), *The Oxford handbook of interdisciplinarity* (pp. 15–30). Oxford: Oxford University Press.
- Kosslyn, S. (1999). If neuroimaging is the answer, what is the question? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 354(1387), 1283–1294.
- MacLeod, M. (2018). What makes interdisciplinarity difficult? Some consequences of domain specificity in interdisciplinary practice. *Synthese*, 195, 697–720.
- Thagard, P. (2010). Cognitive science. En R. Frodeman, J. Thompson-Klein & C. Mitcham (Eds.), *The Oxford handbook of interdisciplinarity* (pp. 234–245). Oxford: Oxford University Press.
- Walker, E. (Ed.) (1978). The Alfred P. Sloan Foundation state of the art report. [http://www.cbi.umn.edu/hostedpublications/pdf/CognitiveScience1978\\_OCR.pdf](http://www.cbi.umn.edu/hostedpublications/pdf/CognitiveScience1978_OCR.pdf)



# Las habilidades prácticas a la luz de la distinción personal-subpersonal

Sofía Mondaca\*.\*

## 1. Introducción

Las explicaciones psicológicas suelen estar divididas, al menos, en dos niveles jerárquicos diferentes: el nivel personal y el nivel subpersonal. La mayoría de los autores acuerdan en que tanto las características como los objetivos explicativos de las teorías que apelan a dichos niveles varían en múltiples puntos. No obstante, sólo algunos han defendido que las diferencias son lo suficientemente determinantes como para que sea posible entender dichos niveles de manera independiente. Esta tesis ha sido llamada “la tesis de la autonomía de la explicación de nivel personal”. Las discusiones al respecto han proliferado, y se la considera actualmente una tesis controvertida en el campo de la psicología y la filosofía.

En el marco de dicha discusión, los fenómenos caracterizados como saber práctico aparecen en disputa. En particular, las “habilidades prácticas”, entendidas en oposición a la caracterización ryleana del saber-qué o saber teórico, han sido un punto de debate. Aun cuando las habilidades prácticas sean reconocidas como un fenómeno paradigmático de nivel personal, algunos autores han sostenido que se vuelve inevitable ofrecer ciertas explicaciones de nivel subpersonal para una explicación completa del ejercicio de las mismas (Bermúdez, 2000).

En consecuencia, en el presente trabajo me propongo analizar el vínculo entre los niveles personal y subpersonal de explicación psicológica en relación al fenómeno de las habilidades prácticas, con el objetivo de mostrar cuál es la contribución de la tesis de la autonomía de la explicación de nivel personal al estudio de las mismas. Para ello, en primer lugar, intentaré sistematizar lo que comúnmente se ha entendido como el nivel personal y el nivel subpersonal de explicación psicológica. En segundo lugar, presentaré la tesis de la autonomía de la explicación de nivel per-

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH) Centro de Investigaciones María Saleme de Burnichon (CIFFyH). Córdoba, Argentina.

\* CONICET, Instituto de Humanidades (IDH). Córdoba, Argentina.  
ssofiamondaca@gmail.com

sonal y la crítica realizada por Bermúdez (2000) a la misma. Por último, señalaré ciertas dificultades que encuentro en la objeción y propuesta de Bermúdez, con el objetivo de argumentar a favor de la posición autonomista, señalando su contribución a la explicación del fenómeno de las habilidades prácticas.

## **2. La distinción personal-subpersonal**

Tanto Ryle como Wittgenstein fueron, probablemente, los primeros filósofos que sostuvieron una distinción entre niveles de explicación psicológica. De formas distintas, ambos autores discutieron con dos enfoques enfrentados, el cartesianismo y el mecanicismo, que disputaban en torno a cuáles eran las verdaderas causas de los estados mentales. El cartesianismo, como conocemos, defendía que los estados mentales no podían ser explicados del mismo modo que los estados físicos ya que eran ontológicamente diferentes. En discusión, los relatos mecanicistas sostenían que el sistema físico resultaba suficiente para explicar nuestra realidad. Suponían que sólo había un elemento (átomos, por ejemplo) que, a través de diferentes relaciones, producía todos los fenómenos experimentados. Por contraposición, se pensaba que la distinción entre niveles de explicación psicológica suponía una distinción entre dos tipos de entidades –las mentales y las físicas– que eran irreductibles unas a las otras. De tal modo, para la corriente mecanicista, dicho esquema impedía que las entidades mentales fueran explicadas científicamente y, por lo tanto, las ubicaba en un oscuro lugar metafísico.

En oposición a ambas posturas, Ryle sostuvo que debíamos distinguir entre un dominio de explicación acerca de lo físico y un dominio de explicación acerca de lo mental, sin con ello extraer implicaciones ontológicas de tal diferencia. De tal modo, asumió que explicar las entidades mentales en términos de entidades físicas implicaba cometer un “error de categoría”<sup>1</sup> y, para evitarlo, debíamos esforzarnos por dar una explicación mental de las mismas (Bermúdez, 2000). En consecuencia, Ryle defendió una comprensión del mundo de lo mental como el conjunto de las

---

<sup>1</sup> Cuando Ryle (1949) sostiene que la llamada “doctrina oficial” (esto es, los enfoques cartesianos que dominaron la filosofía desde el siglo XVII al siglo XX) comete un “error de categoría”, se refiere al malentendido que surge al pensar que la lógica de explicación de lo mental debe ser igual a la lógica de explicación de lo físico. Tal como el autor sostiene, utilizamos los conceptos fuera de su dominio propio, esto es, cuando aislamos los conceptos mentales y los cruzamos con los conceptos del dominio de lo físico, allí, se comete el error de categoría.

disposiciones, esto es, el mundo de los comportamientos que los agentes realizan de manera pública y, como tales, resultan observables y accesibles a observadores externos.<sup>2</sup>

No obstante, fue recién en 1969, con la publicación de *Content and consciousness*, cuando Dennett plasmó formalmente tal distinción epistemológica en el campo de la filosofía y la psicología, distinguiendo, por un lado, entre una explicación de nivel personal y, por el otro, una explicación de nivel subpersonal. Dennett sostuvo que era posible distinguir dos niveles de explicación diferentes sin caer en el nombrado dualismo cartesiano (1969, p. 93). La pregunta ontológica acerca de si hay uno o dos tipos de entidades diferentes fue dejada de lado y suplantada por la pregunta acerca de cómo se relacionan ambos tipos de explicaciones. Así, en sus primeros escritos, Dennett defendió una posición anti-reduccionista<sup>3</sup> y no dualista. La distinción entre niveles parecía radicar en dos tipos de relatos diferentes, con vocabularios diferentes, acerca de un mismo fenómeno. En consecuencia, afirmó que mientras el nivel de explicación subpersonal se ocupaba de describir ciertos procesos en términos físicos –esto es, neurofisiológicos–, el nivel de explicación personal se ocupaba de describirlos en términos mentales. De tal modo, sostuvo, “es sólo a nivel personal que las explicaciones proceden en términos de las necesidades, deseos, intenciones y creencias de un actor en un entorno” (Dennett, 1969, p. 164).<sup>4</sup> Esta distinción llevó a entender la diferencia entre los niveles de explicación como una diferencia entre las partes y el todo. Mientras que la explicación a nivel subpersonal se ocupaba de una de las partes de la persona: su cerebro, sus interacciones neuronales, sus movimientos aislados, etc.; la explicación a nivel personal se ocupaba de la persona como una única entidad total.

A grandes rasgos, desde que Dennett instauró la distinción, las explicaciones de nivel personal y subpersonal han sido interpretadas bajo dos enfoques de explicación diferentes. Por un lado, la explicación de nivel

<sup>2</sup> A raíz de dicha tesis, muchas veces se suele reconocer una tendencia conductista en la obra de Ryle.

<sup>3</sup> Dennett sostuvo que el nivel personal es indispensable para hablar de las mentes y de las acciones humanas, por lo tanto, no es posible reducirlo a las interacciones subpersonales del cerebro, ya que hacer esto implica dejar de hablar de fenómenos tan relevantes como la acción, el dolor, etc.

<sup>4</sup> Años más tarde (post-1969), Dennett sostuvo que el nivel subpersonal no debía ser entendido sólo en términos físicos, sino que también era posible hablar de cognición en dicho nivel.

personal se ha interpretado como una explicación horizontal acerca de nuestros estados mentales –es decir, acerca de nuestras creencias, deseos, intenciones, etc.– y de nuestro comportamiento. Este tipo de explicaciones se caracterizan por ser explicaciones singulares –explican un evento particular– que adoptan un modo temporal de análisis: los estados mentales deben ser explicados en términos de eventos o estados mentales antecedentes (Bermúdez, 2000; Drayson, 2014; Kim, 2005). Por el contrario, las explicaciones de nivel subpersonal son tratadas como explicaciones verticales. Tales suelen interesarse por la siguiente pregunta: ¿cuáles son los componentes o procesos –sean mecanismos cognitivos más bajos o mecanismos físicos– que posibilitan cierto estado mental o cierto fenómeno de nivel personal? Para responder, las explicaciones de nivel subpersonal bajan del nivel personal a las causas posibilitantes del mismo. En este sentido, las explicaciones de nivel subpersonal se suelen basar en los niveles micro de los eventos o fenómenos, es decir, en sus partes o componentes (Kim, 2005). Bermúdez sostiene que a través de las mismas se intenta de explicar cómo las generalizaciones de la psicología a nivel personal pueden ser verdaderas (2000). “El proyecto de explicación vertical puede caracterizarse en términos generales como la explicación de los motivos de las explicaciones horizontales” (Bermúdez, 2004, p. 33).

La división propuesta por Dennett fue rápidamente adoptada por la comunidad psicológica y filosófica. Sin embargo, ha recibido las más diversas interpretaciones, generando desde posturas que defienden la supremacía del nivel de explicación subpersonal (Stich, 1978) hasta quienes defienden la llamada “tesis de la autonomía de la explicación de nivel personal” (Davidson, 1980; Hornsby, 1997, 2000; McDowell, 1994). Dicha variedad de interpretaciones depende, en gran parte, del status bajo el cual es tomada la distinción entre niveles. Por un lado, quienes defienden la supremacía explicativa del nivel subpersonal –a veces identificada bajo el nombre “Tesis de la prioridad”– suelen adoptar una posición (i) reduccionista respecto a los niveles de explicación. Suelen asumir que se trata de una distinción laxa que se corresponde con una división epistemológica que distingue dos tipos de explicaciones diferentes acerca de los fenómenos psicológicos, pero que, en última instancia, descansa en el nivel subpersonal (Dennett, 1969). Por el otro lado, quienes sostienen la tesis de la autonomía de la explicación de nivel personal se caracterizan por defender una posición (ii) anti-reduccionista respecto a los niveles de explicación y, por lo tanto, suelen defender que existe una distinción ta-

jante entre ambos niveles. Así, sostienen que mientras que los fenómenos físicos deben ser explicados descriptivamente, limitándose a leyes naturales, los fenómenos mentales deben ser analizados desde explicaciones normativas basadas en la inteligibilidad racional (Drayson, 2014).

### **3. La tesis de la autonomía de la explicación del nivel personal: su origen y sus críticas**

Quienes defienden la tesis de la autonomía de la explicación de nivel personal sostienen que las explicaciones personales pueden ser comprendidas con independencia de cualquier hecho de nivel subpersonal, a tal punto que los estados subpersonales no forman parte de las explicaciones de nivel personal (Davidson, 1980; Hornsby, 1997, 2000; McDowell, 1994). Tales autores defienden fuertemente una posición anti-reduccionista con respecto a las explicaciones: la explicación de nivel personal no puede ser reducida en términos de una explicación de nivel subpersonal. De tal modo, las explicaciones de nivel subpersonal resultan relevantes para señalar las condiciones empíricas que hicieron posible el fenómeno de nivel personal, pero nunca para la explicación del fenómeno personal en sí mismo.

Los llamados autonomistas de la explicación suelen defender la racionalidad como principal criterio para distinguir entre fenómenos de nivel personal y fenómenos de nivel subpersonal. De tal modo, los fenómenos de nivel personal son de carácter normativo en tanto deben poder ser evaluados en función de criterios de la racionalidad. Davidson señala al respecto lo siguiente:

Si somos inteligibles para atribuir actitudes y creencias, o útiles para describir los movimientos como comportamiento, entonces estamos comprometidos a encontrar, en el patrón de comportamiento, creencia y deseo, un alto grado de racionalidad y consistencia. (Davidson, 1980, p. 237)

Este enfoque acerca de la autonomía de la explicación del nivel personal ha generado más rechazo que aceptación por parte de los filósofos y psicólogos de los últimos años. Probablemente, la razón de tal desestimación reside en que se suele asumir que hay un paso muy estrecho entre tomar una posición autonomista y defender un enfoque dualista acerca de las entidades mentales y físicas. Y, de más está decir, que las posiciones dualistas, y sus consecuencias ontológicas al afirmar que existe más de un tipo de realidad, se encuentra muy asociada a viejos debates del siglo XVII

que hoy en día causan un notable rechazo en la comunidad filosófica.

En discusión con Hornsby, una de las máximas defensoras de la tesis de la autonomía, Bermúdez (2000) ha elaborado una crítica en contra de la tesis de la autonomía de la explicación de nivel personal que intenta mostrar que el nivel personal no resulta suficiente para explicar el fenómeno de la acción. Hornsby ha defendido que, para explicar la mayoría de los comportamientos, no necesitamos acudir a una explicación de nivel subpersonal. Pensemos, nos dice, en cuando encendemos la luz apretando una perilla. Toda la explicación de dicho comportamiento consiste en evaluar nuestra intención de encender la luz, nuestra representación de la perilla justamente como una perilla para lograr prender la luz, nuestro movimiento en consecuencia de dicha intención, etc. No hay nada a nivel subpersonal que sea necesario añadir a nuestra explicación de dicha acción.

En respuesta, Bermúdez parece asumir que el ejemplo utilizado por Hornsby refiere a un tipo de comportamiento simple, y que, por lo tanto, no requiere muchos elementos en su explicación. En oposición, sostiene que existe un tipo de comportamiento particular, el comportamiento hábil, que no puede ser explicado con éxito si no recurrimos a una explicación subpersonal. De acuerdo con Bermúdez, las habilidades prácticas expertas involucran una cierta complejidad de la cual sólo podemos dar cuenta con lo que él llama una explicación “horizontal mixta”.

Para defender tal tesis, Bermúdez analiza la habilidad de jugar al tenis. Nos dice: supongamos que nos encontramos jugando a un partido de tenis y nuestro oponente nos gana haciendo un revés perfecto. Desde una explicación de nivel personal, sólo podremos decir que el sujeto ganador es más hábil que nosotros, ha entrenado más, estaba más decidido a ganar. Y todo esto puede ser cierto. Sin embargo, tales razones son compatibles con haber perdido el punto. Mi oponente podría haber entrenado más que yo, se podría haber encontrado más decidido a ganarme y, aun así, no realizar correctamente el revés. Entonces, ¿cómo explicamos el hecho de que mi oponente me ganó realizando aquel revés en particular? Bermúdez sostiene que para responder a la pregunta “¿por qué mi oponente ganó el juego?” en un grado más fino, es necesario brindar una explicación horizontal –como la de nivel personal– que introduzca “detalles de la integración de la visión y la acción a nivel subpersonal” (2000, p. 79) –como la de nivel vertical. De tal modo, concluye, las explicaciones subpersonales

son necesarias para dar cuenta de, al menos, algunos fenómenos de nivel personal tales como la acción realizada por un experto especialmente habilidoso. Bermúdez sostiene:

Ningún procedimiento de explicación puramente en términos de eventos de nivel personal será lo suficientemente detallado como para explicar por qué se jugó ese tiro en ese momento –para explicar por qué mi oponente extendió su raqueta precisamente esa distancia en ese ángulo. (2000, p. 79)

Bermúdez parece pensar que las explicaciones solo de nivel personal resultan muy gruesas para explicar la acción. De lo contrario, las explicaciones horizontales mixtas que propone, nos pueden indicar, por un lado, la relación exitosa entre la habilidad del jugador y su intención al realizar un revés ganador. Pero también, nos pueden indicar cuáles son las razones por las cuales el experto jugador de tenis realizó determinado movimiento en tal momento en particular –la posición específica de sus brazos, cómo sus músculos reaccionaron, etc.–, el cual terminó configurándose como la causa del revés ganador.

#### **4. Una defensa de la autonomía de la explicación de nivel personal**

Siguiendo la argumentación de Bermúdez, podemos notar que el autor selecciona casos de habilidades específicas sofisticadas, tal como jugar al tenis y realizar un buen revés, y sostiene que dichos comportamientos inteligentes se corresponden con movimientos corporales específicos que permiten explicarlos. Parece, pues, que la acción habilidosa –el saque ganador de nuestro tenista experto– tiene una correspondencia uno a uno con ciertas percepciones y representaciones específicas y movimientos particulares que dicho agente realizó. Si el jugador hubiera hecho cualquier otro patrón de movimientos, el suyo no hubiera sido el tipo específico de saque que le permitió ganar la partida. Las explicaciones horizontales mixtas reflejan dicha conexión al asumir que para un fenómeno particular de nivel personal siempre tendremos su correlación particular a nivel subpersonal que nos permitirá dar cuenta de los rasgos característicos que lo hacen peculiar, en este caso, que lo hacen ser el revés ganador. En consecuencia, si tomamos el ejemplo propuesto por Bermúdez, podemos reconocer que indagar acerca de la información perceptual registrada por el sujeto ganador, los movimientos musculares específicos realizados,

etc., resulta muy útil a la hora de responder cuáles fueron *las causas* que llevaron a realizar el revés, esto es, resulta útil a la hora de determinar cuál fue el mecanismo específico que llevó al movimiento ganador. No obstante, si nuestro objetivo en cambio es explicar ¿por qué mi oponente realizó dicho revés?, esto es, porque el jugador se comportó de tal manera para ganar el juego, la relevancia de tal información no resulta tan evidente. Analicemos con mayor detalle.

Si analizamos la pregunta con profundidad podremos ver que el jugador ganó el juego realizando un revés perfecto o bien porque se entrenó lo suficiente para tener la habilidad específica de realizar revés o bien porque en ese momento particular la situación lo llevó a realizar ese revés. El nivel personal de explicación nos permite, entonces, responder a la pregunta planteada atendiendo a ciertos factores como: mi oponente quiso dirigir la pelota hacia un lugar que me dificultara recibirla, quiso hacer un movimiento inesperado para mí, etc. De tal modo, parece que su acción resultó acertada o no, no en función de si efectivamente el revés tuvo como consecuencia convertirlo en el ganador del partido, sino más bien, en función de que fue una buena respuesta ante las circunstancias y la intención presente. Y esto, en principio, no depende de los movimientos particulares que determinaron que dicho revés fuera como fue, sino más bien, de los componentes intencionales propios del fenómeno de la acción.

Cuando en la bibliografía filosófica y psicológica se hace alusión a las habilidades prácticas se las identifica en general con un tipo de comportamiento inteligente que realiza un agente de manera intencional.<sup>5</sup> La noción se encuentra íntimamente relacionada con el concepto de *know-how* propuesto por Ryle (1946, 1949). Este concepto refiere a un tipo de saber práctico que el sujeto adquiere a través de un proceso de adiestramiento constante que le permite guiar su comportamiento bajo ciertas reglas.

---

<sup>5</sup> Son muchas las interpretaciones acerca de cómo debemos entender el fenómeno de las habilidades prácticas. Resulta imposible analizarlas con detenimiento en el presente trabajo, pero es importante señalar que las interpretaciones suelen identificarse, en última instancia, o con enfoques intelectualistas (McDowell, 1994, 2013), que suelen exigirle al sujeto una serie de capacidades cognitivas sofisticadas para poder afirmar que posee cierta habilidad práctica, o, por el contrario, con enfoques anti-intelectualistas, que suelen deflacionar las exigencias al punto de sostener que en el ejercicio práctico habilidoso el sujeto no se presenta como un agente intencional, sino más bien como un elemento más entre otros fluyendo en el mundo fenoménico (Dreyfus, 2007a, 2007b, 2013).

Bermúdez parece asumir, como mencionamos, que existe una relación uno-a-uno entre un comportamiento inteligente y habilidoso particular –digamos, el revés ganador– y una serie de movimientos musculares sumados a representaciones perceptuales específicas –digamos, los movimientos y representaciones que tuvo el jugador ganador a la hora de realizar el revés. Me gustaría defender, contra Bermúdez, que dicha implicación no nos permite dar cuenta de la particularidad del fenómeno de las habilidades prácticas mencionada anteriormente. Por consiguiente, quisiera señalar que la relación de identidad implicada en el argumento de Bermúdez falla en dos sentidos.

En primer lugar, parece que cuando nos referimos a una habilidad práctica en particular –la habilidad de jugar bien al tenis– no encontramos un patrón de movimientos y representaciones correctas que se correspondan exactamente con el ejercicio de la misma. Cuando decimos que un sujeto tiene la habilidad práctica de jugar al tenis y de realizar exitosamente un revés queremos indicar que tal agente, luego de un proceso de adiestramiento, ha adquirido cierto conocimiento sobre cómo realizar una acción particular. Dicho proceso de adiestramiento consistió, probablemente, en el aprendizaje de ciertas reglas acerca de cómo debía pegarle la pelota, cómo debía aprender a calcular la fuerza con respecto a la distancia y dirección pretendida, etc. Sin embargo, es muy probable que no haya incluido el aprendizaje acerca de qué información debía percibir del entorno y cómo, ni tampoco el modo de relacionarla con el control del movimiento de cierto músculo, en correspondencia con cierta posición particular de cada una de las partes de su cuerpo. Tales elementos, que en la propuesta de Bermúdez son explicados desde el nivel subpersonal, no parecen ser parte de un contenido que el agente pueda o no aprender y, por lo tanto, probablemente, no formaron parte de las instrucciones asimiladas. Este punto ya nos señala que la correspondencia entre un movimiento muscular particular y la habilidad de realizar un revés no se encuentra presente, al menos, en el proceso de aprendizaje. Pero avancemos más.

Cuando el agente ya adquirió la habilidad y decimos que ya sabe cómo jugar al tenis y hacer un buen revés parece ser que no queremos referirnos con ello a que el agente ya sabe cómo realizar un movimiento muscular particular. De lo contrario, queremos referirnos justamente a que, con cierta independencia sobre cómo lo hace –qué músculo mueve, a qué velocidad exacta, etc.–, puede jugar bien al tenis y hacer un buen revés. Esto parece tener una consecuencia en la relación que de hecho hay entre los

movimientos particulares y el ejercicio habilidoso de la acción. Para un mismo fenómeno práctico –para la habilidad de jugar bien al tenis y hacer un buen revés– podemos encontrar una gran cantidad de movimientos particulares posibles y efectivos: hay muchas formas –muchos patrones de movimientos musculares– para realizar correctamente un revés experto. Este hecho nos señala que hay cierta independencia entre la acción habilidosa y los comportamientos o movimientos específicos que pueden generarla.<sup>6</sup>

Asimismo, en segundo lugar, si analizamos la relación inversa entre movimiento particular y habilidad, parece que el movimiento muscular particular sumado a las representaciones perceptuales indicadas no nos asegura un ejercicio habilidoso. Un sujeto podría realizar exactamente dichos movimientos bajo las mismas representaciones perceptuales y, sin embargo, no efectuar un buen revés, mucho menos, un revés ganador, por no encontrarse en un contexto de juego. El hecho de que el revés de nuestro tenista haya sido un revés ganador depende más bien de que haya sido realizado intencionalmente, en determinado momento, por parte de un agente adiestrado en el aprendizaje de la habilidad de jugar al tenis. Esto es, el hecho de que el revés de nuestro tenista haya sido un revés ganador se justifica al identificar que tal comportamiento se encuentra inserto en el dominio de una práctica más amplia que el sujeto sabe dominar, en este caso, la práctica de jugar al tenis.

En consecuencia, cuando analizamos el fenómeno de la acción no queremos poner el foco en cuáles fueron los mecanismos que llevaron a que se diera un comportamiento *x*, sino más bien, cuáles fueron los mecanismos que llevaron a que un sujeto *y* realizara un comportamiento inteligente *x*. El interés en la acción está puesto, justamente, en su carácter inteligente e intencional, a diferencia de otros tipos de comportamientos o movimientos que tanto nosotros como animales humanos, como así también animales no humanos y otros tipos de sistemas, podrían tener (disposiciones, impulsos, etc.). De tal punto, una vez que consideramos a un comportamiento *x* como una acción, también reconocemos a un agente *y* como un agente intencional inserto en una práctica específica.

---

<sup>6</sup>En relación a tal distinción, Hornsby (1997) identificó la llamada “concepción disyuntiva del comportamiento”. La misma plantea que debemos diferenciar entre lo que podríamos llamar como movimientos corporales que son efecto de eventos neurales y movimientos corporales que son efecto de eventos de acciones.

## 5. Consideraciones finales

La tesis autonomista parece ser atractiva a la hora de explicar el fenómeno de las habilidades prácticas. Las explicaciones mixtas tal como la de Bermúdez, que intentan añadir elementos del nivel subpersonal a una explicación a nivel personal de la acción básica, parecen pasar por alto rasgos fundamentales de la acción habilidosa, sin los cuales la acción no sería tal. Por el contrario, la tesis de la autonomía nos permite comprender ciertos comportamientos –realizar un buen revés– como comportamientos intencionales insertos en un contexto habilidoso más amplio que el agente, cuando es experto, domina. Esto nos permite elaborar una explicación de nivel personal que diferencia significativamente el fenómeno de la acción práctica habilidosa de los casos como el dolor propuestos por Dennett.

## Referencias

- Bermúdez J. L. (2000). Personal and sub-personal. A difference without a distinction. *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action*, 3(1), 63-82.
- Bermúdez, J. L. (2004). *Philosophy of psychology: A contemporary introduction*. Londres: Routledge.
- Davidson, D. (1980). *Essays on action and events*. Oxford: Clarendon Press.
- Dennett, D. C. (1969). *Content and consciousness*. Londres: Routledge.
- Drayson, Z. (2014). The personal/subpersonal distinction. *Philosophy Compass*, 9(5), 338-346.
- Dreyfus, H. (2007a). Response to McDowell, *Inquiry*, 50(4), 371-377.
- Dreyfus, H. (2007b). The return of the myth of the mental, *Inquiry*, 50(4), 352-365.
- Dreyfus, H. (2013). The myth of the pervasiveness of the mental. En J. K. Schear (Ed.), *Mind, reason, and being-in-the-world: the McDowell-Dreyfus debate* (1st ed., pp. 15-40). Nueva York: Routledge.
- Hornsby, J. (1997). *Simple mindedness: In defense of naive naturalism in the philosophy of mind*. Cambridge: Harvard University Press.

- Hornsby, J. (2000). Personal and sub-personal; A defence of Dennett's early distinction. *Philosophical Explorations*, 3(1), 6-24.
- Kim, J. (2005). *Physicalism, or something near enough*. Oxford: Princeton University Press.
- McDowell, J. (1994). The content of perceptual experience. *The Philosophical Quarterly*, 44(175), 190-205.
- McDowell, J. (2013). The myth of the mind as detached. En J. K. Schear (Ed.), *Mind, reason, and being-in-the-world: the McDowell-Dreyfus debate* (1st ed., pp. 41-58). Nueva York: Routledge.



# Los parámetros normativos de corrección racional implícitos en las investigaciones en filosofía experimental

## El caso del experimento que correlaciona la receptividad al bullshit con un déficit cognitivo

María Natalia Zavadvivker\*

### Introducción

En este artículo me propongo introducir una serie de reflexiones en torno del peso normativo de los supuestos axiológicos y epistémicos implícitos en las investigaciones en filosofía experimental. Para tal fin, he tomado como ejemplo paradigmático el experimento llevado a cabo por Pennycook, Cheyne, Barr, Koehler y Fugelsang, publicado en un *paper* de 2015 bajo el título “On the reception and deception of pseudoprofound bullshit”. Brevemente, el objetivo del experimento de Pennycook et al. (2015) era hallar correlaciones entre la tendencia a aceptar o rechazar frases “bullshit”, y la posesión de ciertas habilidades cognitivas asociadas a dos estilos de razonamiento diferenciales: uno más intuitivo, consistente en una excesiva apertura mental y una alta predisposición a dotar de significatividad, profundidad y trascendencia a cualquier declaración; y otro más analítico, consistente en la adopción de una posición escéptica inicial que favorecería la detección de falacias e inconsistencias en el discurso como resultado de su sometimiento a un escrutinio que involucraría procesos deliberativos lentos. El término “bullshit” es un neologismo traducible como “basura pseudoprofunda”, y alude a aquellos mensajes que circulan asiduamente en redes sociales, consistentes en frases grandilocuentes y presuntamente profundas con palabras de moda asociadas a la autoayuda que pretenden ser estimulantes, trascendentes y motivadoras pero que en rigor carecen de significado y valor de verdad. Los autores pretendían analizar los factores que intervienen tanto en la receptividad como en la sensibilidad al bullshit, entendiendo por ‘receptividad’ la aceptación y valoración positiva de tales frases, y por ‘sensibilidad’ la disposición opuesta, vale decir, la tendencia a someterlas a un escrutinio crítico que permita

\* Universidad Nacional de Tucumán (UNT), CONICET.  
zavadvivker@gmail.com

detectar su carácter falaz. Para tal fin solicitaron a los participantes que evalúen frases pseudoprofundas en una escala de profundidad, con el propósito de establecer una medida legítima de receptividad al “bullshit” (BSR). El objetivo era correlacionar esa clasificación con diferencias individuales en la posesión de dos estilos diferenciales de pensamiento: uno más analítico (asociado a una actitud de escepticismo inicial que favorecería el reclutamiento de habilidades cognitivas asociadas a la detección de conflictos), y otro más intuitivo, que según los autores podría estar asociado, por un lado, a una fuerte tendencia a aceptar a-priori cualquier información recibida como verdadera o significativa (excesiva apertura mental que predispondría a encontrar profundidad y significatividad en cualquier discurso) y, por el otro, a un déficit cognitivo conocido como “fallas en el monitoreo de conflictos”, que contribuiría a que el receptor confunda vaguedad o falta de claridad con profundidad. También conjeturaron que las personas más receptivas al bullshit serían más propensas a las confusiones ontológicas (errores categoriales usuales en el pensamiento *folk* en los que se confunden los planos animado e inanimado, físico y mental, etc. –por ejemplo, cuando se comparten cadenas de oración para velar por la salud de alguien, asumiendo que las oraciones tienen capacidades curativas–), y a las creencias epistémicamente sospechosas (aquellas que entran en conflicto con concepciones naturalistas comunes sobre el funcionamiento del mundo, como las creencias religiosas y paranormales, las teorías conspiracionistas y la aceptación de terapias medicinales alternativas). Los autores esperaban hallar correlaciones entre este tipo de creencias y la receptividad al bullshit.

En el primer experimento se pidió a 280 estudiantes universitarios que evalúen en una escala del 1 al 5 (1: para nada profunda, 5: muy profunda) diez declaraciones sintácticamente correctas, pero formadas por una selección aleatoria de términos vagos en boga. Las frases fueron tomadas de dos sitios web: uno que mezcla al azar palabras comúnmente utilizadas en los tweets de Deepak Chopra, y otro que genera frases compilando términos que “suenan profundos”. El supuesto era que las calificaciones altas indicarían alta receptividad al bullshit. Los participantes completaron además cinco tareas tendientes a evaluar diferencias individuales en el estilo cognitivo analítico y en las habilidades cognitivas: la prueba de reflexión cognitiva (problemas matemáticos que inducen respuestas intuitivas incorrectas); una batería de “heurísticos y sesgos”, en donde también se inducen respuestas intuitivas incorrectas basadas en heurísticas o prejuicios

comunes; y dos medidas de habilidad cognitiva: la prueba Wordsum sobre inteligencia verbal; y un test sobre habilidades aritméticas globales. Para evaluar las confusiones ontológicas, los participantes debían responder cuán literal o metafórico les parecía el contenido de una serie de frases. Quienes las percibieran como más literales que metafóricas presuntamente estarían evidenciando una mayor confusión ontológica. Finalmente, completaron un cuestionario sobre su nivel de acuerdo/desacuerdo con 8 creencias religiosas comunes (más allá, cielo, infierno, milagros, ángeles, demonio, alma, Satanás).

En el segundo estudio repitieron el experimento utilizando ejemplos reales de bullshit (“tweets” particularmente vagos de la cuenta de Twitter de Deepak Chopra). También incluyeron medidas de auto-reporte sobre la disposición al pensamiento analítico e intuitivo, e incorporaron un test extra de inteligencia fluida.

En el tercer estudio compararon la escala de profundidad atribuida a las frases “bullshit” con la misma escala atribuida a citas genuinamente motivacionales e inteligibles, conjeturando que las diferencias en la atribución de profundidad a frases legítimamente significativas y a frases pseudoprofundas les permitiría obtener una medida fiable de sensibilidad al bullshit. También pidieron a los sujetos que evaluaran frases mundanas con un significado claro pero que no consideraríamos profundo (por ej.: “Mucha gente disfruta con tal tipo de música”).

En el cuarto estudio se correlacionó la BSR con una escala de creencias paranormales, otra de creencias conspiracionistas, y otra sobre el grado de aceptación de diez tipos de medicinas alternativas. También implementaron una escala de personalidad que mide las diferencias individuales en cinco rasgos (extroversión, amabilidad, responsabilidad, estabilidad emocional y apertura).

Entre los resultados más relevantes, los autores hallaron correlaciones significativas entre la propensión a adjudicar profundidad a las declaraciones “bullshit” y la mayoría de las pruebas que medían habilidades cognitivas asociadas al pensamiento analítico: los más receptivos fueron menos reflexivos, evidenciaron una menor inteligencia fluida, verbal y aritmética, y demostraron mayor propensión a las confusiones ontológicas, las ideas conspiracionistas, las creencias religiosas y paranormales y la medicina alternativa. Para medir específicamente la sensibilidad al bullshit (capacidad de detectar ambigüedad, inconsistencias y conflictos cognitivos), los investigadores compararon las evaluaciones que los par-

ticipantes hicieron de frases bullshit y de citas motivacionales legítimamente significativas. Encontraron correlaciones entre dicha sensibilidad, el estilo cognitivo analítico y el escepticismo en relación a los fenómenos paranormales, pero no pudieron correlacionarla con las ideas conspiracionistas ni con la aceptación de tratamientos medicinales alternativos. Sí encontraron correlaciones robustas entre la receptividad al bullshit y las creencias epistémicamente dudosas, y detectaron la propensión de un buen número de personas a una mente excesivamente abierta asociada a un débil espíritu crítico (incluso algunos participantes adjudicaron cierta profundidad a frases completamente mundanas). También hallaron una correlación negativa entre la predisposición a creer en entidades sobrenaturales (creencias religiosas) y dos factores complementarios: una apertura mental excesiva que insta al receptor a adjudicar profundidad y valor espiritual a cualquier estímulo; y la inhibición de la capacidad analítica y reflexiva, ya que el sujeto no quiere “ver” las contradicciones implicadas en las creencias religiosas y las explicaciones paranormales.

Mientras en mi ponencia anterior he sugerido la necesidad de incorporar a este experimento un conjunto adicional de variables de orden pragmático que en situaciones de la vida real podrían influir en una mayor o menor receptividad al bullshit (por ej., la velocidad requerida para la decodificación de las frases, la reputación -intelectual y moral- del emisor, el efecto de sugestión producido por la aceptación, aprobación y viralización de la frase en las redes sociales, el estado anímico del receptor de la frase, su nivel sociocultural, etc.); en esta ponencia me propongo:

- Realizar algunas observaciones críticas a los supuestos teóricos y axiológicos implícitos en el diseño experimental de Pennycook y colaboradores.
- Utilizar al presente estudio como un ejemplo de los escollos que deben enfrentar los filósofos experimentales cuando procuran obtener evidencia objetiva sobre el pensamiento de sentido común (intuiciones filosóficas *folk*, entendiendo por tales las creencias explícitas de la gente y los procesos cognitivos subyacentes), pero bajo la asunción de posiciones axiológicas y normativas implícitas en relación a cuáles son las creencias epistémicamente “correctas” y los mecanismos cognitivos adecuados para obtenerlas.

## Algunas observaciones al diseño experimental

En su trabajo, los autores definen aproximadamente el “bullshit” como frases consistentes en conjuntos aleatorios de palabras motivacionales de moda (las cuales suelen despertar reacciones emocionales asociadas a la profundidad y la trascendencia) con una estructura sintáctica correcta, pero que carecen de sentido y, por ende, en principio serían difíciles de decodificar. A su vez, procuran correlacionar la receptividad al bullshit con la propensión a incurrir en confusiones ontológicas y sustentar creencias epistémicamente dudosas. Ahora bien, si bien es posible que personas propensas a la religiosidad, a creer en fenómenos paranormales, inclinarse por terapias alternativas o teorías conspiracionistas, etc., estén al mismo tiempo más abiertas a adjudicar cierta profundidad a frases ininteligibles; no es lo mismo una declaración perfectamente inteligible cuyo contenido denote una confusión ontológica (Por ej.: “La fuerza de la oración salvará al mundo”), o creencias epistémicamente dudosas (“Todos tenemos un ángel de la guarda que nos protege”), que una frase carente de sentido por haber sido elaborada apelando a una conjunción aleatoria de términos (Por ej.: “Si llegamos a maximizar nuestras experiencias transparentes estaremos preparados para monetizar las nuevas sinergias globales” –extraída de un generador de bullshit en español–). Por lo general en la vida real las personas tienden más a compartir, postear, viralizar y darle “like” a frases motivacionales claramente inteligibles pero cuyas afirmaciones desafían las leyes naturales y el conocimiento de sentido común, que a frases difícilmente decodificables que en última instancia no terminan sirviendo al fin supuestamente perseguido: generar estados motivacionales positivos y provocar efectos de sanación espiritual apelando a las emociones del receptor (cabe señalar además que su objetivo no es proporcionar información objetiva y con valor de verdad sobre cierto estado de cosas). Probablemente los autores, para justificar su presunción de que mucha gente es muy receptiva a frases pseudoprofundas ininteligibles, supusieron que lo que puede estimularlas es el impacto emocional que desencadenan algunos términos motivacionales de moda, los cuales podrían por sí solos disparar sentimientos de profundidad y trascendencia, aun cuando la frase completa carezca de sentido. Sin embargo, a mi juicio cuanto más clara e inteligible es una frase, mejor cumple su fin de fomentar la autoayuda y la psicología positiva. En otras palabras,

Pennycook et al. (2015) asociaron implícitamente la ininteligibilidad y oscuridad de las frases con el valor de verdad de su contenido, determinado por la inclusión de confusiones ontológicas y creencias epistémicamente dudosas. Esto sugiere que los autores adhieren implícitamente a ciertas ramas de la filosofía analítica ligadas al Positivismo Lógico (por ejemplo, la filosofía del primer Wittgenstein) que identifican el significado de un término o proposición con sus condiciones de verdad (vale decir, con sus posibilidades, directas o indirectas, de contrastación empírica). De allí que correlacionen la apertura y aceptación positiva de frases “carentes de sentido”, con la propensión a las confusiones ontológicas y a las creencias epistémicamente dudosas. No queda muy claro si los autores sugieren que estas frases carecen de sentido porque no sería lícito o legítimo suponer la posibilidad de que, por ej., los planos natural y sobrenatural, o material y espiritual se superpongan en los hechos (con lo cual estarían asociando sentido o inteligibilidad a valor de verdad), o si simplemente quieren indagar si las personas propensas a las creencias sobrenaturales (por tener una apertura mental excesiva, y, por ende, una menor predisposición al pensamiento analítico) también son más propensas a aceptar el bullshit. Sea como sea, al procurar ofrecer una definición aproximada de “bullshit”, los autores parecen asociar “significado” a “valor de verdad” cuando afirman que tales frases carecen de ambas cosas, siendo que dicha asociación es privativa de cierto posicionamiento filosófico. Por ej., esta posición sería compatible con la filosofía analítica del primer Wittgenstein, mientras que dicho autor en su segunda etapa asume un criterio mucho más amplio, según el cual el significado de los términos está determinado por su uso en ciertos contextos y por la atribución intersubjetiva de sentido que le adjudica un grupo determinado, siempre en relación a ciertas formas de vida. Mi propósito no es cuestionar la posición de los investigadores, sino poner de manifiesto en qué medida los filósofos experimentales inevitablemente deben partir de supuestos metafísicos, axiológicos y normativos que les permitan establecer de antemano ciertos parámetros de “corrección” en virtud de los cuales determinan, de manera inconsciente e implícita, qué tipo de intuiciones y procesos cognitivos recabados experimentalmente serán correctos y cuáles serán “irracionales”, por alejarse de sus propios estándares. En otras palabras, las investigaciones de la filosofía experimental no se agotan, como es su pretensión originaria, en la mera descripción neutral de las intuiciones filosóficas *folk* de la gente, sino que dicha descripción presupone como telón de fondo juicios norma-

tivos y de valor implícitos que involucran criterios acerca de lo correcto e incorrecto (en este caso en términos cognitivos), lo que es incompatible con esa especie de retorno al empirismo ingenuo que los filósofos experimentales sustentan con bastante convicción. El supuesto de base (que aunque a juicio personal me resulte sensato, es como mínimo discutible) es que existe un ínfimo sector de la población –el de los filósofos o intelectuales formados en el pensamiento crítico– que estaría en condiciones de determinar qué tipo de creencias y procesos cognitivos serían correctas o lícitas, y cuáles serían erróneas y fundadas en razonamientos falaces. En el fondo no se trata de una asunción tan descabellada, pues del mismo modo podríamos decir que son los conceptos teóricos de los físicos, y no las intuiciones físicas *folk* de la gente común, los que explican adecuadamente las leyes y variables del mundo físico. Incluso para Wittgenstein (1921) la tarea específica de la filosofía es precisamente la de esclarecer y depurar los conceptos del lenguaje ordinario, bajo el supuesto de que este proceso nos habilitaría para razonar correctamente acerca del mundo. Pero estos supuestos deberían ser al menos enunciados explícitamente en las investigaciones en filosofía experimental, dejando en claro la posición metafísica, epistémica, axiológica e ideológica de los autores, de modo tal que éstos asuman su punto de partida subjetivo y circunscripto a una perspectiva determinada, en lugar de presentar supuestos datos crudos analizados desde una presunta posición neutral.

### **¿Es posible evitar asunciones normativas implícitas en la descripción de las intuiciones filosóficas *folk*?**

Si bien los autores del experimento son psicólogos, su trabajo puede ser enmarcado dentro de la línea de investigación conocida como “filosofía experimental”. Los filósofos experimentales procuran recabar datos empíricos sobre el pensamiento espontáneo de las personas en relación a ciertos tópicos filosóficos (posibles intuiciones filosóficas *folk*, y los procesos cognitivos responsables de su fijación), pero para evaluar tales intuiciones (por ej., para ponderar su grado de corrección e incorrección), inevitablemente deben establecer ciertos parámetros sobre la forma correcta de razonar filosóficamente (procesos cognitivos fiables) y sobre la corrección de las respuestas emitidas, lo que implica que el abordaje descriptivo sólo puede llevarse a cabo asumiendo ciertos supuestos normativos fijados por los propios filósofos. Esto implica la suposición de que son los especialis-

tas en filosofía, y no la mayoría estadística de las personas, quienes estarían legítimamente autorizados para juzgar el valor de las intuiciones y procesos cognitivos del resto de los sujetos. Tal como lo mencionamos anteriormente, en algún sentido esto se justifica, ya que los filósofos son precisamente esa ínfima franja de la población dedicada al entrenamiento profesional y sistemático de habilidades cognitivas asociadas al pensamiento analítico, con lo cual poseen un estilo cognitivo más reflexivo que el promedio. Sin embargo, una cosa es tener una mayor disposición al pensamiento reflexivo, y por ende, un mayor entrenamiento en las habilidades cognitivas asociadas; y otra es suponer que dicho entrenamiento garantiza la adquisición de las creencias “correctas”, o al menos mucho más fiables que las obtenidas por otras vías (por ej., las emociones, dado que estas últimas tenderían a ser sesgadas y falaces). Si bien en lo personal adhiero a esta posición, y considero adecuada la asociación entre una excesiva apertura mental, escasa deliberación crítica y reflexiva; y el pensamiento “mágico” (creencias religiosas y paranormales, supersticiones, etc.); es precisamente el racionalismo crítico del filósofo el que nos insta a cuestionar la validez de nuestras propias creencias y a asumir las dificultades en la búsqueda de parámetros valorativos universales y objetivos para juzgarlas (de hecho, nada garantiza la falsedad de las creencias religiosas y paranormales, ya que en sentido estricto son irrefutables al ser empíricamente incontrastables). Pero, al mismo tiempo, el filósofo experimental debe adoptar algún parámetro con el cual cotejar el grado de “corrección” de las creencias y procesos cognitivos de los sujetos experimentales, aun cuando el propósito de la investigación sea meramente descriptivo. En el experimento analizado la propia adopción del término “bullshit” (traducible al español aproximadamente como “pura mierda”, vocablo con una carga axiológica profundamente despectiva) sugiere un claro posicionamiento ideológico, de modo que los criterios para identificar frases “bullshit” no pueden agotarse en un conjunto de rasgos claramente demarcables y meramente denotativos capaces de evitar la subjetividad del evaluador. Esta incorporación de valoraciones subjetivas en el diseño experimental es un claro síntoma del “giro pragmático” de la filosofía (influido además por sus cruces interdisciplinarios con dominios tradicionalmente pertenecientes a otras ciencias): los filósofos van abandonando cada vez más sus recaudos profesionales ligados a la neutralidad axiológica y el análisis puramente lógico de los conceptos, para asumir abiertamente ciertas posiciones (de manera consciente o no) amparados en criterios pragmáticos

que los instan a no suspender excesivamente el juicio, a fin de no disociar la investigación metateórica de sus asunciones ideológicas y valorativas sustentadas de hecho. Si bien los autores del estudio son psicólogos cognitivos y no filósofos, lo que podría explicar su menor compromiso con la neutralidad axiológica y la imparcialidad en el uso de los términos, los filósofos experimentales también tienden a adoptar (más o menos implícitamente) posiciones axiológicas y normativas como parámetros para evaluar las evidencias obtenidas, aunque algunos asumen su problematización y las controversias epistemológicas implicadas.

En síntesis, ninguna evidencia empírica puede dar cuenta *per se* de cuán apropiadas o inapropiadas son las intuiciones *folk* si no se adopta de antemano un criterio normativo que permita valorarlas, y es el experimentador quien se arroga tal potestad. El análisis de los resultados consiste entonces en cotejar la evidencia obtenida con parámetros previamente establecidos sobre tipos de respuestas adecuadas o inadecuadas, correctas o erróneas. En el experimento analizado, por ej., lo correcto es sostener una posición escéptica y crítica en relación a ciertas declaraciones, no sólo por su vaguedad, ambigüedad y presunta ininteligibilidad, sino porque su contenido supone la aceptación de creencias que desafían las leyes naturales y el conocimiento avalado empíricamente. Pero la valoración de los datos empíricos depende también de ciertas asunciones teóricas más globales sobre el status de la cognición humana. Por ej., una posición pragmática de corte evolucionista supone que nuestras creencias y herramientas cognitivas son adaptaciones instrumentalmente útiles para lidiar con presiones ambientales específicas, incluyendo desafíos relativos a la necesidad de lograr un equilibrio emocional y espiritual, lo que abonará a favor de la búsqueda de explicaciones y la fijación de creencias capaces de dotar de sentidos y significados trascendentes a nuestra existencia. Desde esta perspectiva, la tendencia a aceptar confusiones ontológicas y creencias epistémicamente dudosas podría considerarse “apropiada” por resultar eficaz para el logro de algún fin adaptativo. Pero si asumimos desde una perspectiva racionalista que las intuiciones y procesos cognitivos deben cumplir con ciertos estándares de racionalidad independientemente de su valor adaptativo o utilidad práctica, entonces el dictamen de la mayoría no influirá en absoluto en el veredicto a priori respecto de qué intuiciones o procesos cognitivos son más adecuados que otros.

## **Referencias**

- Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. (2015). On the reception and deception of pseudoprofound bullshit. *Judgment and Decision Making*, *10*(6), 549–563.
- Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72.
- Wittgenstein, L. (1921). Logisch-philosophische Abhandlung. *Annalen der Naturphilosophie*, *14*, 185–262.





---

## 4. Historia y filosofía de las ciencias sociales y humanas





## El ángulo facial, una investigación cefalométrica publicada por Alda Mercante en 1912

Luis M. Guimarey\*

### Marco teórico

El positivismo, un giro radical de la mirada sobre el mundo, se expandió entre los siglos XVII y XIX y las primeras décadas del XX. Se caracterizó, particularmente, por la paulatina pérdida de “autoridad” de la Biblia como explicación de la naturaleza, que fue reemplazada por leyes científicas desentrañadas mediante diferentes “programas de investigación”. Uno de los aspectos metodológicos más relevante del mismo es, lo que podríamos llamar, la “*cuestión métrica*”, en este sentido, las *ciencias naturales* son las primeras en abandonar las observaciones puramente descriptivas y comenzar a “medir” los fenómenos que caen en su campo de interés. Por su parte las *ciencias sociales* siguen igual camino. Auguste Comte (1798-1857), en el *Curso de filosofía positiva* que dicta en París durante doce años, entre 1830 y 1842, sienta las bases de esta nueva forma de “ver, interpretar y medir los fenómenos sociales”. En el campo de la pedagogía, siguiendo dicho paradigma, se establece el concepto de *antropología pedagógica* (De Alcántara García, 1880; García Castaño, 1994) el que comporta un conocimiento físico/biológico del alumno alcanzable mediante la antropometría, entre otros procedimientos, y sirve de base para las clasificaciones raciales y los estudios de género, tan en boga en la época. A partir de las investigaciones antropométricas, en particular cefalométricas, se realizan inferencias relacionadas con la capacidad de aprendizaje, habilidades, inteligencia, etc. (Gould, 1981/2007).

### La “escuela antropométrica” de La Plata

Adscribiendo al positivismo, un grupo de educadores/investigadores desarrollaron en las primeras dos décadas del siglo XX, en el ámbito de la

\* Universidad Nacional de La Plata (UNLP), Facultad de Ciencias Médicas (FCM). La Plata, Buenos Aires, Argentina.  
lguimarey@gmail.com

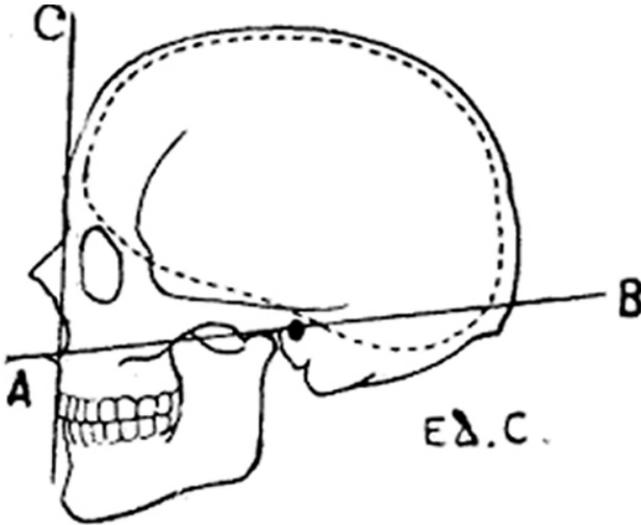
Sección de Pedagogía de la Facultad de Ciencias Jurídicas y Sociales de la UNLP, una intensa labor de investigación antropométrica publicada en sendos trabajos, entre los años 1906 y 1918, en los: *Archivos de pedagogía y ciencias afines*, órgano oficial de la mencionada sección. Uno de estos trabajos, cuya autora es Alda Mercante (1912), tiene como objeto el ángulo facial. En la presente ponencia lo analizamos por su interés histórico y epistémico, dado que su autora representa un caso poco frecuente de autoría femenina para la época (Ostrovsky, 2018), por un lado, y por otro, debido a que en el mismo se contraponen el paradigma de Camper al de Cuvier respecto del valor epistémico del ángulo facial.

### **¿Quién fue Alda Mercante?**

Nacida en San Juan –Víctor Mercante se trasladó a esa provincia en 1890, para dirigir la escuela normal, donde se casó con la pianista Julia Pozo–Alda fue la primera mujer, de los siete hijos que engendró el matrimonio. En 1906, junto a su familia, se mudó a La Plata, donde su padre fue convocado para organizar la *Sección Pedagógica de la Facultad de Derecho y Ciencias Sociales* de la universidad local (la que en 1916 se transforma en *Facultad de Ciencias de la Educación*). Allí completa su formación alcanzando el grado académico de profesora. Con veinte años de edad, y cursando las asignaturas en la sección dirigida por su padre, muestra su capacidad para el trabajo de campo y estadístico que se refleja en la investigación motivo de este análisis. Sin embargo, esta fue la única publicación conocida de la autora dentro de esta área de conocimiento.

### **Reseña de la investigación**

El objeto de la investigación, como su título lo indica, es el: ángulo facial (Fig. 1) –al que la autora llama de Cuvier (1769-1832)– aunque, en la introducción, hace una extensa referencia a Petrus Camper (1722-1789), quién en el siglo XVIII fue el “primero (sic) en conceptualizarlo”. Al respecto dice: “Camper, en suma, trataba de fijar un ángulo estético, trabajaba para los artistas que lo recompensaron dando su nombre al ángulo facial. . .” (Mercante, 1912 p. 187). Seguidamente menciona que con esta investigación se propone, “fijar un patrón estético”, racial y de especie (sic) de la población, y continúa:



**Fig. 1:** Ángulo facial, formado por la línea facial (A-C), definida por la parte más prominente de la frente y el punto más anterior del reborde alveolar del maxilar superior, y el “Plano de Camper” (A-B), formado por una línea horizontal que va desde el oído, hasta el ala de la nariz. Indica la inclinación de la frente.

*Nota:* Imagen tomada de la fig 56 de Artistic Anatomy of Animals de Édouard Cuyer, Proyecto Gutenberg. Recuperado de <https://www.gutenberg.org/files/38315/38315-h/38315-h.htm>. Obra de dominio público.

Mi propósito al determinar . . . el ángulo facial . . . es considerarlo desde los siguientes puntos de vista: *de la edad y de los sexos, de la inteligencia, del tipo étnico, del parentesco, de las relaciones con el desarrollo craneano* [cursiva nuestra] (Mercante, 1912, p. 191).

Agregando: “. . . con nuestras estadísticas; nos proponemos resolver algunos problemas, como los de la edad, el sexo y la inteligencia . . .” (Mercante, 1912, p. 191)

A la “inteligencia” –cuya medición era una preocupación primordial de la época– intenta inferirla a partir de los promedios anuales de las calificaciones obtenidas por los sujetos de investigación en las asignaturas: matemática, lengua y ciencias naturales, comentando, textualmente.

Hemos confrontado las clasificaciones (sic) de cada alumno con su angulación y podemos asegurar que no hay relación alguna, pues la estadística ofrece casos de angulaciones altas y clasificaciones (sic) bajas y

vice-versa. . . Sin embargo, es curioso anotar como *la mayor parte de los sujetos, al medirlos, preguntaban si con ello se averiguaba la inteligencia y creían que al bajo correspondía una inteligencia inferior y se regocijaban cuando era elevado* [cursiva nuestra] (Mercante, 1912 p. 198)

En relación al tipo étnico/racial manifiesta:

. . . está tomado de acuerdo con algunos rasgos que caracterizan el tipo: color del cabello, color de los ojos, grosor de los labios y color de la piel, y particularmente por la nacionalidad de los padres . . . lo consideramos, no desde el punto de vista de la raza misma, sino de las nacionalidades, puesto que *no caben razas en una población cuyo origen es eminentemente europeo* [cursiva nuestra].

En la determinación del ángulo es un factor poderoso la nacionalidad de los sujetos. El ángulo facial tiene positivamente un valor étnico.

La estadística revela, para la ciudad de La Plata, la convergencia de una gran variedad de tipos étnicos. Siendo la masa favorable a un tipo futuro, si consideramos que la angulación más alta corresponde a la *raza blanca, más activa*, y que el tiempo destinado a producir por la fusión una condensación en el sentido de la homogeneidad . . . nos dará una angulación . . . superior a la de cualquier nación grecolatina.

En los cuatro o cinco casos de una angulación inferior a 60° los sujetos ofrecen *caracteres degenerativos y no étnicos*.

Como promedio de la formación platense, vuelvo a repetir, [corresponde] 68°6, *lo que del punto de vista de la euritmia étnica es de un significado halagador, porque se considera ángulo estético desde los 65°* [cursiva nuestra]. (Mercante, 1912, p. 197)

Las diferencias de género también son particularmente consideradas con comentarios como el siguiente:

. . . deducimos que la angulación debería de variar en el sentido agudo, en la mujer; pero no ocurriendo así y estando la abertura del ángulo en relación con el desarrollo del maxilar, esto indica que el maxilar en la mujer se desarrolla mucho menos en el sentido prognata que en el hombre *-lo que hace que en la mujer conserve sus caracteres infantiles-* es decir, un ángulo estético elevado [cursiva nuestra]. (Mercante, 1912, p. 193)

Por fin, el último párrafo, correspondiente a las conclusiones, constituye una síntesis del paradigma con que la autora aborda el tema:

Nuestra estadística muestra una angulación [del ángulo facial] cuya amplitud no la ofrece ninguna otra; este rasgo . . . *demonstraría nuestras excelentes condiciones para la civilización*, pues . . . *a medida que una raza se civiliza, los índices craneanos de sus individuos se diferencian más y más*, lo que, por otra parte, explica que no es hacia la igualdad intelectual a donde (sic) conduce la civilización, sino a la desigualdad más profunda [cursiva nuestra]. (Mercante, 1912, p. 199)

## ¿Camper o Cuvier? Raíces históricas del ángulo facial

La “paternidad” del ángulo facial, sin duda, le es reconocida a Camper, como la propia autora lo hace en la introducción de la publicación, “. . . trataba de fijar (dice) un ángulo estético, trabajaba para los artistas que lo recompensaron dando su nombre al ángulo facial . . .” (Mercante, 1912, p. 187), es decir que el enfoque que Camper daba al tema –en palabras de A. Mercante– estaba más relacionado al arte que a la ciencia, lo que hasta cierto punto fue verdad; de hecho, desarrolló su teoría en nueve conferencias dictadas, alrededor de 1770, en la Academia de Dibujo de Ámsterdam. Sin embargo, si bien él mismo fue un eximio dibujante, también fue uno de los científicos más destacados del siglo XVIII, desarrollando su actividad docente y de investigación principalmente en las áreas de biología, antropología física y medicina, siendo relevante su labor como anatomista comparado. Fue uno de los introductores del *empirismo newtoniano* en el continente europeo y, epistemológicamente, adhirió a una postura *físico/teológica*, opuesta al cartesianismo imperante en la época y coherente con su religiosidad (consideraba que el conocimiento del mundo físico –la naturaleza– creado por Dios, se adquiere por los sentidos y es analizado/interpretado por la razón). Esto lo llevó a postular la posición ontogénica de la *estricta unidad de la humanidad* dentro de la “creación”. Defendió, apoyándose en un considerable número de disecciones cadavéricas, la igualdad entre negros y blancos, y su diferencia taxonómica absoluta con los grandes monos –orangutanes y chimpancés– (Meijer, 1999, 2015). En síntesis, su teoría del ángulo facial nunca tuvo, en su origen, una connotación racista y si una forma matemática de objetivar las variaciones dentro del “reino” animal junto a un ideal de belleza. Gould (1981/2007) comenta a este respecto que Camper en sus escritos sobre el ángulo facial le dedicó más espacio a las diferencias entre hombres jóvenes y viejos, que a las diferencias raciales.

Contrariamente al pensamiento de Camper, el *ateísmo ilustrado* de los intelectuales del cambio de siglo –Camper murió a finales del s. XVIII– desactiva el prurito de considerar una gran proximidad entre primates y el hombre, llegando a postularse la posibilidad de inter fertilidad entre orangutanes y humanos. De hecho, llegó a considerarse, por parte de algunos científicos –v.g. Blumenbach–, al orangután como perteneciente al género homo en un estadio primitivo previo al desarrollo del lenguaje. En este ambiente intelectual que permeaba occidente, Cuvier, uno de los naturalistas más destacado de su tiempo, quien tuvo relación personal con Camper, toma con “entusiasmo” la teoría del ángulo facial pero, a diferencia de las ideas de su creador, lo transforma en un “instrumento apto” para la clasificación jerárquica de la humanidad, partiendo de los primates (grandes monos), pasando por los negros y llegando al hombre europeo, en una escala en la cual, las dos primeras categorías están muy próximas entre sí (podríamos decir “*casi confundidas*”) y el hombre blanco europeo muy por encima. Y va más allá postulando, según el zoólogo germano/holandés Visser, que el ángulo facial es un indicador de las facultades mentales, relacionando al *commensurable* ángulo facial con la *incommensurable* inteligencia. Estas teorías cobran gran fuerza a lo largo del s. XIX a favor de la expansión sin precedentes del colonialismo y su correlato, la esclavitud. En la época, negro y esclavo se hacen sinónimo de gran comercio transatlántico y entre las ideas expuestas y la justificación moral de la esclavitud, hay solo un paso.

## Conclusión

Entre el desarrollo de la teoría por parte de Camper y el trabajo de A. Mercante, pasa algo más de un siglo, periodo en el cual se produce la reconfiguración del paradigma con que se considera al ángulo facial, para transformarlo en una teoría racial y de género, que coloca al *hombre blanco europeo* en la cima de una escala jerárquica, por arriba tanto de los negros como de las mujeres. Al mismo tiempo como instrumento apto, también, para selección individual de sujetos más o menos inteligentes. Con este espíritu, epistemológicamente hablando, la autora desarrolla su investigación. La “elección”, desde el título del trabajo –ángulo de Cuvier– a su contenido y enfoque, –textualmente dice: “Mi propósito al determinar . . . el ángulo facial . . . es considerarlo desde . . . *la edad y de los sexos, de la inteligencia, del tipo étnico, del parentesco, de las relaciones con el desarro-*

*llo craneano* [cursiva nuestra]" (Mercante, 1912, p. 191)– no es azarosa, por el contrario, expresa una preocupación de época para nuestro país. De hecho, el ambiente intelectual argentino de principios del siglo XX se encontraba fuertemente influido por el *naturalismo filosófico positivista*. Sobre esta forma de ver el mundo se apoya la práctica política, educativa y científica de la época, llevando a intelectuales representantes de esta corriente de pensamiento –Ramos Mejía (1849-1914), Bunge (1875-1918) o Ingenieros (1877-1925), entre otros– a elaborar, basándose en la llamada *psicología de las masas*, diferentes representaciones de la sociedad permeadas por el *determinismo biológico* (Gomez di Vincenzo, 2013; Teran, 1987). Es esta la mirada teórica sobre la sociedad argentina, que los intelectuales mencionados consideran en formación, la que sustenta la hipótesis de trabajo en la investigación analizada.

Por último, señalamos la autoría femenina como un hecho singular para la época. Según Ostrovsky (2018), los *Archivos de Pedagogía y Ciencias Afines* fue de las pocas revistas del momento en publicar trabajos cuyas autoras eran mujeres, esto a pesar de que la "psicología" de la mujer fuera uno de los tópicos recurrentes del momento debido al surgimiento de los movimientos feministas acompañados por las consecuentes preguntas acerca de las capacidades femeninas.

## Referencias

- De Alcantara García, P. (1880). *Prolegómenos a la antropología pedagógica*. Madrid: English y Gras Eds.
- García Castaño, F. J. (1994). Antropología pedagógica. En C. Ortiz García & L. Sánchez Gómez (Eds.), *Diccionario histórico de la antropología española* (pp. 91-97). Madrid: Consejo Superior de Investigaciones Científicas.
- Gomez di Vincenzo, J. (2013). *Biotipificar al soberano: Biotipología, psicotécnica, educación y prescripción de roles sociales en Argentina (1930-1943)*. Buenos Aires, Argentina: Rthesis.
- Gould, S. J. (2007). *La falsa medida del hombre* (R. Pochtar & A. Desmonts, trads.). Barcelona, España: Crítica S.L. (Obra original de 1981)
- Mercante, A. (1912). Investigación del ángulo facial de Cuvier, por sexos y edades, en el vivo. *Archivos de Pedagogía y Ciencias Afines. La Plata*, 11(32),187-211.

- Meijer, M.C. (1999). *Race and aesthetics in the anthropology of Petrus Caper (1722-1789)*. Amsterdam, Netherlands: Editions Rodopi.
- Meijer, M.C. (2015). Bones, law and order in Amsterdam. Petrus Camper's morphological insights. En K. van Berkel & B. Ramakers (Eds.), *Petrus Camper in context* (pp. 187-213). Hilversum, Netherlands: Verloren.
- Ostrovsky, A.E. (2018). Mujeres en los archivos de pedagogía y ciencias afines (1906-1914). *Estudos e Pesquisas em Psicologia Rio de Janeiro*, 18(3), 984-999.
- Terán, O. (1987). *Positivismo y nación en la Argentina*. Buenos Aires, Argentina: Puntosur Editores.



## Nuevas interpretaciones de lo humano ante el desafío tecno-científico

Elizabeth Padilla\*

### I

Desde la segunda mitad del Siglo XX a esta parte somos testigos y partícipes de cambios acelerados que afectan distintos ámbitos de la vida, tanto individuales como colectivos, provocados por los adelantos tecnológicos. Estos han modificado desde las formas de interacción social, la accesibilidad a la información, las posibilidades de discriminar entre lo real y lo virtual, hasta llegar a conmover nuestras expectativas más íntimas acerca de la enfermedad, la vejez y la muerte. Esta breve lista sirve sólo a modo de ilustración para referirnos a alguna de las tantas dimensiones sobre las cuales se han extendido las innovaciones. Incluso estos cambios han inaugurado estilos de vida impensados para las sociedades de siglos pasados, vale por caso, aquellos derivados de la eliminación de los límites entre lo privado y lo público. Muestra cabal de la desaparición de esta frontera es el espectáculo que ofrecen de continuo las redes, en donde las personas exponen sin tapujos los aspectos más íntimos de sus vidas, poniéndolos a consideración de un auditorio anónimo de los cuales reciben comentarios de todo tipo.

Al respecto, algunos sostendrán que las variaciones en los estilos de vida en relación a la aparición de dispositivos tecnológicos no es un rasgo distintivo de nuestra contemporaneidad, puesto que toda época tuvo innovaciones que trastocaron de una u otra manera el status quo establecido. No obstante, lo novedoso de la nuestra radica en que los adelantos se instalan de modo rápido y con una penetración tal que se configuran prontamente en elementos imprescindibles para la vida, atravesando todas las condiciones sociales y económicas. Pensemos, por caso, el uso y difusión del celular. Tal dispositivo, en breve tiempo, se convirtió para casi todos los sectores sociales en el medio indispensable de comunicación, de entretenimiento y de acceso rápido a información.

\* Universidad Nacional del Comahue.  
horizontesfilosóficos@gmail.com

Ahora bien, aun cuando pareciera que todas las modificaciones tecnológicas alteran los modos de vida y suponen riesgos a sopesar, no todas son de igual trascendencia, pues algunas de ellas despiertan, con justa razón, toda una gama de emociones que van desde la esperanza hasta el temor creciente, como es el caso de las alteraciones que introducirían las biotecnologías en nuestro genoma. En aras de ordenar las biotecnologías de acuerdo a la peligrosidad potencial que implican, distinguimos las que ofrecen una terapia génica somática, de aquellas que proporcionan un tipo de ingeniería de la línea germinal. Las consecuencias de la aplicación de una u otra, son abismales. En efecto, la primera representa la modificación del genoma pero sólo a nivel del individuo, la segunda, en cambio –propiciada por los transhumanistas–, posee la particularidad de que la alteración introducida es hereditaria, es decir, su influencia alcanza a las generaciones futuras. Como es de esperar, la factibilidad de este tipo de intromisión en el genoma de células germinales suscita además de sentimientos encontrados, un sin número de cuestiones urgentes acerca de la justificación y la necesidad de discutir colectivamente los límites a proyectos experimentales orientados en esa dirección. Por otra parte, tales posibilidades actualizan preguntas filosóficas, que no por antiguas han perdido vigencia, por ejemplo, aquellas que tematizan si hay algo así como una caracterización o determinación de la identidad humana, que se vería amenazada por los avances mencionados.

En síntesis, la revolución biotecnológica nos coloca ante el umbral de la realización de una transformación mucho más radical de la que tenemos memoria como especie, pues ahora el objeto de experimentación somos nosotros. ¿Cómo hemos llegado a ese estado de situación y qué decisiones sobre tales proyecciones futuras podríamos tomar como sociedad global de aquí en adelante?

## II

Si bien creemos que tal asunto no se encuentra en la agenda cotidiana de los ciudadanos, tal vez debido a que los cambios fueron incorporándose paulatinamente y sin percatarnos de su poder coercitivo, no obstante, llegará el momento en que nos enfrentemos ante la contundencia de los hechos. Sostenemos que en parte esta situación se explica por el vínculo conflictivo que se entabla entre tecnología y sociedad. En principio, porque la gran velocidad de los cambios se encuentra con una sociedad no

preparada ni conceptual ni normativamente para evaluar la incidencia de los mismos, a mediano y a largo plazo. En palabras de Stiegler:

El devenir técnico es originariamente un desgarramiento. Pero la tecnogénesis va estructuralmente por delante de la sociogénesis –la técnica es invención y novedad–, y el ajuste entre evolución tecnológica y tradición social conoce momentos de resistencia porque, dependiendo de su alcance, el cambio técnico conmociona más o menos los parámetros definidores de toda cultura. (Stiegler, 1996/2002, p. 8)

A nivel social tiene lugar, entonces, un inevitable desfasaje cultural entre los desenvolvimientos de uno y otro ámbito. Una muestra es el tratamiento conceptual que han merecido los objetos técnicos en su vínculo con el hombre. Se los considera habitualmente como meros instrumentos o medios puestos al servicio de los propósitos humanos y, por tanto, neutrales, es decir, ni buenos ni malos, pues eso depende del uso que hagamos de ellos. Sin embargo, esta caracterización resulta poco eficaz, pues surge como respuesta a problemáticas ajenas al desafío tecnológico actual y no permite arribar a respuestas adecuadas. Hay un matiz a esto, pues también es cierto que la noción de medio o instrumento puede ser útil, siempre y cuando el análisis no se reduzca a él. Un buen ejemplo es el tratamiento que sobre esta noción ofrece Heidegger en “La pregunta por la técnica” (1954/2001). Allí sostiene que ese tipo de respuesta oculta la auténtica esencia de la técnica, pues depende de una representación antropológica corriente, según la cual ella es un medio y producto de un hacer del hombre. La interpretación de la técnica como medio para el logro de un fin pertenece al problema de las causas, a la causalidad. Ahora bien, eso supone para Heidegger un modo del des-ocultar del ser, de lo que es, de ahí que la técnica, por estar emparentada con aquella, también lo es, pero desde una peculiaridad que deberá ser examinada con detenimiento. Más tarde, establecerá diferencias entre la técnica antigua y la contemporánea. En la última, la naturaleza se ve provocada u obligada por la acción del hombre a convertirse en un insumo o recurso a ser explotado, cuantificado y distribuido. Podríamos agregar que a partir de las biotecnologías la naturaleza provocada como insumo, y de un modo más contundente que el pensado en la época de Heidegger, es la humana. Según su descripción, considerar a la técnica sólo como medio nos hace perder de vista el objetivo de nuestra indagación, que es la búsqueda de su esencia auténtica, extraviándonos en la exploración inútil de modos adecuados de manejar o

dominarla. Su análisis muestra finalmente que la esencia de la técnica no es algo asible técnicamente sino más bien consiste en un modo peculiar de develamiento del ser en el que el *Dasein* está fuertemente involucrado. En síntesis, buscar la esencia de la técnica es buscar la del hombre.

Ahora bien, ¿qué ocurre si la sociedad toma como propia y única la caracterización de la tecnología como medio o instrumento para el logro de un fin? Atribuirle neutralidad nos hace creer que los problemas se resuelven a partir de la evaluación de los usos de la tecnología, lo cual, por supuesto, es inconducente. Otra perspectiva, en cambio, la juzga como la sustancia fundamental que informa todos los aspectos de la vida. De ahí que lo mejor que podemos hacer por nuestro bien, ya que todavía contamos con márgenes de acción en algunos ámbitos donde los desarrollos son todavía provisionales, es limitarla para evitar males futuros mayores, asumiendo, en algunos casos, una actitud abiertamente tecnófoba.

Al respecto, consideramos que ambas actitudes son igualmente peligrosas, por tratarse de planteos extremos que impiden tomar decisiones apropiadas ¿Cómo superarlas? Un buen modo es recurrir nuevamente a la filosofía, esta vez, para revisar algunas caracterizaciones de lo humano que siguen operando en nuestra dilemática relación con la tecnología. Ello nos permitirá esclarecer las preguntas a las que deberíamos atender para no caer falazmente en posicionamientos estériles.

### III

Uno de los modos habituales de caracterizar la identidad humana es por el reconocimiento de diferencias en relación al resto de los seres vivos. Este se apoya, por lo común, en algún rasgo distintivo no poseído por los animales, por ejemplo, el lenguaje, la capacidad de crear herramientas u obras de arte, la capacidad de vivir en sociedades y/o la reflexión de los fundamentos últimos de las cosas en virtud de poseer racionalidad. Otras posiciones, en cambio, encuentran el rasgo distintivo en la creencia de una peculiar filiación de los hombres con la divinidad o con los dioses. En el caso de la religión católica, el hombre fue creado por Dios a su imagen y semejanza, es su hijo y en virtud de ello ha heredado la administración y cuidado de la naturaleza. En los mitos griegos, como el descrito por Platón en el Protágoras, el hombre creado por los dioses junto a los demás seres recibe de Prometeo, como único atributo para paliar su indefensión, el dominio del fuego que este robó a los dioses. Por tanto, la única

posesión con que cuenta es una donación resultado de un robo, dado que el hermano de Prometeo, Epimeteo, ha cometido el terrible error de dejar a los hombres sin atributos, al habérsele agotado los que poseía en el reparto a los otros seres. El mito ofrece en su análisis una interesante caracterización del hombre, en donde se conjugan la incompletud esencial con la necesidad de atenuar la falta mediante la creación de artificios que lo ayuden a afrontar los peligros del mundo, es decir con la técnica. Por otro lado, está el tema de la culpa originaria, ya que si lo único que posee es resultado de un robo, deberá expiar ante los dioses por lo que les ha sustraído. Entre los que retoman el carácter de incompletud del hombre, en forma muy similar a lo expresado en el mito, se encuentra el humanista renacentista Pico della Mirandola, quien en su *Discurso sobre la dignidad del hombre* sostiene que este fue creado por dios sin atributos, con el propósito de que encuentre libremente su propia forma. En palabras del autor:

Adán no te he dado ni un lugar determinado, ni un aspecto propio, ni una prerrogativa peculiar con el fin de que poseas el lugar, el aspecto y la prerrogativa que conscientemente elijas y que de acuerdo con tu intención obtengas y conserves. La naturaleza definida de los otros seres está constreñida por las precisas leyes por mí prescritas. Tú, en cambio, no constreñido por estrechez alguna te la determinarás según el arbitrio a cuyo poder te he consignado. (Pico della Mirandola, 1486/2010, p. 4)

Considerar a la identidad humana desde esta perspectiva, no implica aceptar un defecto sino asumir que dicha situación ofrece una oportunidad incomparable: la de hacerse a sí mismo, pues la carencia lleva consigo la necesidad continua de subsanarla, y por qué no, rebasarla. Pero dicha oportunidad se funda en que el hacer humano no está constreñido, pues, a diferencia de los otros seres, siempre puede modificar la determinación natural. Por ende, si el hombre es el que se auto-crea a través de la fabricación de distintos dispositivos y procedimientos, entonces la posibilidad de realización a partir de la carencia esencial es infinita y el uso de la técnica contribuiría al logro de tal fin. Por otra parte, el hombre se encuentra en una relación tensa con la naturaleza, pues si bien está imposibilitado de inaugurar con su accionar una nueva legalidad natural, el conocimiento cada vez más acabado que posee del funcionamiento de la vida, a través de las ciencias, le permite intervenir técnicamente para provocar o desacelerar procesos “naturales”, en otras palabras, usar a su favor ese conocimiento tecno-científico. Según esta línea de justificación, pareciera que las biotecnologías no hacen más que contribuir a la transformación

permanente de ese ser incompleto, proponiendo la modificación o el mejoramiento constante de las condiciones de vida, tanto en lo que respecta al entorno como de aquellas que hacen a la constitución misma del ser psico-biológico tal como hoy lo conocemos.

A esas modificaciones por vía artificial, se suma el dato inequívoco de que como especie somos resultado de otra modificación previa y natural, la de la evolución biológica, acaecida a lo largo de millones de años por mutaciones y adaptaciones. No obstante, según los datos provistos por la antropología esta evolución se ha detenido hace unos miles de años y coincide con la estabilización de ciertas condiciones del medio ambiente, por ejemplo, menor número de catástrofes masivas a nivel planetario, como el caso del retroceso de las glaciaciones o la separación de los continentes. Así y todo, reconocer que la evolución biológica se ha detenido no significa que no pueda continuar por otros medios. En efecto, algunas teorías, tanto filosóficas como científicas, sostienen que la técnica constituye el medio por el cual la humanidad se sigue modificando a sí misma y esto desde hace mucho tiempo, como lo testimonian el descubrimiento de las primeras herramientas líticas. Así, una teoría de tinte filosófico describe a la técnica como un fenómeno ínsito a lo humano. No es algo creado por el hombre a los efectos de procurar comodidad a la vida y por tanto no trastoca en lo más mínimo una suerte de esencia poseída sino, por el contrario, constituye lo que somos. Prueba de eso son los resultados de la arqueología, que ha encontrado entre los objetos de la vida cotidiana restos de utensilios del periodo prehistórico que muestran el empleo de prótesis de dedos o miembros realizados en distintos materiales, en donde podemos inferir que en vida la víctima fue intervenida para subsanar la amputación de modo artificial. Si esto es así, es decir si la técnica es consustancial a lo que somos, la diferencia de la técnica actual en relación a la del pasado es su aceleración. Nos encontramos así con transformaciones que no consisten sólo en aumentar nuestras fuerzas físicas mediante herramientas o máquinas, o en ampliar el rango de nuestras percepciones mediante el uso de instrumentos, o en conservar la salud o detener ciertos procesos morbosos mediante fármacos de diseño, sino que la modificación trastocaría de modo sustancial lo que somos como especie. Este fenómeno es totalmente novedoso respecto a los ocurridos en nuestra historia pasada. Por primera vez, nos encontramos asistiendo al alumbramiento de un acontecimiento de profundas repercusiones para toda la humanidad: la aparición de una singularidad.

## IV

Ahora bien, aun cuando afirmáramos que la pregunta por la esencia es una cuestión filosófica y por tanto siempre abierta, el dato fáctico de que en un futuro no muy lejano se modificaría la línea germinal de nuestro código genético y con ello los rasgos genotípicos de nuestra especie, nos enfrenta a la prueba insoslayable de que este modo de ser “tan apreciado” de lo humano y del cual nos preguntamos por su esencia, no obstante, es resultado de una contingencia evolutiva. Por tanto, muchos se verían tentados a argumentar (como lo hacen los transhumanistas) que dicho modo muy bien podría ser reemplazado por otro, pero esta vez en virtud de una evolución dirigida, que inauguraré una etapa a la que denominaríamos “post-humana”, irreconocible respecto a una etapa anterior, la “humana”. Algunos autores proponen como ilustración de una situación de estas características a aquélla que se produciría si nos encontráramos con un antepasado nuestro, por ejemplo un neardentalense; seguramente este encuentro nos provocaría extrañamiento y desconcierto, tal como le sucedería a un post-humano respecto a nosotros.

Ante la factibilidad de concretarse este panorama, otros sostienen, como el caso de Hans Jonas, que tales modificaciones deben ser impedidas y para ello es necesario elaborar una ética preventiva. Su análisis parte del dato inequívoco de que todas las éticas conocidas han supuesto que la condición humana resultante de la naturaleza del hombre y de las cosas permanece fija de una vez y para siempre. Sobre esa base es posible determinar con claridad y sin dificultades el bien humano y, por último, el alcance de la acción humana; por ende, la responsabilidad humana está estrictamente limitada. Sin embargo, a la luz de los resultados de las biotecnologías, aun en su fase experimental, no podemos desconocer que tales afirmaciones se han vuelto obsoletas. A partir de este dato insoslayable, Jonas se pregunta de qué modo afecta esa técnica a la naturaleza de nuestras acciones, en qué medida hace que las acciones se manifiestan de modo distinto a cómo lo han hecho a lo largo de todos los tiempos, puesto que en ninguna de las épocas careció el hombre de técnicas. Responder estas cuestiones es fundamental para el autor, pues de ello dependerá la elaboración de una ética que proteja esa herencia para las generaciones futuras. De lo que se trata, finalmente, es de conservar la herencia de una evolución precedente, la cual “no puede ser tan mala, ya que parece haber

dotado a sus actuales titulares de la capacidad –que ellos mismos se han atribuido– de juzgar sobre lo bueno y sobre lo malo. Mas esa herencia puede perderse” (Jonas, 1979/1995, pp. 73-74). Así y todo, esta

“propuesta no viene a modificar conceptos tradicionales asociados a la moral, como el bien y el mal, lo permitido y lo prohibido, pero sí supone la irrupción de un nuevo criterio para el juicio moral: la temporalidad, los efectos en el futuro de decisiones que pueden tomarse hic et nunc”. (Giner Comín, 2005, p. 24)

La ética de Jonas postula que el sujeto moral debe considerar objeto de su responsabilidad a seres humanos futuros, es decir, seres que todavía no son y que, por tanto, desde una óptica tradicional, no tienen derechos ni pueden ejercer sus deberes. En ese sentido, las biotecnologías de la línea germinal suponen un riesgo muy grande, aun cuando ofrezcan buenas razones de mejoramiento de lo humano, y ninguna perspectiva de éxito justificara tal riesgo, pues estaríamos decidiendo sobre las vidas de seres humanos que no existen todavía.

Ahora bien, determinar razones por las cuales la herencia de la evolución debería ser conservada tal cual, es tan difícil como especificar razones que justifiquen modificaciones heredables. Según Jonas, la conservación de la primera está bien justificada, pues:

la sencilla verdad, ni sublime ni humillante, pero que respetuosamente debemos aceptar, es que el ‘hombre auténtico’ estuvo ahí presente desde siempre, con su nobleza y su bajeza, con su grandeza y su miseria, con su felicidad y su tormento, sus justificaciones y sus culpas, en definitiva, con la ambigüedad que le es indisoluble. Pretender suprimirla equivale a acabar con el hombre en su insondable libertad. (Jonas, 1979/1995, p. 348)

Si bien comprendemos el sentido de su afirmación, no podemos ignorar que esa herencia valorada que nos constituye en la humanidad que actualmente somos es el resultado azaroso de la evolución y, más importante aún, no es tan fácil circunscribirla, dado que hace mucho tiempo venimos interviniendo en ella a través de las técnicas, en el ambiente, en la alimentación, en el uso de fármacos, etc. No sostengo por ello que debemos dar rienda suelta a lo que podemos hacer sin meditar previamente sobre los límites del accionar biotecnológico, lo que intento mostrar son las dificultades de justificar la defensa de una herencia que debe permanecer inalterable. En posiciones como estas en donde se defiende la permanen-

cia de un estado al que nos llevó la evolución como bueno en sí mismo, se advierte una valoración positiva por lo natural *per se* en detrimento de todo aquello que es de procedencia artificial. Al respecto, es habitual considerar que los cambios producto del azar, aquellos en los que se deja actuar libremente a la naturaleza, son preferibles a los que resultan del control vía intervención tecnológica. En el primer caso, la preferencia se basa en que si se deja actuar al azar se produce algo así como una “justicia natural” que operaría en todos los individuos de igual modo desde la misma concepción, por lo menos, en lo que concierne al patrimonio genético que poseemos. Sabemos que el patrimonio genético de todo individuo es heredado, pues resulta de la combinatoria de los patrimonios poseídos por los progenitores; el resultado de esa combinatoria es un genoma propio, único para cada individuo. Por supuesto, muchos considerarán que esa “justicia”, resultado de una naturaleza ciega, a partir de la cual cada uno de nosotros es constituido, inaugura otras formas de injusticia, por ejemplo, desigualdades resultado de enfermedades congénitas incapacitantes que se hubiesen impedido de haber controlado la combinatoria que constituye el patrón genético. Sin embargo, aun así, estarían dispuestos a correr los riesgos, pues son esos rasgos los que nos iguala en nuestra condición humana, en lugar de intervenir para obtener individuos “a medida” del gusto de los progenitores, de la sociedad o del mercado. En efecto, una consecuencia de la intervención, es que los individuos se encontrarían desde su concepción determinados por otros, y no en virtud del material genético heredado, sino por un diseño previo que respondería a las intenciones circunstanciales, sean cuales fueran, de sus progenitores. En otras palabras, una vez que se logró controlar el azar, el nuevo individuo es resultado de un deseo aún más fuertemente delineado que el referido por el psicoanálisis, es decir, un deseo que busca modificar de modo sustancial el tipo de individuo humano que debería ser engendrado.

Las otras corrientes, enroladas en el transhumanismo, sostienen, en cambio, que no se debe poner límites a las modificaciones. Aducen a su favor que, si se evitaran, se coartarían las libertades de los individuos –dicho sea de paso, los transhumanistas se olvidan que la autodeterminación es uno de los logros de la herencia evolutiva– a participar activamente en esos cambios, desde el momento en que los mismos se encuentran garantizados por la tecnología. Esta propuesta reconoce, por supuesto, que no todos los individuos estarían en condiciones a acceder económicamente a tales posibilidades. De resulta de esto, tendríamos una desigualdad social

más esencial que las ya conocidas, pues coexistirían individuos que poseerían distintos “estados de humanidad” por intervención tecnológica, lo cual nos conduce a un escenario como el descrito por la ciencia ficción, tal el caso de “El mundo feliz” de Huxley o “1984” de Orwell.

## V

Como podemos advertir, y a modo de síntesis de lo tratado hasta ahora, la perspectiva que nos ofrecen los desarrollos biotecnológicos desestabiliza los presupuestos más profundos que poseemos y nos lleva a preguntarnos sobre los límites éticos que como sociedad debemos aplicar sobre sus transformaciones. Como ya señalamos, esto es difícil de resolver dado que toda vez que una tecnología habilitó una transformación fue llevada a cabo, tanto en los entornos ambientales, como en los procesos orgánicos o en la constitución psíquica de los individuos, y todas ellas con el propósito manifiesto de mejora. Ahora bien, ¿cómo justificar cuáles de esos cambios deben ser permitidos y cuáles admitiríamos como cuestionables a sabiendas de que los primeros modifican la herencia dada? pregunta que nos conduce al punto inicial de este trabajo, ¿qué entendemos por lo humano?

## Referencias

- Giner Comín, I. (2005). Introducción. En H. Jonas, *Poder o impotencia de la subjetividad* (I. Giner Comín, trad.). Barcelona. Paidós.
- Heidegger, M. (2001). La pregunta por la técnica. En E. Barjau, (trad.), *Conferencias y artículos* (pp. 9-37). Barcelona: Ediciones del Serbal (Obra original de 1954)
- Jonas, H. (1995). *Principio de responsabilidad: Ensayo de una ética para la civilización tecnológica* (J.M. Fernández Retenaga, trad.). Barcelona: Herder. (Obra original de 1979)
- Pico della Mirandola, G. (2010). *Discurso sobre la dignidad del hombre* (A. Ruiz Díaz, trad.). México: UNAM. (Obra original de 1486)
- Stiegler, B. (2002). *La técnica y el tiempo II: La desorientación* (B. Morales Bastos, trad.). Gipuzkoa: Hiru. (Obra original de 1996)



## Características principales para una cartografía de las epistemologías del sur

Gabriel Paravano\*

### Introducción

Este artículo forma parte del proyecto de investigación: “Epistemologías del Sur. Cartografía de la producción teórica latinoamericana del siglo XXI”, perteneciente al Instituto de Filosofía de la Universidad Nacional de San Juan (UNSJ).

Desde principios del siglo XXI han surgido una serie de reflexiones sobre la producción del conocimiento que están orientados, por un lado, a responder a la crítica relativista que dejó Feyerabend, Kuhn y Hanson y, por otro lado, a mantener una distancia con ciertas formas de comprender al conocimiento o bien de manera universal y segura, o bien de acuerdo con una posibilidad específica de intelección de la realidad. De allí que se encuentren las perspectivas epistemológicas de la complejidad, del naturalismo, del feminismo, etc.

Dentro del marco de los análisis sobre las Ciencias Sociales en Latinoamérica, y siendo heredera del camino de reflexión sobre el conocimiento certero mencionado, cada vez más se mencionan estudios, trabajos, lecturas, jornadas, etc., respecto a lo que se denomina “Epistemología del Sur”, una propuesta iniciada por el jurista, sociólogo y epistemólogo Boaventura de Sousa Santos que, principalmente, subraya la importancia que adquiere la reflexión epistemológica contextualizada, buscando una mejor comprensión de las experiencias sociales, al tiempo que alerta del carácter engañoso de la formulación de epistemologías neutrales. Una simple búsqueda en catálogos y repositorios de revistas, libros y proyectos da cuenta de la gran cantidad de producciones, reflexiones, posgrados, etc. que se auto-designan como participantes de la producción teórica de las Epistemologías del Sur.<sup>1</sup> La cuestión que inmediatamente surge es clara:

---

<sup>1</sup> Para no abundar con ejemplos innecesario, y como tampoco se pretende establecer un ejemplo resolutivo sino, más bien, simplemente ilustrativo respecto a la producción teó-

\* Universidad Nacional de San Juan (UNSJ), Instituto de Filosofía. San Juan, Argentina. paravano69@gmail.com

¿cuáles son las características de las Epistemologías del Sur?, ¿poseen una flexibilidad teórica tal que permiten su aplicación, solicitud o amparo para la amplitud registrada de estudios científicos críticos?

## 1. Condiciones de una cartografía de las epistemologías del sur

No se pretende desvalorizar el esfuerzo sostenido de grupos de investigadores en la creación de saberes desde las perspectivas de la Epistemología del Sur sino, en todo caso, detallar en qué consideraciones, análisis y críticas están de acuerdo o poseen cercanías teóricas y prácticas, ¿cuál es la terminología compartida?, ¿qué ideas forman el piso en común de esta manera de entender a la producción del conocimiento y a los sujetos cognoscentes?, ¿qué pretenden responder, resolver o modificar desde esa posición? Para ello, el grupo de investigación ha diseñado un “Glosario” que sirva como una *cartografía conceptual* de lo que es, de lo que supone, cuáles son límites y relaciones, etc., de las Epistemologías del Sur. La cartografía, en principio, se entiende aquí como un mapeo teórico, i.e., la identificación de paisajes, linealidades, depresiones, relieves, periferias, que operen como cartas de orientación sobre lo que es el territorio de las Epistemologías del Sur, y que sirvan para el (re)conocimiento de los límites fronterizos que implican de por sí la convivencia de producciones de conocimiento plurales.

Las primeras líneas claras que sirven como eje de coordenadas es una oposición binaria, que constituye y articula a la propuesta o, al menos, el sentido de la reflexión de la propuesta que debería seguir todo análisis que la tenga como base, a saber: la “Epistemología del Norte” frente a la “Epistemología del Sur”. En primer lugar, “Norte” y “Sur” no son términos estrictamente geográficos, pues sirven como “metáforas generadoras de conciencia” (De Sousa Santos, 2007, p. 43) al permitir el reconocimiento de subjetividades que, de otro modo, quedan al margen de las reflexiones epistémicas. En segundo lugar, y en términos generales, la Epistemología del Norte se la caracteriza como toda búsqueda básica del conocimiento en términos de progreso, ganancia, acumulación, extensión, profundiza-

---

rica de esta perspectiva epistemológica, se remite al motor de búsqueda de la Red de Revistas Científicas de América Latina y el Caribe, España y Portugal que cuenta, a la fecha (26/06/19), con 17.665 artículos clasificados en la entrada “Epistemología del Sur”. (<http://www.redalyc.org/busquedaArticuloFiltros.oa?q=Epistemolog%C3%ADa%20del%20Sur&a=&i=&d=&cvePais=&idp=17665#panel>)

ción, etc., mientras que la Epistemología del Sur se la caracteriza como toda búsqueda del conocimiento en la que prima el bienestar y el crecimiento inter-subjetivo de la comunidad total a la que estará destinado tal conocimiento (Paravano, 2018). Se comprende que cada sección teórica y práctica ocupada por la Epistemología del Norte y la Epistemología del Sur abarcan una “globalidad”, i.e., la construcción de una de la diversidad de un territorio específico. Cada epistemología se hace visible en su globalidad, en los filtros y límites que establece para definir qué pertenece y qué no pertenece a su concepción de lo real. Estas formas de entender a la realidad y de determinar quién puede habitarla, cómo debe ser conocida y qué prácticas son las adecuadas según sus características inherentes, dan lugar a un parteaguas de cómo se comprenden nociones ontológicas, gnosológicas, sociológicas, etc.

Las características que se han de enunciar y especificar a continuación, y que permiten constituir la estructura de un Glosario sobre las Epistemologías del Sur, busca ofrecer una forma de resolución de este problema sin necesidad de caer en un simple “informe de avance” si no, al contrario, proponer las condiciones mínimas, de carácter normativo, para producir saberes desde las Epistemologías del Sur.

## 2. Presentación general del contenido del glosario

El Glosario que se está conformando dentro del proyecto de investigación: “Epistemologías del Sur. Cartografía de la producción teórica latinoamericana del siglo XXI” tiene, actualmente, proyectado 177 expresiones con sentidos propios y/o con usos específicos dentro de las Epistemologías del Sur. Claramente, no se pueden reproducir los sentidos, significados y definiciones construidos a partir de las investigaciones del *corpus teórico* ya que, la mayoría de ellos, ocuparían la extensión de un artículo. Pero, sí se puede señalar que la *constelación de expresiones* que utiliza la Epistemología del Sur tiene el interés de dar cuenta de cómo se caracterizan los criterios de demarcación del saber contemporáneo, p. ej., reconstruir cómo se establece, define, propone, sostiene, o divulgan las oposiciones de “ciencia y no-ciencia”, “verdadero y falso”, “coherente y no-coherente”, etc. Tal interés yace en la siguiente interpretación: los criterios de demarcación que dan lugar a la configuración contemporánea del saber “legítimo”, terminan estableciendo o dando lugar a prácticas de discriminación, como sean las relaciones entre científico e ignorante, ciencia y popular, autoridad y

prejuicio, progreso y estancamiento, bueno y malo, útil e inútil, etc.

La *constelación de expresiones*, que habrán de ser recuperadas en la cartografía de las Epistemologías del Sur, se configura mediante un objetivo específico o una problemática, la cual puede resumirse en las siguientes interrogantes:

1. ¿Realmente los otros-saberes son no-ciencia?
2. ¿No es casualidad que un número de regiones específicas sean las no-científicas y las no-desarrolladas?
3. ¿No hay conexión alguna entre la idea de inferioridad científica socio-política y la historia-economía de una región?
4. ¿Por qué se concluye la disputa epistémica entre una división de “Nosotros” contra “Ellos”?

Cada expresión se presentará, consecuentemente, como parte de una inmensa estructura de análisis epistemológico sobre las *condiciones de posibilidad de los saberes contemporáneos*, una serie de instrumentos conceptuales que deberían orientar o seguir (desde su génesis) a las consecuencias de tales o cuales concepciones del saber junto con las prácticas, sujetos y/u objetos que se le relacionan (directa e indirectamente).

### 3. Relaciones de las epistemologías del sur

En el apartado anterior se dio cuenta de qué ideas constituían el eje de coordenadas de esta perspectiva: las Epistemologías del Norte y las Epistemologías del Sur. Con esta organización discursiva, se obtuvo un sencillo mapa para ubicar a las Epistemologías del Sur. Estas se conciben como una propuesta alternativa a un modelo de producción de saber, que denominan “Epistemologías del Norte”, y sobre las cuales se pueden ubicar a las Epistemología Neopositivistas, Racionales y de la Complejidad, todas ellas comprenden el conocimiento en términos translocales, transnacionales y/o transcendentales. Entonces, en principio, las Epistemologías del Sur serían parte de las Epistemologías Alternativa y Críticas, en donde se encontrarán las Epistemologías Relativistas, Pragmatistas, Naturalistas, Feministas, Sociales y Críticas (propiamente dichas).

Pero, en realidad, se hace una segunda separación. Por un lado estarían las Epistemologías Alternativas al modelo hegemónico de concebir a

las ciencias: las “Relativistas” (p. ej., Kuhn, Feyerabend), las “Pragmatistas” (p. ej., Dewey, Addams) y las “Naturalistas” (p. ej., Giere, Kitcher), las cuales no serían propiamente “del Sur”, debido a que realizan sus discusiones en el seno de la ciencia europea; y, por otro lado, estarían las Epistemologías del Sur, con los estudios Feministas (p. ej., Harding, Ambroggi) y Socio-Críticos (p. ej., Dussel, Mignolo), que estarían produciendo saberes propiamente locales y contra-hegemónicos. En consecuencia, se presentan tres grandes grupos de pensamiento epistemológico: del Norte, Alternativas y del Sur. Cada grupo estaría atravesado por cinco grandes conjuntos o linealidades.

Para graficarlo, se ha decidido presentar a un par de las nociones de forma vertical, pues se las considera como parte de un eje desde el cual se dan las principales diferencias entre cada grupo; las otras nociones se presentan de manera horizontal pues, en un primer acercamiento, se comprenden como el piso en común que da lugar al debate de cómo se comprende al conocimiento en general y al conocimiento legítimo (o científico) en particular.



Fig. 1: Distintas nociones que atraviesan a los grandes grupos epistemológicos

De este modo, en primer lugar, y de forma vertical, se presentan las nociones de “sujeto” y de “naturaleza”, es decir, según cómo se comprenda al sujeto productor de conocimiento y el objeto sobre el cual trabaja, se comprenderá las posibilidades del saber científico, el tipo de problemáticas, prácticas y metodologías que constituyen su horizonte y qué beneficios, progresos o soluciones cabe esperar de ello.

En segundo lugar, y de forma horizontal, se presentan las nociones de “historia”, “discurso” y “saber”, es decir, el saber posible de la relación sujeto-naturaleza tendrá una posibilidad (o un conjunto de posibilidades) de ser verbalizado, transmitido, publicado, divulgado y hasta aplicado lo cual, a su vez, juega con las condiciones contextuales desde las cuales se ha producido/creado, como sean los rasgos culturales de la época, las influencias socio-políticas, los cambios institucionales, etc.

A partir de estos conjuntos o linealidades, el primer mapa que se obtuvo y, con ello, la primera serie de nociones a definir y estudiar, se puede apreciar en la Figura 1.

#### **4. Límites de las epistemologías del sur**

Como se aprecia en la Figura 1, se establecen las pautas para un cuadro comparativo de doble entrada, por un lado la nomenclatura “epistemología de...” y, por el otro lado, algunas de las categorías mencionadas. Útil para establecer las primeras distinciones necesarias para identificar las características de las Epistemologías del Sur pero inútil para dar cuenta de si poseen tal flexibilidad teórica que permita su aplicación para cualquier estudio crítico. De allí la necesidad de constituir propiamente el “Glosario”. Generar un código de lectura, con un orden específico (en este caso, alfabético) que dé cuenta de qué son, y qué suponen, las Epistemologías del Sur, especificando cuáles son las expresiones comunes de su campo, los usos y sentidos que encierra y en qué obras pueden ser encontradas.

Sobre las primeras dos constelaciones de nociones conformadas: sujeto-naturaleza e historia-saber-discurso, y en base al desarrollo de las lecturas, surgieron subconjuntos o pequeñas agrupaciones de ideas, de términos o conceptos pertenecientes a cada noción en particular lo cual, en parte complejizaba el sistema de análisis, diagnóstico y concepción de las Epistemologías del Sur, y en parte disminuye el espectro de estudios a considerar para comprender a cualquier análisis epistemológico como

“del Sur”. Así, el eje vertical de sujeto-naturaleza este se desglosa en cada aspecto y en distintos subniveles de significado.

Por un lado, la noción de sujeto puede ser entendida desde la relación “individuo-comunidad”, “humano-etnia” y/o “familia-sociedad”, permitiendo apreciar no sólo términos corrientes de las Epistemologías del Norte (p. ej., “comunidad científica”) sino de las Epistemologías Alternativas (p. ej., “sociedades científicas”) y, obviamente, del Sur (p. ej., “saberes étnicos”). Pero estas tres subdivisiones suponen, a su vez, distintos campos o territorios de aplicación, uso o problematización, a saber: “derechos y deberes”, “dominio, emancipación y autonomía” y “riqueza, política e institución”.

Por otro lado, la noción de naturaleza puede ser entendida desde la relación de “finito-infinito”, “verdad-mentira”, “real-irreal-fantasia”, “metafísica-física” y/o “fijo-dialéctica-progreso”, permitiendo también apreciar términos corrientes de las Epistemologías del Norte (p. ej., “progreso científico”), de las Epistemologías Alternativas (p. ej., “dialéctica negativa”) y del Sur (p. ej., “infinitud de saberes”). Al igual que la noción de “sujeto”, la de “naturaleza” comprende también una serie de campos (espacios, territorios, zonas, etc.) sobre los que se pone en juego la posibilidad de aplicación, de uso o de problematización de estos términos, a saber: “orden y desorden”, “utopía, localidad y lugar”, “movimiento y matriz” y “aquí y allá”.

A partir de estas subdivisiones y sus posibles campos de aplicación, ya se presentó un segundo mapa y, con ello, la segunda serie de nociones a definir y estudiar; puede decirse que poseía el aspecto que se aprecia en la figura 2. Allí los campos no se corresponden uno-a-uno con cada subdivisión de cada una de las nociones analizadas sino, más bien, permiten tanto un estudio o bien delimitado así mismo o bien solapado con otro territorio. De este modo se intenta por mantener la dinámica o movilidad defendida en el propio discurso de las Epistemologías del Sur, sobre la cual se asume que la infinidad de experiencias posibles en la realidad es garantía para una infinidad de prácticas y teorizaciones y, por lo tanto, de constitución de saberes dedicados a resolver problemáticas localizadas en modos de vida de colectividades de subjetividades.

A su vez, se dio una inversión de las nociones “naturaleza” y “sujeto” con fines representativos, es decir, dar cuenta la agrupación de términos y sus campos posibles de aplicación. Cabe remarcar que esta estructura que aquí se está desarrollando no ha sido presentada por ningún pensador de

las Epistemologías del Sur, ha sido creada por el equipo de investigación con el fin de identificar las características de este pensamiento: cuáles son sus núcleos teóricos, cuáles es el sentido de sus expresiones, y cuáles son las posibilidades y los lugares de uso de las mismas. Logrando, con lo anterior, especificar el campo de nociones para crear tanto el Glosario como la Cartografía de la Epistemologías del Sur. De forma más sencilla, dar cuenta de qué dicen y desde dónde lo dicen.

Queda, en esta instancia, dar cuenta de las relaciones y los campos que comprenden a las nociones de “historia”, “discurso” y “saber”. Quizás convenga iniciar con la noción de “historia”. Esta supone un único nivel de relación: el movimiento de “pasado-presente-futuro”, cuyas interacciones entre cada uno de los elementos presentan, para la Epistemología del Sur, una adhesión a cierta valorización de la ciencia. La relación “pasado-presente-futuro” desde su clara ambigüedad y amplitud de interacciones, presenta un abanico de campos de aplicación, uso o desarrollo, a saber: “globalidad y cultura”, “hegemónico y contrahegemónico”, “imperio, fascismo y democracia”, “colonial, decolonial y poscolonial”, “modernidad y posmodernidad”, “progreso y beneficio”.

La otra noción que presenta una sola relación es la de “saber”, lo cual no quita su complejidad de análisis y de comprensión, pues se presenta desde la interacción “poder-fundación-paradigma”. Supone una revisión de los parámetros a partir de los cuales se estudia el “saber” en tanto que científico, legítimo, objetivo, útil, etc. Así, ¿es el poder lo que le da status a las ciencias?, ¿responde a su carácter fundacional (sea en términos institucionales, sea en términos gnoseológicos)?, o ¿es la consecuencia de un consenso de un grupo específico de subjetividades destinado a resolver una problemática localizada que luego fue extrapolada a otros modos de vida? Los campos de las interrogantes anteriores (las cuales nunca se agotan en las que este artículo puede ejemplificar) son los siguientes: “dicotomía y cosmopolita”, “bueno, malo y crudo”, “abismal y postabismal”, “útil e inútil”, “conocimiento, ignorancia y ausencia”.

Las interrogantes se multiplican constantemente en razón de las variables a tener en cuenta, pero dicha multiplicidad también se presenta como un criterio cada vez más estricto, pues las condiciones a cumplir son cada vez más. Se genera la presunción de que no muchas investigaciones, estudios, propuestas, etc.- lograrán recorrer toda la topografía de las Epistemologías del Sur y, por lo tanto, ser admitidas como pensamientos que contribuyen a las mismas.

Por último, la noción de “discurso” se comprende desde cuatro relaciones, a saber: “consciente e inconsciente”, “racional e irracional”, “particular, universal y regular”, “traducción, transfiguración y transición”. La presencia de estas relaciones es una declaración de las Epistemologías del Sur de su relación íntima con las Ciencias Sociales y de su aspiración a configurarse finalmente como una Epistemología de las Humanidades. En este punto, entran en juego las cuatro tesis de las Epistemologías del Sur: todo el conocimiento científico natural es científico social, todo el conocimiento es local y total, todo el conocimiento es autoconocimiento, y todo el conocimiento científico busca constituirse en sentido común (De Sousa Santos, 2008; López, Ponce, & Paravano, 2019).

A partir de estas sub-divisiones y sus posibles campos de aplicación, se configura un tercer mapa y, con ello, la tercera serie de nociones a definir

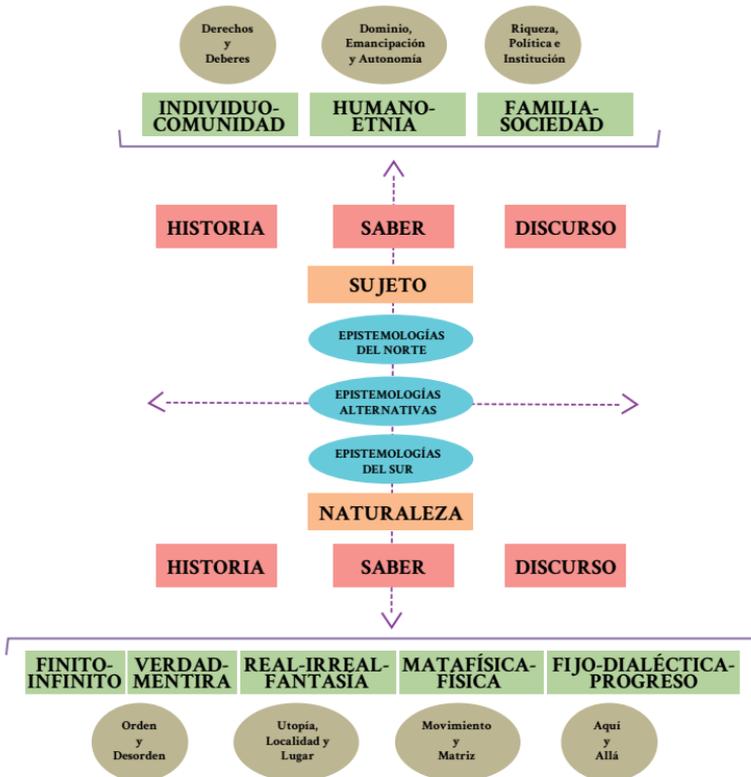


Fig. 2: Ampliación de los conceptos analizados en las subdivisiones de la figura 1.

y estudiar, de este modo puede decirse que el mapeo tiene el aspecto que se observa en la figura 3. Por cuestiones de exposición, se ha dejado de lado el eje transversal, para que pueda apreciarse directamente el eje horizontal. Con este último esquema de nociones principales, “sujeto”, “historia” y “discurso” se da cuenta del piso en común que tienen las Epistemologías del Sur respecto a las formas de entender a la producción del conocimiento y a los sujetos cognoscentes, dando las pistas necesarias para identificar sus pretensiones de estudio.

Luego, lo expuesto en los puntos anteriores, se podría sintetizar del siguiente modo:

- La “ciencia” se entiende como una propuesta de globalidad cognitiva, es el resultado de la interacción de agentes, lugares y producciones contemporáneas, simultáneas (históricamente hablando), diversas y singulares (dada la infinitud de experiencias humanas).
- La “verdad” se entiende desde un “saber”, i. e., como un producto de una serie de consideraciones pragmáticas, en donde se suponen modos específicos de relacionarse con el mundo, por lo que, al no existir una forma única de relación sujeto-mundo, no existe un saber único, sino que existen infinitos saberes.
- El “progreso” se entiende como todo aquello que contribuya a la co-existencia beneficiosa de los “saberes otros”, de allí que se incurra en la expresión “ecología de saberes”, i. e., la producción de saberes que conviven y se interrelacionan de forma continua y dinámica sin perder su propia autonomía (en este aspecto funciona el criterio, p. ej., ético-cívico del cosmopolitismo).
- El “método” se subdivide en cuatro formas de estudio, consistentes con la demanda de conversión de las Ciencias Naturales en Ciencias Sociales y estas en Humanidades.
- La “unidad de análisis” puede identificarse en los conjuntos de “metaforas generadoras de conciencia” (p. ej., los citados “Fascismo Epistemológico”, “Pensamiento Abismal”, “del Norte”, “del Sur”, etc.) y los “principios organizadores de totalidades” que representan.
- El “lenguaje” se entiende como una “globalidad discursiva”, es decir,

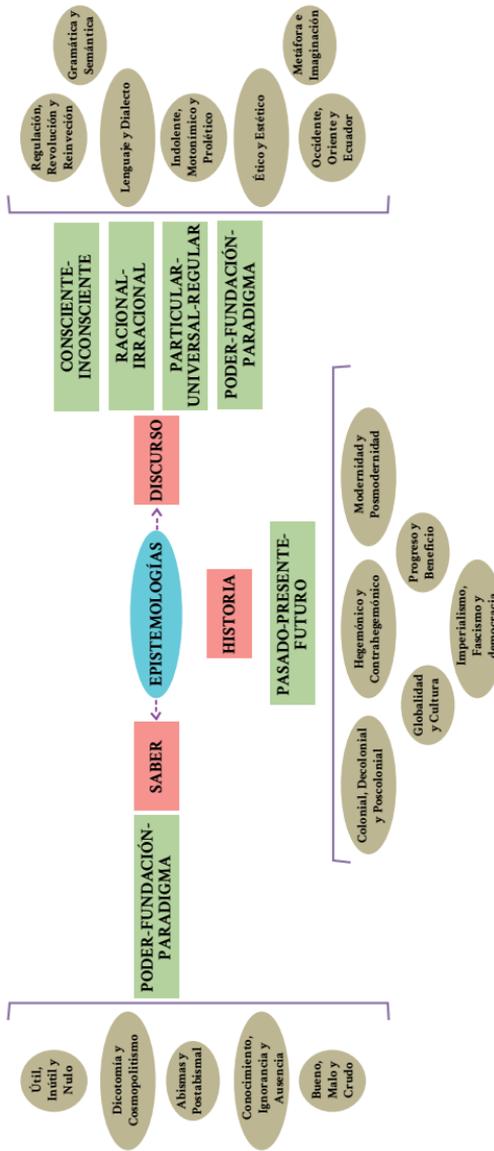


Fig. 3: Serie final de nociones interrelacionadas en el estudio de las Epistemologías del Sur que muestran del piso común con respecto a las forma de entender la producción del conocimiento.

un universo de objetos y sujetos sistematizados según una forma particular de relacionarse con el mundo.

- El “sujeto productor de conocimiento” es una temporalidad, cargado de diferencias, constituido a partir de múltiples escalas y producciones territoriales, por lo que, no hay “sujetos en sí” sino, más bien, “subjetividades”, i. e., emergencias de individualidades a partir de saberes propios que responden a un modo de producción y una necesidad particular.

## 5. Reflexión final

Se presenta una dinámica en la perspectiva epistemológica del sur que, probablemente, sea parte de su carácter atractivo de los estudios socio-críticos contemporáneos respecto a la construcción y legitimación del conocimiento pues, a modo de un *Cubo Rubik*, se poseen una serie de casillas, espacios, límites que comparten una dimensión en común, la “naturaleza”, el “sujeto”, el “discurso”, la “historia” o el “saber”, y en la medida en que se relacionan van produciendo tanto una mixtura como una imagen más clara de dicha dimensión.

Si las respuestas se configuran dentro de las condiciones que demanda el eje del Norte (es decir, translocal, transnacional y trascendente), pues la producción teórica no podrá ser reconocida como un aporte a las Epistemologías del Sur, por más que así se lo proponga. En cambio, si lo hacen dentro de las condiciones del eje del Sur (es decir, local, cosmopolita y ecológica), pues se comprenderá que tienen posibilidad de ser reconocida como un aporte a las mismas.

Ahora, no deja de manifestarse que la flexibilidad teórica pareciera no existir como tal, y que necesariamente debe ser así, pues si se demanda un tipo de reflexión epistemológica que no sea una reproducción de las prácticas hegemónicas del saber, es necesario una toma de consciencia de qué posición se toma respecto al mundo y respecto a los sujetos de conocimiento (que en las Epistemologías Alternativas se pone de manifiesto), como también de las consecuencias de su convivencia con otros saberes, ignorancias, modos de relación y de creación de subjetividades.

Cualquier investigación que no pueda cumplir estos requisitos, por más que se declare participante de las Epistemologías del Sur, no lo será. Podrá ser una epistemología alternativa, pero no una alternativa a las al-

ternativas epistemológicas del s. XXI como así se proponen las Epistemologías del Sur.

## Referencias

De Sousa Santos, B. (2007). *Conocer desde el Sur*. CLACSO.

De Sousa Santos, B. (2008). *Un discurso sobre las ciencias*. Cortez.

López, A., Ponce, A., & Paravano, G. (2019). ¿Cómo se articula el Fascismo Epistemológico con el Fascismo Social? Reconstrucción y reflexión crítica desde el marco de las Epistemologías del Sur. *3° Congreso de la Asociación Argentina de Sociología (Pre-ALAS Perú 2019). 2° Jornadas provinciales de Sociología de San Juan. «Reflexiones y oficios de una ciencia crítica: Medio siglo de sociología en Argentina»*, 4 a 6 de Septiembre, CLALSO-UNSJ.

Paravano, G. (2018). La dicotomía Norte-Sur del globo desde la perspectiva epistemológica. *8° Encuentro de Investigadores de Ciencias Sociales de la Región Centro Oeste y 5° Binacional con la IV Región de la República de Chile*, 17 al 19 de Mayo, UNSJ-FACSO.



# De la filosofía política a la economía política

## La especificidad de la economía política como disciplina

Ruth Pustilnik\*

**E**n este artículo mostraremos que la filosofía política comparte con la economía política la misma problemática. Y que esa problemática es planteada por ambas disciplinas en los mismos términos.

Lo que cambiará de manera significativa entre ambas disciplinas es la estrategia explicativa o la manera de responder la misma pregunta. Mientras que en el primer caso el lazo social se explica dentro del campo de la filosofía, de allí su denominación filosofía política, en el segundo caso dará lugar al comienzo de la economía política, una disciplina perteneciente al campo de la ciencia.

### La problemática de la filosofía política

En *Leviatán o la materia, forma y poder de una república eclesiástica*, Hobbes (1615/2017) muestra que en estado de naturaleza el hombre vive en estado de guerra. Estado de guerra no significa que está efectivamente luchando, sino que existe la amenaza permanente a ser atacado. La razón en estado de naturaleza le muestra al hombre que le conviene buscar la paz, sin embargo, sus pasiones lo llevan a desoír a la razón y actuar de acuerdo a sus pasiones incumpliendo los pactos.

Por ese motivo el autor sostiene que la manera de lograr la paz en la sociedad es que cada hombre por propio consentimiento renuncie al derecho que tiene sobre todas las cosas en estado de naturaleza y lo done al soberano. El súbdito consiente a aceptar toda ley que emane del soberano. Ese consentimiento se da a través de un contrato.<sup>1</sup>

---

<sup>1</sup> Hobbes dice: "Autorizo y transfiero a este hombre o asamblea de hombres mi derecho de gobernarme a mí mismo, con la condición de que vosotros transferiréis a él vuestro derecho, y autorizareis todos sus actos de la misma manera. Hecho esto la multitud reunida en una sola persona se denomina Estado, en latín, *civitas*. Esta es la generación de aquel gran Leviatán o más bien (hablando con más reverencia), de aquel dios mortal, al cual debemos, bajo el Dios inmortal, nuestra paz y nuestra defensa. Porque en virtud de esa autoridad que se le confiere por cada hombre particular en el Estado, posee y utiliza tanto poder y fortaleza

\* Centro de Epistemología de las Ciencias Económicas/Universidad de Buenos Aires.  
58rp23327647@campus.economicas.uba.ar

## Contrato como estrategia explicativa del lazo social

La estrategia explicativa del lazo social se basa en la autoridad, el poder y el consentimiento.

El soberano debe ser el juez último que dirima los conflictos y haga cumplir los pactos. Para ello, el soberano debe tener poder absoluto. El súbdito consiente en obedecer al soberano. Una vez que el soberano asume como tal, lo que él establezca es ley.<sup>2</sup> Aun cuando el súbdito no esté de acuerdo con una determinada ley debe cumplirla (fue exactamente eso lo que pactó), pues no cumplir con lo que el soberano establece conduciría a un estado de guerra que sería mucho más penoso que cumplir cualquier ley.

En tal sentido, el soberano tiene discrecionalidad para dictar leyes<sup>3</sup> y esas leyes establecen el orden social.

No se especifica nada sobre el contenido específico de esas leyes y por tanto no se establece un contenido específico del lazo social.<sup>4</sup>

### El contrato

En *Leviatán*, el lazo social se explica a través de la figura de un contrato,<sup>5</sup> de allí que se denomine contractualismo o corriente contractualista a los filósofos representativos de la filosofía política (Hobbes, Locke y Rousseau). Pareciera que para esos autores ese contrato no algo que suceda realmente. Sin embargo, sí se explica de manera detallada y precisa la necesidad de que la autoridad del soberano sea total o absoluta.

Podemos sostener que, a diferencia de los contractualistas, Adam Smith (1776/2012) propone un tipo de vínculo social –el mercado– del que

---

que por el terror que inspira es capaz de conformar las voluntades de todos ellos para la paz, en su propio país y para la mutua ayuda contra sus enemigos, en el extranjero. Y en ello consiste la esencia del Estado, que podemos definir así: una persona de cuyos actos se constituye autora una gran multitud mediante pactos recíprocos de sus miembros con el fin de que esa persona pueda usar la fuerza y medios de todos como juzgue conveniente para asegurar la paz y defensa común” (Hobbes, 1651/2017, p. 144).

<sup>2</sup> En el capítulo XXIX de *Leviatán* Hobbes explica que la causa principal causa que tiende a la desintegración del estado, es la falta de poder absoluto por parte del soberano.

<sup>3</sup> Aunque Hobbes, a lo largo de *Leviatán*, recomienda ciertos límites que debe tomar en cuenta el soberano a la hora de dictar leyes, como por ejemplo que las leyes no se cambien frecuentemente. Esos límites dejan de todos modos un gran margen de discrecionalidad.

<sup>4</sup> Si bien el capítulo XXVI del *Leviatán* se titula “De las leyes civiles” allí se especifican cuestiones de forma y no cuestiones del contenido específico

<sup>5</sup> Específicamente el contrato es entre cada uno de los súbditos con el resto de los súbditos y una donación de sus derechos al del soberano.

pregona que sí es algo que sucede realmente, le da contenido específico, pretende describirlo y lo detalla.

### **Mercado como estrategia explicativa del lazo social**

De acuerdo con Hobbes, la razón, en estado de naturaleza le muestra al hombre que le es conveniente respetar los pactos y buscar la paz, sin embargo, sus pasiones lo llevan a desoír a su razón y actuar siguiendo su propio interés incumpliendo los pactos o contratos.

Adam Smith mostrará que el hombre escuchando a sus pasiones establece un vínculo social que conduce a una sociedad integrada y que el resultado de ese vínculo social es como si hubiera usado la razón, aunque en realidad sea una consecuencia no planeada. A diferencia de las leyes que emanan del soberano de Hobbes, ese vínculo tiene un contenido específico, no es una ficción ni una convención. Lo que gobierna la vida en sociedad, para Adam Smith, es un conjunto de leyes mecánicas naturales. En ese sentido no existe un corte o discontinuidad entre la naturaleza y la sociedad.

La explicación de la sociedad es un tipo de explicación regulada por leyes objetivas. También las leyes de la naturaleza en Hobbes son objetivas, pero no conducen a la integración de la sociedad, motivo por el que se requiere un contrato.

Otra característica de la explicación social en Adam Smith es que las leyes que regulan el entramado social son leyes causales y mecánicas que se relacionan entre sí a través de un sistema de leyes que se retroalimentan entre sí como veremos más adelante. Esta característica es esencial para que la economía política se convierta en ciencia a partir de Adam Smith.

Una reconstrucción del libro I de *Investigación sobre la naturaleza y causa de la riqueza de las naciones* –mejor conocido como *La riqueza de las naciones*– (Smith, 1776/2012) nos muestra un claro ejemplo de cómo funciona la explicación a través de un sistema de leyes causales y mecánicas que se retroalimentan entre sí: el ser humano posee una propensión a intercambiar,<sup>6</sup> y una tendencia a desear mejorar su propia condición,<sup>7</sup> hasta aquí el tipo de explicación son tendencias y pasiones

---

<sup>6</sup> Smith (1776/2012, p. 16) nos aclara que no es su propósito investigar si el origen de esta tendencia es un principio innato del que no puede darse una explicación ulterior o es la consecuencia de las facultades discursivas del lenguaje. La cuestión que queremos hacer notar es que no son principios que el autor intente siquiera justificar.

<sup>7</sup> Smith (1759/2013, p. 105) lo denomina como pasión egoísta. También aquí la cuestión que

que no se justifican dentro de la obra, esta explicación es muy similar al tipo de explicación de Hobbes al principio de *Leviatán* respecto de las pasiones,<sup>8</sup> pues refiere a ciertas leyes de la naturaleza humana presentadas como simples tendencias comunes a todos los seres humanos.<sup>9</sup>

El individuo tiene una tendencia al intercambio y nota que, si se dedica a un solo tipo de actividad la cantidad que bienes que produce es mayor, por tanto, debido a su tendencia a desear mejorar su propia condición, intentará llevar a cabo esa especialización. Pero para poder producir un solo tipo de bien debe poder sobrevivir, al menos, hasta que pueda adquirir a través del intercambio todos los bienes que le son necesarios para su subsistencia, ya que, si produce, por ejemplo, flechas, durante el lapso de tiempo que las produce no se dedica a cazar ni a recolectar ni a proveerse ninguna de las cosas que necesita para su supervivencia inmediata. Para ello necesita contar con aquellas provisiones que le permitirán subsistir. En ese primer período rudo y primitivo de la sociedad, esas primeras provisiones son un primer y pequeño capital. Qué pueda contar con tales provisiones dependen de la frugalidad del individuo, una característica similar a la prudencia.<sup>10</sup> Para contar con esas provisiones tuvo que abstenerse de consumirlas en pos de un beneficio futuro. Una vez que el individuo adquiere ese pequeño capital intercambia través del trueque, de esa forma se especializa produciéndose la primera división del trabajo.

Esa primera división del trabajo aumenta la cantidad de bienes producidos y por tanto la cantidad de intercambios que se realizan, ya que esos bienes deben adquiridos por más personas, esto hace que se busquen instrumentos para agilizar el intercambio y ese instrumento es la moneda. La división del trabajo genera entonces un aumento de bienes producidos y por tanto una mayor cuantía de riqueza; y los individuos frugales podrán

---

queremos hacer notal es que el autor no desarrolla una explicación respecto de esas pasiones.

<sup>8</sup> Según Hobbes (1651/2017, p. 57): apetito, deseo, amor, aversión, odio, alegría, pena, esperanza, desesperación, temor, valor, cólera, confianza, desconfianza, benevolencia, indignación, bondad, codicia, ambición. Estas pasiones tampoco son explicadas, sino que están dadas.

<sup>9</sup> No desarrollaremos este tema aquí; diremos simplemente que una ley causal tiene la forma universal condicional por ej.: si es metal se dilata con el calor y esto vale para todos los metales y una tendencia no tiene un formato condicional, por ejemplo, el producto bruto del Argentina crecerá, pero no dice si pasa tal evento crecerá solo dice que crecerá.

<sup>10</sup> Según Adam Smith, "En la constancia de su laboriosidad y frugalidad, en su incesante sacrificio de la paz y el disfrute del presente en aras de la expectativa probable de una holgura y gozo mayores en un período de tiempo más distante pero más duradero" (1759/2013, p. 375).

entonces invertir una mayor cantidad tanto porque la cantidad absoluta es mayor, como por tener un mayor incentivo para hacerlo, ya que espera la riqueza futura será mayor aún. Como el individuo busca mejorar su condición en el mismo proceso de división del trabajo, procurará producir cada vez más y esto lo hará mediante el aprendizaje. Creará modos de producir más eficientes como el famoso ejemplo de la fábrica de alfileres y también creará herramientas y maquinarias con ese mismo fin. Esto dará lugar a un nuevo aumento de la cantidad de bienes producidos. En este nuevo ambiente, ya con la división del trabajo desarrollada, seguramente muchos más individuos de inclinarán hacia la frugalidad, pues el premio o recompensa es cada vez mayor. La división del trabajo aumenta el capital y el capital aumenta la división del trabajo de forma continua y acumulativa estimulando un mercado cada vez más amplio.

Sin embargo, a partir de aquí, el tipo de explicación cambia. Será una explicación que utiliza leyes causales mecánicas que se interrelacionan formando un sistema.

Este mercado tendrá un comportamiento que se independiza de la intención o voluntad o razón de quienes participan. Como veremos a continuación en la explicación que aparece en el libro II de *La riqueza de las naciones*. Esto es una explicación que difiere claramente de la explicación a través de un contrato.

Podemos observar explicaciones de tipo causales mecánicas en el segundo libro de *La riqueza de las naciones*. Veamos con detenimiento, como Adam Smith explica el precio de la tasa de interés, la cantidad de capital y la tasa de beneficio:

En la medida que aumenta el capital que se puede prestar en rédito, el interés o sea el precio que se paga por el uso de aquél, disminuye necesariamente, no sólo por aquellas causas generales que hacen bajar el precio de todas las cosas con la abundancia de las mismas, sino por otras peculiares de este caso. A tenor que aumentan los capitales en un país los beneficios resultantes de su empleo disminuyen necesariamente. Cada vez se hace más difícil encontrar un procedimiento para colocar de forma rentable el nuevo capital. Surge por tal causa, la competencia entre los diferentes capitales, porque los dueños de estos buscan aquella colocación que ya está ocupada por otros. No solo es necesario vender algo más barato, sino que en ocasiones se compra a veces más caro. La demanda de trabajo aumenta diariamente con el incremento de los fondos que se destina a su sostenimiento [...]. La competencia encarece los salarios del trabajo y disminuye los beneficios del capital.

Por lo tanto, cuando disminuye de esa forma las ganancias que pueden derivarse del uso del capital, forzosamente tiene que disminuir también el precio que se paga por su uso, o sea el interés que se paga del capital; como si ambas cosas actuaran de consuno, persiguiendo el mismo fin. (Smith, 1776/2012, p. 320)

En esta extensa cita vemos que al aumentar la cantidad de fondos destinados al capital disminuyen las ganancias que se obtiene del capital y eso produce que el precio del capital baje. Esto ocurre sin que los empresarios se lo propongan. En términos causales lo podemos enunciar de esta forma: si aumenta la cantidad de fondos prestables para capital disminuye la tasa de ganancia capitalista. Si disminuye la tasa de ganancia capitalista disminuye la tasa de interés. Si disminuye la tasa de interés disminuyen los fondos prestables y por tanto vuelve a aumentar la tasa de ganancia. La tasa de ganancia se situará en el nivel más conveniente para la sociedad independientemente de la voluntad del capitalista. De allí que el resultado sea igual al que se hubiera obtenido si se aplicara la razón para establecer el vínculo social.

## Comentarios finales

En este artículo mostramos que la economía política tiene un tronco común con la filosofía política. Mostramos también el momento y la forma en que la economía política se separa de la filosofía política y por tanto la especificidad de la economía política como disciplina científica. Esto sucede específicamente cuando Adam Smith postula un sistema de leyes causales y mecánicas que regulan los fenómenos económicos.

## Referencias

- Hobbes T. (2017). *Leviatán o la materia, forma y poder de una república eclesiástica* (M. Sánchez Sarto, trad.). México: Fondo de cultura económica. (Obra original de 1651)
- Smith, A. (2012). *Investigación sobre la naturaleza y causa de la riqueza de las naciones* (G. Franco, trad.) México: Fondo de cultura económica. (Obra original de 1776)
- Smith, A. (2013). *La teoría de los sentimientos morales* (C. Rodríguez Braun, trad.). Madrid: Alianza. (Obra original de 1759)



## Mediación e investigación histórica

Edgar Rufinetti\*

### I

La reflexión sobre el proceso histórico o más precisamente sobre la “historia del efecto”, se encuentra condicionada de distintas maneras. La elección del tema y el interés a él vinculado, por un lado, y la del marco teórico y metodológico por otro, constituyen sin duda elementos relevantes al respecto. Pero también es necesario tener presente la estructuración simbólica del proceso socio cultural. La *res gestae* en tanto producto de las actividades de los seres humanos se encuentra ya siempre estructurada por las acciones e interacciones que estos llevan adelante en el marco de su mundo de la vida compartido.

El carácter simbólico de este “objeto” fuerza al investigador social a tener en cuenta el concepto de sentido; el sentido y el sistema de significaciones que operan en el mundo social de la vida y en el que están inmersos los sujetos sociales, ponen en juego a su vez una *mediación simbólica*. Otro de los elementos centrales que se pone en juego es la instancia de constitución del propio sujeto reflexivo (o de la conciencia) en y por el proceso histórico social. Esto es, que el sujeto, el actor, el investigador es también parte de un mundo de la vida y que en este “trasfondo” y como efecto de él se constituyó o conformó como tal.

Comenzamos señalando que la relación entre presente y pasado, entre el investigador que se vuelve al pasado y su objeto, se encuentra mediada. ¿Cuál es el elemento que media, en qué consiste tal mediación? En este sentido, cuando H. G. Gadamer (1960/1977; 1986a) se propone contrarrestar la concepción de F. Schleiermacher y W Dilthey, según la cual para comprender hay que proceder a una *reconstrucción* de la experiencia vivida o de la estructuración de la expresión, apela al concepto hegeliano de *integración* [*Integration*] y concibe el proceso del *Verstehen* como *mediación* de pasado y presente.

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH) Centro de Investigaciones María Saleme de Burnichon (CIFYH). Córdoba, Argentina. edgarrufinetti@hotmail.com

En esta dirección, nos advierte que más allá de la actitud del historicismo de atenerse a lo dado, de no permitir que ninguna opinión previa sobre el sentido de la historia determine la investigación de la misma, operan en él dos presupuestos. Uno, que esta investigación conforma o “forma una unidad”; y dos, que la “idea de la unidad de la historia del mundo implica la continuidad ininterrumpida del desarrollo histórico universal”. Gadamer señala:

El que la historia del mundo, a lo largo de un desarrollo continuo, haya producido este mundo cultural occidental no es un mero hecho de la experiencia que comprueba la conciencia histórica, sino una condición de la conciencia histórica misma. [Vale decir] . . . sólo porque la historia del mundo ha hecho este camino, puede una conciencia de la historia universal plantear en general la pregunta por el sentido de la historia, y referirse a la unidad de su constancia. (Gadamer, 1960/1977, pp. 265-266 [1986a, p. 196])

Mas, por otro lado, la perspectiva teórico metodológica de la escuela histórica “se apoyó de hecho –sostiene Gadamer– en la teoría romántica de la individualidad, y en su correspondiente hermenéutica” (Gadamer, 1960/1977, p. 253 [1986a, p. 185]). Y, por ello, el esquema fundamental según el cual entiende la investigación histórica es el “esquema del todo y las partes”.

Este viejo principio hermenéutico con su *feedback* lleva ínsito una mediación; sin embargo, con la reelaboración de Schleiermacher, la comprensión se dirige más hacia la expresión que hacia la verdad de lo dicho, los textos pueden considerarse como “puros fenómenos de la expresión, al margen de sus pretensiones de verdad” (Gadamer, 1960/1977, p. 252 [1986a, pp. 184-185]). De este modo, aquello que se trata de entender no es ya una cosa o asunto [*Sache*] sino una individualidad.

Esta última cuestión, la consideración de un texto, de una acción o suceso al margen de sus pretensiones de verdad, nos remite por un lado a Hegel por cuanto este considera que “rehaciendo el contexto histórico” de las obras de arte o de los “frutos del devenir”, “no se adquiere ninguna relación vital con ellos sino solo el poder de imaginarlos” (cf. Gadamer, 1960/1977, p. 221). Por otro lado, nos reenvía a la crítica que realiza E. Husserl (1936/1991) de lo dado y a la descripción de Heidegger (1927/1951) de la estructura circular de la comprensión.

Apoyándose en estos desarrollos, Gadamer cuenta con un nuevo punto de partida. El mundo histórico ya no aparece como mundo objetivado opuesto y extraño al sujeto. Desde G. Vico, uno de los argumentos centrales para sostener la particularidad del campo de estudio de las ciencias humanas y su método, refiere a la “homogeneidad” óptica entre sujeto y objeto, entre conciencia y objeto. Desde ahora, “ni el conocedor ni lo conocido «existen» [*vorhandensind*] «ópticamente» sino «históricamente», esto es, participan del modo de ser de la historicidad” (Gadamer, 1960/1977, p. 327 [1986a, p. 247]).

Gadamer muestra aquí, por un lado, que la historicidad y su forma de darse, la pertenencia, son condición del interés y del conocer histórico; y que por tanto en toda comprensión y en todo interpretar histórico están siempre presentes no solo los momentos de la estructura existencial del estar-ahí –*Geworfenheit* (estar-arrojado, arrojamiento) y *Entwurf* (proyecto)–, sino también la pertenencia a tradiciones. Esto implica asimismo que la estructura de anticipación de la comprensión está inevitablemente anclada en tradiciones transmitidas, aun cuando en la práctica pretendamos leer “lo que se pone” y tomarlo “como realmente ha sido”.

Ahora bien, con el esbozo de esta nueva perspectiva, podemos adelantar algunas de las mediaciones que operan en la investigación histórica. En primer lugar, que toda comprensión histórica es un proyectar posibilidades siempre situado; de modo que toda comprensión de algo significativo presupone que aportamos un conjunto de prejuicios. En segundo, que el significado del pasado no se descubre “mediante la distancia de la comprensión (o el distanciamiento metódico), sino introduciéndonos en el contexto de efecto [*Wirkungszusammenhang*] de la historia” (Gadamer, 1986/1992, p. 40 [1986b, p. 34]). Esto es, lo históricamente significativo es accesible en la realización o ejecución activa. Ya Nietzsche (1874/1998, p. 89) nos lanzaba en “Sobre utilidad y prejuicio de la historia para la vida” que “sólo como artífices del futuro, como conocedores del presente” entenderemos el pasado. En tercer lugar, que “la comprensión sólo alcanza la realidad limitando su libertad” (Gadamer, 1986/1992, p. 40); esto será elaborado por Gadamer al abordar la experiencia que hacemos en la comprensión del tú (incluido el objeto de la historia) cuando no intentamos entenderlo desde nuestro sentido y opinión, sino atendiendo sus propias pretensiones de la validez (cf. Ruffinetti, 2018, pp. 119-126).

## II

Veamos ahora estas instancias con algo de detenimiento. Gadamer sostiene que en la comprensión histórica alcanza su concreción la estructura de la comprensión elaborada por Heidegger, puesto que en ella están “operantes las vinculaciones concretas de costumbre y tradición transmitida y las correspondientes posibilidades del propio futuro” (Gadamer, 1986a, pp. 249-250). Esto significa que el prejuicio que opera en las ciencias histórico hermenéuticas está relacionado con las realidades históricas en las que “siempre ya” estamos inmersos (esto es, lenguaje, familia, sociedad y Estado) y que conforman el modo de comportamiento respecto al mundo, a los otros y a nosotros mismos.

Estas realidades históricas en las que “siempre ya” estamos inmersos son precisamente las tradiciones transmitidas. Estas tradiciones vivas, que constituyen el fundamento de la validez de las costumbres y que tienen poder [*Gewalt*] sobre nuestras acciones y sobre nuestro comportamiento, no nos determinan “mudamente”, sino que “necesita de la afirmación, la aprehensión y el cuidado”. Lo central aquí es la preservación [*Bewahrung*] pero esta no debe entenderse como mero dejar estar, sino como un acto [*Tat*] de la razón relacionado con las costumbres y los valores, la forma de organización social, política e institucional históricamente transmitidos. En todo caso, la preservación “representa una conducta tan libre como la transformación y la innovación”, señala Gadamer (1960/1977, p. 350), y precisamente por ello contiene un momento de crítica hacia lo heredado.

Sin embargo, nuestra relación con el mundo histórico social de la vida no es un comportamiento objetivador que piensa como extraño o ajeno lo que dice la tradición, por el contrario, el momento de la tradición está siempre operando y operante en cuanto pretendemos comprender. En este sentido, la pertenencia a las tradiciones transmitidas es una condición para toda comprensión, por ello la «crítica» no puede meramente entenderse como, ni partir de, una oposición abstracta entre razón y tradición.

Ahora bien, ¿sucede lo mismo en el ámbito de las ciencias histórico hermenéuticas? ¿Rompe realmente los lazos que nos ligan a lo transmitido cuando adoptamos una actitud metódica de distanciamiento, o no será que ese conocimiento libre de prejuicios comparte esa “recepción y reflexión ingenua en la que viven las tradiciones y en las que está presente el pasado”?

Para aclarar esto es preciso recurrir al supuesto que la comprensión en las ciencias humanas comparte con la “pervivencia” de las tradiciones, esto es, el “verse aludido o interpelado [*angesprochen*] por la tradición transmitida [*Überlieferung*]” (Gadamer, 1960/1977, p. 350 [1986a, p. 266]).<sup>1</sup> Esta es la condición hermenéutica principal, ya que esta interpelación no sólo nos proporciona el complemento (semántico y pragmático) necesario para entender el significado de un juicio, aquello de lo que nos habla la tradición, sino que ella es una pregunta dirigida a nosotros. En este sentido le otorga una significatividad al pasado, «pone» aquello que nos resulta relevante y digno de ser tenido en cuenta y se instaure por tanto no sólo como un límite sino como condición de posibilidad de toda comprensión de aquello que participa del modo de ser histórico.

Podemos ver ahora que, dada la estructura circular de la comprensión, el significado de los objetos de la investigación no lo encontramos solamente al final de la indagación, sino también al comienzo pues la «significación» del verse aludido o referido atraviesa la propia investigación del tema (Jay, 2009).

De allí que Gadamer sostenga que en el punto de partida de toda hermenéutica histórica debe disolverse o resolverse la abstracta oposición entre la tradición transmitida y la investigación histórica, entre el proceso histórico y el conocimiento del mismo. “El efecto de la tradición [*tradition*] que pervive y el efecto de la investigación histórica [*historischen*] forman una unidad de efecto [*wirkungseinheit*] cuyo análisis puede hallar siempre sólo una red o entrelazamiento de efectos recíprocos” (Gadamer, 1960/1977, pp. 350-351 [1986a, p. 267]).

Esta unidad de efecto que recoge los efectos de la historia y de la investigación encuentra su correlato en la interpretación del círculo hermenéutico que describe la comprensión como

la interpenetración o el juego entrelazado [*Ineinanderspiel*] del movimiento de la tradición y del movimiento del intérprete. La anticipación de sentido que guía nuestra comprensión de un texto –nos dice Gadamer– no es un acto de la subjetividad, sino que se determina desde la comunidad [*Gemeinsamkeit*] que nos une con la tradición. (Gadamer, 1960/1977, p. 363 [1986a, p. 277])

---

<sup>1</sup> Y continúa: “¿No vale pues para los objetos de su investigación –lo mismo que para el contenido de la tradición– que solo entonces se hacen experimentables o cognoscibles en su significado?” (Gadamer, 1960/1977, p. 350 [1986a, p. 266]).

En la comprensión interviene algo más que la construcción o reconstrucción histórica del mundo pasado al que perteneció la obra o el acontecimiento; en la acción de comprender se da siempre también una co-pertenencia [*Mitzugehörigkeit*] en una doble dirección, la de nuestro sabernos ligados a ese mundo y la del vínculo de la obra o el acontecimiento con nuestro propio mundo. A nuestro entender, esto conduce precisamente a la cuestión de que en todo comportamiento histórico opera esta *mediación* histórica del pasado con el presente.<sup>2</sup>

La historia del efecto [*Wirkungsgeschichte*] no debe entenderse entonces sólo como la historia de la recepción, como la historia de las sucesivas interpretaciones a través del tiempo, sino a la vez como el efecto o la acción de la historia, como “eficacia histórica” (Grondin, 1999/2003). Aquello que obra o efectúa esta acción de la historia viene a conformar el mundo de la vida en el que nos encontramos, las tradiciones transmitidas que constituye nuestros prejuicios. En ese sentido su acción rebasa la conciencia que podamos tener de ella y determina las anticipaciones que guían nuestro comportamiento. Consecuente con esto, tenemos que sostener que esta acción no puede ser objetivada del todo y que en última instancia

El comprender debe pensarse menos como una acción [*Handlung*] de la subjetividad que como un desplazarse uno mismo hacia un acontecer de la tradición, en el que el pasado y el presente se hallan en continua mediación. (Gadamer, 1960/1977, p. 360 [1986a, pp. 274-275])

El autor recoge esta cuestión a través del polémico concepto de anticipación de perfección [*Vorgriff der Vollkommenheit*] que deriva del círculo del comprender. Este concepto no sólo expresa un presupuesto formal que guía toda comprensión, también se refiere siempre a algún contenido. En toda comprensión de lo transmitido, sostiene, no sólo suponemos una unidad inmanente de sentido, sino que ella está guiada constantemente por expectativas de sentido provenientes de nuestra relación previa con la cuestión. En este sentido, el prejuicio de perfección no solamente contiene la cuestión formal de que un texto debe expresar perfectamente su opinión, sino que implica también que aquello que dice es una perfecta verdad.

---

<sup>2</sup> Este es precisamente el sentido que adquiere en Hegel el concepto de *integración*.

Para Gadamer esta “apertura” a la posibilidad de que el otro, el pasado, puedan tener razón, viene a confirmar que comprender significa ante todo «entenderse en la cosa» y sólo en segundo lugar comprender la opinión del otro. De allí que la primera condición hermenéutica sea la “precomprensión que surge del tener que ver [*Zu-tun-haben*] con el mismo asunto” (Gadamer, 1960/1977, p. 364 [1986a, p. 278]).

Esto es, el punto de partida de la hermenéutica es que aquél que comprende “está vinculado al asunto que se expresa en la tradición, y que tiene o logra una determinada conexión con la tradición desde la que habla lo transmitido” (Gadamer, 1960/1977, p. 365 [1986a, p. 279]). Ya indicamos que el momento de la tradición en el comportamiento histórico-hermenéutico se realiza a través de los prejuicios comunes que nos sostienen, agreguemos ahora que la tradición se cumple en y como lenguaje con lo cual el entenderse con otro en la cosa se realiza siempre para Gadamer sobre la base de un consenso de fondo ya alcanzado.

Antes señalamos que el efecto de la historia rebasa la conciencia que podamos tener de ella. Como puede verse, esto plantea la cuestión de cómo es posible esclarecer las condiciones bajo las cuales se realiza la comprensión, si el intérprete y también el investigador, no pueden «disponer» libremente de los propios prejuicios y opiniones previas que “ocupan su conciencia”, y si la distinción entre los prejuicios que hacen posible la comprensión de aquellos otros que la obstaculizan únicamente puede tener lugar en la comprensión misma y como un proceso de historia del efecto.

Al plantear estas cuestiones entramos ya en el concepto de conciencia de la historia del efecto [*wirkungsgeschichtliches Bewusstsein*]. Ciertamente, el principio de la *historia del efecto* que Gadamer reclama no es una exigencia metodológica nueva, sino que va dirigido a la conciencia metódica de la investigación. El autor no desconoce aquellas disciplinas que se ocupan de las continuidades e influencias de textos y obras de arte, y cuyo fin es distinguir el «original» de la historia de su recepción para liberarnos así de los efectos y deformaciones que ésta pudiera arrastrar. Lo que Gadamer reclama ahora en su crítica es que el investigador tenga presente que en la “aparente inmediatez con que se orienta hacia la obra o la tradición transmitida está siempre en juego este otro planteamiento, aunque no reconocido y consecuentemente no controlado” (Gadamer, 1960/1977, pp. 370-371). Esto significa que al tratar de comprender nos encontramos “siempre ya” bajo los efectos de la *historia del efecto*. Y es ella la que

“determina de antemano” lo que nos resulta problematizable y digno de investigación. La exigencia es, en definitiva, que las ciencias históricas, y las ciencias humanas en general, aprendan a conocerse más correctamente y reconozcan que en toda comprensión actúa esta *Wirkungsgeschichte*.

Ahora bien, justamente porque tal poder de la historia opera más allá de nuestro reconocimiento es que resulta necesario y urgente «hacer consciente» la historia del efecto. Sin embargo, esta exigencia, indeclinable para toda reivindicación de cientificidad, no implica que pueda ser completamente realizable; hay que descartar la posibilidad de que esta historia del efecto pueda objetivarse enteramente y hacerse completamente consciente (Ricoeur, 1986/2001).

Y esto precisamente porque no se trata sólo de la conciencia que podamos hacernos de esta historia del efecto, sino también y sobre todo de que la conciencia es un «producto» de esta acción de la historia. Hay que entenderla en las dos formas del genitivo, como acción y como producto. Esto significa que no podemos eludir el acontecer histórico ni enfrentarnos a él convirtiendo el pasado en un objeto; pues siempre estamos *en* la historia y es en este *estar dentro* que hacemos verdadera experiencia de la historia. Y, sin embargo, tenemos que tratar de “formar siempre de nuevo en nosotros una conciencia de esa efectualidad [*Bewirktheits*], de manera que todo pasado del que tenemos experiencia nos obliga a hacernos cargo, a asumir en cierto modo su verdad” (Gadamer, 1986b, pp. 142-143 [traducción propia]).

Para dar cuenta de la circularidad y constante actualización que implica la realización de esta conciencia de la efectualidad, el autor recurre al concepto de *lingüisticidad* presente en toda comprensión. En esta realización nuestra conciencia histórica termina conjugando siempre diferentes “conceptualidades”, nuestros propios conceptos y los de mundos ajenos que pretende tener separados. En este sentido, “No es solo que la distancia histórica quede salvada por la lingüisticidad, sino que esa mediación es previa a cualquier conciencia específicamente histórica” (Gadamer, 1986/1992, pp. 141-142).

### Cierre provisorio

Dado que este concepto resulta de central importancia, para su esclarecimiento nos resulta necesario vincular “la conciencia expuesta a los efectos

de la historia” como la llama P. Ricoeur (1986/2001, pp. 312, 319)<sup>3</sup> a cuatro problemáticas; estas cuestiones marcan asimismo el curso de nuestras futuras indagaciones.

En primer lugar, es necesario considerarla en tensión con el concepto de distancia histórica y el de distanciamiento como condición del conocimiento objetivo. Por otro lado, hay que relacionarlo con el concepto de situación hermenéutica; cuestión esta que se desprende de la circunstancia de que “no hay *sobrevuelo*” (Ricoeur, 1986/2001, p. 320) y que se refiere, como antes indicamos, a que es imposible objetivar y conocer el conjunto de los efectos de la historia. En tercer lugar, y como contracara de lo anterior, es necesario considerar la noción de horizonte y lo que significa ponerse en el punto de vista del otro. Estas instancias conducen, por último, al concepto de fusión de horizontes y a la concepción gadameriana de la comprensión como mediación integradora o fusión de horizontes.

## Referencias

- Gadamer, H. G. (1977). *Verdad y método* (A. Aparicio & R. de Agapito, trads.; 4ta ed.). Salamanca: Sígueme. (Obra original de 1960)
- Gadamer, H. G. (1986a). *Hermeneutik I (Wahrheit und Methode: Grundzüge einer philosophischen Hermeneutik)*. En *Gesammelte Werke* (vol. 1). Tübingen: J.C.B. Mohr (Paul Siebeck).
- Gadamer, H. G. (1986b). *Hermeneutik II (Wahrheit und Methode: Ergänzungen, Register)*. En *Gesammelte Werke* (vol. 2). Tübingen: J.C.B. Mohr (Paul Siebeck).
- Gadamer, H. G. (1992). *Verdad y método II* (M. Olasagasti, trad.). Salamanca: Sígueme. (Obra original de 1986)
- Grondin, J. (2003). *Introducción a Gadamer* (C. Ruiz-Garrido, trad.). Barcelona: Herder. (Obra original de 1999)
- Heidegger, M. (1951). *El ser y el tiempo* (J. Gaos, trad.). México: F.C.E. (Obra original de 1927)
- Husserl, E. (1991). *La crisis de las ciencias europeas y la fenomenología trascendental* (J. Muñoz & S. Más, trads.). Barcelona: Crítica. (Obra original de 1936)

<sup>3</sup> Así es como propone traducir Ricoeur el concepto de *wirkungsgeschichtliches Bewusstsein*.

- Jay, M. (2009). *Cantos de experiencia*. Buenos Aires: Paidós.
- Nietzsche, F. (1998). *Consideraciones intempestivas: Sobre utilidad y perjuicio de la historia para la vida* (O. Caeiro, trad.). Córdoba: Alción. (Obra original de 1874)
- Ricoeur, P. (2001). *Del texto a la acción: Ensayos de hermenéutica II* (P. Corona, trad.). Buenos Aires: FCE. (Obra original de 1986)
- Rufinetti, E. (2018). *La racionalidad práctica en el debate Habermas-Gadamer*. Villa María: Eduvim.



## Ciencias sociales: “imbricación” de valores epistémicos y sociales

Eduardo Sota\*

La preceptiva weberiana de la ciencia “libre de valores” constitutiva del *L*ethos científico ha gozado de un acendrado consenso dentro de su respectiva comunidad abonado por las corrientes epistemológicas hegemónicas. En efecto, y a pesar de no ser un movimiento homogéneo, el positivismo lógico en su manifiesto fundacional *La concepción científica del mundo* (Carnap, Hans & Neurath, 1929/2002), propiciaba la unidad de la ciencia, la identificación de una estructura lógica común del lenguaje científico, el análisis lógico de los conceptos que permite discriminar entre los genuinos problemas empíricos de los problemas pseudocognitivos como los planteados por la metafísica y la ética que no proveen ninguna información acerca del mundo empírico: “esta actitud delimitadora entre conocimiento auténtico del mundo y afirmaciones sin sentido va a ser sacralizada posteriormente por medio del criterio empirista del significado, uno de los núcleos ‘trágicos’ de las posturas del Círculo” (Gómez, 2014, p. 21). Y es este criterio el que va a introducir un foso ontológico entre el mundo de los hechos y el mundo de los valores aherrojando estos al dominio del sinsentido:

Lo positivistas lógicos introdujeron una célebre clasificación tripartita de todos nuestros juicios dividiéndolos en “sintéticos” (verificables o falsables empíricamente), “analíticos” (verdaderos o falsos con arreglo sólo a las reglas lógicas), y, por último, –y esta categoría incluye especialmente todos nuestros juicios éticos, metafísicos y estéticos– “carentes de valor cognitivo”. (Putnam, 2002/2004, p. 24)

En esta división tripartita de los juicios dos de ellos –sintéticos o de hechos y analíticos– tienen cabida legítima como ciudadanos de la ciudadanía científica mientras que los juicios éticos y estéticos carecen de valor cognitivo y este es el denominado “tercer dogma del positivismo”

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH).  
Córdoba, Argentina.  
eduardomsota@gmail.com

por Putnam. Este, junto a otros dogmas –esto es, supuestos injustificados– hicieron que, finalmente, esta corriente denominada genéricamente Concepción Heredada entrara en crisis en cuanto a que sus tesis principales o bien dejaban de dar respuestas a los problemas para las que habían sido planteadas o bien eran abandonadas como caminos infecundos. Es así como durante la década de los años sesenta se gestan concepciones alternativas a la Concepción Heredada inscribiéndose en lo que se denominó nueva filosofía de la ciencia con autores como Toulmin, Kuhn, Hanson, Feyerabend y Laudan, entre otros. En lo que al tema que nos ocupa, uno de sus epónimos, Kuhn, nos dice que en las elecciones que hacen los científicos en relación a las teorías intervienen valores tanto cognitivos como no cognitivos y que las decisiones que se adopten dependen menos de resoluciones mediante pruebas que mediante técnicas de persuasión ya que la decisión nunca es de científicos aislados sino que, en la medida en que no hay criterios neutros y válidos para todo contexto, para cada científico individual no puede haber mejor criterio que “la decisión del grupo científico” (Kuhn, 1977, pp. 321-322). Son varios los criterios que el autor distingue como características para que una teoría sea una buena teoría, entre otras, la adecuación empírica, la consistencia, la simplicidad y el alcance teórico. Estos criterios epistémicos funcionan menos como reglas que como valores relevantes que tendrán distinto peso de acuerdo con las circunstancias y son una mezcla de factores subjetivos como objetivos por lo que en las elecciones intervienen tanto factores contextuales (políticos y sociales) como cognitivos. Veamos cómo podríamos caracterizar esta distinción de valores epistémicos o internos y sociales o externos. Lacey nos señala que en el discurso ordinario cuando nos referimos a un valor personal aludimos a situaciones tan variadas como un bien fundamental que uno persigue consistentemente, una práctica o cualidad que da significado a la vida que uno lleva, una cualidad que es parcialmente constitutiva de la identidad de uno y un “objeto de valor”, entre otras. Los objetos de valor pueden incluir objetos de arte, recursos tecnológicos, objetos sagrados, instituciones, teorías científicas, etc. Veamos como lo precisa el autor: “Hay varios tipos de valor. Un valor es sostenido por un agente o agentes. Cuando un agente (X) sostiene un valor (v), la expresión fundamental es: ‘X valora que  $\emptyset$  sea caracterizado por v’” (Lacey, 1999, p. 27).

Los diferentes tipos de valores corresponden a distintas instanciaciones de  $\emptyset$ ; por ejemplo, cuando  $\emptyset$  son relaciones entre personas tenemos valores morales, cuando  $\emptyset$  es una sociedad valores sociales, cuando  $\emptyset$  es una

obra de arte valores estéticos, cuando  $\emptyset$  son teorías científicas o cuerpos sistemáticos de creencias valores cognitivos y así. A pesar de que gramatical y lógicamente los valores cognitivos o epistémicos tienen mucho en común con los otros valores, sostener un valor cognitivo  $v$  involucra una actitud proposicional de segundo orden, "una creencia acerca de creencia, una creencia de segundo orden que siendo caracterizada por  $v$  contribuye a hacer a una creencia (de primer orden) racionalmente aceptable" (Lacey, 1999, p. 46). Tenemos, pues, hasta aquí, la discriminación de distintos tipos de valores, pero podríamos preguntarnos, cuál es la relación que es posible establecer entre ellos. En la línea de defensa de concebir una ciencia libre de valores  $o$ , a lo sumo, aceptar a los cognitivos como pertinentes al conocimiento relevante como tal, Lacey distingue tres momentos o contextos de actividad científica: la adopción de una estrategia  $-M1-$ , la aceptación de teorías  $-M2-$  y la aplicación del conocimiento científico  $-M3-$ . En efecto, en  $M2$ :

T es correctamente aceptada para un dominio específico de fenómenos si y solamente si manifiesta los valores cognitivos en grado elevado y si, dados los "patrones" corrientes para "medir" el grado de manifestación de los valores cognitivos no existe una perspectiva plausible para obtener un grado mayor. Dado  $A$  como fin de la ciencia y dado que son las teorías las que expresan un entendimiento de los fenómenos, no existe papel racionalmente destacado para los valores sociales en  $M2$ ; el hecho de que  $T$  pueda manifestar en grado elevado algún valor social no cuenta racionalmente ni a favor ni en contra de su aceptación correcta. (Lacey, 2003, p. 143)

Así, los valores cognitivos son relevantes para el contexto de aceptación y/o justificación  $-M2-$  de las teorías mientras que los valores sociales (bajo los cuales se subsumen todos los otros) son adscriptos al viejo contexto de descubrimiento y al contexto de aplicación,  $M1$  y  $M3$ , respectivamente, y no juegan rol relevante alguno a la hora de la justificación de las teorías evaluadas y aptas para aceptar.

La libertad de valores reposa sobre un trípode: la neutralidad de valores, la autonomía y la imparcialidad. La primera implica la legitimidad de tres supuestos: las teorías científicas no tienen juicios de valor entre sus implicaciones lógicas; aceptar una teoría no tiene consecuencias cognitivas en todo lo concerniente a los valores que uno sostiene y las teorías científicas están disponibles para ser aplicadas con el objetivo de promover proyectos vinculados con cualquier valor. Por otra parte, la au-

tonomía, como segundo principio que sostiene la libertad de valores, no sólo concede autoridad únicamente a la comunidad científica respecto a la definición de los problemas y evaluación de las teorías sino también para determinar las cualificaciones requeridas para la pertenencia a dicha comunidad y decidir el contenido de la educación científica. Por último y como vimos, la imparcialidad coincide con el contexto M2 por el que una teoría es completa o correctamente aceptada solo si sus enunciados están basados únicamente en evaluaciones que incluyan juicios de valor cognitivo de las teorías y con exclusión de valores sociales, morales o políticos. Así, el de imparcialidad es el principio fuerte que está por detrás del desiderátum “la ciencia libre de valores” puesto que estima cuán bien las teorías manifiestan ciertos valores cognitivos tales como el poder explicativo y predictivo y que juegan un rol esencial en la aceptación o rechazo de teorías y presupone que los valores cognitivos deben ser claramente distinguidos de otros tipos de valores. En definitiva, la imparcialidad, en términos de Lacey, nos asegura el ideal de objetividad por el que la contrastación de las teorías en términos de evidencia empírica y criterios cognitivos es incompatible con la asunción de compromisos éticos y sociales. Seguramente no nos equivocamos si señalamos que es, precisamente, la noción de objetividad la que funge como la otra cara de la moneda de la libertad de valores; es decir, la ilusión de una suerte de reproducción cognitiva *in toto* del objeto, sin interferencia valorativa excepto los valores cognitivos como ya señalamos. Veamos las diversas acepciones que señala Lloyd (1995) a propósito de “objetividad” u “objetivo”: a) objetivo significa desinteresado, impersonal, no sesgado; b) objetivo significa públicamente accesible, observable; c) objetivo significa “existiendo independientemente de nosotros” o separadamente de nosotros; d) objetivo significa “existiendo realmente”, la manera en que las cosas realmente son. Naturalmente (a)-(d) no son equivalentes ya que objetivo es predicado de diferentes entidades: (a) es una propiedad del sujeto del conocimiento, mientras que (b) y (c) supone una relación entre el conocimiento y la realidad y (d) es el estatus de lo que es a pesar de su relación con el sujeto; en efecto, ellas no son equivalentes en el sentido de, por ejemplo, Dios, si tal ser existe, es real y objetivo (en el sentido de (d)) pero no siempre se presume que sea público. Ahora bien, cuando a estas diversas acepciones se las hace converger para otorgar un sentido único al vocablo objetivo nos deslizamos en lo que Lloyd denomina “tiranía ontológica”:

Examinemos la afirmación fuerte que la realidad 'objetiva' –la realidad convergente a través de la aplicación de métodos objetivos– es igual a toda la Realidad Real. Una tal ecuación parece ser un juicio que la Realidad Real puede, en su totalidad, ser alcanzada o conocida a través de su ser públicamente accesible en el modo propio, combinado con el tipo correcto de desinterés del que conoce. Llamo a esta posición la "tiranía ontológica". (1995, p. 356)

Determinadas tradiciones metafísicas y epistemológicas han configurado estas opresivas propiedades –bajo la restricción de estar reunidas conjuntamente– de la noción de objetividad que es correlativa de las mismas altas e irreales exigencias que tiñen a la noción de ciencia libre de valores. Naturalmente no sólo nos desembarazaremos de la convergencia de estos cuatro sentidos en una caracterización unívoca de objetividad por considerarla artificiosa y a contramano de la misma práctica científica, sino que también discutiremos el modo en que son caracterizados los sentidos mismos de (a)-(d). En efecto, (a) y tal vez (b), remiten abiertamente al principio de imparcialidad de Lacey, la adopción del "punto de vista de ningún lugar" regulado por los valores cognitivos, con la exigencia de que se debe extirpar toda preferencia o sesgo subjetivo proveniente de los valores sociales o políticos. Sin embargo, Longino va rebatir esta visión dicotómica y asimétrica de valores cognitivos, por una parte, y sociales, por otra, ya que lo social no se opone a lo racional puesto que las mismas normas y principios guías que orientan las prácticas científicas presuponen un trasfondo social y las mismas normas cambian y son funcionales al cambiante contexto social. Es decir, no hay una frontera rígida entre valores cognitivos y sociales y la misma noción de objetividad se va a ver modificada. Las ciencias son prácticas sociales que requieren variedad de perspectivas y una crítica transformadora incesante y sistemática al interior de la misma comunidad científica. Ya no se requiere una objetividad absoluta y despersonalizada, sino que es social por naturaleza y supone grados entre dos polos interactivos de un continuum que va del monológico al dialógico; "un método de investigación es objetivo según el grado que permite una crítica transformativa. Su objetividad consiste en la inclusión de crítica intersubjetiva" (Longino, 1990/1998, p. 181). Para transitar en esa dimensión transformadora de la crítica del discurso las comunidades científicas serán objetivas según el grado en que satisfagan las siguientes condiciones: posean los espacios y marcos institucionales para que se efectivice la crítica tales como publicaciones, evaluación de pares,

etc.; estándares compartidos de crítica los cuales incluyen tanto principios substantivos como epistémicos y sociales; la comunidad debe siempre tener capacidad de respuesta frente a la crítica; la autoridad intelectual debe ser compartida equitativamente entre los miembros de la comunidad. Así, “la objetividad de los individuos en este esquema consiste en su participación en el “toma y daca” de la discusión crítica colectiva y no en alguna especial relación (desinterés, obstinación) que puedan sostener con sus observaciones” (Longino, 1990/1998, p. 183).

En suma, tenemos dos modelos claramente divergentes en cuanto a “objetividad”: no se trata ya de una propiedad que se posee o no absolutamente, sino que supone grados en su acceso y no se hace desde un presunto lugar impersonal sino desde la crítica intersubjetiva y donde se ha difuminado la distinción dicotómica de valores cognitivos y sociales ya que las prácticas científicas mismas son de naturaleza social. La versión más débil pero no menos vigorosa de Lacey por la que la ciencia libre de valores admite en realidad valores epistémicos a la hora de la elección de teorías más allá de las reglas lógicas y la evidencia empírica y en el ámbito de la aceptación y justificación de teorías excluye sí, inequívocamente, la admisión de valores no epistémicos. Pero esto es lo que hemos pretendido horadar juntamente con la noción de objetividad por lo que ahora daremos un paso más y exploraremos la tesis semántica de Putnam sobre la distinción hecho-valor o, mejor dicho, la dicotomía esencialista hecho-valor a la que aludimos más arriba como el tercer dogma del positivismo. Putnam se remonta a Hume para hallar allí el origen de tal bifurcación con la “semántica figurativa” y su relación de semejanza por la que un hecho se parece o no a la imagen que de él nos hacemos y así verificamos el valor de verdad de los enunciados mientras que una situación como esta está ausente para determinar la verdad o no de un enunciado tal como “el crimen es malo” ya que no hay un hecho correlativo para el predicado malo. No hay pues “cuestión de hecho” en materia de determinar qué es lo correcto o virtuoso y de allí el aherramiento al sinsentido de estos términos. En esta carencia de contenido cognitivo y de irracionalidad podríamos agregar que de los juicios morales y políticos radica el rechazo a la intromisión de los valores en la declamada libertad de valores de la ciencia.

La dicotomía de referencia presupone la falacia naturalista humana (un “debe” nunca debe ser derivado de un “es”) y obedecen a la visión estrecha según la cual un hecho es algo que corresponda a una impresión sensorial. La posición alternativa de Putnam es que tal dicotomía no es

sostenible y propicia menos una partición que una "imbricación" entre juicios de hecho y juicios de valor tanto epistémicos como morales.

Veamos lo que sucede con los denominados "conceptos éticos densos" (que entremezclan o superponen términos descriptivos con otros de carácter más abstractos como serían los "conceptos éticos finos" tales como "virtud", "deber", "obligación") como "valiente" en el enunciado "El gobernante era valiente". Tener una comprensión intuitiva de este predicado supone adjudicarle hechos (describibles) congruentes y que sostienen dicho valor moral. No podríamos decir que el gobernante encarne dicho valor y, simultáneamente, que huyó vergonzosamente de tal batalla. Siguiendo a Putnam sería el caso, por ejemplo, cuando el historiador emplea la palabra "cruel" para caracterizar descriptivamente a cierto monarca, pero no puede eludir la resonancia ética y normativa: "'cruel" simplemente ignora la presunta dicotomía hecho/valor y se permite el lujo de ser empleado unas veces para propósitos normativos y otros como término descriptivo. En la literatura filosófica, tales conceptos son llamados a menudo "conceptos éticos densos" (2004, p. 50). Es precisamente en el área de las ciencias sociales donde este carácter inescindible hecho-valor se hace ineludible. Supongamos un enunciado sociológico como "Argentina es un país violento"; es difícil asignar este rasgo o bien a una categoría factual o evaluativa ya que por más que nos inclinemos a adoptar un lenguaje técnico para esta caracterización nos interesa preservar, igualmente, el aspecto valorativo ya que nos provee de razones para la acción, por ejemplo, adoptar políticas públicas para reducir dicha violencia o mitigar sus efectos. Veamos que nos dice al respecto Dupré:

Muchos términos del lenguaje ordinario son descriptivos y evaluativos a la vez. La razón de esto es obvia. El lenguaje evaluativo expresa nuestros intereses los cuales, como era de esperarse, son cosas que estamos interesados en expresar. Cuando describimos cosas es, a menudo, en términos que relacionan la relevancia de las cosas para satisfacer nuestro interés. A veces tratamos de imponer criterios más bien precisos para aplicar a las cosas terminología relativa a nuestros intereses. (Dupré, 2007, p. 30)

Una ciencia social como la economía, por ejemplo, cuya tradición ha pretendido separar nítidamente lo que es del orden factual del dominio evaluativo se interesa por formular invariantes y leyes económicas; un mecanismo identificado es el óptimo Pareto –si hay cambios en la asignación de bienes hay al menos una que mejora la situación de un indivi-

duo sin hacer que empeore la situación de los demás— cuyo propósito es la asignación eficiente de recursos. Sin embargo, deja de lado cuestiones propiamente normativas que son las que nos importan como qué es la justicia, la distribución justa de los bienes en este caso.

Una última cuestión que abordaremos tiene que ver con las acepciones (c) y (d) de la noción de “objetividad”, esto es relativas a la existencia de las cosas con independencia de nosotros y la cual supone un compromiso con concepciones metafísicas realistas fundacionalitas. Desde la teoría social y política el “fundacionalismo” implica aquellas teorías que suponen que la sociedad, la política y el sujeto se basan en principios que están exentos de examen y habitan fuera de la misma política, sea la metáfora topológica de la infraestructura económica o el “espíritu de la historia”.

Generalmente esto es seguido de la frecuente situación de plantearse un cuerno dilemático, esto es, la opción de hierro de elegir entre un fundamento último y la negación absoluta de éste, una suerte de antifundacionalismo, ya sea en las variantes del “todo vale” o del posmodernismo. Podemos colegir que, en la medida que uno es la simple inversión del otro, participan del mismo horizonte. Una alternativa a ese cuerno dilemático es horadar las premisas mismas donde opera el fundacionalismo evitando negar cualquier fundamento, aunque sí un fundamento último y esto es lo que caracteriza al posfundacionalismo (Marchart, 2007/2009). Para éste, es la ausencia de un fundamento último, precisamente, lo que hace posible la emergencia de fundamentos *contingentes*.

A esta dirección posfundacionalista vamos a ejemplificarla con el “constructivismo realista” acerca del estatus epistemológico de las ciencias sociales con Bourdieu. La comunidad científica se somete al arbitraje de lo “real” pero, a la vez, es lo producido por el equipo teórico y experimental efectivamente disponible en el tiempo considerado y en esa misma comunidad se plantea la lucha por el monopolio de la representación científicamente legítima de lo “real”:

En el caso de las ciencias sociales, lo “real” es absolutamente exterior e independiente del conocimiento, pero es a su vez una construcción social, un producto de las luchas anteriores que, por esas mismas razones, sigue siendo un objetivo de las luchas actuales [...] Conviene, pues, asociar una visión constructivista de la ciencia y una visión constructivista del objeto científico. Los hechos sociales están contruidos socialmente, y todo agente social, como el científico, [...] tiende a imponer su singular visión de la realidad, su “punto de vista” (Bourdieu, 2001/2003, p. 153).

Las diversas disciplinas sociales toman partido, pues, en las mismas luchas que describen y la ciencia que se produce no es más que una de las fuerzas que se enfrentan en ese mundo. En su esclarecimiento de las luchas por la identidad étnica o regional el sociólogo nos dice que son luchas por el monopolio respecto al poder de hacer ver y hacer, por imponer la definición legítima de las divisiones del mundo social y, a través de esto, hacer y deshacer los grupos. Es en la misma "realidad" social donde se desarrolla una lucha permanente por definir la "realidad" donde la misma ciencia está no menos comprometida, si bien de otra manera, que las representaciones de los agentes sociales; es así como las clasificaciones más "naturales" de esa "realidad" que es absolutamente social se apoyan siempre en rasgos que no tienen nada de natural y que en parte son producto de una imposición arbitraria. El discurso regionalista es, pues, performativo en no menor medida que el de la misma ciencia social que lo crea al describirlo:

Quando es reintegrado en las luchas de clasificaciones que se esfuerza en objetivar, el discurso científico se pone de nuevo a funcionar en la realidad de esas luchas de clasificación: está condenado a aparecer como crítico o como cómplice según la relación cómplice o crítica que el lector mantenga con la realidad escrita. Así, el simple hecho de mostrar puede funcionar como una manera de acusar (Kategoriethai) o, a la inversa, como una manera de hacer ver y de hacer valer. (Bourdieu, 1982/1985, p. 92)

Estamos pues, lejos de la noción de objetividad alusiva a una realidad pre-existente y de la cual debemos alcanzar una representación fidedigna incontaminada de valores sino que la misma representación configura esa realidad social; las mismas. Las mismas categorizaciones científicas ejercen por sí mismo un poder y son, por lo mismo, categorías políticas: "los veredictos más neutros de la ciencia contribuyen a modificar el objeto de la ciencia" (Bourdieu, 1982/1985, p. 93).

## Referencias

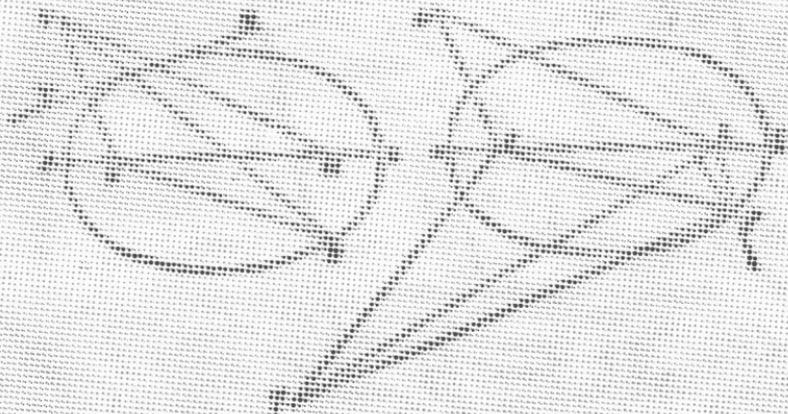
- Bourdieu, P. (2003). *El oficio de científico* (J. Jordé, trad.). Barcelona: Anagrama. (Obra original de 2001)
- Bourdieu, P. (1985). *¿Qué significa hablar?* (E. Martínez Pérez, trad.). Barcelona, Akal. (Obra original 1982)

- Carnap, R., Hans, H., & Neurath, O. (2002). La concepción científica del mundo: El Círculo de Viena (P. Lorenzano, trad.). *Revista de Estudios sobre la Ciencia y la Tecnología*, 18, 103-149. (Obra original de 1929)
- Dupré, J. (2007). Fact and value. En H. Kincaid, J. Dupré, & A. Wylie (Eds.), *Value-free science?: Ideals and illusions* (pp. 27-41). Oxford: Oxford University Press.
- Gómez, R. (2014). *La dimensión valorativa de las ciencias*. Bernal, UNQ Editorial.
- Kuhn, T. (1977). *The essential tension*. Chicago: The University of Chicago Press.
- Lacey, H. (1999). *Is science value free?* London: Routledge.
- Lacey, H. (2003). Existe uma distinção relevante entre valores cognitivos e sociais? *Scientia Studia*, 1(2), 121-149.
- Longino, H. (1998). Values and objectivity. En M. Curd & J. Cover (Eds.), *Philosophy of science: The central issues* (pp. 170-191). Nueva York, W. Norton & Co. (Obra original 1990)
- Lloyd, E. (1995). Objectivity and the double standard for feminist epistemologist. *Synthese*, 104(3), 351-381
- Marchart, O. (2009). *El pensamiento político posfundacional: La diferencia política, en Nancy, Lefort, Badiou y Laclau*, (M. Delfina Álvarez, trad.). Buenos Aires: Fondo de Cultura Económica de Argentina. (Obra original de 2007)
- Putnam, H. (2004). *El desplome de la dicotomía hecho-valor* (F. Forni Argimon, trad.). Barcelona: Paidós. (Obra original de 2002)



---

## 5. Historia y filosofía de las ciencias físicas y naturales







# Secuela de la “Hipérbola de la química cuántica” Una mirada sobre el presente de las simulaciones computacionales en este ámbito

Penélope Lodeyro\*  
Ma. Silvia Polzella\*

## Introducción

En 1965, Pople presentó un gráfico acerca de la relación entre la sofisticación de los métodos que emplean las simulaciones de la química cuántica y el tamaño de los sistemas abordables por los mismos que mostraba la “Hipérbola de la química cuántica” (Pople, 1965). El grado de sofisticación reflejaba el esfuerzo puesto en evitar el empleo de datos empíricos en el cálculo como correlato de la distinción entre métodos ab initio y semiempíricos. Veinticinco años después, en 1990, en una reunión en homenaje al primer artículo de Pople (1950), Karplus (1990) expuso esta vez un gráfico tridimensional proyectado sobre la representación de Pople en el que sumaba la consideración explícita de la precisión de los métodos. El objetivo de este trabajo es analizar las implicancias que tuvo el desarrollo de métodos basados en la teoría del funcional de la densidad (DFT) para el desenvolvimiento de la disciplina, los cuales no fueron considerados ni por Pople (1965) ni por Karplus (1990). Mostraremos que estos métodos, por sus características particulares, cobraron gran relevancia en las últimas tres décadas e influyeron drásticamente en las prácticas constructivas de la química cuántica modificando las tendencias en los gráficos y la relación entre las variables. Presentaremos un análisis de estas prácticas que pone en evidencia que el ideal de rigor teórico se subordinó al imperativo de abordar sistemas moleculares más complejos sin por ello renunciar a los estándares de precisión.

## Criterios constructivos

Las simulaciones computacionales transformaron radicalmente las prácticas de modelado de la química cuántica. Desde el inicio de esta disciplina

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH) Centro de Investigaciones María Saleme de Burnichon (CIFYH). Córdoba, Argentina. lopenelope@ffyh.unc.edu.ar

se establecieron dos criterios constructivos, ab initio y semiempírico. Los métodos denominados semiempíricos son caracterizados como aquellos que involucran supuestos, aproximaciones, generalizaciones y datos empíricos para simplificar el cálculo y producir un resultado con suficiente adecuación empírica (Dewar, 1992; Thiel, 2005). Esta caracterización surge por contraposición a la de los denominados métodos ab initio, es decir, aquellos cálculos que parten de primeros principios o están basados en ecuaciones fundamentales de la física o la química. En este sentido, se considera que estos últimos operan sobre una base puramente teórica. En relación a estas caracterizaciones hubo un disenso, tanto en el terreno científico como en el filosófico acerca de cómo debían elaborarse los métodos de aproximación (Scerri, 2004; Ramsey, 2000). Aunque ambos están formulados dentro del mismo marco teórico-cuántico, la desconfianza en los métodos semiempíricos proviene de que los valores de algunos parámetros en vez de ser calculados se satisfacen con resultados experimentales. En este sentido, se considera que éstos no tienen suficiente rigor teórico (Freed, 1995, p. 26) y que restringen su "contenido inferencial" (Ramsey, 1997), es decir, comprometen su generalidad de aplicación al emplear datos empíricos de sistemas moleculares particulares. Se generó así una tendencia a la dicotomía en los desarrollos.

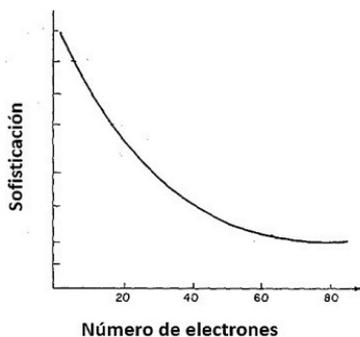
### **La gráfica de Pople: una muestra de las tendencias en los desarrollos de las simulaciones computacionales de la química cuántica**

La hipérbola de la química cuántica mostrada por Pople (1965) es una representación en dos dimensiones de la relación establecida entre el grado de sofisticación de los métodos computacionales elaborados y el número de electrones de los sistemas moleculares posibles de ser abordados por los mismos (Fig. 1). Ésta ponía en evidencia que el empeño de los desarrollos de la química cuántica computacional estaba puesto en dos direcciones marcadamente opuestas. Por un lado, en el desarrollo de métodos de un mayor grado de sofisticación en el sentido de un gran esfuerzo de cálculo para evitar la introducción de parámetros empíricos, es decir, ab initio. Por el otro, en el desarrollo de métodos que por el interés de investigar sistemas más complejos e interesantes para la química requerían disminuir el esfuerzo de cálculo y por ello concedían la introducción de aproximaciones más drásticas y el empleo de parámetros empíricos, es decir, eran semiempíricos. El diagrama de la relación en la forma de una curva

hiperbólica, expuso la frontera de desarrollo de la disciplina y marcó las tendencias hacia ambos extremos en relación a la construcción de métodos bajo los distintos criterios.

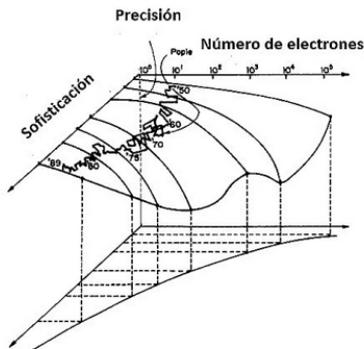
### La gráfica de Karplus: considerando la precisión

Para 1990, los avances en el poder computacional y en el desarrollo de métodos permitían, para el caso de los ab initio, tratar sistemas de tamaño intermedio (alrededor de 20 electrones) a un nivel aceptable de precisión. Sin embargo, lo más desafortunado, señalaba Karplus (1990), es que estas aplicaciones no siempre llegaban al nivel de comprensión que lograban los tratamientos semiempíricos. Estos últimos utilizaban generalmente parámetros que tenían un sentido físico por lo que obtenían resultados que podían contrastarse con datos experimentales y en ese sentido mostrar que las interpretaciones de los datos eran incorrectas o no, mientras que en los métodos más sofisticados en muchas ocasiones para lograr una buena adecuación empírica se apelaba al uso de parámetros ficticios, es decir, sin un significado físico claro. En general, no siempre es posible reducir la complejidad a conceptos simples (Karplus, 1990). No obstante, los partidarios de los métodos semiempíricos intentaban relacionar los resultados numéricos con modelos intuitivos.



**Fig. 1:** Representación la gráfica de Pople (1965).

*Nota:* Imagen adaptada de "Two dimensional chart of quantum chemistry" de J. A. Pople, 1965, *The Journal of Chemical Physics*, 43(10), p. 229.



**Fig. 2:** Representación del diagrama de Karplus (1990).

*Nota:* Imagen adaptada de "Three-dimensional 'Pople diagram'" de M. Karplus, 1990, *The Journal of Physical Chemistry*, 94(14), p. 5436.

En la práctica, siempre hay una demanda creciente por parte de los investigadores de abordar sistemas cada vez más complejos. En la época en que Karplus presentó su diagrama (Fig. 2), los químicos estaban preocupados por resolver problemas que requerían el estudio de sistemas de  $10^3$ - $10^5$  electrones –sólidos y reacciones de biomoléculas en solución. Aunque las simulaciones de grandes sistemas habían progresado notablemente, sin embargo, estos sistemas no podían tratarse con suficiente nivel de precisión por ninguno de los dos métodos. Ello motivó a que Karplus agregara una tercera dimensión al diagrama de Pople. Además de la sofisticación (tipo de método) y el tamaño de los sistemas (número de electrones), consideraba explícitamente la precisión del cálculo para los sistemas estudiados.<sup>1</sup>

La dimensión vertical del gráfico representaba la precisión de los métodos. Ésta delineaba una superficie para la precisión puntualizando un límite para un tipo de cálculo específico y un tamaño de sistema abordado determinado. El borde de esta superficie se proyectaba sobre la hipérbola de la química cuántica, con los dos ejes en el plano correspondiente a los del diagrama de Pople. En el eje de sofisticación se agregaron sólo métodos semiempíricos. En relación al tamaño de los sistemas, Karplus reemplazó la escala lineal de Pople que cubría el rango de 1 a 100 electrones, por una escala logarítmica de 1 a  $10^6$ . Como él mismo señala, los métodos ab initio, mostraban un incremento monotónico en sofisticación y un decrecimiento monotónico en el rango de aplicación. La altura marcaba la precisión del método. Una mayor altura respecto del plano de Pople indicaba un incremento de la misma. La mejor precisión alcanzada por un método ab initio cubría sistemas de 10 electrones y por un semiempírico de  $10^5$ . Los otros métodos estaban comprendidos entre estos límites de sofisticación y número de electrones con la superficie de precisión siendo más bien compleja (Karplus, 1990). La conclusión a la que arribaron en la reunión de 1990, fue que en los 25 años transcurridos entre las gráficas de Pople y Karplus los avances en la metodología de los cálculos químicos cuánticos habían mostrado un incremento de tres órdenes de magnitud en tanto que el tamaño de los sistemas abordables al mayor nivel de precisión, sólo se había incrementado en un orden de magnitud, de 2 a 20 electrones.

Cabe notar que, la evaluación de la dimensión de la precisión muestra

---

<sup>1</sup> La gráfica expone además el derrotero que siguió el trabajo de Pople en cuarenta años de sus desarrollos dentro del campo de la química cuántica computacional, poniendo de relieve el tránsito desde los métodos ab initio hasta los semiempíricos.

que algunos de los métodos semiempíricos menos sofisticados son más precisos que los cálculos *ab initio* y los métodos medianamente sofisticados. Además, en las gráficas, Pople (1965) y Karplus (1990), no consideraron a los métodos basados en la teoría del funcional de la densidad (DFT). Estos métodos surgieron en la década del 60 y revolucionaron el campo de la química cuántica computacional en los 90s. La amplia difusión que experimentaron posteriormente plantea la cuestión del impacto que pueden haber producido en las prácticas constructivas de las simulaciones computacionales en este ámbito.

### **DFT: un enfoque promisorio**

Originalmente tanto los llamados métodos *ab initio* como los semiempíricos se construían sobre la base del marco teórico-cuántico de la función de onda. Todos ellos afrontan el problema de los tres cuerpos y de los  $n$ -cuerpos, por lo cual sólo pueden resolver analíticamente las interacciones de sistemas de dos partículas. A raíz de que la función de onda de una molécula de  $N$ -electrones es una función de  $4N$  variables, puede notarse que la complejidad de la función de onda se incrementa exponencialmente con el aumento del número de electrones.<sup>2</sup> Por ello, con el propósito de abordar sistemas cada vez más complejos, el modo de avanzar ha sido introducir aproximaciones, idealizaciones, datos empíricos y asunciones drásticas.

Otro enfoque teórico alternativo se desarrolló con el objetivo de superar el problema de la complejidad de la función de onda multielectrónica (problema de los  $n$ -cuerpos), la Teoría del funcional de la densidad. Aunque su origen se remonta a la década del 20 del siglo pasado (Fermi, 1927; Thomas, 1927), los desarrollos posteriores de Hohenberg y Kohn (1964) y de Kohn y Sham (1965) terminaron de perfilar este promisorio marco teórico *ab initio*. En principio, mediante DFT se podrían determinar las propiedades de un sistema de muchos electrones, dado que reducía la complejidad de una función de onda de  $N$  electrones y por lo tanto  $4N$  variables por cada electrón (tres coordenadas espaciales más el espín del electrón por cada electrón) a un funcional de la densidad de sólo tres variables en total (coordenadas espaciales).<sup>3</sup>

---

<sup>2</sup> En el caso de la "ecuación de Schrödinger" dependiente del tiempo, considera  $4N$  variables: tres coordenadas espaciales y una coordenada del spin del electrón por cada electrón (una corrección relativista aproximada por considerar el espín del electrón).

<sup>3</sup> En los métodos tradicionales, la función de onda multielectrónica se construye con orbitales de un simple electrón, mientras que en DFT el principal elemento es la densidad

Este enfoque trata a los electrones de alrededor del núcleo como una nube homogénea. El mismo se basa en el funcional que captura la densidad de probabilidad de electrones o densidad de carga –distribución electrónica en el espacio–.<sup>4</sup> A partir de la teoría del funcional de la densidad es posible escribir una ecuación para orbitales de una partícula, de los cuales se obtiene la densidad, dado que la energía es un funcional de la densidad y la densidad del sistema minimiza este funcional. Ello implicaba una enorme ventaja operativa respecto de los enfoques basados en el marco teórico de la función de onda. Además, la función de la densidad electrónica puede determinarse experimentalmente por difracción de rayos x, o bien, por difracción de electrones, lo que permite una contrastación experimental de los resultados. Mientras que la función de onda de una molécula o átomo, hasta el momento, no es un rasgo medible. DFT es una teoría, en principio, exacta de la estructura electrónica del estado fundamental de un sistema, mientras que los métodos basados en la función de onda multielectrónica –Thomas-Fermi, Hartree-Fock-Slater–, son intrínsecamente aproximados. Sin embargo, en estos métodos tradicionales, en principio, se puede obtener el nivel de precisión seleccionado, si se cuenta con recursos computacionales suficientes. En tanto, DFT tiene una limitación intrínseca. Si bien, en su formulación es teóricamente exacta, la teoría no dice cuál es ese funcional, ni cómo encontrarlo (Segall et al., 2002). Además, la complejidad de muchos electrones que interactúan persiste. Por ello DFT en su estado “puro” no es suficientemente precisa, depende de encontrar un adecuado funcional de la energía de intercambio y correlación, y no hay un modo sistemático de alcanzar el nivel de precisión buscado.<sup>5</sup> Se hace necesario encontrar la energía de intercambio y correlación como una función de la densidad, lo que demanda la introducción de aproximaciones.<sup>6</sup>

---

electrónica y en la formulación de Kohn y Sham, orbitales ficticios de una partícula simple (Kohn, Becke, & Parr, 1996).

<sup>4</sup> Designada por, ésta es una probabilidad por unidad de volumen. Si se considera la carga de un electrón como unidad de carga entonces se tienen unidades de carga de electrones.

<sup>5</sup> La precisión de un método refiere a la dispersión entre sí de los datos resultantes en distintas corridas de una simulación. En ese sentido, un método puede ser preciso (*precise*), aunque no exacto (*accurate*). En la práctica, en sentido estricto, sólo se cuenta con un valor exacto para el caso del átomo de hidrógeno, dado que se puede resolver analíticamente a partir de la teoría. Generalmente, se emplean como blanco (*target*) valores obtenidos experimentalmente. Precisión (en mediciones), requiere de estandarización (Norton Wise, 1995, p. 8). Así, precisión significa, qué tanto se aproxima el resultado de un método al valor blanco o establecido como estándar.

<sup>6</sup> La formulación original de Kohn y Sham constituía una aproximación, “aproximación de la

Como señalamos anteriormente, este enfoque comenzó a difundirse a partir de la década del 90 cuando se introdujeron aproximaciones y parámetros empíricos al cálculo, como así también, algunos aspectos híbridos basados en la función de onda, con el propósito de mejorar su nivel de precisión y disminuir el esfuerzo de cálculo. De ese modo consiguieron abordar sistemas más grandes. En este sentido, más interesantes para la química.

Primeramente, se introdujo un gradiente en la distribución de la densidad de los electrones. Como una forma de considerar un gradiente en la distribución electrónica de la densidad, se desarrolló la “aproximación del gradiente generalizado” (GGA) (Becke, 1988). Esto se hace de manera semiempírica, incorporando resultados experimentales.<sup>7</sup> Esta aproximación ha mejorado mucho los resultados calculados, pero introdujo algunos parámetros numéricos de ajuste. Más tarde, se mejoró el enfoque introduciendo un método híbrido parametrizado (Becke, 1993a, 1993b, 1996).<sup>8</sup> Gracias a estas modificaciones, se logró ampliar el tamaño de los sistemas abordables a un nivel de precisión satisfactorio. En general 1 Kcal/mol suele ser la precisión apropiada seleccionada para energías tales como calores de formación o potenciales de ionizaciones.<sup>9</sup>

Los trabajos de Becke generaron la posibilidad de abordar sistemas más grandes y con un alto grado de precisión, lo que provocó una marcada profusión de desarrollos con distintos grados de sofisticación basados en la teoría del funcional de la densidad (DFT). La irrupción de estos métodos, en sus diferentes mixturas y combinaciones, en la química cuántica provocó un marcado cambio en la relación de las variables mostrada por ambas gráficas. Se desvirtuó la tendencia hacia los extremos del gráfico de las elaboraciones de la química cuántica computacional que representa la hipérbola de Pople (1965), también marcada por Karplus (1990). Cabe recordar que la gráfica de Pople data de casi los inicios de esta disciplina y que cuando Karplus realizó la suya, 25 años después, la tendencia aún se mantenía.

Por otra parte, hemos mostrado que las expectativas que generó DFT de abordar sistemas más grandes al disminuir la complejidad del cálculo y

---

densidad local” (LDA). Sin embargo, ésta no era lo suficientemente precisa.

<sup>7</sup> Los valores son determinados mediante la optimización de la precisión de las energías de atomización a partir de un conjunto estándar de moléculas.

<sup>8</sup> El método de funcionales híbridos Hybrid Hartree-Fock/GGA.

<sup>9</sup> Los modelos son cuantitativos en la medida que reproducen y predicen valores experimentales con la precisión seleccionada. En el sentido de cuánto se adecuan empíricamente.

en el espíritu de los métodos *ab initio* en cuanto al grado de sofisticación, sólo cuando resignaron el rigor teórico consiguieron un nivel de precisión aceptable. En la actualidad, hay una amplia gama de simulaciones computacionales basadas en este enfoque con distinto grado de sofisticación y precisión. Un acercamiento a estos paquetes muestra que, en muchos casos, las menos sofisticadas son más precisas. Un ejemplo ilustrativo de problemas relativos a la precisión en DFT, es el caso del hidrógeno iónico y el molecular ( $H_2$  y  $H_2^+$ ). En general, la modelización de los dos sistemas con un funcional de la densidad dado, nunca son correctas ambas. Es decir, siempre sólo uno de los dos se puede describir bien, mientras que, el funcional falla para describir el otro (Cohen, Mori-Sánchez, & Yang, 2012). Así, uno de los principales resultados de nuestro análisis es que no hay una relación directa entre la sofisticación y la precisión de estos métodos, cuestionando uno de los pilares de la argumentación para sostener la preferencia por los métodos *ab initio*.

Cabría preguntarse a qué obedece el que perdure la preferencia por los métodos *ab initio* si, como hemos argumentado, ello no puede ligarse al grado de precisión. Además, en términos de rigor teórico, cómo se podría explicar que las simulaciones construidas sobre la base de la teoría del funcional de la densidad sea el enfoque predominante en los desarrollos de la química computacional, a pesar de que tuvieron que renunciar al más alto nivel teórico e introducir valores experimentales e híbridos para mejorar su precisión.

## **Consideraciones finales**

Tanto la gráfica de Pople (1965), como la de Karplus (1990), marcaban la forma de una curva hiperbólica que exhibía las fronteras de la química cuántica computacional, mostrando las divergencias en los desarrollos de métodos *ab initio* y semiempíricos. En el primer caso, tendían a progresar sobre el eje de sofisticación y en el segundo caso, sobre el otro eje intentado tratar sistemas cada vez más grandes. Hemos mostrado que la consideración de los desarrollos basados en DFT generaron un cambio en las tendencias señaladas y ya no tenemos una hipérbola. El crecimiento en número de estas elaboraciones se ve reflejado en el centro del diagrama, exhibiendo que en las últimas tres décadas se trabajó intensamente en la zona media de precisión, sofisticación y tamaño de sistemas abordables. En cuanto a estos desarrollos, hemos apuntado que sólo lograron una pre-

cisión aceptable cuando renunciaron al rigor teórico y concomitantemente lograron tratar sistemas más interesantes para la química.

En su artículo “Obituary: Density functional theory (1927–1993)”, Gill (2001) profetizaba que la introducción de funcionales híbridos ocasionaría la muerte de DFT, debido a la pérdida de su rigor teórico. A partir de algunas publicaciones actuales como las de Goerigk y Mehta (2019) y de Grimme y Schreiner (2018), notamos que pervive el desiderátum de abordar sistemas cada vez más grandes mediante métodos preferentemente ab initio. En nuestra lectura, analizando las progresiones de las gráficas, la práctica muestra que ello opera casi como un imaginario. Consideramos que la amplia difusión de métodos basados en la teoría del funcional de la densidad en las más variadas combinaciones y mixturas, pone en evidencia el imperativo de abordar sistemas cada vez más grandes, con una mayor precisión, aunque para ello se tenga que renunciar al rigor teórico.

## Referencias

- Becke, A. D. (1988). Density-functional exchange-energy approximation with correct asymptotic Behavior. *Physical Review A*, 38(6), 3098-3100.
- Becke, A. D. (1993a). A new mixing of Hartree-Fock and local density-functional theories. *The Journal of Chemical Physics*, 98(2), 1372-1377.
- Becke, A. D. (1993b). Density-functional thermochemistry. III. The role of exact Exchange. *The Journal of Chemical Physics*, 98(7), 5648-5652.
- Becke, A. D. (1996). Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing. *The Journal of Chemical Physics*, 104(3), 1040-1046.
- Cohen, A. J., Mori-Sánchez, P., & Yang, W. (2012). Challenges for density functional theory. *Chemical Reviews*, 112, 289-320.
- Dewar, M. J. (1992). *A semiempirical life*. Washington DC: American Chemical Society.
- Fermi, E. (1927). Un metodo statistico per la determinazione di alcune proprietà dell'atomo. *Rendiconte della Reale Accademia Nazionale dei Lincei*, 6, 602-607.

- Freed, K. F. (1995). Building a bridge between ab initio and semiempirical theories of molecular electronic structure. En J. L. Calais & E. S. Kryachko (Eds.), *Structure and dynamics of atoms and molecules: Conceptual trends* (pp. 25–67). Dordrecht: Springer Netherlands.
- Gill, P. M. (2001). Obituary: Density functional theory (1927–1993). *Australian Journal of Chemistry*, 54(11), 661–662.
- Goerigk, L., & Mehta, N. (2019). A trip for the density functional theory zoo: Warnings and recommendations for the user. *Australian Journal of Chemistry*, 72(8), 563–573.
- Grimme, S., & Schreiner, P. R. (2018). Computational chemistry: The fate of current methods and future challenges. *Angewandte Chemie Int. Ed.* 57, 4170–4176.
- Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Physical Review*, 136(3B), 864–871.
- Karplus, M. (1990). Three-dimensional "Pople diagram". *The Journal of Physical Chemistry*, 94(14), 5435–5436.
- Kohn, W., Becke, A. D., & Parr, R. G. (1996). Density functional theory of electronic structure. *The Journal of Physical Chemistry*; 100, 12974–12980.
- Kohn, W. & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A), 1133–1138.
- Norton Wise, M. (Ed.). (1995). *The values of precision*. Princeton, USA: Princeton University Press.
- Pople, J. A. (1950). The molecular orbital theory of chemical valency. V. The structure of water and similar molecules. *Proceedings of the Royal Society of London*, 202, 323–335.
- Pople, J. A. (1965). Two dimensional chart of quantum chemistry. *The Journal of Chemical Physics*, 43(10), 229–230.
- Segall, M. D., Lindan, P. J. D., Probert, M. J., Pickard, C. J., Hasnip, P. J., Clark, S. J., & Payne, M. C. (2002). First-principles simulation: ideas, illustrations and the CASTEP code. *Journal of Physics: Condensed Matter*, 14, 2717–2744.

- Thiel, W. (2005). Semiempirical quantum-chemical methods in computational chemistry. En Frenking, G., Kim, K. S. & Scuseria, G. E. (Eds.) *Theory and applications of computational chemistry* (pp. 559-580). Amsterdam: Elsevier.
- Ramsey, J. (2000). Of parameters and principles: Producing theory in twentieth century physics and chemistry. *Studies in History and Philosophy of Science, Part B*; 31(4), 549-567.
- Ramsey, J. L. (1997). Between the fundamental and the phenomenological: The challenge of “semi-empirical” methods. *Philosophy of Science*, 64(4), 627-653.
- Scerri, E. (2004). Principles and parameters in physics and chemistry. *Philosophy of Science*, 71(5), 1082-1094.
- Thomas, L. H. (1927). The calculation of atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society*, 23(5), 542-548.



# Cosas aprehendidas de la física de la gravitación

Oswaldo M. Moreschi<sup>\*,\*</sup>

## 1. Introducción

Con el anuncio en 2016 de la detección de ondas gravitacionales, hemos comenzado una nueva era en el desarrollo de la física de la gravitación. Este hito en la historia del desarrollo científico de la humanidad abre una nueva era en la que comenzamos a obtener información de sistemas distantes de otros métodos de las ya acostumbradas interacciones electromagnéticas.

Esta circunstancia nos provoca que repacemos qué hemos aprehendido en estos pocos siglos desde el comienzo de nuestro entendimiento de teorías de la gravitación.

Aunque intentaremos seguir un orden cronológico en la presentación; la misma no pretende ser una contribución a la historia de la ciencia, sino más bien al estudio de los avances en nuestro entendimiento de los fenómenos asociados a la gravitación.

## 2. De la época clásica

### 2.1. Espaciotiempo galileano y dinámica newtoniana (1600s - 1700s, Galileo - Newton)

Del trabajo de Galileo quedó claro los conceptos de velocidad, aceleración; como así también las nociones de movimiento con velocidad constante y uniformemente acelerado. Él también dedujo la relación cuadrática del tiempo con la aceleración constante. Asimismo, describió adecuadamente el movimiento de proyectiles en un campo de aceleración gravitatoria constante.

\* Universidad Nacional de Córdoba (UNC), Facultad de Matemática, Astronomía, Física y Computación (FaMAF). Córdoba, Argentina.

\* Instituto de Física Enrique Gaviola (IFEG), CONICET, Córdoba Argentina.  
moreschi@famaf.unc.edu.ar



De esta manera, en los comienzos del entendimiento de la mecánica quedó cristalizada una visión del *espaciotiempo*, que hoy denominamos galileana o newtoniana. En ella, sintéticamente, se entiende al tiempo como absoluto y al espacio tridimensional, como relativo a un estado del movimiento. Hoy presentamos estas ideas en torno al concepto de *sistemas inerciales*, y *transformaciones entre sistemas inerciales*.

De los aportes de Newton, quedó claro una manera de expresar la *ley de movimiento* para la mecánica y además la explicación de las *interacciones gravitatorias*; introduciendo en particular la fuerza entre dos partículas.

Hoy presentamos estas ideas en término de principios (Moreschi, 2000).

Supongamos un sistema mecánico que consta sólo de un conjunto de partículas de prueba; ante estas circunstancias postulamos:

**Principio: 2.1. (de existencia de los sistemas inerciales)** Dado un sistema mecánico que consta sólo de un conjunto de partículas de prueba, existe una familia de sistemas de coordenadas para los cuales las trayectorias de las partículas son líneas rectas. A estos sistemas se los denomina *sistemas inerciales*.

Supongamos un sistema mecánico compuesto por un conjunto de partículas; luego se cumple que:

**Principio: 2.2. (de determinación de la mecánica de partículas)** El estado mecánico de un sistema queda determinado por el valor de las posiciones y velocidades de sus partículas en un determinado tiempo. O sea, con esta información es posible predecir su ulterior movimiento.

Este es equivalente a la segunda ley de Newton. Aunque se podría generalizar a la mecánica que admita momento angular intrínseco para las partículas.

## 2.2. Nociones incipientes de relatividad (1600s - 1700s, Leibniz; 1800s, Mach)

Si bien hoy entendemos a la noción de espacio clásico como relativo; inicialmente Newton afirmaba una naturaleza absoluta también del espacio. Esto fue confrontado por Leibniz que afirmaba: “el espacio es algo puramente relativo, como el tiempo; el espacio es un orden de coexistencia como el tiempo es un orden de sucesiones” (Leibniz, 1976, p. 682).

Si bien es muy práctico y sintético postular la existencia de un espaciotiempo galileano, con características de geometría euclidiana; el mismo no tendría, en principio, ninguna relación con la distribución de materia. Leibniz nos invita a asociar la existencia misma del espacio con su contenido de materia. Es natural pensar que la noción de espacio no tiene sentido a no ser que se considere como conteniendo materia.

Estas concepciones de relatividad fueron acentuadas en los escritos de Mach, quien decía (relatividad galileana):

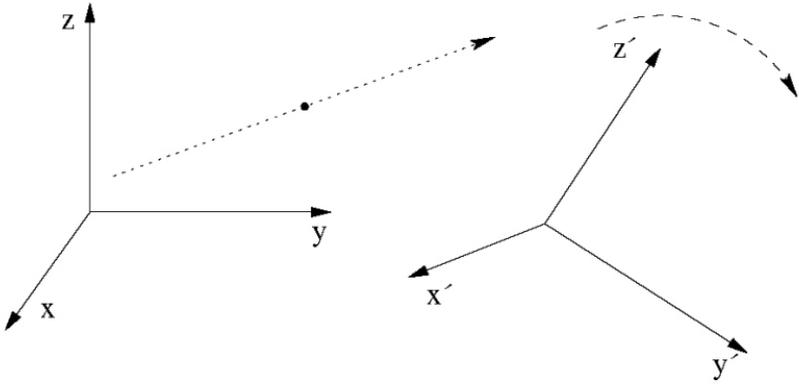
Ninguno es competente para predicar cosas acerca del espacio absoluto y movimiento absoluto; ellas son cosas del pensamiento, puras construcciones mentales, que no pueden ser producidas en la experiencia. Todos nuestros principios de la mecánica son, como hemos mostrado en detalle, conocimiento experimental concerniendo las posiciones y movimientos relativos de cuerpos. (Mach, 1883/1919, p. 229)

Esto está en un todo de acuerdo con lo que se entiende por el principio de la relatividad en mecánica clásica; que se podría expresar por:

**Principio: 2.3. (de la relatividad)** Todo sistema *cartesiano* en movimiento uniforme rectilíneo respecto de un sistema inercial es también un sistema inercial.

Las leyes de la física son las mismas en cualquier sistema inercial de referencia.

Sin embargo, Mach avanza sobre las nociones relativistas, aportando a dilucidar qué determina *la naturaleza de los sistemas inerciales*. Recordemos su discusión del experimento de Newton del balde con agua rotante. En dicho experimento Newton describe la situación de un balde, suspendido por una cuerda larga, que es girado tantas veces que finalmente la cuerda es torsionada fuertemente, luego es llenado con agua, y mantenido en reposo junto con el agua. Luego se lo suelta, y por la acción de la fuerza de torsión comenzará a rotar en el sentido contrario, y continuará, mientras la cuerda se vaya desenroscando, en este movimiento por un tiempo. La superficie del agua estará primeramente nivelada; pero subsiguientemente, el balde irá comunicando gradualmente el movimiento al agua, haciéndola rotar, entonces el nivel del agua bajará en el medio y se elevará en el borde del balde, asumiendo la superficie una forma cóncava. Mach escribió al respecto (semilla de la relatividad 2):



**Fig. 1:** Dos sistemas cartesianos de coordenadas. El sistema con coordenadas  $(x',y',z')$  se mueve rotando respecto del sistema  $(x,y,z)$ .

El experimento de Newton con el balde de agua rotante nos informa simplemente que la rotación relativa del agua con respecto al borde del balde no produce fuerzas centrífugas notables, pero que esas fuerzas son producidas por la rotación relativa con respecto a la masa de la Tierra y de los otros cuerpos celestes. Nadie es competente para decir cómo resultaría el experimento si el borde del balde se incrementase en espesor y masa hasta que tuviese un ancho de varias leguas. (Mach, 1883/1919, p. 232)

Esto nos induce a pensar que:

- Lo que determina que un sistema local sea o no inercial, es la existencia de la distribución de materia en el Universo.
- Es probable que esto haya influenciado a Einstein en su búsqueda de una ecuación de campo apropiada para la relatividad general.

### 3. De la época relativista

#### 3.1. La existencia de la velocidad máxima

La era relativista está asociada a la aparición en escena de una velocidad universal en la descripción de los fenómenos físicos. Si bien históricamente Einstein se refirió a la constancia de la velocidad de la luz en distintos sistemas inerciales (Einstein, 1905/1952); hoy preferimos referirnos a la existencia de una velocidad máxima para las interacciones, lo que formulamos en forma de principio.

**Principio: 3.1. (de la existencia de una velocidad máxima)** Existe una velocidad finita máxima para la velocidad de propagación de las interacciones.

Es posible deducir de este principio y del de la relatividad que las transformaciones entre sistemas inerciales con movimiento relativo, son las transformaciones de Lorentz (Moreschi, 2000).

### 3.2. El principio fundamental de la relatividad general

La existencia de una velocidad máxima para las interacciones fuerza la necesidad de contar con una teoría de la gravedad relativista. Pero, ¿cuál puede ser la idea base que permita construir el marco teórico para una teoría relativista de la gravedad? Esta piedra angular es provista por el llamado *principio de equivalencia*. Para la formulación del mismo es necesario perfeccionar la idea de sistema inercial a *sistema inercial local*.

**Principio: 3.2. (de existencia de los sistemas inerciales locales)** Existe un conjunto de sistemas de coordenadas, definidos alrededor de todo evento  $p$ , para los cuales las trayectorias de partículas de prueba son líneas rectas hasta el segundo orden en la separación de las coordenadas al punto  $p$ . A estos sistemas se los denomina sistemas inerciales locales.

Cuando una observación depende cuadráticamente de las coordenadas rectangulares de un sistema inercial local que tiene como origen un cierto punto  $p$ , diremos que la *observación es local*. Cualquier otra observación se llamará no local.

**Principio: 3.3. (de equivalencia)** No se puede distinguir por medio de observaciones locales entre un sistema de referencia en presencia de un campo gravitatorio y un sistema de referencia que está apropiadamente acelerado con respecto a un sistema inercial local.

Habría que remarcar que, como en el enunciado del principio de equivalencia no se ha restringido su acción a ningún subconjunto de situaciones, su aplicación es a todos los fenómenos físicos que puedan ser observados localmente. En particular esto implica que incluye a todas las otras interacciones.

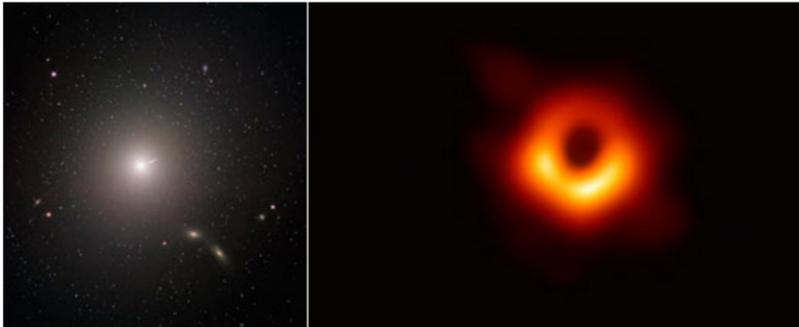
El principio de equivalencia tiene muchas consecuencias. Algunas de ellas son:

- La gravedad es un fenómeno geométrico que se manifiesta en la estructura del espaciotiempo. (Más precisamente: el espaciotiempo es una variedad lorentziana y los efectos gravitacionales se codifican en la curvatura del mismo).
- Las trayectorias de las partículas de prueba son geodésicas del espaciotiempo.
- Si un pequeño sistema tiene energía  $E$ , entonces  $m=E/c^2$ , se comporta como una masa gravitacional; donde  $c$  es la velocidad máxima de las interacciones. (En lenguaje coloquial, uno diría: “la energía pesa”.)

### 3.3. Situaciones esperadas en sistemas de la gravedad relativista

La gravedad relativista implica varias situaciones que se espera observar. Uno de ellos es la posibilidad de que se observen agujeros negros. La pregunta es: ¿se han observado? La respuesta es afirmativa.

Existe mucha evidencia indirecta que indica la existencia de agujeros negros en diversos sistemas astrofísicos. Pero la que ahora parece ser la más directa, es la imagen obtenida por el llamado *Event Horizon Telescope* que reproducimos aquí (Fig. 2).



**Fig. 2:** (izquierda) Messier 87 (M87) es una enorme galaxia elíptica ubicada a unos 55 millones de años luz de la Tierra, visible en la constelación de Virgo. (derecha) Primera imagen del agujero negro supermasivo en el centro de Messier 87 y su sombra. Imagen del *Event Horizon Telescope* (EHT).

*Nota:* Imagen izquierda de “Messier 87 Captured by ESO’s Very Large Telescope”, ESO, 2019. (<https://www.eso.org/public/images/eso1907b/>). CC BY 4. Imagen derecha de “First Image of a Black Hole”, Colaboración EHT, 2019. ESO. (<https://www.eso.org/public/images/eso1907a/>). CC BY 4.

También se debería mencionar que la mayoría de las detecciones de ondas gravitacionales están asociadas a sistemas binarios de agujeros negros en coalescencia.

Otras de las situaciones esperadas en sistemas relativistas, es la posibilidad de observar lentes gravitacionales; de las cuales existe amplia evidencia observacional.

### 3.4. La ecuación de campo de la relatividad general

Al presente, lo que se conoce como la teoría de la relatividad general, hace uso de la ecuación de campo de Hilbert-Einstein; dada por

$$G_{ab} \equiv R_{ab} - \frac{1}{2}g_{ab}R = -\frac{8\pi G}{c^2}T_{ab};$$

donde  $G_{ab}$  es el llamado tensor de Einstein,  $R_{ab}$  es el tensor de Ricci,  $g_{ab}$  es la métrica,  $R$  es el escalar de Ricci y  $T_{ab}$  el tensor energía-momento. La constante  $G$  es la de la gravitación, y  $c$  la velocidad máxima de las interacciones.

Redefiniendo los índices, esta ecuación se la entiende como:

$$G_{\text{eometria}} = m_a T_{\text{eria}}.$$

Lo que en principio podría estar en concordancia con las ideas de Mach; pero de cada lado de la ecuación se observan tensores de naturalezas absolutamente distintas. Esto hace pensar que en realidad esta ecuación de campo es un intento fenomenológico para describir la curvatura del espaciotiempo. Pues es difícil hoy darle un sustento fundamental.

Esto lleva a la pregunta: ¿cuál es el sustento observacional de la ecuación de campo de Hilbert-Einstein? Aquí debemos distinguir entre dos casos principales; que consisten en la presencia o ausencia de materia, que denotamos respectivamente por no vacío y vacío:

Casos	Subcasos	Observado
vacío	Solución de Schwarzschild Solución de Kerr ondas gravitacionales régimen de campo débil	sí sí sí sí
no vacío	régimen newtoniano ( $\equiv$ Física newtoniana) régimen relativista	sí ¿?

Nuestra duda en el último punto está asociada a que, aunque entendemos que sí hay corroboración observacional de la ley de conservación del tensor energía-momento en el régimen relativista, esto es  $\nabla_a T^{ab} = 0$ , somos críticos de las interpretaciones de las observaciones adjudicadas a la ecuación

$$R_{ab} - \frac{1}{2} g_{ab} R = -\frac{8\pi G}{c^2} T_{ab}.$$

Los datos observacionales que testean soluciones de vacío son abrumadores. Sólo mencionaremos aquí que algunos ejemplos son: la precesión del perihelio de Mercurio, desviación de la luz por objetos masivos, mediciones precisas del sistema planetario, descripción del funcionamiento del sistema GPS, observaciones de pulsar binarios, detección de ondas gravitacionales, etc.

Cuando se buscan situaciones donde las soluciones de no-vacío son importantes, usualmente se piensa en dos situaciones. Una es la descripción de la física interna de las *estrellas de neutrones*. La otra es la descripción del espaciotiempo usado en *cosmología*. Respecto de las estrellas de neutrones, la situación es que se tiene demasiada incerteza en la ecuación de estado de la materia nuclear; por lo que es difícil interpretar las deducciones de los cálculos teóricos. En el caso de la cosmología el tema se torna complicado por varios motivos; en particular por que se usa un término adicional en la ecuación de campo con una constante cosmológica. Volveremos a este punto más adelante.

#### 4. Era de la detección de las ondas gravitacionales

Como mencionamos en la introducción, apenas hemos entrado en la era de la detección de las ondas gravitacionales; por lo que merece que le dediquemos un espacio a este tema.

**Los observatorios:** los observatorios de ondas gravitacionales que han participado en detecciones de las mismas son hasta ahora tres. El observatorio LIGO ubicado en Hanford, estado de Washington, EEUU; el de LIGO ubicado en Livingston, en el estado de Louisiana, EEUU; y el observatorio Virgo ubicado cerca de Pisa en Italia. Existen otros observatorios en construcción y en desarrollo que se espera aporten a las observaciones en el futuro.

**Las detecciones:** al presente se han superado la decena de observaciones de la coalescencia de sistemas binarios de agujeros negros o estrellas de neutrones. Estos sistemas presentan la mayor demanda para la descripción de efectos de gravedad relativistas fuertes.

**La observación del espín 2:** para poder detectar la naturaleza del espín 2 de las ondas gravitacionales es necesario comparar la señal en al menos dos detectores. Para ello hemos construido una medida de comparación que denominamos  $\Lambda$ . En una conferencia reciente (Moreschi, 2019) hemos presentado la aplicación de nuestra medida al caso de los datos grabados por los observatorios LIGO para el evento GW150914 y hemos podido detectar la existencia de dos polarizaciones distintas de la onda gravitacional.

Esta detección es una importante confirmación de nuestro entendimiento geométrico de la relatividad general. Dado que confirma que la radiación gravitacional se comporta como una señal de espín 2 (en realidad  $-2$ ). Lo importante es distinguir que no se comporta ni como un escalar ni como una señal vectorial.

## 5. Era de los modelos cosmológicos

Mencionaremos aquí diversas ideas claves que caracterizan el enfoque que la comunidad usa para el estudio del sistema cosmológico.

**Espaciotiempos homogéneos e isotrópicos:** los spaciotiempos homogéneos e isotrópicos pueden ser expresados por el elemento de línea de Robertson-Walker dado por:

$$ds^2 = c^2 d\tau^2 - A(\tau)^2 (d\chi^2 + f_k^2(\chi) d\Sigma^2)$$

donde  $d\Sigma^2$  es la métrica de la esfera unidad y

$$f_k(\chi) = \begin{cases} \sinh(\chi) & \text{para } k = -1, \\ \chi & \text{para } k = 0, \\ \sin(\chi) & \text{para } k = 1. \end{cases}$$

El uso de esta geometría para modelar la cosmología ha sido basado en la observación de que para escalas muy grandes, la distribución de las galaxias parece ser más o menos homogénea e isotrópica; aunque también se observen agrupaciones en conjuntos, como en cúmulo de galaxias.

Si tomamos a las galaxias como la estructura más pequeña sobre la que describiremos a la cosmología, entonces se llega a un problema.

**El problema de los promedios implícitos:** en un modelo cosmológico simple que está compuesto por galaxias, éstas determinan la distribución de masa; entonces los fotones que detectamos de las galaxias estarían viajando por regiones del espaciotiempo donde:

El tensor de Ricci es cero,  $R_{ab} = 0$  y el tensor de Weyl es distinto de cero,  $W_{abc}{}^d \neq 0$ .

Sin embargo, en un universo que se asume homogéneo e isotrópico, como los descritos por la métrica de Robertson-Walker, uno tiene exactamente lo contrario; esto es:

Un tensor de Ricci distinto de cero,  $R_{ab} \neq 0$  y un tensor de Weyl cero,  $W_{abc}{}^d = 0$ .

Sin embargo, la comunidad no tiene inconvenientes de asumir que la geometría de Robertson-Walker proviene de hacer alguna clase de *promedio*; aunque el promedio de cero nunca podría dar distinto de cero.

**El problema de las masas:** dado que los espacios de Robertson-Walker son esféricamente simétricos alrededor de cualquier punto; también se los puede expresar en un sistema de coordenadas de tipo Schwarzschild, de la siguiente forma:

$$ds^2 = (1 - \Phi(t, r)) dt^2 - \frac{1}{1 - \frac{2M(t, r)}{r}} dr^2 - r^2 d\Sigma^2.$$

De donde se deduce que:

$$A(\tau) f_k(\chi) = r.$$

Se puede probar que:

$$\frac{2GM(t, r)}{c^2 r} = k f_k(\chi)^2 + \frac{r^2 H^2}{c^2}, \quad (\text{donde } H \equiv \frac{dA}{d\tau} = \frac{\dot{A}}{A})$$

donde hemos incluido las constantes universales explícitamente. Se concluye entonces:

$$\frac{GM(t,r)}{c^2} = \frac{1}{2}(krf_k(\chi)^2 + \frac{r^3 H^2}{c^2}) = \frac{r^3}{2} \left( \frac{k}{A^2} + \frac{H^2}{c^2} \right) = \frac{4\pi r^3}{3} \left( \frac{G}{c^4} \rho c^2 \right),$$

donde recién en la última ecuación hemos usado la ecuación de campo de Hilbert-Einstein.

Uno diría: “¡Qué bien!” dado que se tiene la simple expresión que parece decir: “masa = densidad constante × volumen” dado que

$$M(t,r) = \frac{4\pi r^3}{3} \rho(\tau);$$

sin embargo, en la geometría de Robertson-Walker un volumen en la hipersuperficie  $\tau = \text{constante}$  debería tener la expresión (tomando  $k = -1$ ):

$$\begin{aligned} V_{\chi_1} &= 4\pi A(\tau)^3 \int_0^{\chi_1} \sinh(\chi)^2 d\chi = \frac{4\pi A(\tau)^3}{2} (\sinh(\chi_1) \sqrt{1 + \sinh(\chi_1)^2} - \chi_1) \\ &= \frac{4\pi A(\tau)^3}{2} \left( \frac{r_1}{A} \sqrt{1 + \left(\frac{r_1}{A}\right)^2} - \operatorname{arcsinh}\left(\frac{r_1}{A}\right) \right); \end{aligned}$$

que tienen expansión en serie de Taylor alrededor del origen dada por:

$$V_{\chi_1} = 4\pi A(\tau)^3 \left( \frac{\chi_1^3}{3} + \frac{\chi_1^5}{15} + \frac{2\chi_1^7}{315} + \mathcal{O}(\chi_1^9) \right),$$

y

$$\frac{V_r}{A} = 4\pi A(\tau)^3 \left( \frac{1}{3} \left(\frac{r}{A}\right)^3 - \frac{1}{10} \left(\frac{r}{A}\right)^5 + \frac{3}{56} \left(\frac{r}{A}\right)^7 + \mathcal{O}\left(\left(\frac{r}{A}\right)^9\right) \right).$$

Resulta ser que  $\frac{4\pi}{3} r^3 > V_r$ ; por lo que si se define la *masa aditiva* del espacio homogéneo por

$$M_h(\tau, \chi) = M_h(A, r) = \rho(\tau) V_r;$$

se tendrá  $M > M_h$ . Esto es un poco confuso, pues se podría pensar que en el sistema basado en coordenadas  $(\tau, \chi)$ , integrar  $\rho$ , es equivalente a integrar  $G\left(\frac{\partial}{\partial \tau}, \frac{\partial}{\partial \tau}\right)$ ; lo que tomaría en cuenta automáticamente la contribución para esferas cada vez mayores de velocidades de alejamiento cada vez mayores.

Así y todo, vale la desigualdad de masas. ¡Pero la masa física es  $M$ ! Dado que, si se reemplazase el interior  $r < r_1$  por una solución de Schwarzschild de vacío, una partícula en la superficie frontera, sentiría la masa  $M$ . (Al presente cosmológico se asume la ausencia de presión).

Por otro lado si al interior de la esfera denotada por  $(t, r)$  ahora asumimos que hay una estrella de densidad constante dada por  $\rho_0 = \rho(\tau)$ ; entonces su masa correspondiente al sistema de observadores estáticos  $t = \text{constante}$  será

$$M_s = 4\pi \int_0^r \frac{\rho_0 r'^2}{\sqrt{1 - \frac{2m(r')}{r'}}} dr';$$

que debería tener en cuenta la interacción gravitacional. Como el elemento de volumen asociado a  $dr$  en el espaciotiempo estático de Schwarzschild es mayor que el asociado al espacio euclídeo, se tiene:  $M_s > M > M_h$ .

Esto señala las sutilezas en la relación entre:

- distribuciones de masa,
- densidad promediada,
- masa deducida de la densidad promediada.

Lo único que tenemos en claro es que la masa física  $M$  no coincide con la masa aditiva  $M_h$ .

**El problema de la materia oscura y la energía oscura:** la aplicación de las ecuaciones de Hilbert-Einstein al caso de espacios homogéneos e isotrópicos, se las conoce como ecuaciones de Friedman. Para construir el modelo se asume una mezcla de materia no relativista, materia relativista y contribución de la constante cosmológica. Se define la densidad crítica por:

$$\rho_c \equiv \frac{3}{8\pi G} \left(\frac{\dot{A}}{A}\right)^2 = \frac{3}{8\pi G} H^2;$$

luego se definen las distintas contribuciones en término de las relaciones:

$$\Omega_m \equiv \frac{\rho_{0,m}}{\rho_{0,c}}, \quad \Omega_r \equiv \frac{\rho_{0,r}}{\rho_{0,c}}, \quad \Omega_\Gamma \equiv \frac{\rho_\Gamma}{\rho_{0,c}};$$

con lo que la ecuación de campo al presente se puede expresar por

$$1 = \omega_k + \Omega_m + \Omega_r + \Omega_\Gamma;$$

donde

$$\omega_k \equiv -\frac{3}{8\pi G} \frac{k c^2}{\rho_{0,c} A_0^2} = \frac{-k c^2}{H_0^2 A_0^2}$$

Si determinamos estos parámetros de las observaciones de la distribución de materia y de la radiación, se tiene:  $\Omega_m = \Omega_b = 0.042569$  y  $\Omega_r = 4,7647 \times 10^{-5}$ ; de donde se deduce

$$\omega_{k,obs} = 1 - (\Omega_m + \Omega_r) = 0,957382909;$$

que a su vez indica  $k = -1$  y

$$A_{0,obs} = \frac{c}{H_0 \sqrt{\omega_{k,obs}}} = 1.31309 \times 10^{26} \text{ m} = 4.2554 \times 10^3 \text{ Mpc};$$

donde estamos usando la medición del parámetro de Hubble de aproximadamente  $H_0 = 72 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

Sin embargo, la comunidad no duda en ajustar un modelo cosmológico con contribuciones muy distintas (Aghanim et al., 2018). Del último análisis de los datos del satélite Planck, los parámetros que el *Planck Collaboration* deduce son:

$\Omega_m = 0.3153 \pm 0.0073$  y  $\Omega_\Lambda = 0.6847 \pm 0.0073$ ; ajustando el parámetro de Hubble a  $H_0 = 67.36 \pm 0.54 \text{ km s}^{-1} \text{ Mpc}^{-1}$  (Aghanim et al, 2018, Tabla 2, penúltima columna). Hay que notar aquí que:  $\Omega_m + \Omega_\Lambda = 1.0000 \pm 0.0146$ ; pues ellos usan la suposición que  $k = 0$ , o sea geometría espacial plana. Su estimación de la materia bariónica es:  $\Omega_b = 0.02237/h^2 = 0,03320962$ , por lo que la materia oscura se estima en:  $\Omega_o = \Omega_m - \Omega_b = 0,2821$ . (Despreciando la contribución de neutrinos).

Por otro lado recientes observaciones comprueban que el valor correcto de  $H_0$  es mayor. En (Riess, Casertano, Yuan, Macri, & Scolnic, 2019), los autores infieren de sus últimas observaciones y de medidas anteriores que el valor es:  $H_0 = 74.03 \pm 1.42 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

En definitiva, notamos un fenómeno sociológico en el ambiente de cosmólogos, donde se puede distinguir:

- En cosmología es aceptado el uso de la ecuación de Hilbert-Einstein.
- En cosmología se usa el llamado modelo  $\Lambda$ CDM [*lambda cold dark matter*], que impone la elección  $k = 0$ .
- La imposición  $k = 0$  obliga ajustar a  $H_0 = 67.36 \pm 0.54 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

•El valor medido observado del parámetro de Hubble es  $H_0 = 74.03 \pm 1.42$  km s<sup>-1</sup> Mpc<sup>-1</sup>.

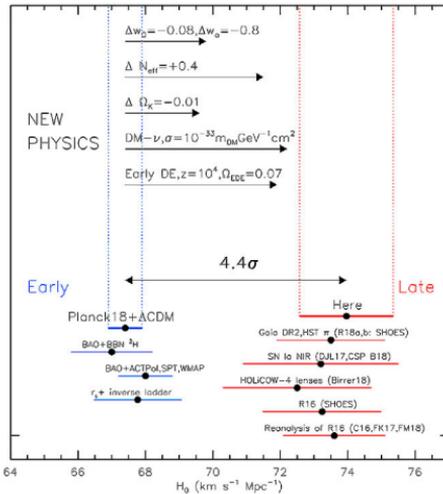
Es curioso que en la literatura a esto le llaman la existencia de una tensión; en vez de aceptar que la elección de  $k = 0$  es incompatible con las observaciones. ¿No deberíamos dudar también del uso de la ecuación de campo de Hilbert-Einstein en este régimen?

### Comentarios finales

Hemos intentado hacer un sintético repaso de las cosas aprehendidas del desarrollo de la física de la gravitación por la comunidad y el autor.

De lo anterior podemos mencionar que de la confirmación observacional del principio de equivalencia se deducen varias cosas. Que el espaciotiempo es una variedad lorentziana y los efectos gravitacionales se codifican en la curvatura del mismo. Que las soluciones de vacío de la ecuación de Hilbert-Einstein han sido corroboradas observacionalmente. Que respecto de las soluciones de no-vacío de la ecuación de Hilbert-Einstein se puede decir que se han corroborado en el límite no relativista, esto es, newtoniano. Pero su entendimiento en el ámbito de la cosmología nos deja con serias dudas sobre el actual uso de ellas.

THE ASTROPHYSICAL JOURNAL, 876:85 (13pp), 2019 May 1



### Referencias

Fig. 3: Tomamos prestado de Riess et al. (2019) su figura 4; donde se muestran los resultados de distintos estudios sobre  $H_0$  y su desacuerdo. Del lado izquierdo aparecen las estimaciones que ajustan los fenómenos observados a tiempos cosmológicos tempranos, y por lo tanto distantes. Del lado derecho aparecen las mediciones realizadas con observaciones asociadas a tiempos cosmológicos tardíos, y por lo tanto cercanas.

Nota: figura 4 de Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. (2019). *Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond  $\Lambda$ CDM*. 876(1), 85. doi: 10.3847/1538-4357/ab1422

- Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., ... Zweigig, J. [LIGO Scientific Collaboration and Virgo Collaboration]. (2016). Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters*, 116(6), 061102. doi: 10.1103/PhysRevLett.116.061102
- Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., ... Zonca, A. [Planck Collaboration]. (2019). Planck 2018 results. VI. Cosmological parameters. *arXiv:1807.06209* [astro-ph]. Recuperado de <http://arxiv.org/abs/1807.06209>
- Akiyama, K., Alberdi, A., Alef, W., Asada, K., Azulay, R., ... Ziurys, L. [T.E.H.T. Collaboration]. (2019). First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. *The Astrophysical Journal Letters*, 875(1), L1. doi: 10.3847/2041-8213/ab0ec7
- Einstein, A. (1952). On the electrodynamics of moving bodies [Zür Elektrodynamik bewegter Körper]. En *The Principle of Relativity*. New York: Dover pub. Inc. (Obra original de 1905)
- Leibniz, G. W. (1976). The controversy between Leibniz and Clarke, 1715-16. En L. E. Loemker (Ed., trad.), *Philosophical papers and letters* (2nd ed., 2nd print, pp. 675-721). Dordrecht, Holland Boston: D. Reidel Pub. Co.
- Mach, E. (1919). *The Science of Mechanics [Die Mechanik in ihrer Entwicklung historisch-kritisch dargestellt]*. Chicago: The Open Court Publishing Co. (Obra original de 1883)
- Moreschi, O. M. (2000). *Fundamentos de la mecánica de sistemas de partículas*. Córdoba: editorial Universidad Nacional de Córdoba.
- Moreschi, O. M. (2019). *Missed astrophysical signal in the GW150914 event*. Presentación oral en la Session C2 of GR22, 22th International Conference on General Relativity and Gravitation/13th Edoardo Amaldi Conference on Gravitational Waves, Valencia, España, julio 7-12, 2019.

Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. (2019). Large Magellanic Cloud cepheid standards provide a 1% foundation for the determination of the Hubble constant and stronger evidence for physics beyond  $\Lambda$ CDM. *The Astrophysical Journal*, 876(1), 85. <https://doi.org/10.3847/1538-4357/ab1422>



## Notas históricas sobre el puente Einstein-Rosen

Víctor Rodríguez\*

Pedro Walter Lamberti\*\*,‡

### Introducción

Recientemente se ha propuesto una conexión entre el entrelazamiento cuántico y la solución de las ecuaciones de Einstein que conducen a lo que hoy se conoce como el “puente de Einstein-Rosen”. Esta propuesta se realiza en el contexto del estudio de una teoría cuántica de la gravedad. Inicialmente fue presentada por Mark Van Raamsdonk (2010) y ha sido reformulada por J. Maldacena y L. Susskind (2013). Esta eventual y sumamente conjetural conexión ha sido denominada EP=EPR.

¿A qué se refieren estas iniciales? Con diferencia de dos meses, en el año 1935 Einstein y sus colaboradores publicaron dos trabajos que con el tiempo adquirieron una importancia crucial en el marco de la física teórica. El primero (EPR), fue publicado en mayo de 1935 en *Physical Review*, bajo el título “Can quantum-mechanical description of physical reality be considered complete?”, con autoría de A. Einstein, B. Podolsky y N. Rosen (1935). El segundo (ER), fue publicado en la misma revista en el mes de julio bajo el título “The particle problem in the general theory of relativity”. En este caso los autores son A. Einstein y N. Rosen (1935), este último un joven colaborador del primero.

El trabajo de EPR ha sido extensamente analizado y reconocido por sus profundas implicancias, tanto en los fundamentos de la mecánica cuántica como en las consecuencias tecnológicas de la teoría. De este trabajo surge el concepto de “entrelazamiento cuántico”, recurso básico para el desarrollo de la computación cuántica. Escapa a nuestros objetivos eva-

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH). Córdoba, Argentina.

\*\* Universidad Nacional de Córdoba (UNC), Facultad de Matemática, Física, Astronomía y Computación (FAMAF), Córdoba, Argentina.

‡ CONICET, Córdoba, Argentina.  
lamberti@famaf.unc.edu.ar



luar este tema dentro del presente trabajo. Nos limitaremos a hacer unas pocas consideraciones sobre el segundo artículo. El primer trabajo refleja, sin lugar a dudas, el estado de desconcierto en materia de interpretación de la mecánica cuántica, al momento de su publicación. No se conoce la verdadera influencia que Einstein tuvo en él. Existen argumentos para concluir que Einstein no se sentía totalmente satisfecho con el resultado del artículo. En una carta a Schrödinger, expresa que el punto principal había sido “sepultado por la erudición”. En esa misma carta da a conocer que el artículo había sido escrito mayoritariamente por Podolsky. Desde su perspectiva, el objetivo del trabajo era presentar un argumento simple para demostrar la “incompletitud de la teoría cuántica”. A la luz de lo ocurrido en los años siguientes a su publicación, el trabajo generó un profundo debate acerca de los cimientos sobre los que se monta la teoría, siendo el realismo y la no-localidad, dos de los aspectos más investigados experimentalmente. La génesis de la no-localidad tiene muchas facetas que están aún lejos de comprenderse.

Desde el punto de vista del desarrollo histórico de la física teórica, 85 años después de la publicación de los trabajos EPR y ER, los grandes temas involucrados siguen teniendo plena vigencia. Si prestamos atención a las motivaciones del trabajo ER, es decir, a la conciliación entre la macro y micro física, si bien se han logrado avances significativos, se está lejos de tener una teoría que permita la descripción del fenómeno gravitatorio a la luz de la mecánica cuántica. Es por ello que estimamos que pequeños abordajes relacionados con aspectos parciales pueden ser útiles, sobre todo atendiendo a un contexto que integre las diferentes facetas que entran en juego. En ese sentido, el presente trabajo tiene pretensiones muy modestas. Consideramos que incursionar en algunos aspectos históricos, puede ayudar a esclarecer algo el panorama conceptual general. En particular, nos limitaremos a elaborar sólo ciertos comentarios sobre el trabajo de A. Einstein y N. Rosen (ER), ya que consideramos que el mismo incluye un conjunto de reflexiones de considerable valor metodológico y epistemológico. Esto toca tanto al propio artículo ER, como a su entorno. La primera impresión que se extrae es que se trata de un período notablemente activo de la física teórica y que constituye una rica fuente de discusiones filosóficas.

## El artículo ER

Este trabajo comienza con el siguiente párrafo:

Más allá del gran éxito en varios campos, la física teórica actual está lejos de ser capaz de proveer una fundamentación unificada sobre la cual pueda basarse el tratamiento teórico de todos los fenómenos. Tenemos una teoría general relativista de los fenómenos macroscópicos, la cual hasta ahora ha sido incapaz de dar cuentas de la estructura atómica de la materia y de los efectos cuánticos, y tenemos una teoría cuántica, la cual permite dar cuenta satisfactoriamente de una gran cantidad de fenómenos cuánticos y atómicos pero la cual, por su propia naturaleza, es inadecuada con respecto al principio de la relatividad. Bajo estas circunstancias no parece superfluo encarar la cuestión de en qué extensión la relatividad general puede tomar en cuenta los fenómenos atómicos.<sup>1</sup> (Einstein & Rosen, 1935, p. 73)

Los autores señalan que el objetivo del trabajo es investigar la posibilidad de una teoría atomística de la materia y de la electricidad que sólo haga uso de variables de la teoría de la relatividad general y de la teoría de Maxwell, aunque señalan explícitamente que excluyen las singularidades del campo. A través de un ejemplo, modifican ligeramente las ecuaciones gravitacionales, las que entonces permiten soluciones regulares para el caso estático esféricamente simétrico. Allí aparece un tema sobre el cual ha corrido mucha tinta. Estas soluciones permiten la representación matemática del espacio físico por medio de un espacio de dos hojas idénticas y consideran a una partícula como representada por un puente que conecta a estas dos hojas. Luego se extienden al caso de un sistema de muchas partículas, considerando que se lo podría representar por dos hojas unidas por muchos puentes, pero no lo desarrollan en detalle.

En el artículo, tras hacer un comentario general respecto del estado de la teoría de la relatividad general y de la teoría cuántica, llegan a la conclusión de que es legítimo e importante plantearse la pregunta sobre el alcance del método de la relatividad general para dar cuenta de los fenómenos atómicos, aunque reconocen que no están en condiciones de dar una respuesta contundente sobre si esta teoría puede dar o no cuenta de los fenómenos cuánticos. No obstante, justifican esta estrategia por tratarse de un método con mínimas suposiciones y que permite operar con un

---

<sup>1</sup> Todas las traducciones de los textos en inglés han sido realizadas por los autores del presente trabajo.

procedimiento claro, donde las eventuales complicaciones se encuentran restringidas a las expresiones matemáticas. En realidad, la pregunta que se hacen es simple: ¿se puede concebir una teoría atomística de la materia y de la electricidad que haga uso solamente de variables del campo gravitacional y de los potenciales vectoriales del electromagnetismo de Maxwell? Aclaran que se supone que las singularidades en el campo están excluidas. Su respuesta a la anterior pregunta es afirmativa y eso es lo que se proponen desarrollar, aunque interpretan que considerando a la solución de Schwarzschild y la extensión de Reissner a la misma, uno podría estar tentado a responder negativamente. Por otro lado, atienden en general a los enfoques previos que intentaron considerar a las partículas como singularidades del campo, pero no comparten esa opinión.

### **Sobre las singularidades en una teoría de campos**

“Una singularidad trae tanta arbitrariedad dentro de la teoría. . . que realmente anula sus leyes. . . . Toda teoría de campo, en nuestra opinión, debe en consecuencia adherir al principio fundamental de que deben excluirse las singularidades del campo” (Einstein & Rosen, 1935, p. 73).

Tomada en sí misma esta cita es potencialmente engañosa. Lo que Einstein y Rosen estaban objetando era no tanto las singularidades del espacio-tiempo en general como la idea de que las partículas materiales puedan ser tratadas como singularidades de un campo gravitacional, una idea que fue explorada por Einstein mismo de vez en cuando. En el artículo en cuestión esta idea fue rechazada por Einstein y Rosen por la razón de que tal singularidad “anula” las leyes de la teoría.

Esta afirmación, que es una declaración de principios, formó parte de un credo que fue discutido por Einstein y otros autores cuando se formuló la teoría general de la relatividad, pero que tiene su historia y se remonta a los albores de la teoría del campo electromagnético de Maxwell. A continuación, y con la única intención de dar una idea del tipo de discusión que originó el concepto de partícula como singularidad, mencionaremos algunas ideas que predominaban en los albores del siglo XX.

En 1904 A.H. Bucherer propuso un electrón no puntual, de volumen constante pero sujeto a la contracción de Lorentz-FitzGerald. En 1908 este autor comunicó a Einstein que él abandonaba su modelo en favor del modelo provisto por la teoría de la relatividad especial. Otros autores de gran influencia en la época, como lo fueron Paul Langevin y Henri

Poincaré propusieron sus propios modelos para una partícula elemental. En un congreso de 1904, Langevin afirmaba que “los hechos experimentales imponen en [las] cargas móviles una estructura granular, discontinua, y llevan a la idea del electrón como una región particular del éter” (Langevin, 1904/1908, p. 124). Poincaré proclamó en 1906 “el final de la materia”. Y explicó que “consecuentemente, esos electrones negativos no tienen masa real; si ellos aparecen equipados con una inercia, es debido al hecho que su velocidad no puede ser cambiada sin una perturbación simultánea del éter luminoso. Su aparente inercia es sólo tomada prestada, no pertenece a ella, sino al éter” (Poincaré, 1906/1917, p. 285). Para este autor francés “hay sólo huecos en el éter” (p. 288).

En el año 1936, Einstein publicó un artículo bajo el título “Física y realidad” (Einstein, 1936), el cual es relevante para conocer los problemas que él consideraba de gran importancia para la física teórica. Y respecto del asunto de la descripción de las partículas en el marco de la teoría electromagnética clásica, dice:

Las ecuaciones de Maxwell en su forma original no permiten . . . la descripción de partículas, debido a que sus soluciones contienen singularidades. Los físicos teóricos han tratado, por largo tiempo, de lograr la modificación de las ecuaciones de Maxwell. Esos intentos, sin embargo, no han sido coronados por éxito. Así ocurre que el objetivo de erigir una teoría puramente electromagnética de la materia, no se ha alcanzado, pero eso no quiere decir que eso no se pueda lograr. (p. 365)

Regresemos al trabajo ER. En su discusión sobre el rol de las singularidades sus autores rescatan un argumento dado a conocer por L. Silberstein en una carta a Einstein y expresan:

Como es bien sabido, Levi-Civita y Weyl han dado un método general para encontrar soluciones estáticas axialmente simétricas de las ecuaciones gravitacionales. Por este método uno puede claramente obtener una solución la cual, excepto para dos singularidades puntuales ubicadas sobre el eje de simetría, es en todas partes regular y es euclidiana en el infinito. En consecuencia, si uno admitiera singularidades como representando partículas uno tendría aquí un caso de dos partículas no aceleradas por su acción gravitacional, lo cual ciertamente debe excluirse físicamente. (Einstein & Rosen, 1935, p. 73)

El intercambio epistolar entre Einstein y Silberstein terminó en una disputa de tono elevado entre ambos científicos, extendiéndose a debates en la prensa, como veremos un poco más adelante (Havas, 1989, 1993).

## El trabajo ER con algún detalle formal

En el trabajo ER, los autores buscan construir un modelo geométrico de una partícula elemental que fuese finita y libre de singularidad. El resultado del trabajo, y que aparentemente los satisface, es que

logramos una solución de las ecuaciones de campo modificadas (de la RG,) que resulta ser una representación matemática del espacio físico a través de un espacio de dos hojas idénticas, estando una partícula representada por un 'puente' que conecta a esos dos hojas. (Einstein & Rosen, 1936, p. 73)

Las ecuaciones de campo de Einstein son

$$R_{ij} - \frac{1}{2}Rg_{ij} = kT_{ij}$$

en las cuales, sin entrar en detalles, el lado izquierdo representa a la geometría del espacio-tiempo y el lado derecho hace referencia a la distribución de materia-energía. El vacío se representa con  $T_{ij} = 0$ . En este caso las ecuaciones de Einstein se reducen a la anulación del tensor de Ricci:

$$R_{ij} = 0$$

Un poco más de un mes después de que Einstein formulara sus ecuaciones en 1915, Karl Schwarzschild encontró la primera solución exacta de las ecuaciones, bajo la condición de simetría esférica. Lo que hoy se conoce como métrica de Schwarzschild se escribe, en coordenadas esféricas  $(t, r, \theta, \varphi)$ , de la forma

$$ds^2 = -A(r) dt^2 + \frac{1}{A(r)} dr^2 + r^2(d\theta^2 + (\sin\theta)^2 d\varphi^2)$$

donde la función  $A(r)$  está dada por

$$A(r) = 1 - \frac{2m}{r}$$

y  $m$  es la masa del objeto ubicado en el origen de coordenadas. Esta métrica tiene dos puntos singulares. Una en  $r = 2m$  y la otra en  $r = 0$ . En esos puntos se anulan los coeficientes de la métrica. La primera no es una singularidad esencial, mientras que la segunda sí lo es. El carácter no esencial de la singular tiene que ver con que ella es debida al sistema de coordenadas en el que se escribe la solución. La región  $r > 2m$  es el exterior de la

solución de Schwarzschild, mientras que  $r < 2m$  corresponde al interior de un agujero negro. Es importante resaltar que en los primeros años de la RG, no se entendía muy bien la idea y el rol que las singularidades tenían en la teoría (Curiel, 2019).

Una modificación que hizo Einstein a sus ecuaciones, y de la cual se arrepintió, fue la inclusión de un término que hoy se conoce como asociado a la constante cosmológica. En el año 1921, Erich Trefftz construyó una solución exacta de las ecuaciones con constante cosmológica con simetría esférica (Trefftz, 1922). Al año siguiente, Einstein demostró que la solución encontrada por Trefftz contenía una verdadera singularidad. Es interesante rescatar un párrafo de este trabajo de Einstein, donde una vez más pone énfasis en las singularidades:

En vista de la importancia del problema en la cuestión cosmológica, es decir el asunto de la estructura geométrica del universo, estuve interesado en saber si las ecuaciones conducen a la posibilidad física de un universo estático cuyas masas materiales están concentradas en sólo dos cuerpos celestes. Resulta obvio, sin embargo que la solución de Trefftz no permite en absoluto esta interpretación física. (Einstein, 1922, p. 448)

En el trabajo ER los autores introducen una nueva coordenada  $u$  tal que

$$r = r_0 + u^2$$

con

$$u \in (-\infty, +\infty)$$

En estas coordenadas, la métrica de Schwarzschild queda

$$ds^2 = -\frac{u^2}{u^2+r_0^2} dt^2 + 4(u^2+r_0) du^2 + (u^2+r_0)^2 d\Omega^2$$

Físicamente esta métrica se puede interpretar como correspondiente a dos copias idénticas del espacio-tiempo de Schwarzschild ( $r > r_0$ ) para  $u > 0$  y  $u < 0$ . Ambos se pegan en  $u = 0$ .

Desde un punto de vista más técnico, esta solución tiene algunos inconvenientes, uno de ellos no reconocido por los autores. Einstein y Rosen se dan cuenta de la singularidad en  $u = 0$ . Pero no mencionan que esta solución no satisface las ecuaciones de campo de vacío, pues al reemplazar la solución de ER en ellas resulta un tensor de energía-momento ( $T_{ij}$ ) singular.

Tras su publicación el trabajo ER recibió atención en la prensa, ciertamente por estar relacionado con el científico más “popular” de ese momento. Fue titular del *New York Times* cuatro días después de su publicación. El cronista encargado de dar la noticia fue W. Laurence, quien también dedicó la primera plana del NYT al artículo EPR.<sup>2</sup> El titular del NYT tenía cierto tono sensacionalista: “En una nueva vasta teoría, Einstein une los átomos con las estrellas, en un sistema unificado”. Pero no sólo en esta oportunidad el trabajo ER terminó en la prensa.

En el número del 1 de febrero de 1936, aparece un artículo en el *Physical Review*, cuyo autor es L. Silberstein. El mismo lleva por título “Solución de dos centros de las ecuaciones del campo gravitatorio, y la necesidad de una teoría de la materia modificada”. El autor de este trabajo cree haber demostrado que la RG era una teoría problemática. Rescatamos la forma en que lo expresa:

Así la solución (1), (10) corresponde a una ausencia total de materia por doquier (tensor de materia), excepto sólo en los puntos A y B, los “centros de masa”. Así (a diferencia de las soluciones de Bach y Weyl) no hay ninguna tensión para mantener los puntos apartados como una barra rígida. Y realmente siendo la solución estacionaria, invariable en el tiempo) los dos puntos son relativamente fijos uno del otro, en lugar de caer uno hacia el otro, lo cual contradice de manera flagrante una de las experiencias humanas más antiguas y primitivas. (Silberstein, 1936, p. 270)

El 10 de febrero de 1936, el periódico *Montreal Gazette* (Canadá) publica una nota titulada “Einstein dice que el ataque a su teoría por parte de un profesor (de la Universidad) de Toronto, estaba basado en un error”.

En marzo de ese año Einstein pide a Silberstein que retire esa publicación. Además, en términos poco amigables, expresa que la noticia del periódico canadiense “contenía la afirmación idiota” de que Silberstein había modificado su teoría de la RG.

Más allá de estas escaramuzas, el trabajo ER no tuvo gran impacto inmediato en la comunidad de físicos. El término “Agujero de Gusano” con que se asocia a la solución de ER, no aparece hasta un trabajo de J.A. Wheeler y C. Misner del año 1957 (Misner & Wheeler, 1957). Ellos, inspirados en un trabajo de H. Weyl de 1928 (quien llamaba a los agujeros de

---

<sup>2</sup> La aparición de la nota periodística dedicada al trabajo EPR provocó un fuerte enojo de Einstein con Podolsky, pues el artículo en el *Physical Review* apareció 11 días después de la nota en el NYT.

gusano, “tubos unidimensionales” [1928, p. 65]), introducen la idea de que “los campos electromagnéticos están atrapados en los ‘huecos de gusano’ de un espacio múltiplemente conexo” (Misner & Wheeler, 1957, p. 525).

Algunos autores ven a un trabajo de L. Flamm (1916) como la primera referencia a un agujero de gusano en el contexto de la RG. Revisando tal trabajo no es obvia tal mención; a lo sumo lo que se puede interpretar del trabajo de Flamm es la idea de lo que hoy llamaríamos “un agujero blanco”, pero ese enfoque escapa a nuestro objetivo.

## Conclusiones

En el trabajo “Física y realidad”, ya mencionado, Einstein dedica extensos párrafos a los dos trabajos del '35. Respecto al ER dice:

Aquí, nuevamente el intento de construir partículas a partir de la teoría de campos, aparentemente conduce a singularidades. Aquí el esfuerzo ha sido realizado para vencer este defecto mediante la introducción de nuevas variables de campo y mediante la elaboración de un sistema de ecuaciones de campo ampliado. Recientemente, sin embargo, yo descubrí con el doctor Rosen, que la arriba mencionada simple combinación de las ecuaciones de campo de la gravedad y la electricidad produce soluciones centralmente simétricas las cuales pueden representarse como libre de singularidades. . . . De este modo parece posible lograr para la materia y sus interacciones una teoría de campos pura libre de hipótesis adicionales. (Einstein, 1936, p. 371)

Hay autores que ven en ambos trabajos, el EPR y el ER, una búsqueda de auto consistencia de ambas teorías: la RG y la mecánica cuántica. En cuanto a la mecánica cuántica, el trabajo EPR habla por sí sólo, pero lo que no se explicita en el trabajo ER, es el tema de las ecuaciones de movimiento en la teoría del campo gravitatorio.

El problema del movimiento estuvo en el centro de las preocupaciones de Einstein y de varios científicos desde el origen de la teoría de la relatividad general. En su formulación inicial había supuesto que las leyes del movimiento son independientes de las ecuaciones de campo que dan cuenta de las interacciones entre los cuerpos. Quizás valga la pena recordar que Eddington llegó a la conclusión de que, para las partículas materiales consideradas como singularidades del campo, las leyes de movimiento de las singularidades deben estar contenidas en las ecuaciones de campo (Eddington, 1918). En particular, el problema de los dos cuerpos fue una

parte importante en este tema. Sin embargo, hubo que esperar hasta 1938, para que el tema de las ecuaciones de movimiento se esclareciera un poco más. Ese trabajo que tiene por autores a A. Einstein, L. Infeld y B. Hoffmann, fue publicado bajo el título “The gravitation field equation and the problem of motion” (Einstein, Infeld, & Hoffmann, 1938).

## Referencias

- Curiel, E. (2019). Singularities and black holes. En E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2019 edition). <https://plato.stanford.edu/archives/spr2019/entries/spacetime-singularities/>
- Eddington, A. S. (1918). *Report on the relativity theory of gravitation*. Londres: Fleetway Press.
- Einstein, A. (1922). Bemerkung zu der Abhandlung von E. Trefftz: ‘Das statische Gravitationsfeld zweier Massenpunkte in der Einsteinschen Theorie’. *Sitzungsberichte der Preußischen Akademie der Wissenschaften Physikalisch-Mathematische Klasse*, 448-449.
- Einstein, A. (1936). Physics and reality. *Journal of the Franklin Institute*, 221(3), 349-382.
- Einstein, A., Infeld, L., & Hoffmann, B. (1938). The gravitational equations and the problem of motion. *Annals of Mathematics*, 39(1), 65-100.
- Einstein, A, Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical review*, 47(10), 777-780.
- Einstein, A, & Rosen, N. (1935). The particle problem in the general theory of relativity. *Physical Review*, 48(1), 73.
- Flamm, L. (1916). Beiträge zur Einsteinschen Gravitationstheorie. *Physikalische Zeitschrift* 17, 448-454.
- Gibbons G. (2015): Editorial note to: Ludwig Flamm, Contribution to Einstein’s theory of gravitation. *General Relativity and Gravitation*, 47, 71.

- Havas, P. (1989). The early history of the “problem of motion” in general relativity. En D. Howard & J. Stachel (Eds.), *Einstein and the history of general relativity* (pp. 234-276). Boston: Birkhäuser.
- Havas, P. (1993). The general-relativistic two-body problem and the Einstein-Silberstein controversy. En J. Earman, M. Janssen, & J. D. Norton (Eds.), *The attraction of gravitation: New studies in the history of general relativity* (vol. 1, pp. 88-125). Springer Science & Business Media.
- Langevin, P. (1908). The relations of physics of electrons to other branches of science. En H. J. Rogers (Ed.), *International congress of arts and science (Vol. VII): Physics and chemistry* (pp. 121-156). London; New York: University Alliance. (Obra original presentada en 1904)
- Maldacena, J., & Susskind, L. (2013). Cool horizons for entangled black holes. *Fortschritte der Physik*, 61(9), 781-811.
- Misner, C. W., & Wheeler, J. A. (1957). Classical physics as geometry. *Annals of Physics*, 2(6), 525-603.
- Poincaré, H. (1917). *La science et l'hypothèse*. Paris: Flammarion. (Obra original publicada en 1906)
- Silberstein, L. (1936). Two-centers solution of the gravitational field equations, and the need for a reformed theory of matter. *Physical Review*, 49(3), 268.
- Trefftz, E. (1922). Das statische Gravitationsfeld zweier Massenpunkte in der Einsteinschen Theorie. *Mathematische Annalen*, 86, 317-326.
- Van Raamsdonk, M. (2010). Building up spacetime with quantum entanglement. *General Relativity and Gravitation*, 42(10), 2323-2329.
- Weyl, H. (1928) *Philosophie der Mathematik und Naturwissenschaft*. En A. Bäumler, & M. Schröter, *Handbuch der Philosophie*. Munich: Leibniz Verlag.



## Aportes al debate sobre la reducción de la química a la mecánica cuántica

Esteban Guillermo Szigety\*

La mecánica cuántica surge en las primeras décadas del siglo XX y se presenta como la condensación de una serie de leyes y experiencias que se realizaron a nivel de las partículas atómicas y subatómicas. Su formalismo comenzó a utilizarse y mostrar una gran capacidad predictiva en problemas relacionados no sólo con los átomos, sino también con las moléculas. Esta teoría despertó en los científicos la esperanza de subordinar otros temas de la física a la mecánica cuántica, y de pronto disciplinas científicas que parecían independientes comenzaron a ser pensadas nuevamente en base a estos nuevos fundamentos. De éstas, la que más parece haber perdido su autonomía es la química, la cual comenzó a ser vista como una disciplina puramente fenomenológica o aplicada. La causa de esto reside en que la estructura molecular es tan fundamental para la química, que comprender sus leyes implica de alguna manera tener una teoría completa de la química.

Entre los filósofos de la ciencia la idea de reducción de la química a la mecánica cuántica ha tenido y continúa teniendo sus partidarios acérrimos: “Hay muchos ejemplos de reducciones que se han logrado. Por ejemplo, una gran parte de la química clásica se ha reducido a la física atómica; y la teoría clásica del calor se ha reducido a la mecánica estadística” (Kemeny & Oppenheim, 1956, p. 7).

En este artículo seminal sobre la reducción interteórica en la física, los autores dejan claramente expresado que ya no hay nada que discutir al respecto. Bader (2003) es un ejemplo más actual de esta postura, fortaleciendo la idea de la física como una ciencia fundamental. En su artículo se enfrenta a las posturas que consideran el orbital como un concepto acuñado por la química, y sobre el cual no existen en la mecánica cuántica referencias teóricas. Bader defiende la mecánica cuántica (desde ahora MC) como la física subyacente de los átomos y moléculas: para él es posible hablar de una física de los enlaces químicos. De alguna manera estos

\* Facultad de Ingeniería, Universidad Nacional de Mar del Plata. Buenos Aires, Argentina. esteszige@gmail.com

autores están tomando un compromiso fisicalista reduccionista, es decir, suponen que las leyes y hechos de la física determinan todas las leyes y hechos de la química.

La pregunta sobre si la química es una rama de la física es demasiado amplia para ser contestada en este trabajo. Con el objetivo de focalizar la atención en un debate de menor extensión, analizaremos las posturas alrededor de una temática más específica: ¿se pueden reducir los átomos y las moléculas, tratados por la química, a la mecánica cuántica? (Weisberg, Needham, & Hendry, 2019). Tanto la química como MC tienen distintos niveles ontológicos y explicativos, por lo tanto, para lograr una reducción adecuada para el caso, ambos niveles deben poder encontrar un modelo de reducción que los vincule adecuadamente.

Existe una reducción tradicional, basada en los escritos de Nagel; más específicamente, en el capítulo 11 de *La estructura de la ciencia* (1961). Para el autor la química admitiría una reducción a la mecánica cuántica del tipo que él denomina heterogénea. Reconoce que hay términos de la química como la valencia, el enlace o las reacciones químicas que no aparecen en la MC, y por lo tanto es necesario conectarlos mediante leyes-puente (bicondicionales). El problema radica en la dificultad de encontrar estas leyes que permitan unificar ambas teorías, correlacionando, por ejemplo, la valencia con los orbitales, y demostrando así que ambos términos refieren a la misma cosa. El reduccionismo tradicional es básicamente un modelo deductivista, y la posibilidad de alcanzarse en el caso de la química presenta una gran complejidad. No es sólo una simple deducción lógica; la reducción que Nagel plantea es poco realista e inaplicable para este caso, debido a los requisitos lógicos que supone. Los filósofos de la ciencia que han tratado de mostrar la irreductibilidad de la química han hecho hincapié en este punto.

Sobre la reducción interteórica de Nagel diremos que es una postura explícitamente jerárquica y acumulativa de la ciencia. El mismo Nagel, en su afán por alcanzar estos objetivos, no logró ver que la reducción interteórica, en este caso, es impracticable. Considero que las ideas de Nagel podrían ser retomadas si en un futuro próximo se produjese un desplazamiento de la MC por otra teoría que permita alcanzar la tan deseada reducción interteórica de la química. A finales del siglo XX y principios del XXI, distintos filósofos de la ciencia han concluido lo mismo: que, tal cual como conocemos a la MC, es imposible considerar la reducción de la química. A continuación, profundizaremos las posturas más significativas.

Autores como Scerri (1997, 2000) han presentado fundamentos cabales contra la reducción de la teoría química de átomos y moléculas a la MC. Para este autor, la ecuación de Schrödinger describe en términos bastante precisos un gran número de fenómenos físicos, y puede especificarse completamente por unas pocas cantidades, como las cargas y las masas de los núcleos y de los electrones, así como la constante de Planck. Se puede resolver con precisión para un pequeño número de partículas (átomos aislados y moléculas pequeñas), y concuerda en detalle minucioso con los experimentos. Sin embargo, como señalan Laughlin y Pines (2000), cuando la cantidad de partículas supera el número diez dejamos atrás el terreno de lo exacto y lo preciso. Es posible realizar cálculos aproximados para sistemas más grandes, que brindan conocimiento de los tamaños atómicos, la elasticidad y demás propiedades macroscópicas. El uso de técnicas de aproximación significa que ya no son deducciones de los primeros principios o de la teoría fundamental, sino que requieren experimentación. Es decir, ingresar detalles específicos de la química, detalles de la ciencia que se quiere reducir. Lo que esto indica es una ruptura del ideal reduccionista de derivar explicaciones de una gran cantidad de fenómenos a partir de unas pocas ecuaciones o leyes simples.

Scerri no es el único que se opone al reduccionismo de la química a la MC. Trabajos anteriores como el de Primas (1981) y el de Woolley (1978) argumentan que, aunque la MC pueda ser muy esclarecedora sobre el mundo de los átomos, no por ello implica la reducción de la química. Sostienen estos autores que existen dificultades conceptuales para acomodar el conocimiento químico de las moléculas a la MC. Uno de esos conceptos que no ha encontrado reducción posible son los de enlace químico, unión y estructura molecular (Primas, 1981), los cuales no pueden expresarse a no ser a nivel químico. Para Woolley (1978) ese último concepto, el de estructura molecular, está ausente en el nivel mecánico-cuántico. Argumenta que muchas estructuras complejas no requieren de la aproximación Born-Oppenheimer para ser calculadas, y hasta se pueden obtener resultados más precisos omitiendo por completo dicha aproximación. Podemos concluir, entonces, que estos autores no niegan que exista una dependencia ontológica entre ambas teorías, pero alcanzar una reducción interteórica resulta imposible. Una vez aceptada la inviabilidad de la reducción, veamos posibles soluciones al problema.

Mario Bunge desarrolla en varios de sus artículos (1982, 2003) el debate de la reducción de la química a la física. En su artículo de 1982 lleva su

análisis a la química cuántica e intenta mostrar que lo que es válido para la física cuántica lo es también para la química cuántica, entonces se dará por exitosa la reducción. Adecuadamente demostrado, Bunge (1982, apéndice I) deja en claro que a través de las premisas de la química molecular se pueden obtener variables macro de la química, como las constantes de equilibrio de una reacción química. En la química clásica, estas constantes son consideradas parámetros experimentales; por medio de la mecánica cuántica se convierten en parámetros teóricos de la química cuántica. La interpretación queda ceñida a procesos de dispersión de colisiones inelásticas. Bunge agudiza el ojo escudriñador y logra detectar que este aparente éxito se ve empañado por el uso inevitable en los pasos deductivos de un supuesto perteneciente a la química clásica (macroquímica): la ecuación de velocidad de la teoría cinética química. Esta ecuación empírica de la velocidad de formación del producto de reacción no se deduce, sino que es postulada cuando la constante de velocidad se calcula en términos teórico-cuánticos. Por lo tanto, la química molecular no se deduce directamente de la teoría cuántica. Bunge escribe:

A primera vista, la química está incluida en la física porque los sistemas químicos parecen constituir una clase especial de sistemas físicos. Pero esta impresión es errónea, ya que lo que es físico acerca de un sistema químico son sus componentes más que el sistema en sí, que posee propiedades emergentes (aunque explicables) además de las propiedades físicas. (Bunge, 1982, p. 210)

La noción de emergentismo que Bunge introduce en esta cita es claramente incompatible con un reduccionismo fuerte al estilo de Nagel. Al no poder lograr una descripción completa desde la física, Bunge está aceptando que la MC continúa restringiendo el comportamiento de los sistemas de átomos y moléculas, aunque no pueda determinarlos completamente. La química pasa a ser una disciplina que estudia hechos autónomos de la física, pero que están asociados a sistemas físicos de orden superior.

Por otro lado, la relación entre la química y la MC ha sido tomada por algunos autores como caso paradigmático de verdadera superveniencia. Tomando la definición que nos ofrece la *Stanford Encyclopedia of Philosophy* (Weisberg, Needham, & Hendry, 2019), la superveniencia es una condición necesaria para alcanzar la reducción. Por ejemplo, consideremos el caso de dos sistemas macroscópicos con propiedades idénticas que se han construido a partir de componentes microscópicos idénticos. La

observación de propiedades macroscópicas idénticas en cualquiera de los dos sistemas no necesariamente implica identidad de propiedades a nivel microscópico. Este tipo de relación no admite una ida y vuelta de un nivel de reducción al otro. Esta relación condicional es la denominada superveniencia. Si la relación entre la física y la química fuera de este tipo, estaríamos hablando de una reducción de estilo débil. Este argumento ha sido ampliamente utilizado en la filosofía de la ciencia y de la mente como una maniobra de rescate en la discusión cuando no se produce reducibilidad interteórica en las ciencias especiales, ya que se considera que la superveniencia no garantiza las consecuencias epistemológicas o explicativas que necesariamente se derivarían incluso de una fuerte dependencia ontológica entre dos niveles descriptivos diferentes.

Existen dos casos interesantes a plantear en esta discusión. El primero es el caso de los isómeros estructurales (moléculas con los mismos átomos, pero con diferentes estructuras moleculares) en los cuales su similitud no implica observar las mismas propiedades. El problema surge cuando consideramos la descripción mecánico-cuántica de isómeros como el dimetiléter ( $\text{CH}_3\text{OCH}_3$ ) y el etanol ( $\text{C}_2\text{H}_5\text{OH}$ ). Ambos comparten un mismo hamiltoniano ( $\text{C}_2\text{H}_6\text{O}$ ), es decir, la misma descripción mecánico-cuántica de sus estados físicos. Sin embargo, sus moléculas tienen comportamientos muy diferentes. El segundo caso, muy similar al anterior es el de los isómeros ópticos o enantiómeros. Estos compuestos comparten casi todas sus propiedades químicas y físicas, sólo se diferencian entre sí por el tipo de interacción que manifiestan con la luz polarizada (Fortin & González, 2013). En ambos casos, las distancias interatómicas son iguales entre los átomos de las moléculas, de modo tal que el hamiltoniano es exactamente el mismo. En consecuencia, la mecánica cuántica brinda la misma descripción para dos estructuras químicas que pueden efectivamente ser diferenciadas en la práctica a través de sus propiedades ópticas, punto de ebullición y otras características macroscópicas. Aunque Hendry (2006) argumenta que una reducción del tipo emergentista es la manera correcta de interpretar estos casos, en mi opinión una superveniencia parece ser la forma más acertada. Dado que ambos ejemplos presentan una asimetría notoria entre los niveles fundamental y especial, no hay forma de reconciliar esta relación si no es mediante este tipo de reducción débil. Ambos ejemplos rompen cualquier relación de simetría, ya que el nivel superior no puede realizar una explicación descendente. Por el momento, para aceptar una reducción fuerte deberían poder explicarse estos casos asi-

métricos y argumentar la inadecuación de la superveniencia en este caso. Creo que estos casos emblemáticos son consecuencia de la ausencia de un instrumento para la detección de variaciones cuantificables en la naturaleza micro de sus componentes. Se desconoce si la escalada de desarrollo tecnológico que la física y la química pueden lograr en un futuro permita resolver los problemas de medición a nivel micro.

Un autor que niega la reducción dura y se inclina por la superveniencia para destrabar el debate de la reducción de las disciplinas especiales es Kim. De él citamos el siguiente extracto:

La superveniencia no es un tipo de relación de dependencia, no es una relación que pueda colocarse junto con la dependencia causal, la dependencia reductiva, la dependencia mereológica, la dependencia basada en la definición o la vinculación, y similares. Más bien, cualquiera de estas relaciones de dependencia puede generar la covarianza requerida de propiedades y, por lo tanto, calificarse como una relación de superveniencia. (Kim, 1998, p. 14)

Kim dice acertadamente que la noción de superveniencia no pretende ser una solución completa para las disciplinas especiales. En un principio, la función de la superveniencia es establecer un patrón de covariación entre la mente y el cuerpo, señalando una dependencia asimétrica entre ellos. Se trata de una relación fenomenológica y explicativa, que ayuda a expresar el problema sin pretender darle una solución. A diferencia de otras tesis, la afirmación de la superveniencia parece decir algo positivo acerca de la relación entre las propiedades físicas y la química. Para ello Kim establece una tesis fisicalista no-reduccionista mediante la superveniencia. Por otro lado, su más alto logro es establecer una dependencia asimétrica entre la mente y el cuerpo, lo cual equivale a un fisicalismo mínimo que se puede aplicar a los ejemplos de los isómeros.

Creo que la idea de la superveniencia fue aceptada por aquellos influenciados por el argumento de la realización múltiple en el debate mente-cuerpo (Fodor, 1974). La tesis de las propiedades mentales como supervenientes de las propiedades físicas parecía ajustarse muy bien a los requisitos metafísicos del funcionalismo: prometía dar un sentido claro a la primacía del dominio físico y sus leyes, y así cumplir con los compromisos físicos de la mayoría de los funcionalistas como Kim (1998). Esta perspectiva es positiva en cuanto que no implica un reduccionismo físico fuerte y permite proteger la autonomía de la química. Cabría preguntarse

si la irreductibilidad interteórica que ha apoyado dentro de la filosofía de la química se transfiere fácilmente a los debates dentro de la filosofía de la mente, dando pie a un entrecruzamiento con corrientes fisicalistas no-reduccionistas de este último tiempo (Kim, 1998). La perspectiva de Fodor es claramente una posición intermedia entre el reduccionismo fuerte y la superveniencia, que parece ofrecer una forma de modelizar la reducción entre la mecánica cuántica y la química para los problemas actuales.

En definitiva, el anhelo de la filosofía de la química es lograr una descripción filosófica general de cómo su disciplina se relaciona con la física básica. Un químico de una forma u otra busca reconfortarse en la idea de estar haciendo legalmente ciencia, independientemente de lo que ocurra en un nivel más fundamental; pudiendo formular leyes y explicaciones con su propio vocabulario. A fin de cuentas, esto es lo que hacen constantemente los científicos especiales, pero necesitan de un sustento filosófico para ello. Está claro que no hay insumos suficientes, ni teóricos ni tecnológicos, para explicar la estructura y las funciones que vinculan una y otra teoría. Las nociones débiles como la superveniencia y la realización múltiple son los modelos más viables en el debate actual de la reducción de la química; concluyendo así que la reducción interteórica entre dos teorías tan fronterizas no está actualmente disponible.

## Referencias

- Bader, R. F. W. (2003). Letter to the editor: Quantum mechanics, or orbitals? *International Journal of Quantum Chemistry*, 94(3), 173-177.
- Bunge, M. (1982). Is chemistry a branch of physics? *Zeitschrift für allgemeine Wissenschaftstheorie*, 13(2), 209-223.
- Bunge, M. (2003). *Emergence and convergence: Qualitative novelty and the unity of knowledge*. Toronto: University of Toronto Press.
- Fodor, J. A. (1974). Special sciences: Or the disunity of science as a working hypothesis. *Synthese*, 28, 97-115.
- Fortin, S., & González, J. C. M. (2013). La relación entre química y física: Isomerismo óptico y la paradoja de Hund. *Revista Colombiana de Filosofía de la Ciencia*, 13(26), 199-224.

- Hendry, R. F. (2006). Is there downward causation in chemistry? En D. Baird, E. Scerri, & L. McIntyre (Eds.), *Philosophy of chemistry: Synthesis of a new discipline* (pp. 173-189). Dordrecht: Springer.
- Kemeny, J. G. & Oppenheim, P. (1956). On reduction. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 7(1-2), 6-19.
- Kim, J. (1998). *Mind in a physical world: An essay on the mind-body problem and mental causation*. Cambridge: MIT Press & Bradford Books.
- Laughlin, R. B., & Pines, D. (2000). The theory of everything. *Proceedings of the National Academy of Sciences of the United States of America*, 97(1), 28-31
- Nagel, E. (1961). *The structure of science*. London: Routledge & Kegan Paul.
- Primas, H. (1981). *Chemistry, quantum mechanics and reductionism: Perspectives in theoretical chemistry*. New York: Springer-Verlag.
- Scerri, E. R. & McIntyre, L. (1997). The case for the philosophy of chemistry. *Synthese*, 111(3), 213-232.
- Scerri, E. (2000). Have orbitals really been observed? *Journal of Chemical Education*, 77(11), 1492-1494.
- Weisberg, M., Needham, P., & Hendry, R. (2019). Philosophy of chemistry. En E. Zalta, *The Stanford encyclopedia of philosophy* (Spring 2019 Edition). Recuperado de <https://plato.stanford.edu/archives/spr2019/entries/chemistry/>.
- Woolley, R. G. (1978). Must a molecule have a shape? *Journal of the American Chemical Society*, 100(4), 1073-1078.





---

## 6. Historia y filosofía de las ciencias biológicas







# El aporte de la ingeniería de ecosistemas y la construcción de nicho a las explicaciones paleontológicas

Susana Gisela Lamas\*

Raúl Montero\*

Vicente Dressino\*

## Introducción

A finales del siglo XX y principios del XXI, debido en gran parte al avance tecnológico, se desarrollaron novedosas teorías en la biología. Uno de los descubrimientos que mayor impacto causó fue el Proyecto Genoma cuyos resultados mostraron conclusiones contrarias a las predicciones de la teoría evolutiva dominante, la llamada Teoría Sintética de la Evolución (TSE). Porque este proyecto reveló, entre otras cosas, que la evolución consistió más en la conservación de los genes que en su variabilidad; que el postulado básico de la genética de un gen-una función ya no se podía sostener en base a la evidencia experimental, etc. Por tanto, se reconoció la necesidad de nuevos postulados para poder explicar la inmensa variabilidad biológica y su heredabilidad. De esta manera comienzan a surgir, en biología evolutiva, nuevas propuestas basadas en supuestos novedosos o en teorías previas, aunque reformuladas; desarrollándose, así, disciplinas o teorías tan diversas como la epigenética, evo-devo (por sus siglas en inglés de *evolution and development*), eco-evo-devo (por sus siglas en inglés de *ecology, evolution and development*), la construcción de nicho, etc. que modificaron profundamente las explicaciones evolutivas. De este modo, los principios explicativos de la anterior teoría evolutiva (TSE), fueron transformados hasta tal punto que un grupo de biólogos y filósofos propusieron una nueva teoría para dar cuenta del fenómeno evolutivo a la que denominaron *síntesis evolutiva extendida* (SEE).

Al analizar ambas teorías podemos observar varias diferencias, pero, a los fines de este trabajo, nos interesa focalizar en que la síntesis moderna

\* Laboratorio de Investigación en Ontogenia y Adaptación (LINO). Facultad de Ciencias Naturales y Museo. Universidad Nacional de la Plata, Argentina.  
sglamas@fncym.unlp.edu.ar

considera a la selección natural como el principal agente de cambio evolutivo y a los organismos o poblaciones como entes pasivos. En la síntesis extendida, en cambio, el organismo es un protagonista activo de su propia evolución, pudiendo modificar su propio proceso evolutivo o el de otras especies. La teoría de construcción de nicho constituye un ejemplo de este nuevo enfoque.

Además, uno de los aportes más relevantes que ha tenido la SEE, ha sido la propuesta de diversos sistemas de herencia, a saber, el genético, el epigenético, el sociocultural o etológico y el ecológico. Esta perspectiva complejiza mucho más el modo de conceptualizar la evolución y cada uno de estos sistemas de herencia supondrá disciplinas, teorías, metodologías y técnicas diferentes. En este trabajo nos focalizaremos sólo en la importancia que implica la herencia ecológica para las explicaciones paleontológicas. De los sistemas de herencia antes expuestos, éste parecería ser el más cuestionado y cuestionable. Sin embargo, al examinar cuidadosamente sus supuestos, consideramos que puede ser de gran utilidad en ciertas disciplinas biológicas, entre ellas la paleontología; disciplina en la cual prácticamente no se ha hecho mención a la teoría de construcción de nicho. De allí nuestro interés en determinar los posibles aportes que la construcción de nicho puede tener al considerar dimensiones novedosas en los estudios de casos de la paleontología.

## **La teoría de la construcción de nicho y la ingeniería de ecosistemas**

La noción de construcción de nicho se encuentra estrechamente relacionada con los conceptos de *herencia ecológica* y de *ingeniería de ecosistemas*. Definiremos a la *construcción del nicho* como el proceso mediante el cual los organismos, a través de su metabolismo, sus actividades y sus elecciones, modifican sus propios nichos. Una consecuencia de ello es que también pueden afectar los nichos de otras especies (Odling-Smee, Erwin, Palkovacs, Feldman, & Laland, 2013). La construcción de nicho puede resultar en cambios que influyan en una o más presiones del entorno externo y puede ocurrir a través de la perturbación física del medio ambiente o mediante la reubicación a un nuevo entorno. Por lo tanto, puede tener efectos tanto positivos como negativos sobre sí mismo o sobre otros organismos. Es importante aclarar que la construcción de nicho comienza con modificaciones fisiológicas respetando el mantenimiento de la homeostasis (el equilibrio interno) del organismo que modifica. En este sentido,

cabe destacar que no resulta posible construirse un nicho adecuado para el organismo y para su posible descendencia, si ese nicho implicara una modificación o alteración significativa de su fisiología. También es preciso señalar que esta última teoría no descarta el accionar de la selección natural, sino que la considera como un agente secundario en el proceso evolutivo.

Por otro lado, siguiendo a Jones, Lawton y Shachak (1994), entenderemos *ingeniería de ecosistemas* como la creación, destrucción o modificación de hábitats y/o modulación de la disponibilidad de recursos para otras especies por parte de organismos. De este modo cuando hablamos de ingeniería de ecosistemas nos referimos sólo a cambios físicos en los materiales bióticos o abióticos (por ejemplo, la construcción de un hormiguero, de un nido, etc.) que influyen en otras especies. Así, la ingeniería de ecosistemas puede equipararse con el componente de construcción de nicho que altera el medio ambiente repercutiendo sobre otras especies; porque para que exista construcción de nicho debe haber modificación de los aspectos físicos del entorno. Sin embargo, podría pensarse en una modificación realizada al medio que no influyera ni positiva ni negativamente en organismos de otra especie. Un ejemplo de este último caso podría ser el comportamiento de desove de las tortugas marinas (*Chelonioidea*, *Sauropsida*, *Testudines*) que modifican el ambiente para desovar realizando un pozo de cierta profundidad en la playa cerca del mar. La actividad de realización del hueco para la deposición de los huevos podría no repercutir significativamente en organismos de otras especies.

De esta manera, podemos concluir que la diferencia entre construcción de nicho e ingeniería de ecosistemas es de tipo metodológica. Porque la construcción de nicho supone una perspectiva vinculada a la modificación del medio biótico y abiótico y su relación y la influencia que tiene sobre la propia especie. En tanto que la ingeniería de ecosistemas, como su nombre lo indica, tiene una mirada ecológica; preguntándose sobre el efecto que las modificaciones anteriores tienen sobre las otras especies. Podríamos afirmar que, en general refieren al mismo fenómeno, aunque con enfoques distintos.

Finalmente, el tercer concepto es el de *herencia ecológica*. Este tipo de herencia se da, a través de un entorno externo, cuando hay presiones del medio que fueron previamente modificadas por otros organismos (Odling-Smee et al., 2013). Es decir, la herencia ecológica podría pensarse como los legados ambientales realizados por los organismos que pueden

afectar a otros organismos de su misma especie o de otras especies que no se relacionan con quienes estén realizando la modificación pero que comparten el mismo ecosistema. También habrá herencia ecológica cuando exista un acto por parte de organismos que conduzca a un cambio en la composición de especies de la comunidad ecológica local. Es decir, tanto la construcción de nicho como la ingeniería de ecosistemas constituyen procesos de herencia ecológica.

Podemos apreciar cómo estos tres conceptos, a pesar de estar estrechamente interconectados, pueden diferenciarse. Y, llevar a cabo esta distinción semántica, nos permitirá reconocer los distintos elementos presentes en nuestros casos de estudio.

### **La ingeniería de ecosistemas, la construcción de nicho y la paleontología**

En esta sección analizaremos dos estudios de caso provenientes de la paleontología. Pero antes de comenzar resulta necesario aclarar que en esta disciplina es muy difícil determinar si ciertos comportamientos terminaron generando o no mayores presiones en el entorno en los organismos o poblaciones. Lo único que podemos inferir son las modificaciones llevadas a cabo en su entorno y la consideración de los cambios materiales con carácter heredable generados por los organismos; es decir, la ingeniería de ecosistemas. No obstante, estimamos que esta perspectiva supone un aporte para las explicaciones paleontológicas. Podríamos ofrecer diversos procesos actuales de ingeniería de ecosistemas, por ejemplo, los diques hechos por los castores (*Castor canadensis*) en el sur de Argentina que cambiaron drásticamente las condiciones para muchas especies.

Otro ejemplo lo constituyen los nidos de aves en desuso que son reutilizados por otras aves que no construyen sus propios nidos y que son un importante recurso disponible que muchas especies aprovechan. Para citar sólo un caso documentado en la Provincia de Córdoba, Argentina, en una misma área, fueron reutilizados los nidos de horneros, cotorras, pájaros carpinteros, etc. Los nidos del hornero (*Furnarius rufus*), por ejemplo, fueron reutilizados por doce especies diferentes, a saber, el Gorrión (*Passer domesticus*), el picabuey (*Machetornis rixosa*), el coludito copetón (*Leptasthenura platensis*), entre otras (Salvador, 2012).

El último caso que señalaremos de ingeniería de ecosistemas es el de las maras patagónicas (*Dolichotis patagonum*) que hacen cuevas que son re-

utilizadas por diversos vertebrados, entre los que podemos enumerar la lechucita de las vizcacheras (*Athene cunicularia*), la liebre europea (*Lepus europaeus*), el peludo (*Chaetophractus villosus*), el zorrino (*Conepatus chinga*), la martineta (*Eudromia elegans*), el zorro gris (*Lycalopex gymnocercus*) y la yará ñata (*Bothrops ammodytoides*) (Roldán & Udrizar Sauthier, 2016).

Con la ayuda del actualismo, supuesto teórico fundamental de las explicaciones paleontológicas, podemos inferir que procesos de ingeniería de ecosistemas similares se han dado en ciertos casos del registro fósil cuyos indicios quedaron representados en las paleo cuevas. Para ello, nos basaremos en datos secundarios provenientes de dos paleo cuevas de la región pampeana en Argentina que datan del Pleistoceno (periodo que abarca desde 2,5 millones de años a 10.000 años antes del presente).

El primer caso hace referencia a la paleo cueva descrita por Cenizo, Soibelzon, & Magnussen Saffer (2015). Este ensamble fósil consta de una estructura –cueva– excavada en el sustrato por un armadillo de gran tamaño *Ringeletia simpsoni*. Esta obra ambiental fue reutilizada a posteriori por otros mamíferos de pequeño y mediano tamaño como, por ejemplo, los marsupiales *Thylophorops chapadmalensis* o el prociónido *Cyonasua lutaria* para resguardarse, alimentarse, defecar o depredar. Esto último puede inferirse a partir de la presencia de coprolitos (fecas fosilizadas) y restos óseos con marcas de la acción abrasiva de los jugos gástricos, producto de los procesos digestivos típicos de carnívoros. Los restos fósiles de pequeños vertebrados además de otros mamíferos de igual o menor tamaño como roedores caviomorfos, encontrados dan evidencia de la actividad metabólica de estos animales dentro de la cueva. Estos datos nos permiten plantear que las cuevas previamente excavadas fueron reocupadas cierto tiempo después por otro grupo específico diferente a la especie que originariamente la construyó. Por lo tanto, esto es un caso paleontológico de herencia ecológica enmarcado en una construcción de nicho conjunta entre varias especies.

El segundo ejemplo corresponde a una paleo cueva hallada en Mar del Plata (Soibelzon, Pomi, Tonni, Rodríguez, & Dondas, 2009). Esta cueva probablemente fue realizada por la acción excavadora de un perezoso terrestre asignada al género *Scelidotherium* o a otro género conocido como *Glossotherium*. La forma de la cueva y su tamaño dan señales de que fueron realizadas por este animal si la comparamos con otros registros de la misma índole (Vizcaíno, Zárate, Bargo & Dondas, 2001). En el interior de la cueva fueron hallados restos de tres individuos de oso extinto de la especie

*Arctotherium angustidens*. Dos de ellos juveniles y un adulto, sin ninguna marca de depredación y parte de los huesos; esto hace suponer que quedaron en la misma posición que en el momento de su fallecimiento. Esto nos permite plantear la hipótesis de que los animales murieron dentro de la cueva debido a algún suceso de corrimientos de tierras que tapó su entrada; ya que en el momento de su hallazgo la paleo cueva se encontraba intacta. Con ello concluimos la acción por parte de los úrsidos de ocupar las cuevas construidas por otros coetáneos, como los perezosos terrestres, ambos taxones representantes de la megafauna extinta pampeana.

## Conclusión

Para analizar el proceso de herencia ecológica se propusieron las nociones de *construcción del nicho* y de *ingeniería de ecosistemas*, concluyendo que esta última es más satisfactoria para las investigaciones paleontológicas. Cabe destacar que en la bibliografía especializada se utiliza la noción de *tafoconosis* –conjunto de restos y señales de organismos que están enterrados en un área delimitada– y de *ensamble ecológico* –como un conjunto de especies que comparten un lugar. Pero es casi nula la perspectiva teórica de herencia ecológica que supone la ingeniería de ecosistemas.

Otra consideración, estrechamente relacionada con la anterior, es la revalorización de la información de las paleo cuevas. Así, restos fósiles a los que se les otorgó una importancia secundaria en la bibliografía consultada, pueden ser reinterpretados al reconocer que ofrecen información acerca del uso y la reutilización de las obras ambientales, mostrando una interrelación previamente no considerada entre especies. Esto se puede inferir claramente en el caso de la segunda paleo cueva analizada.

Finalmente, concluimos que la ingeniería de ecosistemas y la construcción de nicho suponen dos enfoques metodológicos diferentes al fenómeno de modificación del entorno. Por tanto, consideramos que el análisis de la ingeniería de ecosistemas representa un aporte tanto teórico como metodológico para las explicaciones paleontológicas.

## Agradecimientos

Queremos agradecer al Dr. Esteban Soibelzon por sus aportes y comentarios que fueron de gran utilidad para este trabajo. Esta investigación fue llevada a cabo gracias al Subsidio Automático del Proyecto de Incentivos a la Investigación 11/N907 de la Universidad Nacional de La Plata.

## Referencias

- Cenizo, M., Soibelzon, E., & M. Magnussen Saffer. (2015). Mammalian predator-prey relationships and of burrows in the Pliocene of the Pampean Region (Argentina): new ichnological and taphonomic evidence. *Historical Biology*, 28, 1026-1040.
- Jones, C.G., Lawton, J.H., & Shachak, M. (1994). Organisms as ecosystem engineers. *Oikos*, 69, 373-386.
- Odling-Smee, F.J., Erwin, D.H., Palkovacs, E.P., Feldman, M.W., & Laland K.N. (2013). Niche construction theory: A practical guide for ecologists.. *The Quarterly Review of Biology*, 88, 3-28.
- Roldán, V.A., & Udrizar Sauthier, D. (2016). Madrigueras de *Dolichotis patagonum* como recurso para otros vertebrados en Península Valdés. *Mastozoología Neotropical*, 23, 515-520.
- Salvador, S. (2012). Reutilización de nidos por aves en el área central de Córdoba, Argentina. *Nótulas Faunísticas*, 91, 1-9.
- Soibelzon, L. H., Pomi, L. H., Tonni, E. P., Rodríguez, S., & Dondas, A. (2009). First report of a South American short-faced bears' den (*Arctotherium angustidens*): Palaeobiological and palaeoecological implications. *Alcheringa: An Australasian Journal of Palaeontology*, 33(3), 211-222.
- Vizcaíno, S., M. Zárate, S. Bargo, & Donda, A. (2001). Pleistocene burrows in the Mar del Plata area (Argentina) and their probable builders. *Acta Paleontologica Polonica*, 46, 289-301.



# Transgénicos y agrotóxicos: el rol de la bioética en el sistema agroindustrial argentino

Nahuel Pallitto\*  
Guillermo Folguera\*

## Introducción

El presente trabajo parte de un reconocimiento, un diagnóstico y una pregunta. El reconocimiento se vincula con la adopción del Estado argentino de un modelo de producción agrario basado en un tipo de paquete tecnológico, el cual se ha consolidado en el país en las últimas dos décadas. Dicho paquete tecnológico combina el uso de organismos genéticamente modificados y agrotóxicos, pudiéndose reconocer su presencia en diferentes variedades de cultivos que se destinan tanto para el consumo interno como para su exportación.

El diagnóstico es que se trata de un modelo con fuertes impactos sociales y ambientales, entre los cuales pueden mencionarse la adopción de un modelo único de agricultura, el desplazamiento forzado de comunidades de sus territorios, la pérdida de biodiversidad, la contaminación de suelos y cuerpos de agua y los graves impactos en la salud de las personas que entran en contacto con los agrotóxicos. Además, resulta necesario destacar los problemas relacionados con la dependencia económica en el valor del precio de ciertos cultivos y el creciente monopolio por parte de unas pocas empresas multinacionales de la comercialización de las semillas y los insumos (Francese & Folguera, 2018).

La pregunta se vincula con el rol específico que ha asumido el campo de la bioética en las problemáticas sociales y ambientales asociadas al modelo agroindustrial argentino, en un contexto general en que fueron pocas las voces críticas al seno de los campos profesionales que señalaron los problemas sociales y ambientales vinculados con los transgénicos y los agrotóxicos. Por ello, en este trabajo procuraremos visibilizar algunas características del posicionamiento que en particular tuvo la bioética

\* Instituto de Filosofía “Dr. Alejandro Korn”, Facultad de Filosofía y Letras, Universidad de Buenos Aires – CONICET.  
nahuelpallitto@gmail.com

hegemónica al respecto en nuestro país en los últimos años, brindando elementos de comprensión que relacionan nuestra propuesta con aspectos epistémicos y epistemológicos de las áreas científicas involucradas.

### **La tecnociencia y su expansión acelerada**

La relación entre la ciencia y la tecnología ha sido descrita de múltiples maneras por la historia y la filosofía de la ciencia. Una de las caracterizaciones más recientes propone que una y la otra se encuentran actualmente imbricadas, sosteniendo con ello que resulta inadecuado concebir la una sin la otra (Echeverría, 2003; Linares, 2008). El término que se ha acuñado para describir este estado de situación es el de «tecnociencia» (Hottois, 1990/1991). Entre otros aspectos, la imbricación propuesta por el concepto de tecnociencia apunta a reconocer que ya no es posible analizar y comprender alguno de los polos de la relación sin tomar en cuenta el otro. Por ejemplo, en términos de racionalidades involucradas, se sugiere que la ciencia no se encuentra exclusivamente dominada por un tipo de racionalidad logocéntrica (justificada sobre la base de la obtención de conocimiento), sino que también le resulta propia e inextricable un tipo de racionalidad pragmático-utilitaria (justificada sobre la base de ofrecer intervenciones exitosas sobre nuestras realidades). De este modo, conocer y hacer conforman una díada que requiere ser pensada en su completitud e interdependencia.

A su vez, el concepto consigue traer a la mano un conjunto de distinciones que resultan sumamente valiosas para reflexionar en torno a las implicancias éticas y políticas de los desarrollos científicos y tecnológicos. Por un lado, invita a explicitar cuáles son los intereses y finalidades que se ponen en juego al momento de producir conocimiento y/o aplicarlo en la resolución de problemáticas de interés social y ambiental. Si conocer el mundo no es el único objetivo posible, entonces quienes se involucren con las prácticas científicas deben dejar en claro qué finalidades persiguen.

Por otro lado, la noción de tecnociencia permite visibilizar cuáles son las lógicas que subyacen a esas producciones y aplicaciones tecnocientíficas. En este último caso, cabe destacar que en las últimas décadas la ciencia y la tecnología han tendido a reproducir lógicas de tipo empresarial en las que predominan valores económicos y eficientistas, expresados en discursos publicitarios que prometen siempre mayores cuotas de bienestar y felicidad. En ese sentido, conocer y hacer se valoran fundamentalmente por

su capacidad de generar réditos económicos e innovaciones que puedan fácilmente ser introducidas en los mercados.

Por último, debe ser dicho que la diada conocer-hacer presenta tensiones de tipo epistémicas que comúnmente no son reconocidas. Con ello nos referimos a la omisión de saberes y a las simplificaciones teóricas que en nombre de la eficiencia y la eficacia suelen realizarse de los saberes científicos. Debido a que lo que se espera es que un producto de la ciencia tenga un éxito práctico (ya sea técnico o económico), las exigencias epistémicas muchas veces se flexibilizan al punto de hacer casi irreconocible los conocimientos que se ponen al servicio de una intervención. O, dicho de otro modo, se establece una suerte de “distancia epistémica” significativa entre los marcos teóricos que buscan describir nuestras realidades y los mismos marcos teóricos que son aplicados para resolver alguna problemática. Dicha “distancia” suele adoptar la forma de una simplificación. En este esquema la propia noción de riesgo es modificada, en tanto la misma depende de los elementos del conocimiento que sean reconocidos e incorporados en una evaluación determinada.

Las características asociadas a la tecnociencia, su expansión acelerada desde la década de 1970 y los riesgos que conlleva han contribuido a la emergencia y consolidación de lo que hoy en día se conoce como el campo de la bioética y que pasaremos a describir sucintamente a continuación.

### **La bioética y sus orígenes**

En parte como consecuencia de los efectos percibidos de la tecnociencia en expansión, en la década de 1970 se dio un cambio que modificó, de aquel tiempo a esta parte, la relación entre la ciencia, la tecnología y la ética. En aquellos años, surgió en Estados Unidos el campo de la bioética, extendido luego durante los ochenta por Europa y, posteriormente, por los países latinoamericanos (Saada, 2008). Dicho campo emergió como respuesta a la acelerada expansión de las ciencias de la vida y de la salud que, junto con los desarrollos biotecnológicos asociados, han suscitado problemas éticos hasta entonces desconocidos. Tal es así que se llegó a sugerir a la naciente bioética como “la ciencia de la supervivencia”, dando a entender por ello que su aceptación y difusión eran de suma importancia para la continuidad del ser humano en nuestro planeta (Potter, 1970). La bioética pasó entonces a ser considerada una urgente necesidad para afrontar los desafíos que trajera consigo la tecnociencia.

Debido a la magnitud y la importancia muchas veces global de los efectos de los productos tecnocientíficos, la bioética comenzó como un campo de gran amplitud que contemplaba todo aquello que involucraba una relación ética con lo vivo. En ese sentido, Van Rensselaer Potter destacó que “[t]enemos una gran necesidad de una ética de la tierra, una ética de la vida silvestre, una ética de la población, una ética del consumo, una ética urbana, una ética internacional, una ética geriátrica, etc. Todos estos problemas requieren acciones basadas en valores y hechos biológicos. Todos ellos implican a la bioética” (Potter, 1970, p. 127). Concebida de dicha manera, la bioética fue pensada como un campo de amplia acción cuyo alcance sólo quedaba limitado por las fronteras de la vida misma.

Por otro lado, cabe destacar que, al margen de cómo fue pensado su dominio de reflexión y de acción, la bioética fue originalmente concebida como un campo profesional intermedio capaz de tender puentes entre diversos actores. Así, en el caso de las disciplinas científicas y tecnológicas vinculadas, fundamentalmente, con la biología y la medicina, el profesional de la bioética se consideraba capaz de entender y hacer entender los distintos saberes que requerían ser puestos en diálogo en la resolución de diferentes problemáticas sociales y ambientales.

### **De las pretensiones originales de la bioética a sus acciones concretas**

Pese a que la bioética fue originalmente pensada como un campo de reflexión acerca de todo lo concerniente a lo vivo, diversos autores han manifestado que su ámbito de injerencia ha quedado en los hechos reducido a una única dimensión: la biomédica. Con estas palabras se ha referido Gilbert Hottois a esta cuestión: “Pero este éxito de la palabra [bioética] no fue recompensado en relación con las intenciones originales según las cuales fue creado. El mundo médico se había apoderado de él” (2015, p. 24). Tanto su modo de existencia profesional como social quedaron de este modo configurados por el marco de referencia de la biomedicina y el horizonte de sentido que la caracteriza. Como también comenta Alya Saada cuando se refiere particularmente a la situación en los países latinoamericanos: “resulta procedente observar que algunos investigadores, universidades y países comprenden el bio de la (bio)-ética como más relacionado con la biomedicina” (2008, p. 16).

Asimismo, entre los principales elementos de la bioética aparece como principal respuesta la formulación y puesta en práctica de extensas nor-

mativas y regulaciones. De esta manera, la bioética se asocia fuertemente con los aspectos legales. Si bien algunos autores consideran que este es un aspecto positivo de la práctica, otros sostienen que promueve la adaptación de las distintas instituciones y actores a un sistema social ya constituido en vez de ofrecer vías de reflexión hacia soluciones alternativas (Diligio, 2017). Tal pareciera ser el caso, por ejemplo, de las discusiones que giran en torno a la regulación de ciertas tecnologías médicas en las que no se cuestiona si es deseable o no que ciertas tecnologías sean desarrolladas, sino que se discute qué tipo de usos son adecuados y legítimos desde los puntos de vista ético y legal.

Un tercer aspecto vinculado con lo dicho hasta el momento se relaciona con la idea de que la reflexión ética puede ser “transferida” a otro sujeto sin sufrir pérdidas esenciales. Este problema puede ser reconocido a partir de uno de los instrumentos privilegiados que asume el campo a la hora de intervenir: los comités de ética. Dichos comités están conformados por un grupo de expertos y representantes de las comunidades locales y su objetivo principal radica en evaluar la dimensión ética de la tecnociencia, ya sea en etapas vinculadas con la formulación de un proyecto de investigación o de la relación entre un investigador y su objeto de estudio. Con la presencia de los comités, la bioética asume un modelo de acción en el cual los investigadores son dispensados de pensar en las consecuencias de sus acciones, dado que han “transferido” esa capacidad y responsabilidad a los miembros de los comités.

Por último, cabe destacar la marcada indiferencia recíproca que existe entre la bioética y el campo de la epistemología. Si bien gran parte de la filosofía de la ciencia del siglo XX ha sostenido el ideal de una ciencia libre de valores y ha, a partir de ello, mantenido separadas las esferas de la ética y la ciencia, aportes recientes problematizan esta postura y exhiben distintos modos de relación entre ambas (ver Gómez, 2014). Por ejemplo, algunos autores han trabajado el modo en que distintos valores, tanto epistémicos como no epistémicos, inciden en la elaboración y justificación de hipótesis y teorías (Douglas, 2009; Lacey, 2005), mientras que otros se han centrado en cómo distintas consideraciones epistemológicas del saber científico inciden en las correspondientes evaluaciones éticas de las consecuencias de dicho saber (Pallitto, 2020). La referida reorientación en las reflexiones todavía es incipiente, por lo que la bioética y la epistemología aún se mantienen como esferas separadas y en los hechos prácticamente desvinculadas.

Agenda reducida, vínculo estrecho entre lo ético y lo legal, “transferencia” de las reflexiones éticas y separación entre epistemología y ética son cuatro de las características que asume la bioética en la actualidad, incluso en los países latinoamericanos.

## La bioética y el modelo agroindustrial argentino

En las últimas dos décadas, el Estado argentino adoptó un modelo de producción agrario basado en un tipo de paquete tecnológico. Con la combinación de transgénicos y agrotóxicos, rápidamente diferentes tipos de cultivos fueron incorporados tanto para el consumo interno como para su exportación. Si bien fueron pocas las voces críticas que se han levantado al seno de los campos profesionales, se ha hecho referencia en múltiples oportunidades a los problemas sociales y ambientales que acarrea dicho paquete tecnológico. Como ya fue señalado, entre algunas de sus consecuencias se señalan, por ejemplo, la pérdida de biodiversidad, el desplazamiento forzado de comunidades de sus territorios, la contaminación de suelos y cuerpos de agua, los daños ambientales y en la salud humana y los aspectos vinculados con los derechos de la comercialización de las semillas y la dependencia económica en el valor de mercado de unos pocos cultivos (Francese & Folguera, 2018).

En ese sentido, incluso cuando se adopte una definición de vida restrictiva, quedan pocas dudas de que los elementos previamente señalados refieren a algún aspecto de lo vivo, con lo cual, *a priori*, el modelo agroindustrial argentino debería ser parte de la agenda de la bioética, tal como fue pensada y elaborada en sus orígenes. Sin embargo, las características señaladas en la sección precedente respecto a la bioética muestran que la misma no incorpora como tema central de su agenda las cuestiones vinculadas con los organismos modificados tecnológicamente y los agrotóxicos. En los casos en los que sí lo hace, lo hace de manera tardía, una vez que el horizonte de posibilidades ya ha sido aceptado. A modo de ilustración de esto último, puede señalarse la discusión en torno a las distancias mínimas de aplicación de los agrotóxicos. Colocar en el centro de la disputa a las distancias de aplicación mínimas implica asumir al modelo como algo incuestionable. Asimismo, exhibe el vínculo estrecho que la bioética con las vías legales y normativas, operando exclusivamente mediante esos dominios de acción.

Un último aspecto insoslayable consiste en pensar qué saberes científicos y cómo son considerados por la bioética cuando se presenta como un campo intermedio entre la ciencia y la ética. En esta consideración reside uno de los puntos de encuentro entre la epistemología y la ética que urge ser explorado por la filosofía de la ciencia. Sin dudas, diversas ciencias se relacionan con el modelo agroindustrial. Por solo mencionar las relativas a las ciencias naturales, es posible reconocer áreas tales como la biología molecular, la ecología en sus diversas subdisciplinas, la biología evolutiva, la agronomía y las ciencias ambientales, entre otras. Esta pluralidad de saberes implica que hay diversas maneras de representar fenómenos vinculados con el modelo. Se puede tomar como ejemplo la medición de consecuencias ambientales y para la salud humana del uso de agrotóxicos. ¿Qué saberes son relevantes a la hora de responder a esa pregunta? Ciertamente, las respuestas no son en absoluto evidentes. Por ejemplo, considerar los conocimientos provenientes de la ecotoxicología puede brindar percepciones de las consecuencias ambientales diferentes a las que podría otorgar la ecología de ecosistemas. Por lo tanto, los riesgos asociados al modelo dependen, en parte, de qué saberes son considerados y cuáles son omitidos, tendiendo un puente directo entre la epistemología y la ética. A su vez, en función de las racionalidades en juego, distintas simplificaciones teóricas pueden estar operando, tales como las que refieren a la omisión de factores complejizantes que son considerados por la biología molecular cuando la finalidad principal consiste en brindar representaciones epistémicamente virtuosas de los mecanismos moleculares presentes en las plantas (Francese & Folguera, 2018). En función de lo dicho, la bioética precisa contemplar las diversas formas de representar que se relacionan con el modelo agroindustrial para ejercer una reflexión crítica y adecuada de la dimensión ética involucrada.

## **Cierre y apertura**

Frente a un escenario de expansión acelerada y acrítica de la tecnociencia, surge el interrogante acerca de si las estrategias asumidas por la bioética en la Argentina (y sus presupuestos) permiten repensar y modificar sustancialmente la situación actual. En tanto la bioética argentina, y en mayor escala la latinoamericana, no incluyan un análisis crítico y propositivo de la estructura, tanto axiológica como epistemológica, de las consecuencias adversas de las prácticas agrícolas regionales, la respuesta a tan crucial interrogante se cubre de un manto de negatividad.

## Referencias

- Digilio, P. (2017). Conocimiento y ética. *Revista Debate Público. Reflexión de Trabajo Social*, 13/14, 21–33.
- Douglas, H. (2009). *Science, policy, and the value-free ideal*. Pittsburgh: University of Pittsburgh Press.
- Echeverría, J. (2003). *La revolución tecnocientífica*. Madrid: Fondo de Cultura Económica.
- Francese, C., & Folguera, G. (2018). Saberes simplificados, tecnociencia y omisión de riesgos. El caso de los organismos genéticamente modificados. *Runa*, 39(2), 5–27.
- Gómez, R. (2014). *La dimensión valorativa de las ciencias: Hacia una filosofía política*. Bernal: Universidad Nacional de Quilmes.
- Hottois, G. (1991). *El paradigma bioético: un paradigma para la tecnociencia* (M.C. Monge, trad.). Barcelona: Anthropos. (Obra original publicada en 1990)
- Hottois, G. (2015). Defining bioethics: back to the sources. En D. Meacham (Ed.), *Medicine and society. New perspectives in continental philosophy* (pp. 15–38). Dordrecht: Springer.
- Lacey, H. (2005). *Is science value free? Values and scientific understanding*. Londres: Routledge.
- Linares, J. (2008). *Ética y mundo tecnológico*. México: Fondo de Cultura Económica.
- Pallitto, N. (2020). Pluralidad científica y evaluación de riesgos: nuevos argumentos en el debate monismo vs. pluralismo en Filosofía de la Ciencia. *Revista Colombiana de Filosofía de la Ciencia*, 19(39). En prensa.
- Potter, V.R. (1970). Bioethics, the science of survival. *Perspectives in Biology and Medicine*, 14, 127–153.
- Saada, A. (2008). Prólogo. En J. C. Tealdi (Director), *Diccionario Latinoamericano de Bioética* (pp. xix–xxii). Bogotá: UNESCO - Red Latinoamericana y del Caribe de Bioética.



## Redescubriendo a Claude Bernard

Abigail Prchal\*

Jesús Alberto Zeballos\*

### Introducción

Claude Bernard (1813-1878) ocupa un merecido lugar en la historia de las Ciencias Biomédicas y en ellas su nombre resulta familiar. Sin embargo, muy pocos conocen los descubrimientos y las ideas por las cuales trascendió en la historia de la ciencia. Sus principales hallazgos (la función glucogénica del hígado, los nervios vasomotores, la utilización de sustancias tóxicas para diseccionar funciones, el medio interno) fueron acontecimientos científicos más que suficientes para que su nombre quedara en la historia de las ciencias biomédicas.

Se ha dicho de C. Bernard que más que ser un fisiólogo es la Fisiología (González Recio, 2017; Lami, 1911). Es innegable que hubo fisiólogos antes de Bernard, pero no había una Fisiología como ciencia: sólo hechos dispersos, sin conexión teórica, sin ninguna idea directriz (Lami, 1911). En tal sentido, Claude Bernard crea un programa de investigación para la fisiología vigente hasta nuestros días, instituyéndola como disciplina independiente (Caponi, 2001).

También es cierto que antes de Bernard hubo experimentadores. Sin embargo, no existía el método experimental como tal. John Locke, David Hume, F. Bacon y S. Mill establecieron algunas pautas lógicas del conocimiento empírico desde una perspectiva filosófica. Pero Claude Bernard es quien instaura el método experimental como forma necesaria de investigación y quien establece las leyes del método (Houssay, 1941; Lami, 1911).

Su contribución más importante se halla magistralmente expuesta en el libro *Introducción al estudio de la medicina experimental* de 1865 (Bernard, 1865/2005). Allí Bernard sistematiza la metodología experimental y funda el programa de investigación para la Fisiología. Mostraremos que se trata de una obra extraordinaria en varios sentidos.

Las contribuciones de Bernard han sido poco estudiadas en la historia y filosofía de la ciencia, a pesar de ser un tratado de neto corte epistemológico.

\* Universidad Nacional de Tucumán. Tucumán, Argentina.  
aprchal@gmail.com

gico (Lorenzano, 1980). Por ello nos pareció relevante volver a este libro en el que Bernard postula algunas ideas que anticipan la epistemología del S XX (Lorenzano, 1980; Malherbe, 1981).

Comenzaremos por un breve repaso de las ideas de Bernard, para luego mostrar las similitudes y diferencias entre ellas y el falsacionismo de Popper, el verificacionismo del Círculo de Viena y las orientaciones actuales de la epistemología.

## La ciencia bernardiana

Para C. Bernard, la ciencia es “un salón soberbio todo resplandeciente de luz, al que no se puede llegar más que pasando por una larga y espantosa cocina” (Bernard, 1865/2005, p. 176).

En Bernard se conjugan el típico sabio del siglo XIX, ampliamente informado del pensamiento filosófico, con el investigador de laboratorio, conocedor al detalle de los procedimientos experimentales. Por ello, al decir de B. Houssay (1941), Bernard, más que Bacon o S. Mill, es el verdadero legislador del método experimental. Efectivamente, en su libro de 1865 (Bernard, 1865/2005) expone con claridad meridiana los preceptos lógico-filosóficos de la “cocina” científica.

La ciencia contemporánea ha heredado dichos preceptos y los implementa sistemáticamente como la más efectiva metodología científica. No sin razón se ha afirmado que C. Bernard fue un pionero de la epistemología de las ciencias de la vida, en la misma medida (o tal vez más) en que Newton fundó la epistemología de la Física (Normandin, 2007).

Los pilares de la investigación científica de Bernard se asientan en los siguientes supuestos:

- a) ontológicos: estratificación ordenada de la realidad (física, química, biológica) y una ínsita legalidad causal, determinismo;
- b) metodológicos: observación de los hechos, intuición para formular hipótesis explicativas, deducción de consecuencias observacionales, comprobación y/o refutación de hipótesis.
- c) epistemológicos: sólo podemos conocer progresivamente y en continuo perfeccionamiento las leyes causales de la realidad como relaciones fenoménicas relativas, no absolutas.

A continuación, describiremos, brevemente, estos supuestos.

## Supuestos ontológicos de la ciencia bernardiana: determinismo y reduccionismo

Para un desarrollo racional de las ciencias empíricas, Bernard (1865/2005) supone un determinismo ontológico y lo pone en el corazón mismo de su fisiología (Conti, 2001). Las causas finales y la esencia de las cosas son inaccesibles para el conocimiento científico. Por ello, afirma, la ciencia sólo puede (y debe) ocuparse de las causas necesarias y/o eficientes y próximas. Este rechazo de todo principio metafísico le permite erradicar el vitalismo<sup>1</sup> de las ciencias biológicas y establecer la experimentación como el método de investigación adecuado para la Fisiología.

El determinismo bernardiano es un principio absoluto que gobierna tanto a la materia inorgánica como a la materia viva (Caponi, 2016; Conti, 2001). Esta afirmación, a mediados del S. XIX, suponía algunos cambios radicales en el modo de considerar al mundo, sus hechos y cosas. No solo diluye la frontera entre física y biología, sino que borra también el límite entre lo vegetal y lo animal (límite que, por otra parte, Bernard ya había trascendido al demostrar los procesos de glucogénesis hepática) (González Recio, 2017). Además, Bernard extiende el determinismo a la esfera psíquica que no podría expresarse sin un sustrato físico químico (Bernard, 1872; Conti, 2001; Lami, 1911).

Otra consecuencia importante del determinismo es que pone a la Fisiología en el centro mismo de la medicina y la terapéutica. En primer lugar, Bernard establece que no hay nada radicalmente diferente entre los procesos que rigen la salud y aquellos que determinan la enfermedad o, para decirlo en términos bernardianos, entre el estado fisiológico (normal) y el patológico. Los estados fisiológicos y los estados patológicos obedecen a las mismas leyes y sólo difieren en sus condiciones iniciales. Por tanto, la patología deberá descubrir cuáles son las alteraciones de las condiciones que producen la enfermedad; la terapéutica, por su parte, deberá actuar sobre esas condiciones de modo de retrotraerlas a aquellas que producen el estado fisiológico (Bernard, 1872; Lami, 1911).

---

<sup>1</sup> Se ha discutido bastante hasta qué punto Bernard fue o no vitalista (Canguilhem, 2002/2004; Caponi, 2016; Normandin, 2007). El tema excede los objetivos del presente trabajo. Si bien nos parece que en esa discusión se ha sido injusto con Bernard, en tanto no siempre se ha tenido en cuenta el influjo del ambiente intelectual de la época del cual nadie puede ser totalmente ajeno. En cualquier caso, es claro que para Bernard el vitalismo es una enfermedad que paraliza el progreso científico de la medicina (Lami, 1911).

El determinismo, para Bernard es físico-químico y debe buscarse a nivel celular. Bernard considera que toda la fisiología debe reducirse a mecanismos físico-químicos de la célula, que es el elemento fundamental de la vida tanto vegetal como animal (Caponi, 2001; González Recio, 2017; Houssay, 1941; Lami, 1911). Como programa de investigación, estas ideas aún rigen en la Biología Molecular contemporánea.

### **Supuestos metodológicos de la ciencia bernardiana: experimentación**

No puede decirse que Claude Bernard haya descubierto la experimentación; sin embargo, es quien sistematiza los conocimientos y experiencias de los siglos anteriores y agrega algunas características de su propia cosecha epistemológica (Conti, 2001; Houssay, 1941).

El determinismo, como único principio incuestionable, le permite a Bernard establecer la experimentación como criterio de demarcación entre la ciencia y el conocimiento no científico (“pseudociencia”, diríamos hoy).

Para Bernard, el método experimental conjuga un modo de razonar, un modo de observar y una experiencia. La experimentación es, entonces, un método para contrastar (o controlar) una idea o hipótesis por medio de dos o más observaciones, que permitirán establecer (conocer o verificar) las leyes que rigen la aparición de un fenómeno. Se trata de una metodología comparativa con tres momentos y una regla fundamental:

- Formulación de la hipótesis explicativa del fenómeno.
- El experimento (o prueba) que busca probar si, dadas ciertas condiciones iniciales, se sigue el efecto previsto; *posita causa, ponitur effectus*.
- La contra prueba: si se eliminan las condiciones iniciales, no aparece el efecto. *Subleva causa, tollitur effectus*. Esta contra prueba es más importante que la prueba. El experimentador debe esmerarse más en tratar de demostrar que está errado que en tratar de demostrar que tiene razón.
- Dudar de todo, menos del determinismo. Si el experimento no verifica la hipótesis, si la prueba y la contra prueba no son consistentes, deben buscarse fallas en el razonamiento, en el experimento

mismo, en la observación o en los instrumentos utilizados. De ningún modo estas “fallas” en la verificación experimental de las hipótesis pueden poner en duda la determinación de los fenómenos de la vida.

Claude Bernard dedicará la primera parte de su libro de 1865 a describir y fundamentar el método experimental refiriéndolo constantemente a la fisiología. En tal sentido, buscará demostrar que no hay diferencias importantes entre observar y experimentar, ya que el experimentador debe necesariamente observar los resultados de su experimento.

En la misma línea de pensamiento, Bernard destaca la importancia de las hipótesis ya que todo experimento se realiza para verificar alguna idea. Si bien Bernard admite la posibilidad de realizar experimentos para “ver qué pasa”, considera que, en general, el método es más fructífero para controlar hipótesis.

### **Supuestos epistemológicos de la ciencia bernardiana: conocimiento provisorio**

Para C. Bernard la ciencia se diferencia de la metafísica y de la escolástica en que es un conocimiento relativo, provisorio e independiente del principio de autoridad.

Según Bernard, el método experimental prueba una y otra vez que las causas primeras y la realidad objetiva de las cosas permanecerán para siempre ocultas y que sólo se pueden conocer relaciones entre los fenómenos. En tal sentido, el conocimiento científico o los resultados experimentales no son más que interpretaciones más o menos probables de la Naturaleza. En el ámbito de la ciencia, no hay verdades absolutas y eternas. La verdad es siempre relativa a la cantidad de observaciones y/o experimentos realizados. Y es también una verdad provisoria: “Si hasta el presente ninguna observación ha desmentido la verdad en cuestión, no por ello concibe el espíritu la imposibilidad de que las cosas puedan ser de otro modo” (Bernard, 1865/2005, p. 194).

Por otra parte, al ser el método experimental el *criterium* que demarca lo científico de lo no científico se trata, claramente, de un conocimiento que se somete a la evidencia de los hechos y no a la autoridad personal.

## Claude Bernard y la epistemología del siglo XX

La filosofía de la ciencia, se ha preocupado poco por la medicina. Con excepción de Canguilhem, Foucault y recientemente Bunge y Samaja, existen relativamente pocos desarrollos epistemológicos acerca del conocimiento médico. Según Lorenzano (1980) esto se debería a que a) la medicina ha sido considerada más como una práctica que como una disciplina independiente, b) como conocimiento, ha sido incluido en la Biología, c) la epistemología clásica se ha ocupado de la física, y sólo recientemente se han creado campos de reflexión específicos para otras ciencias. A estas razones habría que agregar que los autores de origen anglosajón tradicionalmente ignoran los desarrollos continentales y que la mayoría de la bibliografía sobre Bernard no sólo proviene principalmente de médicos y fisiólogos, sino que además está escrita en idiomas latinos.

Cualquiera que haya leído la *Introducción al estudio de la medicina experimental* encontrará en la *Lógica de la investigación científica* de Karl Popper muchos pasajes familiares. Se ha dicho que Bernard fue un “popperiano *avant la lettre*” (González Recio, 2017; Lorenzano, 1980; Malherbe, 1981), lo que es un anacronismo flagrante. Sería más adecuado, cronológicamente, afirmar que Popper fue bernardiano.

Popper admitió haber leído a Bernard recién en 1970 (35 años después de la primera edición de la *Lógica de la investigación científica*), como consecuencia de que Sir P. Medawar le señalara las similitudes entre su obra y la de Bernard (Popper, 1974). Sin embargo, hasta donde hemos podido averiguar, Popper nunca abundó en el análisis de estas coincidencias (Lorenzano, s. f.).

### El método hipotetico deductivo

Bernard no utilizó nunca esta denominación, sin embargo, su formulación metodológica coincide punto por punto con la postulada por Popper 70 años después. Lorenzano (1980) y Malherbe (1981) sistematizaron las coincidencias entre los textos de Bernard y de Popper. Nosotros nos limitaremos a resumir la metodología bernardiana en tres aspectos principales: 1) El problema y la hipótesis, 2) Deducción y contrastación y, 3) La contraprueba. Las coincidencias con Popper, como se verá, saltan a la vista.

## 1) El problema y la hipótesis

Bernard afirma que lo que inicia la investigación experimental es un problema, una incógnita que se puede originar tanto de la observación de un hecho como de la teoría misma. Para resolver el problema, despejar la incógnita se postula una hipótesis.

En este punto Bernard hará una clara distinción entre lo que luego (Reichenbach, 1970) se denominarían contextos de descubrimiento y de justificación: “no pueden darse reglas fijas para hacer nacer en el cerebro una idea exacta y fértil. El método experimental no puede dar ideas nuevas y provechosas a los hombres que carecen de ellas” (Bernard, 1865/2005, p. 200).

Para Bernard, la formulación de una hipótesis tiene un papel fundamental en el método:

La idea anticipada o hipótesis, es entonces el punto de partida necesario de todo razonamiento experimental. Sin ella no podríamos hacer ninguna investigación, sólo podríamos ir amontonando observaciones estériles. (Bernard, 1865/2005, p. 195)

La razón o razonamiento no sirve más que para deducir consecuencias de esta idea y para someterlas a la experiencia. (Bernard, 1865/2005, p. 195)

## 2) Deducción y contrastación

Como se sabe, en el método hipotético deductivo las hipótesis se ponen a prueba deduciendo de ellas con el mayor rigor posible consecuencias observacionales, que ratifican o refutan la hipótesis en cuestión:

El experimentador, más modesto, plantea al contrario su idea como una pregunta, como una interpretación anticipada de la naturaleza, más o menos probable, de la que deduce lógicamente consecuencias que confronta a cada instante con la realidad por medio de la experiencia. (Bernard, 1865/2005, p. 193)

Bernard discute con algún detalle la inducción y la deducción para llegar a la conclusión de que siempre se trata de deducciones:

Ahora bien, nos será fácil demostrar más tarde, que esta idea a priori que surgió en nosotros a propósito de un caso particular, encierra siempre

implícitamente, y en cierto modo a pesar nuestro, un principio al que queremos referir el hecho particular. De suerte que cuando creemos ir de un caso particular a un principio, es decir inducir, deducimos realmente. (Bernard, 1865/2005, p. 215)

Y podemos concluir el camino o método hipotético deductivo popperiano, con los términos que utilizaba Claude Bernard:

En fin, como conclusión de este párrafo podemos decir que, en todo razonamiento experimental, hay dos casos posibles: o bien la hipótesis del experimentador queda invalidada, o bien queda confirmada por la experiencia. Cuando la experiencia invalida la idea preconcebida, el experimentador debe rechazar o modificar esta idea. (Bernard, 1865/2005, p. 220)

### 3) La contraprueba

Hasta aquí el empirismo lógico y Karl Popper pudieron acompañar a Claude Bernard en su metodología. Los algoritmos lógicos de que disponían al momento se lo permitían: *modus ponendo ponens* para el círculo de Viena; *tollendo tollens* para Karl Popper, luego de advertir la falacia de afirmar el consecuente.

No pudieron en cambio seguirlo cuando Claude Bernard avanza sobre lo que denominó “contraprueba”:

La contra-prueba deviene, pues, el carácter esencial y necesario de la conclusión del razonamiento experimental. Ella es la expresión de la duda filosófica llevada a su máximo alcance. Es la contra-prueba la que dictamina si la relación de causa a efecto que se busca en los fenómenos ha sido encontrada. (Bernard, 1865/2005, p. 224)

Limitándonos a la sola prueba podríamos erróneamente considerar relaciones de causa a efecto, cuando no hay más que simple coincidencia o secuencia meramente temporal de circunstancias. La contraprueba “suprime la causa admitida a fin de ver si el efecto persiste. . . . Es un contra-juicio que se refiere directamente a la conclusión experimental y forma uno de sus términos necesarios” (Bernard, 1865/2005, p. 224).

En otro pasaje Bernard, afirma que “la verdad no aparece jamás al espíritu más que bajo la forma de una relación, de una conexión absoluta y necesaria” (Bernard, 1865/2005, p. 193).

Esta afirmación nos pone en la pista de la lógica que estructura las argumentaciones explicativas y predictivas de su epistemología. Obviamente, no es la lógica clásica que concluye en las consabidas paradojas de la implicación material. Tampoco, la lógica de la implicación estricta de Lewis y Langford (1932) que, a su vez, culmina también en paradojas. Agreguemos, por otra parte, que estas formas lógicas no son, no pueden ser la representación formal del concepto de inferencia semántica del lenguaje ordinario, que utilizan las ciencias empíricas. La relación entre causa y efecto, propuesta por Claude Bernard se aproxima más a la concepción relevante de Anderson y Belnap (1975), que interpreta más adecuadamente la noción de “consecuencia lógica” del lenguaje común.

Por supuesto, Claude Bernard no alcanza la claridad formal que le dieron estos lógicos; pero la contraprueba es un claro ejemplo de sustitución en lenguaje común de fórmulas relevantes. Recordemos que Anderson y Belnap (1975) propusieron esta lógica como una alternativa a la lógica clásica para resolver las paradojas de la implicación estricta. Su exigencia básica es que el antecedente se use para probar el consecuente. En Bernard esta exigencia es mayor: no sólo el antecedente se usa para probar el consecuente, sino que la contraprueba muestra que el antecedente de la prueba es esencial y necesario para concluir el razonamiento experimental, la condición *sine qua non*.

## Misceláneas

En la Introducción de este ensayo, adelantamos que este tratado de medicina experimental “es una obra extraordinaria en varios sentidos”.

Pues bien, además de los apuntes lógicos y metodológicos referidos, hay perlas epistemológicas a lo largo de la *Introducción al estudio de la medicina experimental* que abonan nuestra aserción inicial y muestran que la epistemología contemporánea es heredera de los postulados de Claude Bernard.

Citamos dos ejemplos tomados casi al azar:

El método por sí mismo no engendra nada, y es un error de ciertos filósofos haber acordado demasiado poder en este sentido. (Bernard, 1865/2005, p. 200)

Si se nos presenta una idea, no debemos rechazarla sólo porque no esté de acuerdo con las consecuencias lógicas de una teoría dominante . . . nunca

se debe dar un valor absoluto a estas teorías . . . grandes físicos hicieron descubrimientos de primer orden, con motivo de experiencias instituidas de una manera ilógica con relación a las teorías. (Bernard, 1865/2005, p. 201)

La primera frase bien podría pertenecer a Fayerabend y la segunda a Khun.

## Conclusión

En *Introducción al estudio de la medicina experimental* Bernard establece un método acorde con la epistemología actual, a partir de las reflexiones sobre su propia actividad de investigación. Así, fue un pionero de la epistemología de las ciencias de la vida, en la misma medida (o tal vez más) en que Newton fundó la epistemología de la Física (Normandin, 2007).

## Referencias

- Anderson, A. R., & Belnap, N. D. (1975). *Entailment: The logic of relevance and necessity*. Princeton University Press.
- Bernard, C. (1872). *Leçons de pathologie expérimentale par Claude Bernard*. J. B. Baillière.
- Bernard, C. (2005). *Introducción al estudio de la medicina experimental* (A. Espina y Capó, Trad.). Crítica Fundación Iberdrola. (Obra original publicada en 1865)
- Canguilhem, G. (2004). *Escritos sobre la medicina* (I. Agoff, Trad.). Amorrotu. (Obra original publicada en 2002)
- Caponi, G. (2001). Claude Bernard y los límites de la fisiología experimental. *História, Ciências, Saúde-Manguinhos*, 8(2), 375-406.
- Caponi, G. (2016). Claude Bernard y el lugar de la fisiología en un mundo físicamente determinado. *Ludus Vitalis*, 23(44), 43-68.
- Conti, F. (2001). Claude Bernard: Primer of the second biomedical revolution. *Nature Reviews Molecular Cell Biology*, 2, 703.
- González Recio, J. L. (2017). Who killed histological positivism? An approach to Claude Bernard's epistemology. *Ludus Vitalis*, 12(22), 61-82.

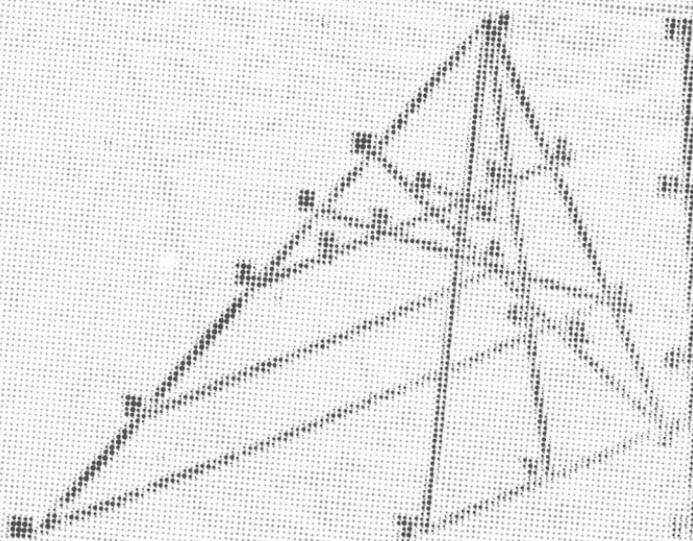
- Houssay, B. A. (1941). Claude Bernard y el método experimental. *Revista de la Universidad Nacional de Córdoba*, (9/10).
- Lami, J. (1911). *Claude Bernard et la methode expérimentale*. Paris: Medicina
- Lewis, C. L., & Langford, C. H. (1932). *Symbolic Logic*. The Century Co.
- Lorenzano, C. (s. f.). *La epistemología médica de Claude Bernard*. [https://www.researchgate.net/profile/Cesar\\_Lorenzano/publication/268407194\\_LA\\_EPISTEMOLOGIA\\_MEDICA\\_DE\\_CLAUDE\\_BERNARD/links/5c066ca1299bf169ae315d4a/LA-EPISTEMOLOGIA-MEDICA-DE-CLAUDE-BERNARD.pdf](https://www.researchgate.net/profile/Cesar_Lorenzano/publication/268407194_LA_EPISTEMOLOGIA_MEDICA_DE_CLAUDE_BERNARD/links/5c066ca1299bf169ae315d4a/LA-EPISTEMOLOGIA-MEDICA-DE-CLAUDE-BERNARD.pdf).
- Lorenzano, C. J. (1980). Dos racionalismos críticos: Claude Bernard y Karl Popper. *Teoria, Anuario de Filosofía*, 1, 223-245.
- Malherbe, de J.-F. (1981). Karl Popper et Claude Bernard. *Dialectica*, 35(4), 373-388.
- Normandin, S. (2007). Claude Bernard and An Introduction to the Study of Experimental Medicine: «Physical Vitalism,» Dialectic, and Epistemology. *Journal of the History of Medicine and Allied Sciences*, 62(4), 495-528.
- Popper, K. (1974). Replies to my critics, 16. Reply to Medawar on hypothesis and imagination. En P.A. Schilpp, (Ed.) *The philosophy of Karl Popper* (Vol. 1, pp. 1030-1037). Open Court.
- Reichenbach, H. (1970). *Experience and prediction: An analysis of the foundations and the structure of knowledge*. University of Chicago Press.





---

## 7. Filosofía del lenguaje y teoría del conocimiento







## **“Es como si nuestros conceptos estuvieran condicionados por un andamiaje de hechos” (Wittgenstein sobre la formación de conceptos)**

Gustavo Arroyo\*

**E**n esta ponencia me propongo dos objetivos. El primero es comentar y desarrollar una tesis apenas esbozada por Wittgenstein a través de una metáfora (la metáfora del “andamiaje”) en un pasaje de *Zettel* (1967/2007). En segundo lugar, me propongo extraer algunas consecuencias de esa tesis para la práctica del análisis conceptual. El pasaje donde aparece la metáfora dice así:

Es como si nuestros conceptos estuvieran condicionados por un andamiaje de hechos. Esto significaría: Si imaginas, si describes ciertos hechos de otro modo que como son, entonces ya no puedes imaginar la aplicación de ciertos conceptos, porque en las nuevas circunstancias no hay nada análogo para las reglas de su aplicación. (Wittgenstein, 1967/2007, § 350)

Wittgenstein parece exponer una tesis similar en el siguiente pasaje de las *Investigaciones filosóficas*:

Sólo en los casos normales nos es claramente prescrito el uso de una palabra; sabemos, no tenemos duda, qué hemos de decir en este o aquel caso. Cuanto más anormal es el caso, más dudoso se vuelve lo que debemos decir entonces.

Y si las cosas fueran totalmente distintas de cómo efectivamente son –si no hubiera, por ejemplo, ninguna expresión característica de dolor, de miedo, de alegría; si la regla se convirtiera en excepción y la excepción en regla; o si ambas se convirtieran en fenómenos de aproximadamente la misma frecuencia– entonces nuestros juegos de lenguaje normales perderían con ello su *quid*. (1953/1988, § 142)

Wittgenstein sostiene aquí que las condiciones empíricas inciden en la formación de los conceptos del lenguaje natural. Hay varios ejemplos históricos que respaldan esta posición. Uno de ellos es el concepto

\* Universidad Nacional de General Sarmiento.  
gjarroyo@yahoo.com.ar

de "muerte". Hasta no hace mucho tiempo atrás, el cese de las funciones cardio-respiratorias en una persona conducía invariablemente al colapso irreversible del sistema nervioso y con él al colapso de todos los otros sistemas del cuerpo. Por ejemplo, era empíricamente imposible que el sistema nervioso de una persona colapsara y que sus funciones cardio-respiratorias continuaran en actividad. Como he argumentado en otro lugar (Arroyo, 2020), en esas condiciones, era perfectamente razonable que el concepto de muerte estuviera ligado conceptualmente al cese de las funciones cardio-respiratorias. Si una persona respiraba y su corazón latía, estaba viva. En caso contrario se la declaraba muerta y entraban en vigencia todas las disposiciones éticas, médicas y legales aplicables al cuerpo de un fallecido. Como bien sabemos, a mediados del siglo pasado, la irrupción de ciertas tecnologías modificó radicalmente este estado de cosas. Con la aparición de las técnicas de resucitación, el cese de las funciones cardio-respiratorias no siempre conducirá al colapso general del organismo y con la invención de los respiradores artificiales podemos mantener respirando a una persona cuyo sistema nervioso ha colapsado irreversiblemente (lo que hoy denominamos "muerte cerebral").

Comenzó en los años cincuenta una intensa discusión dentro y fuera de la medicina respecto de si los conceptos de "vida" y "muerte" debían continuar siendo definidos como hasta entonces, si nuevas maneras de definirlos debían buscarse y si eventualmente nuevas categorías debían ser creadas. El debate, que se cerró ya hace un par de décadas, concluyó con el desdoblamiento del antiguo concepto de muerte en los conceptos de "muerte clínica" y "muerte biológica", donde el segundo concepto hereda todas las consecuencias éticas, médicas y legales del antiguo concepto de muerte.

Un ejemplo similar está dado por el concepto "madre". Hasta principio de los años 80 del siglo XX, la madre de un individuo era la mujer que lo había gestado hasta el momento de su nacimiento. Pero la posibilidad, surgida gracias a la fertilización *in vitro*, de que una mujer geste un embrión a partir de un óvulo cedido por otra, condujo a una revisión del concepto de madre. Al igual que con el concepto de "muerte", nuevas categorías fueron creadas (madre gestante, madre genética, madre legal, etc.) para discriminar y visibilizar clases de fenómenos que hasta ese momento eran coextensivos.

El hecho de que muchas categorías y distinciones conceptuales sean forjadas sobre el trasfondo de condiciones empíricas que se dan por sen-

tadas en un determinado momento, implica que el significado de esas categorías y distinciones deberá ser rediscutido toda vez que las condiciones empíricas se vean afectadas. La aplicación de las viejas categorías a las nuevas condiciones podría arrojar decisiones injustas o irrazonables. Por ejemplo, hay buenos motivos para no considerar “vivas” a las personas con muerte cerebral conectadas a un respirador artificial. En las condiciones actuales, darles ese estatus implicaría invertir valiosos recursos en mantener funcionando un organismo que no tiene ninguna posibilidad de recuperación y privar a muchas personas de usar sus órganos. Calificar de “madre” a una mujer por el solo hecho de haber gestado a un individuo, cuando el óvulo a partir del cual el individuo gestado se desarrolló proviene de otra, podría llevarnos a pasar por alto muchos matices de importancia ética y legal. En síntesis, el ámbito de aplicación de las categorías lingüísticas es modalmente acotado. Las categorías del lenguaje son creadas para ser utilizadas dentro de un contexto empírico específico y necesitan ser reajustadas cuando dicho contexto cambia. No podemos esperar que sean aplicables a cualquier situación concebible (lógicamente posible).

Aceptar este resultado supone revisar de manera drástica la manera en que se practica el análisis conceptual en el ámbito de filosofía. Me refiero a la búsqueda de las condiciones que guían la aplicación de ciertas categorías básicas del lenguaje natural; conceptos como “conocimiento”, “identidad personal”, “percepción”, “causa” y muchos otros más. El análisis de esos conceptos y otros similares ha estado en discusión desde tiempos platónicos. Pero en el intento de establecer las condiciones de aplicación de esos términos, se suele asumir precisamente aquello que los dos ejemplos anteriores sugieren que no debe ser asumido.

Tomemos como ejemplo el concepto de percepción. Es indudable que para poder decir que alguien ha *percibido* un objeto, la persona en cuestión debe tener una experiencia (visual) y, en segundo lugar, debe existir en las inmediaciones del sujeto un objeto que corresponda a esa experiencia; de otro modo no diríamos que está “viendo” el objeto, sino que lo está recordando, imaginando o alucinando. “Percibir” tiene tanto un componente subjetivo (la experiencia de un objeto) como un elemento objetivo, la presencia en el entorno del sujeto del objeto experimentado. Pero ¿diríamos que alguien *ve* algo, se pregunta el filósofo Paul Grice (1961/1985), si la experiencia del objeto no fuera producida a través de la vista, sino, por ejemplo, a través de unos electrodos conectados a su cabeza? ¿Debemos categorizar esa experiencia como un caso de “ver”?

Tomemos ahora el concepto de identidad personal. ¿Qué significa que una persona ubicada en un tiempo  $t_1$  es numéricamente idéntica a una persona ubicada en un tiempo  $t_2$ ? En contextos normales nuestro criterio para identificar personas es la identidad corporal. ¿Pero cómo deberíamos aplicar el concepto de identidad personal, se pregunta Derek Parfit (1984/2005), en el caso de que fuera posible realizar un trasplante de cerebro? Si el cuerpo de una persona A recibiera el cerebro de una persona B y el individuo resultante (con el cuerpo de A y el cerebro de B), empezara a comportarse de la manera en que se comportaba B, ¿diríamos que esa persona es A o es B?

Cuando se propone este tipo de escenario en filosofía no solo se asume que podemos dar una respuesta a la pregunta de categorización, sino que esa respuesta puede ser justificada en términos del significado pre-teórico de los conceptos de "percepción" y "identidad". Stephen Stich expone con particular claridad este supuesto al hablar de (lo que él denomina) "proyectos mentalistas" en el campo del análisis conceptual:

En los proyectos mentalistas, las intuiciones son usadas como evidencia acerca del concepto o de la teoría tácita que supuestamente juega un rol central en generar la intuición. El supuesto que subyace en el uso de intuiciones como evidencia en estos proyectos es que las intuiciones de una persona son una fuente confiable de evidencia respecto de si sus conceptos se aplican a los casos planteados en los experimentos mentales o respecto de lo que lo que su teoría presupuesta implica acerca de esos casos. (Stich & Tobia, 2018, p. 375)

De acuerdo a este punto de vista, nuestras intuiciones acerca de la aplicación de un concepto (incluso en contextos contrafácticos) deben poder ser explicadas en término del significado tácito de dicho concepto. Esto es como creer que los conceptos "muerte" y "madre" nos permitirán tomar decisiones (justificadas) en cualquier situación concebible. Supongamos que a personas del siglo XIX se les hubiera pedido que imaginaran que fuera posible mantener funcionando el sistema cardio-respiratorio de una persona con una severa e irreversible lesión cerebral o la posibilidad de que una mujer geste un óvulo proveniente de otra. Si a esas personas les hubieran preguntado si las personas en el primer escenario califican como "vivas" o si las mujeres del segundo escenario califican como las "madres" del individuo en gestación, es probable que muchas de esas personas hubieran dado una respuesta (probablemente afirmativa) a esas preguntas.

Pero es claro que esas respuestas no podrían ser justificadas a través del significado que tenían entonces esos conceptos. Si alguno de ellos hubiera intentado justificar su intuición alegando que “vivo significa que respira y la persona conectada al hipotético aparato respira”, la maniobra habría chocado con la objeción señalada antes. En épocas en que resultaba tecnológicamente imposible mantener funcionando el sistema respiratorio de una persona con daño cerebral severo, hubiera sido una pérdida de tiempo y de energía intelectual decidir cómo esos casos debían ser conceptualmente legislados.

Pasa con los conceptos lo que pasa con las normas jurídicas. En nuestra legislación, los familiares directos de una persona tienen derecho a sus bienes cuando esta fallece. ¿Pero qué ocurriría si la persona fallecida ha sido asesinada por uno de sus familiares directos? ¿Tiene ese familiar derecho a los bienes del difunto? ¿Qué ocurriría si la persona es declarada muerta por error? ¿Estarían los familiares obligados a devolver enteramente los bienes que han recibido como herencia? No sería sorprendente que la norma en vigencia no especificara qué se debe hacer en casos como esos. Pero esa omisión no significaría que la norma debe ser aplicada sin más a cualquier caso que satisfaga el antecedente. Hay muchas situaciones que podrían “derrotar” la aplicación de una norma jurídica y que no figuran entre las cláusulas de excepción. Y esto no se explica necesariamente por la incompetencia de quien la promulgó. Dado que la inclusión de todas las situaciones exceptivas haría la formulación de la norma demasiado engorrosa, es de esperar que sean incluidas solo aquellas que gozan de una saliencia empírica apreciable.

Creo que el punto que intenté ilustrar tomando como ejemplo las distinciones “vivo-muerto” y “madre-no madre” puede ser generalizado para alcanzar a muchos conceptos que interesan a la filosofía y cuyo significado es dependiente de condiciones empíricas muy generales. En cierto sentido, nos encontramos frente a los escenarios imaginados por Grice y Parfit, en la misma posición en que se encontraban las personas del siglo XIX respecto de las tecnologías de respiración artificial y fertilización *in vitro*. Los conceptos de percepción e identidad personal no han sido forjados con vistas a legislar (tomar decisiones) en el tipo de escenario imaginado por esos autores. La única respuesta justificada a la pregunta de cómo esas extrañas ocurrencias deben ser categorizadas consistiría en decir que las mismas están más allá del alcance de nuestras categorías. Como ya se ha dicho, las previsiones conceptuales son siempre de naturaleza post-hoc.

Si esos escenarios se volvieran realidad, deberíamos buscar distinciones que los representen.

## Referencias

- Arroyo, G. (2020). Categorizing imaginary objects. *Kriterion-Journal of Philosophy*, 34(1), 1-20.
- Grice, P. (1985). La teoría causal de la percepción (A. Caso, trad.). *Cuadernos de Critica*, 41. (Obra original de 1961)
- Parfit, D. (2005). *Razones y personas* (M. Rodríguez González, trad.). Madrid: Editorial Visor. (Obra original de 1984)
- Stich, S., & K. Tobia. (2018). Intuition and its critics. En M. Stuart, Y. Fehige, & J. Brown. (Ed.), *The Routledge companion to thought experiments* (pp. 369-384). Abingdon: Routledge.
- Wittgenstein, L. (1988). *Investigaciones filosóficas* (A. García Suárez & U. Moulines, trads.). México: UNAM. (Obra original de 1953)



## De las metáforas de los gestos al lenguaje

Irene Audisio\*,\*

### Recuperar el cuerpo

Luego del intento de diferentes enfoques teóricos por identificar y superar el dualismo mente “cuerpo” en la concepción de lo mental y del lenguaje, surgieron inquietudes tanto teóricas como empíricas que llevaron a reconsiderar el papel del cuerpo en los abordajes lingüísticos y filosóficos. “Cuerpo” –entrecomillado– porque no hay un único concepto que haya irrumpido en la discusión. Tenemos el clásico cuerpo materialista que, en nuestro caso, podría traducirse *grosso modo* como el cerebro y el sistema nervioso central y su actividad. Sin embargo, no es ese el único cuerpo que reclama su lugar en el esquema explicativo de lo mental y el lenguaje. Hay un cuerpo, constituido básicamente por herencia genética. Pero este, tomado aisladamente, parece conducir a un desarrollo tan reduccionista como el primero. Hay un cuerpo como “esquema corporal”, como posibilidad de acción, que implica la potencialidad sensorio-motriz. Hay un cuerpo que en clave evolutiva encarna el acoplamiento con el ambiente. Hay cuerpos, que se enlazan socialmente y se construyen en el tejido del lenguaje. Hay cuerpos que constituyen sistemas de procesamiento de información que se extienden hacia otros cuerpos y dispositivos. Estos son algunos entre otros cuerpos que tienden un puente entre mente, lenguaje y cuerpo. Han sido considerados a veces de modo aislado, otras de modo complementario.

Asimismo, a medida que se ha desplazado, ampliado, renovado el rol de estos cuerpos en las reflexiones teóricas sobre el lenguaje, también se han desempolvado y para ser reconsiderados, conceptos vinculados. El significado, la representación, la experiencia, el gesto, hasta afectar el lenguaje mismo.

Teóricos de los últimos diez años, de diversas áreas disciplinares – lingüistas, psicólogos, etólogos, filósofos– han replanteado la cuestión

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH). Centro de Investigaciones María Saleme de Burnichon (CIFFyH). Córdoba, Argentina.

\* Universidad Nacional de Córdoba (UNC), Facultad de Lenguas. Córdoba, Argentina.  
irene.audisio@unc.edu.ar

semántica de si el significado está en la mente o en algún otro espacio (virtualmente hablando) de lo corporal. Será el caso que la posibilidad de comprender y comunicar trazos significativos se deba a ciertos contenidos mentales, conceptuales, desencarnados, privados, que sean expresados mediante sonidos, movimientos y acciones corporales visibles. O se deberá a las mismas acciones corporales en las que directamente se perciban los significados considerados inmanentes (Johnson, 2006).

Consideremos aquí el comentario de Johnson (2006) respecto a que no se han podido dar ni podrían darse, tal vez, argumentos completos y definitivos ni evidencias decisivas sobre el carácter inmanente, no conceptual del significado. A pesar de la parcial coincidencia con ese comentario –efectivamente siguen vigentes y eficaces los marcos teóricos que anclan el significado en aspectos puramente conceptuales, abstractos, desprovistos de consideraciones corporales–, asumimos que se desarrollaron investigaciones que dan indicios de la incidencia potente de la corporización del significado, más allá del cerebro.

En ese sentido, se han realizado avances en estudios sobre la adquisición del lenguaje que centran el interés en los gestos y las acciones corporales. Las conclusiones de esas investigaciones no afectan solamente a la evolución filo y ontogenética del lenguaje, sino a los marcos teóricos más amplios desde los cuales se elabora la definición misma de significado y de lenguaje. De ese modo, se ha dirigido hacia definiciones multimodales del lenguaje.

## En retrospectiva

Podría preguntarse si recién ahora se comienza a considerar el cuerpo y los gestos, en particular, para hablar del lenguaje.

Desde la Antigüedad los gestos fueron foco de atención. Cicerón y Quintiliano desarrollaron estudios de las formas en las que los gestos intervienen en la presentación retórica del discurso. El mismo interés revivió en el siglo XVII cuando se releen a los maestros romanos de la retórica. En el siglo XIX, Darwin en *La expresión de las emociones en el hombre y en los animales* (1872/1984), introduce un giro naturalista con respecto a las expresiones faciales y corporales.

En el siglo XX, con la aparición de la psicología dinámica, resurge el interés por los gestos desde la búsqueda de lo que los movimientos corporales tenían para revelar sobre la personalidad.

Ahora bien, a pesar del interés que despertaron, no se consideraba que esos movimientos corporales visibles de los gestos debían revelar un tipo de vehículo particular para la expresión del pensamiento, no se esperaba que los gestos jugaran algún rol en la representación de las ideas. En 1940, David Efron es uno de los primeros en investigar el patrón que vincula gestos manuales con el habla. Por ejemplo, puso atención en cómo los hablantes pueden usar sus manos para organizar sus ideas y dar representaciones gráficas de algunos aspectos de sus pensamientos. Sin embargo, estos estudios fueron opacados ante el enfoque que reducía los procesos cognitivos exclusivamente al lenguaje verbal.

Recién hacia los años 70 del siglo XX, surgieron diferentes investigaciones sobre la vinculación entre lo que el hablante piensa y los gestos que realiza. Se reconoció una vinculación semántica compleja entre los movimientos de las manos y el habla. Movimientos que en el contexto comunicativo constituyen acciones (Gallagher & Zahavi, 2012).

La vinculación entre acción y pensamiento ha sido también conquistada en el plano teórico en las últimas décadas. Históricamente, la mente y el cuerpo han sido estudiados como entidades separadas. Este dualismo cartesiano condujo a pensar la cognición y la acción como dominios independientes. Sin embargo, según recientes investigaciones, la mente no es un sistema aislado, sino más bien está integrada al sistema sensorio-motriz y que nuestras representaciones de objetos y eventos están vinculadas con nuestra experiencia de las acciones en el mundo (desde Merleau-Ponty hasta los enfoques corporizados actuales).

En los '80, David McNeill publica investigaciones dirigidas a mostrar evidencia empírica y marco teórico acerca de la unidad dialéctica entre gesto y palabra. También Genevive Calbris publica una completa investigación sobre el modo como los gestos de los hablantes o gestos co-verbales se organizan en tanto movimientos físicos.

Dadas teorías clásicas que siguen prevaleciendo como, por ejemplo, la de Saussure, el viraje hacia los aspectos corporizados, implica una revisión de qué será estudiado en el campo del lenguaje. Deja de justificarse la selección de la lengua como objeto privilegiado de estudio y se abre la investigación a las múltiples formas de comunicación en tiempo real. No se trata solo de analizar estructuras permanentes de un sistema convencional, sino de visibilizar la acción comunicativa en escenas reales entre hablantes ampliando el estudio a múltiples aspectos y enfoques multimodales intervinientes en la comprensión lingüística.

El estudio de los gestos provee evidencia para las bases corporizadas del pensamiento. Los gestos ofrecen un campo relevante para la lingüística cognitiva porque manifiestan físicamente el principio de las metáforas conceptuales, cuyo dominio base son las acciones corporales. Los gestos pueden desplegar en el espacio los elementos del dominio fuente de las metáforas, cosa que no es posible para las expresiones metafóricas verbales (Cienki, 2008).

En ese marco de discusión, nos dedicaremos a desarrollar algunos aspectos centrales de la gestualidad en la comunicación que conducen a considerar el significado como un elemento inmanente de las experiencias corporales.

Nos concentraremos especialmente en la hipótesis de que es central para el estudio del papel de los gestos en el campo del significado, el detenernos en los “gestos metafóricos”. Tal tópico de la clasificación de los gestos fue remarcado por Adam Kendon (2004), retomado por David McNeill (2005, 2012), pero principalmente estudiado por Cornelia Müller (2004; Cienki & Müller, 2008). Como esperamos mostrar, tal tipo de gestos nos resultan relevantes para la identificación del proceso por el cual un movimiento corporal adquiere carácter gramatical y semántico en el plano comunicativo. De ahí la relevancia de esa clase de gestos para el estudio corporizado del lenguaje.

## **Gestos naturales**

Tomasello (2008) clasifica los gestos comunicativos más primitivos, naturales, en deícticos e icónicos. De estos gestos proviene el lenguaje tanto filogenéticamente como en el desarrollo y adquisición del lenguaje en los niños.

- Deícticos: los que hacemos para guiar en el espacio la atención de un receptor hacia algo ubicado en el entorno perceptual inmediato (Kendon, 2004; Kita, 2003; Tomasello, 2008).
- Icónicos: las acciones que buscan orientar la imaginación de un receptor hacia algo que, por lo general no está en el entorno perceptual inmediato, lo que se consigue con ciertos comportamientos: simulación de una acción, relación u objeto. Es decir, representan simbólicamente cosas reales, pero en su ausencia (Tomasello, 2013).

Según este autor, estos gestos naturales poseen gran potencialidad comunicativa porque ambos buscan orientar la atención del receptor o su imaginación hacia algo determinado. En ese sentido, constituyen actos de referencia que se proponen conseguir que el receptor infiera la intención social del comunicador, lo que este quiere que el receptor haga, sepa o sienta. Es por ello que pueden considerarse los principios básicos de la comunicación humana.

Tomasello sostiene la tesis según la cual los móviles comunicativos más básicos, no difieren entre los gestos naturales y los de las lenguas codificadas. En ambos casos pueden resumirse en:

- *Pedir*: quiero que hagas algo para ayudarme.
- *Informar*: quiero que sepas algo porque pienso que te ayudará o te interesará.
- *Compartir*: quiero que sientas algo de modo que los dos podamos compartir actitudes o sentimientos (experimentar emociones o actitudes en común con otros).

Esto nos plantea una cuestión. Los gestos naturales tienen carácter prelingüístico; que contienen mínima información en sí mismos, presentándose sumamente dependientes del contexto intersubjetivo compartido.

En el caso de Tomasello, postula que esto dependerá fundamentalmente del terreno conceptual común y del marco atencional conjunto del comunicador y el receptor en ese momento. Tanto los gestos espontáneos como la comunicación gramaticalizada, dependen de la misma infraestructura de intencionalidad compartida. Pero además de ese *background* psicológico evolutivo, en algunos casos, puede desarrollarse una justificación de orden cultural. Es decir, este autor hace depender la potencia comunicativa de los gestos de la capacidad de cooperación de la especie.

Más específicamente, podríamos seguir un curso que va de las experiencias corporales, vía metáforas, al contenido cognitivo que se jugaría en la comunicación ya gramaticalizada. Cuando referimos a metáforas, seguimos la tradición de Lakoff y Johnson (1980, 1999). Teniendo en cuenta las investigaciones de Cienki (2008), sabemos que las metáforas conceptuales provienen de mapeos conceptuales entre dominios. Tienen, por eso, bases que son de dominios cruzados. Este carácter metafórico del

conocimiento, abriría la posibilidad a pensar que el carácter de metáfora conceptual puede partir de dominios base que sean gestuales y que proporcionen contenido semántico a la comunicación verbal, es decir, que proporcionen significados que se perciban en el gesto mismo y no estén comprendidos en la comunicación verbal.

## Gestos metafóricos

Tomaremos a Cienki y Müller (2008) quienes consideran ciertos gestos como metáforas en el sentido de que pueden ser vistos como una expansión de ciertos signos en la modalidad visual del enfoque sobre las metáforas conceptuales de Lakoff y Johnson (1980), que trata las metáforas conceptuales como una forma de cognición corporizada. En el mismo orden, también se estudia el gesto con relación al pensamiento.

Cabe aclarar que nos referimos a los gestos co-verbales, en general, movimientos espontáneos de manos y brazos (Kendon, 1988, los clasifica como gesticulación), no a aquellos convencionales que ya han adquirido un significado fijado culturalmente (emblemas).

La gesticulación puede brindarnos importante información acerca de los procesos no conscientes del proceso de generación de significados en el habla. En especial, aquellos denominados gestos metafóricos.

McNeill (2005) clasifica los gestos espontáneos co-verbales en cuatro categorías: rítmicos (*beats*), déicticos, icónicos y metafóricos.

Los gestos metafóricos podrían ser ejemplificados con el caso de la metáfora del discurso como un conducto, según la cual, cuando hablamos de ideas y cuestiones referentes al discurso, nuestras manos hacen un movimiento como si estuvieran mostrando un tubo por el que se desplaza algo.

Otro ejemplo de gesto metafórico, tiene que ver con las nociones temporales. La ubicación del pasado y del futuro en una línea imaginaria de izquierda a derecha o el delante y atrás. La ubicación de las manos según esa línea indicando la temporalidad. Como estudió Calbris (1990), en algunas ocasiones los gestos metafóricos ocurren sin que tengan correlato en la verbalidad.

Estos autores afirman que hay ocasiones cuando se señala el tiempo con el gesto, dando a entender si es antes o después, pasado o presente, en las que el gesto en sí mismo está ofreciendo la información temporal que no aparece en el discurso oral. Este tipo de gesto provee evidencia de un mapeo entre dominios que podría ser la fuente de los gestos.

Consideramos que el gesto metafórico tiene su especificidad, a pesar de compartir con los gestos icónicos su carácter pictórico (McNeill, 1992), es que representan un dominio abstracto. En los gestos icónicos, el referente es una acción, entidad o relación concretas, o ciertos aspectos de estas. De modo más general, podría pensarse que la metáfora conceptual constituye un procedimiento cognitivo consistente en entender una cosa en términos de otra. Ello implicaría que los gestos metafóricos serían aquellos que incorporan un mapeo multidominio (Cienki & Müller, 2008). Entonces, los gestos metafóricos son definidos (Cienki & Müller, 2008) como acciones corporales que realizan un mapeo multidominio para expresar pensamientos abstractos o sentimientos.

## Conclusiones

Los gestos metafóricos podrían considerarse así, un eslabón entre la experiencia corporal y la conceptualización. Esto es viable ya que, como hemos desarrollado, las metáforas gestuales están atravesadas por la multimodalidad. A través de diferentes vehículos representan un concepto abstracto. El contenido pictórico de estas acciones espontáneas co-verbales, representan una idea abstracta, logrando representar “una imagen de lo invisible” (McNeill, 1992, p. 14). Las metáforas conceptuales se plasman, en este caso, en combinaciones multimodales, como es el caso de las metáforas verbal-gestuales (Müller, 2004) que nos interesan, incluso puede que se presenten solo en el registro gestual.

Se trata de representaciones de conceptos abstractos a partir de una modalidad perceptual. Más precisamente, movimientos de las manos que indican o representan el dominio fuente de una metáfora conceptual.

El enfoque de la metáfora conceptual que se funda en la experiencia del cuerpo plantea que el pensamiento es el resultado de las experiencias corporales, el lenguaje debería hacer un uso extensivo de estas experiencias y procesos sensorio-motrices. En este sentido, entre todos los gestos que realizamos para comunicarnos, los gestos metafóricos tendrían la potencialidad de mostrar cómo desde una experiencia corporal se representan conceptos abstractos. Así, podrían considerarse un eslabón cuya relevancia es innegable para incorporar el conocimiento y recuperar así al cuerpo en el lenguaje.

## Referencias

- Calbris, G. (1990). *The semiotics of French gestures*. Bloomington: Indiana University Press.
- Cienki, A. (2008). Why study metaphor and gesture? En A. Cienki & C. Müller. (Eds.), *Metaphor and gesture* (pp. 5-26). Amsterdam, Netherlands: J. Benjamins Publishing co.
- Cienki, A., & Müller, C. (2008). Metaphor, gesture, and thought. En R. W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 483-501). Cambridge, UK: Cambridge University Press.
- Darwin, C. (1984). *La expresión de las emociones en el hombre y en los animales* (T. R. Fernández Rodríguez, trad.). Madrid: Alianza. (Obra original de 1872)
- Gallagher, S. & Zahavi, D. (2012). *The phenomenological mind*. London; New York: Routledge.
- Johnson, M. (2006). Mente encarnada: de Dewey a Damasio. *Dédalo*, 135(3), 46-54.
- Kendon, A. (1988). How gestures can become like words. En F. Poyatos (Ed.), *Crosscultural perspectives in nonverbal communication* (pp. 131-141). Toronto: C. J. Hogrefe, Publishers.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, UK: Cambridge University Press.
- Kita, S. (Ed.) (2003). *Pointing: Where language, culture, and cognition meet*. Mahwah, NJ: Erlbaum.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago; Londres: The University of Chicago Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Nueva York: Basic Books.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

- McNeill, D. (2005). *Gesture and thought*. Chicago; London: University of Chicago Press.
- McNeill, D. (2012). *How language began: Gesture and speech in human evolution*. Cambridge: Cambridge University Press.
- Müller, C. (2004). *Metaphors, dead and alive, sleeping and waking: A cognitive approach to metaphors in language use* [professorial dissertation]. Freie Universität Berlin.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.



## Alcances y límites epistemológicos de la metáfora, según Ricoeur

Carlos Emilio Gende\*

### I

Independiente de la teoría que intente explicarla, sabemos que para la interpretación de las metáforas se requiere de algún proceso distinto y/o adicional al empleado para dar cuenta de las expresiones literales. Esto se debe a que la impresión de absurdo, contradicción, falsedad evidente, expresión defectuosa, entre otras caracterizaciones, es lo que se manifiesta en la primera recepción de una frase, ante la cual, a la vez, en una segunda recepción (que no lo es en el orden del tiempo sino de los procesos que re-construimos a posteriori) intentamos interpretarla como metafórica. Así, si ante la expresión literal procuramos que el proceso interpretativo empareje expresión y recepción del mensaje según el contenido de lo dicho, ante la metafórica partimos de la presunción de que si persistiésemos en ello no lograremos otra cosa que declarar el error o sinsentido.

Y dado que el intento es entenderla, sabemos también que el modo de superar el infortunio inicial es atribuirle sentido. No podemos asignarle un sentido unívoco, como sí lo intentamos con la expresión literal, pues, como señalamos, de ese modo sólo podríamos declarar el absurdo, pero tampoco nos resignamos a que carezca de él, pues de ese modo persistiríamos en detectar sólo el error, en este caso como resultado de una evaluación que sólo reconoce la expresión mal formada. De allí que, si no es sólo uno, ni ninguno, la recepción procura algún sentido, producido por la metáfora, y en eso distinto al que proporciona lo literal, pero sentido al fin. Es usual entonces suponer que la recepción de las metáforas se predisponga a dar con varios sentidos, múltiples, diversos, incluso muy distintos; sin embargo, ¿cualesquiera, incluyendo en esto aquellos sentidos que no coincidirían con el propuesto por su autor? Esta última posibilidad, a modo de pregunta, es algo compartido con la interpretación de las expresiones literales, pues también ante ellas podemos dudar si la

\* Universidad Nacional del Comahue.  
soutopo@gmail.com



justeza de nuestra interpretación debiera o no medirse con la intención del autor, con su “querer decir”. Sin embargo, es sabido que no es eso lo que la legitimaría, si por tal cosa suponemos la posibilidad de captar el sentido por relación con la psicología del autor. Lo que se procura es más bien dar cuenta de la intención de la obra, es decir, del sentido en principio auto-contenido, que exhibe la elaboración de un autor que con su obra innova, pero que como configuración debe revelar sus estrategias, modalidades y estilos de composición desde ella misma. Ahora bien, si no corresponde emplear como criterio de interpretación la remisión al querer decir del autor en una obra elaborada según principios de configuración literal (un ensayo, un tratado, un informe, etc.), pareciera menos aún poder recurrirse a ello ante una expresión metafórica, dado el desajuste insalvable entre el querer decir y lo dicho.

No obstante, tal vez sea ese el caso en las expresiones metafóricas, del que sostuvimos hay que tener en cuenta dos procesos: el que lee lo literal y lo declara absurdo y el que a la vez lo lee como metafórico y entonces supera al primero. Si ante lo literal el contenido de lo dicho se autonomiza del querer decir porque suponemos que coincide, en lo metafórico se relacionan ambos precisamente porque no coinciden. El emisor de una metáfora, como sostiene John Searle (1979/2005), dice “S es P” queriendo decir “S es R”, por lo cual todo indica que interpretarla exige la reconstrucción de los pasos que pudieran reenviarme como receptor al querer decir de su autor a partir de lo que ha dicho, como único modo de tratar con ella; e insisto, interpretarla no es decir lo que me venga en ganas sino esforzarme por dar cuenta de lo que allí quiso decirse.

Pero, aquello que quiso decirse, lo intentado, no está presente como tal ni disponible para el receptor que la lee. ¿Cómo hacer entonces si acaso lo que procuramos es vencer la resistencia impuesta por el absurdo y ahora potenciado por la ausencia de un querer decir no recuperable como instancia que procure asir el sentido?

Ante todo, habría que distinguir si de lo que estamos hablando es de condiciones de hecho o de derecho. Si lo primero, es un hecho la ausencia del emisor y por ende sus contenidos de conciencia; si lo segundo, es de derecho la presunción del querer decir como requisito. En tanto receptor, presupongo al menos que hubo una intención declarada, precisamente, de no haber dicho lo que literalmente se expresa, pues de otro modo sólo cabe atribuir error, pero no porque pretenda reconstruir de hecho la intención tal o cual que estuvo por detrás y antes de la emisión metafórica.

De allí que el trabajo del receptor consistiría en eso, en un trabajo, la interpretación como elaboración paciente y esforzada en pos de un sentido.

La presunción de un querer decir distinto a lo dicho cuando leo una expresión como metáfora se resuelve como trabajo del que interpreta. Mal encaminado si va en busca del querer decir, ya inasible de hecho, que satisfaga su demanda de sentido, pero bien orientado si descubre en eso una invitación a elaborar las paráfrasis que se lo procuren. Desde las mismas condiciones, ahora para el receptor, vemos que su trabajo puede llevarlo de hecho ante la situación de encontrarse con múltiples interpretaciones, e incluso contradictorias. Sin embargo, de derecho, aceptar la invitación a entenderla exige una evaluación que le permita detectar cuál/es de ellas logra hacerlo pensar un asunto de modo novedoso, creativo.

## II

Hasta aquí, una suerte de *desiderátum* en mi descripción: reconozco una relación tensiva, no reductiva, entre lo literal y lo metafórico, pues objeto la pretensión de agotar esta última por traducción a la primera, así como rechazo las declaraciones poco interesantes que subsumen genéticamente el empleo literal en el metafórico.<sup>1</sup> De allí que, propongo, lo provechoso de su examen reside en admitir la tensión no resoluble que exige una suerte de trato bicéfalo: leo en la oración una contradicción o un absurdo y a la vez leo un sentido posible que me predispone a ver las cosas de otro modo. Sin embargo, bien puede ser todo esto mera expresión de deseos que ante la metáfora se niega a reducirla a lo que literalmente podría haber enunciado vía una traducción, y tampoco acepta quedarse inerte como si fuera una oración bella pero inefable, de la que nada puedo decir. Es cierto, quiere algo más, interpretarla; pero, ¿cómo, bajo qué condiciones?

Sostendré que la tensión reductiva entre lo literal y lo metafórico puede superarse vía una tensión productiva entre lo conceptual y lo metafórico, pues allí se verá el trabajo de la interpretación.

Para mi examen me valgo del aporte de Paul Ricoeur en *La metáfora viva* (1975/2001); allí su exigencia es radical y de consecuencias contrarias a las que suelen extraer algunos estudiosos de estos temas. Se trata de la asignación a la metáfora de una relación con el sentido específica, inconfundible y de la que puede obtenerse una experiencia de la lingüis-

<sup>1</sup> Reviso esta posición en detalle para el caso de la teoría de Lakoff y Johnson en el capítulo 12, de la segunda parte de Gende (2019).

tividad no subsumible a otras; pero por eso mismo inhabilitada para absorber toda posición de sentido, es decir, no es ubicable en el lugar del fundamento último que, en ese caso, operaría por traslado e invención ex nihilo. Y no sería así, porque hay al menos otra relación con el sentido, que también involucra procesos de interpretación, pero con recursos y objetivos distintos: el concepto.

Podríamos decir que la metáfora se “mide” en esta obra con el concepto, que ambos se necesitan y son mutuamente irreductibles. Del intercambio entre ellos obtendremos resultados muy provechosos, por ejemplo, sobre las nociones de verdad y referencia, nociones que no solo no quedarán disueltas por *more metafórico*, sino que, al contrario, exigen una redescrición que nos permita aplicarlas a ámbitos respecto de los cuales no esperamos una teoría en términos de pretensión de verdad.

El enfoque que oficia de punto de partida y sobre el que Ricoeur vuelve de modo cada vez más elaborado, es el que se conoce como teoría descriptiva de la referencia. En principio, con esto solo ilustra un funcionamiento interno a la estructura del enunciado, pues dependerá de cuánto y cómo aporte el sentido para “iluminar” aquello que enuncia. Ricoeur destaca de esta versión que la referencia es lo presupuesto, diría mejor, es lo que se ambiciona; es un presupuesto, pero entendido como pretensión de trasfondo que justifica el procedimiento general del lenguaje: relacionarse con lo extralingüístico. Por cierto, no necesariamente la referencia estará garantizada en su efectividad por el sentido, pero tampoco puede desconocérsela en una caracterización cabal del enunciado; es algo sin lo cual dejaría de entenderse por qué y para qué hablamos.

Así, por un lado, el funcionamiento del enunciado se sostiene en una pretensión de verdad, por otro, su estructura dispone las condiciones de posibilidad para satisfacerla (función identificante y función adscriptiva). Claro, insistimos, esto no significa que las garantice, y no porque repitamos el clásico *dictum* “todo signo tiene designado, pero no por ello tiene denotado”. Si por principio hablar es hablar de algo a alguien, según ciertas características, designar es denotar. Pero como tampoco podemos ignorar la experiencia del discurso vacío, declarada o intencionadamente vacío, persiste el problema acerca de cómo delimitar el mobiliario del mundo de modo que podamos establecer el alcance del sentido sin confundirlo con el sinsentido.

### III

Por ello, es decisivo tener en cuenta la estructura del enunciado según lo descrito, pero ahora para recuperar su rendimiento gramatical desde dos operaciones que Ricoeur destaca: la identificación singularizante para el sujeto del enunciado y la caracterización universalizante, propia de su predicado. Estas operaciones revelan un dinamismo “doble y cruzado” en el acto de enunciar, “en el que todo progreso en dirección al concepto tiene como contrapartida una exploración más eficiente del campo referencial” (Ricoeur, 1975/2001, p. 393). Si admitimos esto, vemos que la función predicativa, por sí sola, actuaría de modo abstracto y genérico adjudicando un significado estabilizado en el campo semántico del vocabulario, pero en el contexto de un enunciado permite focalizar un aspecto determinado de un referente, el que a su vez ha sido singularizado por la función identificante. De modo que, si vemos el funcionamiento desde la predicación, podemos aprender a detectar variaciones de sentido al emplearla en referentes distintos. Y si lo vemos desde la singularización, la aparición de nuevos referentes obliga a realizar descripciones más ajustadas. El dinamismo en este intercambio entre sentido y referencia, como “trabajo inacabado” de la significancia, sugiere la posibilidad de extender el campo referencial “más allá de las cosas que podemos mostrar, y aún más allá de las cosas visibles y perceptibles” (Ricoeur, 1975/2001, p. 393). A la vez, como proceso de formación del significado se consigue una estabilización semántica, pero transitoria, pues está disponible a continuas transformaciones fruto de su empleo en situaciones nuevas.

Cabe agregar que el empleo del enunciado, orientado tanto a la abstracción como a la concretización, supone un acto por parte de quien enuncia, un acto de enunciación al cual cabe describir como acontecimiento de sentido. De este modo, aun el empleo habitualizado del sistema de la lengua, en la medida que requiere el concurso de un locutor muestra al acto de significar como una iniciativa.

Ahora bien, esta caracterización del funcionamiento del lenguaje en general se intensifica en la enunciación metafórica, pues pretende operar sobre dos campos de referencia desde dos niveles de significación. Con el primer nivel realizamos un acople más o menos habitualizado, el cual, según cómo resulte, puede eventualmente impactar en una ampliación del sentido o en una especificación de la referencia, o en ambos, pero siempre

como resultado de un dinamismo que se asegura su estabilidad. De ese modo, vuelve previsible y controlable el proceso de ajuste del lenguaje con el mundo. Con el segundo nivel, en cambio, caracterizamos desde una “red de predicados que funcionan ya en un campo de referencia familiar” (Ricoeur, 1975/2001, p. 394) otro campo, todavía desconocido, y por ello carente de descripción propia. El significado, entonces, proviene del campo conocido, que contribuye a la configuración de la referencia desconocida.

Como vemos, la relación entre sentido y referencia queda ahora mejor descrita en términos de retroalimentación, pues el mismo sistema de significación habilita, dada su disponibilidad a recibir los resultados de nuestros intentos de nuevas descripciones y nuevas referencias, a comprender mejor, expandiendo la capacidad descriptiva del sentido o precisando el referente. Ahora bien, si de lo que se trata, como ocurre en la enunciación metafórica, es de transferir un campo a otro, el trabajo con el significado consistirá también en aprovechar la disponibilidad y apertura que por su propia naturaleza tiene el sistema semántico para dejarse apropiar por una referencialidad anunciada, presunta, pero no asegurada.

Insistimos, en toda actividad enunciativa la lingüística va del sentido a la referencia, como pretensión de referencia, ambición que puede satisfacerse por un ajuste cognitivo con el mundo de los objetos. Pero en el caso de la enunciación metafórica, esa pretensión queda detenida y postergada: detenida por la referencia conocida que se corresponde con el significado literal (y obliga al que enuncia la metáfora a decir “no es”) y postergada hasta dar con la referencia que se desdobra a partir de la primera (y lo provoca a decir “es”). El sistema de significaciones, por su dinamismo y apertura, le muestra que eso siempre es posible, que toda estabilización es transitoria y sujeta a cambios para cualquier lenguaje. Por eso no se trata de una suplantación de significados, de un reemplazo según criterios retóricos o estéticos, pues no es un trabajo con el vocabulario, del que ya conocemos su plasticidad. Se trata de un trabajo con la referencia; sin embargo, no con la ya dicha sino con la presunta. Pero que se presupone, a su vez, no como ambición generalizable para todo lenguaje, pues, como vimos, es el principio orientador de toda relación entre sentido y referencia, sino que se presupone ya habiendo satisfecho una relación referencial, la que se da a instancias de ajustar el lenguaje literal con el mundo.

Si para todo lenguaje, entonces también para el literal, es cierto que establecemos una relación que ambiciona referir; cuando esto se cumple,

cuando queda satisfecha la relación de ajuste tendemos a esclerosar ese resultado y tomarlo como lo dado. Y ahí viene a cumplir su propósito el enunciado metafórico, pues toma la “energía” del trabajo realizado entre el significado literal y el mundo del que habla, para referir de otro modo, otra cosa que la que se presupone, y empleando para ello la significación resultante del trato literal.

Sin embargo, esta “vehemencia ontológica” que Ricoeur recupera para la metáfora no debería ser confundida con ingenuidad, porque la enunciación metafórica como tal, prepara, predispone, “esboza” la determinación conceptual, sostiene el autor, pero no la constituye. Con una paráfrasis un tanto provocativa, sugiero que para Ricoeur todo conocimiento conceptual comienza con la metáfora, pero no por eso está fundamentado en ella. Y eso repercute también en su descripción del concepto, pues no sólo no está fundado de modo metafórico, sino que tampoco vía lo producido por la imagen o el percepto. Más aún, lo especulativo no es producido de ninguna manera, pues es la condición de articulación y sistematización de nuestro conocimiento de todo lo producido. Ahora bien, la metáfora sí queda asociada a la producción de imágenes y por eso, del mismo modo que ellas, no puede ser constitutiva. La justificación es directa y con consecuencias gnoseológicas: “si la *imaginatio* es el reino de lo ‘semejante’, la *intellectio* es el de lo mismo” (Ricoeur, 1975/2001, p. 397).

Y a su vez por una razón a mi juicio fundamental: “significar es siempre distinto de representar” (Ricoeur, 1975/2001, p. 397). Así, los límites de la metáfora están dados por su mismo modo de realización, pues todo el trabajo de asimilación y esquematización la vuelve productiva pero no constitutiva. Su eficacia está en el plano representativo, “redescriptivo” lo va a llamar Ricoeur en relación con los modelos. Por eso, si bien puede recrear “lo semejante” y ambicionar desde esa actitud una nueva referencia, no puede fundar “lo mismo”, pues depende de este último concepto primitivo para dar con las semejanzas. ¿Pero esto quiere decir que no es significativa? No, quiere decir que el concepto no es representativo. La metáfora es significativa, precisamente porque promueve el dinamismo del sistema de significados. Pero también es representativa y entonces queda ligada a la producción de la imaginación.

Cabe destacar que Ricoeur no apunta tanto a mostrar los límites de la metáfora, como a establecer la discontinuidad entre niveles del discurso. Discontinuidad que, como vimos, tampoco descansa en una descripción indiferente que sólo la reconoce y deja las cosas como están. Por eso le

interesa destacar también que el orden conceptual no destruye ni anula el orden metafórico, así como antes sostuvo que el segundo no produce al primero.

Y aquí aparece un desplazamiento interesante en su estrategia para justificar la tensión, relacionado con la lingüisticidad y la interpretación.

En efecto, la autonomía y relación de los distintos niveles de discurso quedan ahora subsumidos en una diferencia entre “lucha por la univocidad”, búsqueda de lo claro y distinto –propio del concepto– versus peligro de equivocidad, propio de la metáfora? Si se tratara de un asunto de orden exclusivamente semántico tal vez podría presentárselo de ese modo: lo unívoco y su opuesto, lo equívoco. Pero no es de orden estrictamente semántico, sino conceptual, aunque descripto como significado; de ahí la lucha por lo unívoco que está detrás de lo conceptual aportado por lo especulativo. Y no lo es, porque no nombra el peligro de lo equívoco como su opuesto, sino que redefine todo el asunto en términos de teoría de la interpretación.

#### IV

Al respecto, así como recién veíamos insinuarse una limitación para el enunciado metafórico, ahora habrá otro tanto para el conceptual. De modo que ante la pregunta sobre las razones por las cuales el orden conceptual no logra destruir el orden metafórico, contestamos que se debe a una imposibilidad interna a la lucha por conseguir lo unívoco, que proviene de un específico impedimento interpretativo: el orden conceptual se relaciona con interpretaciones racionalizantes y reduccionistas.

Es cierto que cuando todo indicaba que la presentación de la metáfora y su productividad creativa alcanzaría sus cotas más altas, pareció detenerse su alcance al oponerle el concepto. Y cuando la descripción semántica de este último lo exhibió como la condición de lo pensable, se evalúa su dependencia de un modelo reduccionista de interpretación. Pero esto, que parece una inconsistencia en la relación de metáfora y concepto, que destruye ambas partes de los términos y los devalúa, deja de serlo si superamos la caracterización semántica, que opera como si fuera asunto del sistema de la lengua, y logramos entenderlo como resultado de interpretaciones, pues las auténticas rivales que interactúan entre modalidades distintas de discurso son las interpretaciones, es decir, las realizaciones lingüísticas de inferencia de sentido.

Así, el trabajo desde el concepto siempre podría ser restrictivo para la interpretación, pues le va de suyo intentar la anulación de la experiencia que le llega desde el enunciado metafórico. Mientras esta relación se mantenga como tensión, ninguna absorbe a su contraria, pero si el concepto logra su propósito es porque está operando una interpretación racionalizante. Se trata de la que sólo decodifica. La anulación completa de la experiencia promovida por la metáfora únicamente se consigue por una inversión del proceso interpretativo abductivo, pues en el racionalizante se procede desde la ley –o código–, hacia el caso, que o bien se subsume en aquella o bien queda descartado. Y denomino códigos a los sistemas de explicación que funcionan en la cultura como modelos muy sedimentados, tanto que clausuran toda iniciativa de sentido, paradójicamente, desde la sospecha de que se trate de un engaño de la conciencia que interpreta. Todo aquello que parezca querer decir algo inédito, bien revisado, como describe Ricoeur, no es para la interpretación racionalizante más que “tal posición del deseo, tal pertenencia de clase, tal grado de fuerza o de debilidad del querer fundamental” (Ricoeur, 1975/2001, p. 399). Así, el modelo interpretativo reduccionista es el de la sospecha ante la novedad, pero porque presupone una clave de sentido sobre la cual no se sospecha, al contrario, funge como dadora de sentido, aunque su objetivo no sea otro que desenmascarar el presunto falso sentido.

El trabajo con el concepto librado a su suerte, es decir, empleado como clave de interpretación apriorística, suele relacionarse sólo con los productos, con el sedimento de interpretaciones anteriores, y deja de ver el momento de producción y relación con la experiencia que reclama ser dicha. A su vez, en clave semántica, esto refuerza la ambición de univocidad, pues los códigos explican siempre y cuando las diferencias queden absorbidas o anuladas, nunca como posibilidades de transformación. Se sospecha de toda iniciativa de sentido porque no se sospecha jamás del modelo explicativo, unívoco, claro y distinto, que se emplea. No obstante, debería admitirse que el modelo de explicación actúa por los interpretantes, no sólo por el concepto que lo articula; es decir, actúa desde el contenido estabilizado, fuertemente sedimentado, de tratos codificados con el mundo.

La interpretación metafórica, en cambio, si bien responde al concepto, también lo hace ante “la intención constitutiva de la experiencia que intenta manifestarse según el modo metafórico” (Ricoeur, 1975/2001, p. 399). Es decir, también cabe conceder al trato metafórico una ambición

en el plano constitutivo, siempre y cuando se sostenga en un tipo de interpretación que opere en la intersección de lo conceptual y lo metafórico.

Para concluir, desde una descripción como la que ofrezco, esto se explica mostrando que el concepto fija e inmoviliza en relación con los interpretantes, de los cuales el intérprete no duda y por ello los emplea como códigos explicativos de la experiencia. Por su parte, la metáfora preserva el dinamismo de la significación, porque con ello el intérprete pretende devolverle a los interpretantes su momento de invención, y con ello exhibirlos como resultado de estabilizaciones transitorias.

## Referencias

- Gende, C. (2019). *La interpretación de las metáforas: Ensayos filosóficos del lenguaje desde el lenguaje*. Buenos Aires: Prometeo.
- Ricoeur, P. (2001). *La metáfora viva* (A. Neira, trad.). Madrid: Trotta; Cristiandad. (Obra original de 1975)
- Searle, J. (2005). Metáfora (L. Valdés Villanueva, trad.). En L. Valdés Villanueva (Ed.), *La búsqueda del significado* (pp. 583-617). Madrid: Tecnos. (Obra original de 1979)



## Tres virtudes del conservadurismo de la suposición

Rodrigo Laera\*

### Introducción

El objetivo de esta exposición consiste en presentar tres virtudes elementales del conservadurismo de la suposición y argumentar que tales virtudes son lo suficientemente beneficiosas en el desarrollo de nuestro conocimiento para que sea aceptado como una tesis epistemológica legítima, aún con sus defectos. Si bien hay variedades del conservadurismo epistémico (Vahid, 2004), la tesis conservadora que se defenderá hace hincapié en la noción de suposición, por eso se llamará conservadurismo de la suposición (CS) y puede definirse de la siguiente manera: *S* está justificado en mantener suposiciones que no están justificadas a menos que haya una evidencia disponible que la contradiga.

### Tres virtudes

El conservadurismo de la suposición (CS), en parte, fue mantenido por Sherman y Harman (2011) al sostener que el conocimiento descansa en suposiciones que uno no sabe, pero que justificadamente asume o da por sentado. Un ejemplo de ello es el supuesto siempre presente de que hay un mundo externo, incluyendo coches o casas, y que uno no es un cerebro en una cubeta. Esto no quiere decir que el valor de verdad de una atribución de conocimiento sea relativo a las cosas se suponen para concederlo, sino que la verdad de la atribución se funda sobre la base de una garantía positiva en la que se podría depositar confianza en forma de evidencias; una suerte de plausibilidad a priori con la que mantenemos nuestras creencias, sin necesidad de revisarlas constantemente. De este punto de partida se desprende que no es razonable dejar de suponer la verdad de una proposición simplemente a causa de la existencia de supuestos alternativos incompatibles, cuya garantía no es mayor que los supuestos que se encuentran en funcionamiento.<sup>1</sup>

<sup>1</sup> Esto último podría ser visto como una variante más exacta de lo que Kvanvig (1989) llamó “con-

\* Instituto de Investigaciones Filosóficas, CONICET. Buenos Aires, Argentina.  
rodrigolaera@gmail.com



El CS tiene tres virtudes elementales. Estas virtudes son: (i) permite explicar por qué avanzamos en muchas de nuestras reflexiones; (ii) pone especial énfasis en el carácter social y diacrónico de la justificación y (iii) puede considerarse como una seria respuesta al escepticismo radical.

La primera virtud puede considerarse metodológicamente como la más importante y abarca tanto los contextos de la vida ordinaria como los contextos más específicos o rigurosos. En ambos contextos, revisar constantemente nuestros supuestos sin razón es cognitivamente ineficiente. Siguiendo a Harman (1986) uno no recuerda las evidencias o las fuentes de donde surgen todos nuestros supuestos, de la misma manera que no podemos citar la justificación original para todas las creencias que constantemente están operando en la vida diaria, aunque nuestra vida dependiera de ello. Hay un número sorprendente de supuestos que dan lugar a creencias más específicas y no podemos saturar nuestro cerebro recordando la evidencia de todo lo que creemos. Pero ¿hay que decir que estas creencias son todas irracionales, y aconsejar a la gente a renunciar a ellas? ¿La gente pasa a ser igualmente racional si simplemente cambia de bando, y adopta las creencias opuestas? Intuitivamente, la respuesta es “no”. Y esto es una clara manifestación de la economía del conservadurismo. Porque no revisamos constantemente los fundamentos de nuestras creencias, tomando la forma de suposiciones, es que solemos avanzar eficazmente en nuestras conversaciones; es decir, podemos mantener nuestras creencias fácilmente para centrarnos en otras creencias más críticas o que son consideradas más importantes en determinado momento. Esta virtud puede ser ampliada al carácter de cierto quehacer epistemológico en general. Sobre todo, cuando, desde el análisis epistemológico, se pretende pensar en el conocimiento como la exclusiva posesión de los sujetos críticamente reflexivos, en el que ser “críticamente reflexivo” consiste en la tendencia a pedir, y la capacidad de dar respuesta a preguntas sobre qué es lo que justifica las creencias de uno o cuáles son marcas que las vuelven razonables. Pero ninguno de nosotros satisface esta condición con respecto a todas las creencias. Esto viene a indicar que un examen de la situación epistémica de las creencias es el resultado de un ejercicio altamente sofisticado que presupone una fundación masiva que en la práctica no se da.

Si bien la primera virtud parece claramente descansar en un anti-intelectualismo, la segunda caracteriza al CS de modo anti-individualista. Si

---

servadurismo no restringido”: basta con que  $S$  crea  $P$  para que  $P$  esté justificada por  $S$ .

se parte desde el punto de vista que los sujetos no están constantemente maximizando ventajas y minimizando desventajas para arribar a sus convicciones, sino que para sus prácticas epistémicas tienen estándares internalizados que guían las conductas cognitivas. Resulta bastante simple observar que los supuestos que se ponen en funcionamiento en las creencias cotidianas están socialmente aceptados. Si el anti-individualismo epistémico, a grandes rasgos, es la idea de que los estados doxásticos están determinados por hechos contingentes del entorno, entonces podemos considerar que parte de los hechos son otras creencias, es decir que implican interdependencia epistémica. Lo mismo ocurre con los supuestos que participan de las creencias; los supuestos epistémicos se comparten. Y no solo eso, parecen necesitar del hecho que se compartan para que una creencia salga a la luz. El punto importante es que los supuestos son contingentes y pueden cambiar con el tiempo. Muchos de los supuestos que funcionaban bastante bien en la época medieval, hoy ya no funcionan. El carácter dinámico del conservadurismo epistémico se enfoca en la idea de que los sujetos mantienen sus supuestos un período de tiempo, pero que puede surgir una evidencia tan fuerte que los haga saltar por los aires. Los supuestos de un sujeto no poseen un estatus especial solamente sobre la base de su sola pertenencia al sujeto. Solamente alguien que ha interactuado doxásticamente con otra persona y el mundo, bajo supuestos adquiridos que comparten, podría llegar a tener conocimiento y mantenerlo como tal.

Este punto de vista abre una relación con la epistemología social, es decir aquella posición en la que los sujetos pueden perseguir el conocimiento solo con la ayuda o en colaboración con otros sujetos. Si, por ejemplo, tomamos el camino rortyano, rechazando a la concepción tradicional del conocimiento como “precisión de la representación”, y se lo reemplaza con una noción de “justificación social de la creencia”, entonces podemos encontrar en el conservadurismo de la suposición un puente hacia el anti-reduccionismo en el que el testimonio es, en sí mismo, una fuente básica de evidencia o justificación. No importa cuán poca evidencia positiva tenga un oyente sobre la confiabilidad y la sinceridad de un hablante, o de los oradores en general, ella tiene un incumplimiento o una garantía *prima facie* de creer lo que dice el orador y de conservarlo en virtud de suposiciones tácitamente aceptadas. Así, de la misma manera que los sujetos tienen derecho a aceptar como verdadero algo que se presenta como verdadero y que es inteligible para ellos, a menos que existan

razones más contundentes para no hacerlo, también tienen el derecho a mantener dicha aceptación y que ella se transforme en un supuesto con el paso del tiempo para aceptar otras creencias justificadamente. Este quizás sea el logro más importante del conservadurismo.

La tercera virtud del CS quizás necesite de una explicación un poco más exhaustiva. Según el escéptico, *S* no sabe que tiene dos manos, porque bien puede ser un cerebro en una cubeta. Pero en contextos cotidianos a *S* le basta con percibir que tiene dos manos para saber que las tiene, pues no considera la posibilidad de que sea un cerebro en una cubeta. Ahora bien, como la creencia de que posiblemente tenga dos manos es anterior a la creencia de que posiblemente no las tenga (por ser intuitiva), entonces no hay una razón suficiente para cambiar de opinión. Por lo tanto, el escéptico debe demostrar, y no meramente enunciar una posibilidad, que somos cerebros en una cubeta para afirmar que no sabe si tiene dos manos. Como no puede dar una demostración de esto, entonces es razonable seguir atribuyendo el conocimiento de que tiene dos manos. Puesto que el contexto escéptico es diferente al contexto de la vida diaria, las creencias del sentido común están justificadas porque no hay nada en su contexto que incite a sospechar de ellas. Análogamente ocurre con los supuestos. Si el supuesto de que *S* tiene dos manos funciona para avanzar en el conocimiento que tiene del mundo, no es necesario revisar ese supuesto sin tener una buena razón para ello. Como la razón del escéptico consiste simplemente en aludir que posiblemente no tenga dos manos, genio maligno o cerebros en una cubeta mediante, entonces el supuesto de que tiene dos manos seguirá siendo correcto. Y uno puede mantener dicho supuesto siempre que sea necesario para justificar otras creencias con algún tipo de evidencia empírica.

## Dos objeciones

Se ha presentado tres virtudes de CS, ahora es momento de plantear al menos dos objeciones generales junto con sus respectivas respuestas, la primera se encarará en la noción de responsabilidad epistémica, mientras que la segunda en el anti-intelectualismo.

Se puede pensar que esta aproximación implica que la noción de creencia justificada se compromete con cierta irresponsabilidad epistémica, pues alcanza con creer algo para que uno sea responsable de su creencia. Sin embargo, este compromiso resulta insuficiente por dos motivos.

El primer motivo radica en la confusión de dos sentidos de “responsabilidad”. La noción de “responsabilidad” puede referirse tanto a que *S* se responsabiliza de lo que cree, como a que la creencia misma es responsable. Cuando la noción de responsabilidad se refiere al sujeto, se pone en juego la discusión filosófica sobre la libertad para creer lo que se cree y el derecho a estar equivocado. Por lo tanto, si somos responsables de mantener ciertas creencias, también somos responsables de su racionalidad. En cambio, cuando la noción de responsabilidad connota una propiedad de las creencias, se pone en juego una serie de atribuciones características de la concepción deontológica de la justificación, pues encierran determinadas obligaciones intelectuales (ej., no rechazar una evidencia por conveniencia). En relación con las obligaciones intelectuales, al buscar creencias verdaderas y evitar las falsas, CS reduce la justificación a meros estados doxásticos. Si las creencias se justificaran simplemente porque son sostenidas por los sujetos, entonces la noción de creencia justificada se reduciría a la de mera creencia. En consecuencia, la noción de creencia justificada se convierte en espuria y el conservadurismo se refuta a sí mismo, careciendo de sentido. En ambos casos, cobra relevancia que el concepto de justificación sea un concepto evaluativo adherido a la creencia, haciendo que esta tenga un valor epistémico adicional en el objetivo de alcanzar la verdad. No obstante, si toda justificación ocurre en el contexto de un conjunto de trasfondo de creencias aceptado y no es discutido, entonces la noción de responsabilidad se aplicaría a la creencia que surge del trasfondo y no al trasfondo. Es decir, la responsabilidad radica en cómo evaluamos tanto la justificación como las evidencias correspondientes y no al trasfondo, a no ser que tengamos fuertes razones para ello. Si la responsabilidad no se aplica a los supuestos, entonces el CS estaría a salvo de esta crítica.

La segunda objeción radica en la necesidad de preservar información no implica el CS. Por ejemplo, imagínese que las creencias de *S* se basan en el supuesto de que su pareja está embarazada. A partir de ello, formó la creencia de que es una niña. Pero no alcanzó a ver la imagen de ultrasonido, ni hay antecedentes familiares en el que predominan mujeres. No tiene ninguna justificación para la formación de esta creencia, la formó de manera caprichosa, en la base de la expresión de deseos. En tal caso, la creencia no refleja ningún aprendizaje de su parte, porque surge del falso supuesto de que está embarazada. De un falso supuesto pueden surgir falsas creencias. Por lo tanto, se pierde la intuición de que, si tengo un

supuesto, entonces tengo una buena razón para mantenerlo. Esto sucede especialmente cuando se pierde el antecedente (los supuestos) y solo se recuerda el consecuente (las creencias correspondientes). El antecedente pudo adquirirse de manera irracional, por un mero deseo, en consecuencia, también será irracional lo que surja de ello. En casos como estos, la propiedad anti-intelectualista del CS, lejos de ser una gran virtud, es un gran defecto. La respuesta a esta objeción deriva del hecho de que el CS maximiza los recursos de los creyentes y evita la pérdida de tiempo, preservando el conocimiento. Los recursos que tenemos para atribuir conocimiento, principalmente en la vida diaria, son recursos limitados, en muchos casos escasos. Estos recursos deben ser manejados de manera que nos permitan avanzar en nuestras creencias. Por lo tanto, si un supuesto es antojadizo como en el ejemplo anterior, nuestras creencias se estancarán y chocarán con las creencias de los otros, a no ser que pensemos en una suerte de esquizofrenia epistémica. De la misma manera, nuestro tiempo al igual que nuestra vida es muy limitado, indagar constantemente en los fundamentos de nuestras creencias llevaría a una pérdida de tiempo que nos dejaría a mitad de camino en nuestras empresas diarias. Es decir, desde un punto de vista práctico es inviable. Por último, cabe aclarar que el hecho de que los supuestos se adquieran de manera irracional, no dice nada acerca de que la creencia sea útil o inútil. En efecto, los supuestos irracionales pueden perfectamente conducir al desiderátum epistémico de alcanzar un número mayor de verdades que de falsedades y, desde ese punto de vista, resulta racional conservarlos, pues lo que el CS quiere afirmar no es una tesis sobre la racionalidad de la adquisición, sino, en todo caso, sobre la racionalidad de la preservación.

## Conclusión

A lo largo de este escrito se ha indagado en tres virtudes del conservadurismo epistémico, estas virtudes son suficientes para considerar que su pertinencia no solo en la teoría del conocimiento, sino también en la filosofía práctica. Su pertinencia se debe principalmente a la escasez de recursos y de tiempo con el que contamos para extender nuestro conocimiento del mundo y para manejarnos exitosamente, de acuerdo a nuestras posibilidades, en él.

## Referencias

- Harman, G. (1986). *Change in view: Principles of reasoning*. Cambridge, MA: MIT Press.
- Kvanvig, J. (1989). Conservatism and its virtues. *Synthese*, 79, 143-163.
- Sherman, B., & Harman, G. (2011). Knowledge and assumptions. *Philosophical Studies*, 156, 131-140.
- Vahid, H. (2004). Varieties of epistemic conservatism. *Synthese*, 141, 97-122.



# Atribución de creencias a animales

## Un análisis de las posturas de Davidson y Wittgenstein a la luz de la evidencia empírica

Juana Regues\*

### Introducción

¿Pueden los animales no humanos tener creencias? La psicología popular parece responder implícitamente de manera afirmativa a esta pregunta. De hecho, en nuestra interacción cotidiana con los animales necesitamos atribuirles deseos y creencias para dar cuenta de la complejidad de sus comportamientos y responder en consecuencia. Sin embargo, es necesario preguntarse: ¿Qué tipo de creencias es posible atribuirles? ¿Cuál es la justificación de dicha atribución? La principal dificultad que se presenta en torno a estos interrogantes es que, a diferencia de los seres humanos, las demás especies animales carecen de capacidad lingüística para expresar creencias y sus contenidos. Esto abre un debate filosófico entre diversas lecturas, dentro de las cuales es posible identificar dos posturas paradigmáticas: la de Donald Davidson y la de Ludwig Wittgenstein.

La postura de Davidson, también denominada lingüista (Glock, 2000), plantea que sólo se le pueden atribuir creencias a criaturas capaces de manifestarlas en un comportamiento lingüístico (Davidson, 1982/2001, 1984). Esta tesis depende de la forma en que Davidson caracteriza la creencia: ésta es eminentemente conceptual y holista. Tal noción de creencia limita su atribución a criaturas cuyo comportamiento exhiba capacidades lingüísticas y meta-representacionales, a saber, únicamente a los humanos lingüísticos.<sup>1</sup> Por otro lado, la postura de Wittgenstein, también denominada postura intermedia (Glock, 2000), sostiene que es legítimo atribuir creencias de tipo simple a criaturas sin lenguaje. Esta tesis depende del carácter público y comprobable según el cual Wittgenstein concibe los juegos del lenguaje de atribución de creencias y de atribución

---

<sup>1</sup> Davidson no incluye a los niños prelingüísticos dentro de las criaturas a las cuales considera legítimo atribuir creencias.

\* Universidad Nacional del Sur. Bahía Blanca, Buenos Aires, Argentina.  
reguesjuana@gmail.com

de contenidos conceptuales. Dicha caracterización resulta en una noción de creencia cuyo contenido es un hecho objetivo y no una representación mental interna. Esto habilita la adscripción de creencias a animales, aunque con ciertas restricciones: sólo podemos atribuirles estados mentales que pueden manifestarse en un comportamiento no lingüístico.

En el presente trabajo argumentaré que el holismo conceptual sobre el que descansa la tesis de Davidson resulta inadecuado para dar cuenta del comportamiento animal. Para apoyar este planteo presentaré los resultados experimentales de una serie de estudios realizados con diversas especies de simios. Mostraré que, aunque los argumentos de Wittgenstein resultan más adecuados en términos explicativos, la noción de creencia que dicho autor considera válida para la atribución también resulta insuficiente. Finalmente, brindaré una definición de creencia simple que satisfaga los criterios wittgensteinianos de atribución y que permita dar cuenta de la evidencia empírica presentada.

### **La postura lingualista de Davidson**

En su artículo “Rational animals” (1982/2001), Davidson sostiene que sólo podemos atribuir creencias a criaturas con habilidades lingüísticas. Esta tesis depende de tres argumentos: el argumento del carácter holista de la creencia, la preeminencia de la creencia entre las actitudes proposicionales y el argumento del carácter conceptual de la creencia.

El argumento del holismo plantea que no se puede tener sólo una creencia, sino que tener una significa tener un amplio entramado de creencias en la que esta se inscribe. Según esta perspectiva el contenido y la identidad de la creencia dependen del lugar que ocupa en la red en la que se encuentra. Para Davidson la relación entre las creencias es de tipo lógico, es decir que se implican mutuamente. En consecuencia, tener una creencia equivale a tener un gran patrón de creencias lógicamente coherentes.

Un ejemplo clásico de la literatura puede resultar útil para comprender mejor los argumentos de Davidson: vemos un perro perseguir un gato hasta un árbol. El gato trepa el árbol, salta al techo del vecino y se va. El perro, que no vio al gato escapar por el techo, continúa ladrando hacia la copa del árbol. En una situación de este tipo, explicamos el comportamiento del perro a partir de la atribución de la creencia “*el gato está en el árbol*” (Malcom, 1973). Sin embargo, de acuerdo con el argumento del ho-

lismo, para atribuirle al perro esa creencia, este debería ser capaz de tener muchas otras creencias sobre lo que es un gato, un árbol, etc.

El postulado de la preeminencia de la creencia entre las actitudes proposicionales plantea que no sólo son las creencias las que deben estar enmarcadas en un entramado de creencias, sino que toda actitud proposicional depende en su particularidad de un *background* de creencias de distintos tipos. Según Davidson, para ser capaces de desear o de intentar algo, es necesario tener tres tipos de creencias: creencias generales (“*el gato es un ser vivo*”), creencias particulares (“*el gato está en el árbol*”), y creencias lógicas (“*si el gato no está en el árbol entonces está en el techo*”). De acuerdo con este argumento, a menos que estemos dispuestos a admitir que el perro es capaz de realizar un razonamiento inferencial del tipo que demanda la creencia lógica o que este es capaz de generalizar, no resulta probable atribuirle ningún tipo de actitud proposicional.

Finalmente, el argumento del carácter conceptual de la creencia se divide en dos premisas: (i) Para tener una creencia es necesario tener el concepto de creencia y (ii) Para tener el concepto de creencia es necesario tener lenguaje. Con (i) Davidson se refiere a la posibilidad de tener una creencia sobre una creencia, esto es, a la capacidad para el pensamiento reflexivo. La criatura en cuestión debe ser capaz de evaluar epistémicamente sus creencias y de corregirlas. Únicamente de esa forma podemos decir que la criatura domina el contraste entre lo subjetivo y el mundo objetivo. Por otra parte, (ii) alude al carácter intersubjetivo de la posesión de creencias: la criatura debe ser capaz de comunicar e interpretar las propias creencias sobre el mundo objetivo y las de los demás, y esto sólo es posible a través de la comunicación lingüística.

### **La postura intermedia de Wittgenstein**

A diferencia de Davidson, Wittgenstein no abordó el problema de la atribución de estados intencionales a animales de forma específica. Sin embargo, en varias de sus obras podemos encontrar argumentos para una teoría de la atribución, así como también se puede observar el interés por su aplicación a los animales no lingüísticos (Wittgenstein, 1953/2009a, 1967/2009b). En su obra *Investigaciones filosóficas* plantea que es legítimo atribuir creencias de tipo simple a los animales no humanos.<sup>2</sup> Esta tesis

<sup>2</sup> Véase: “Podemos imaginarnos a un animal enojado, temeroso, triste, alegre, asustado. Pero ¿esperanzado? ¿Y por qué no?” (Wittgenstein, 1953/2009a, p. 531). También §25, §281, §357, §650.

depende de la forma en que caracteriza el juego del lenguaje consistente en atribuir conceptos y del carácter público de la noción de creencia.

Para Wittgenstein la atribución de contenidos es un juego del lenguaje específico que jugamos de forma cotidiana en nuestra sociedad. Este juego consiste en atribuir ciertos conceptos vigentes en una comunidad (que dependen de un conjunto de prácticas determinado) cuando observamos cierto patrón de comportamiento regular que nos parece semejante a estos. Este juego depende de tres factores: la concordancia, la comprobabilidad, y los criterios de comprobación (Kripke, 1982/2006).

La concordancia remite al hecho de que, si las prácticas de una comunidad no coincidieran, no existiría la base para la formación de determinados conceptos y, por lo tanto, para su atribución. Por otro lado, la comprobabilidad es la posibilidad de demostrar en la práctica si hay concordancia. Finalmente, Wittgenstein pone el énfasis en que los criterios que comprueban la concordancia deben ser externos. Esto quiere decir que, para juzgar que alguien lleva a cabo un proceso interno como “creer que P” o “pensar que P”, hace falta observar determinadas actitudes que den cuenta de ello.

El carácter público de los criterios de comprobación juega un papel central para la teoría de la atribución de creencias y para la noción misma de creencia que plantea esta postura. Según estos criterios, el juego de atribución de creencias consiste en un conjunto específico de acciones y actitudes que un individuo manifiesta en su conducta. Retomando el ejemplo del apartado anterior, para poder atribuirle al perro la creencia “*el gato está en el árbol*” es suficiente con observar que éste ladra al árbol con determinación, que dirige toda su atención a ese punto, etc.

Por otra parte, la noción de creencia también se encuentra atravesada por criterios públicos, dado que Wittgenstein considera que el contenido de la misma es un hecho objetivo del mundo y no un proceso o una representación mental interna. En el caso de los humanos lingüísticos, el hecho de expresar proposicionalmente sus creencias da cuenta de la complejidad de las mismas. No obstante, en el caso de los animales no es necesario demandarles la capacidad de representarse un hecho y comunicarlo lingüísticamente, sino sólo la capacidad de percibirlo y actuar en consecuencia. De acuerdo con esta postura, entonces, el único tipo de creencias que cabe adscribir a criaturas no lingüísticas son aquellas conformadas por meros estados informacionales del entorno.

## **Razonamiento inferencial, clasificación y metacognición en grandes simios**

De acuerdo con lo desarrollado en la primera sección, Davidson postula tres requisitos generales para la atribución intencional: incluso para acreditar creencias particulares la criatura debe ser capaz de articular lógicamente los contenidos que conforman sus creencias, debe poder generalizar y debe ser capaz de evaluar la verdad y la falsedad de sus creencias. Aunque Davidson concluye que estos requisitos sólo pueden ser reunidos por una criatura lingüísticamente competente, diversos estudios realizados en el área de etología cognitiva sugieren la presencia de tales capacidades en varias especies de simios.

Una forma de evaluación doxástica que involucra pensamiento reflexivo es la metacognición. Un estudio diseñado para evaluar las capacidades metacognitivas de los chimpancés indicó que estos animales son capaces de acceder a la información que poseen sobre su percepción visual para diseñar estrategias que les permitan resolver la tarea dada (Call & Carpenter, 2001). Durante el experimento se colocó una plataforma con tres tubos de PVC situados de forma perpendicular frente a la jaula donde se encontraba el sujeto. Se le mostró un trozo de fruta que luego se colocó en uno de los tubos y se aguardó durante cinco segundos. En la condición control, este procedimiento se realizó a plena vista del sujeto. En cambio, la condición experimental se llevó a cabo detrás de una pantalla opaca que ocultaba la plataforma. Por último, se acercó la plataforma a la jaula para que el sujeto pudiera observar los tubos y elegir uno. Si el tubo seleccionado contenía el trozo de fruta, este se le entregaba al sujeto como recompensa.

Los simios miraron más dentro de los tubos en la condición experimental que en la condición control. Este comportamiento estuvo directamente relacionado con el éxito en la tarea: los simios eligieron el tubo lleno, en mayor proporción, cuando miraron dentro que cuando no lo hicieron. Esto indica la presencia de cierto tipo de reflexividad en su pensamiento: estos animales son capaces de acceder y monitorear la información perceptual que poseen para actuar en función de ello. En este caso, los chimpancés supieron cuándo les faltó cierta información y respondieron a dicha carencia buscando información adicional. Otros estudios mostraron que tanto los chimpancés como los monos rhesus y los delfines pueden

afrontar dicha falta de información escapando de la situación (Hampton, 2001; Smith et al., 1995).

Adicionalmente, un análisis de las estrategias de búsqueda empleadas durante el experimento reveló que los simios son capaces de realizar un razonamiento inferencial por exclusión. Este hecho ha sido apoyado por otros estudios específicos (Call, 2006; Premack & Premack, 1994) que han permitido ampliar esta habilidad a gorilas, orangutanes y bonobos. Si bien la mayor parte las búsquedas fueron sistemáticas (los simios miraron dentro de los tubos hasta encontrar la recompensa), también se emplearon estrategias de selección por exclusión: en varias oportunidades eligieron el tubo lleno luego de mirar únicamente dentro de los dos tubos vacíos. La evidencia acerca de las habilidades inferenciales en estos simios constituye un antecedente significativo de la articulación lógica de las creencias requerida por Davidson. Debido al escaso nivel de abstracción involucrado en este razonamiento, es preciso considerarlo como una proto-inferencia más que como una inferencia auténtica.

Por último, un estudio de cognición social reveló que los babuinos son capaces de clasificar a los individuos simultáneamente según el rango que ocupa en su grupo y sus relaciones familiares (Bergman, Beehner, Cheney, & Seyfarth, 2003). Estos animales viven en grupos organizados de acuerdo a una jerarquía interna en la que las hembras heredan los rangos de dominancia a sus hijas (A1 y sus hijas superan en rango a A2 y a A3) y a una jerarquía externa, en la que un grupo matrilineal supera en rango a otro (A3 y sus hijas superan en rango a B1, la cual supera a C2). Así pues, el experimento consistió en reproducir ante un grupo de babuinos hembra una grabación que simulaba una inversión de rango, con el fin de medir su reacción en el tiempo: en la interacción entre dos hembras, una dominante y una subordinada, la hembra subordinada emitía gruñidos de amenaza a la hembra dominante, la cual respondía con gritos de miedo. En la condición control se reprodujo una interacción normal, consistente con la jerarquía establecida. En la condición familiar se reprodujo una secuencia que simulaba una inversión de rango donde B1 le gruñía a B4. Finalmente, en la condición inter-familiar, la secuencia simulaba una inversión de rango donde C3 le gruñía a A4.

Los resultados indican que los simios mostraron mayor perplejidad ante los dos tipos de inversión de rango que ante la condición control. A su vez, entre las condiciones de inversión de rango, la condición inter-familiar tuvo mayor efecto en los babuinos que la condición familiar. La

sorpresa que mostraron estos animales parece sugerir que, ante la presencia de un estímulo ocasionado por otro individuo de la misma especie, no buscan en una base de datos masiva y desestructurada para establecer asociaciones, sino que organizan su conocimiento de acuerdo a categorías determinadas, a saber, rango y parentesco (Seyfarth & Cheney, 2015). Estas categorías son similares a los conceptos humanos ya que no pueden reducirse unas a las otras, ni pueden reducirse a atributos perceptuales. Asimismo, su existencia es independiente de los objetos que la componen, esto es, excede al individuo particular que emitió el sonido. Finalmente, tales categorías son causalmente eficaces, esto es, determinan el comportamiento de los individuos.

### **Creencias animales: una propuesta wittgensteiniana**

Toda esta evidencia muestra que las nociones necesarias para la atribución de creencias se encuentran presentes de forma débil en una variedad de especies de animales no humanos. Esto permite poner en cuestión ciertos aspectos de los argumentos de Davidson: especialmente aquellos que tienen que ver con la posesión de conceptos y el holismo doxástico.

Con respecto al argumento del carácter conceptual de la creencia, el aspecto problemático recae en la primera premisa, la cual sostiene que para tener una creencia es necesario tener meta-creencias. Davidson concibe la creencia como un estado que puede ser verdadero o falso. Entonces, tener una creencia implica, por un lado, la posibilidad de estar equivocado o tener razón acerca de un hecho del mundo y, por otro, la capacidad de corregir la creencia falsa. Dicha corrección se daría a partir de una meta-creencia del tipo “P” (siendo P una proposición cualquiera) es falsa”. Sin embargo, este proceso no es el único a través del cual podría darse la corrección doxástica. La evidencia no da cuenta de este tipo de mecanismo en los simios, pero sí de una gran flexibilidad en su comportamiento. En efecto, estos animales no parecen actuar disposicionalmente, sino que poseen habilidades metacognitivas para adaptar su conducta al tipo de situación que enfrentan. Por tanto, sería posible que adquieran una nueva creencia o que pasen de la creencia P a la Q sin que este cambio involucre la creencia explícita de que P era falsa (Glock, 2000, p. 56), sino en función de sus capacidades proto-inferenciales. De esta forma, el animal no tiene que volverse necesariamente sobre sus creencias para evaluar su estatus epistémico, sino simplemente ser capaz de pasar de un estado

informativa a otro a través de una operación inferencial basada en la búsqueda de nueva información de su entorno (Danón, 2011).

Por otra parte, podemos estar de acuerdo con la naturaleza holística del pensamiento sin postular un holismo tan radical que demande necesariamente contenido proposicional. La idea de que no se puede tener una creencia sin tener una red inmensa de creencias que la respalden entra en conflicto con la evidencia empírica sobre la posesión de conceptos sociales en grandes simios. Otros experimentos apoyan la idea de que estos animales también poseen conceptos que les permiten comprender las relaciones causales básicas entre los objetos (Call, 2006a). Esto nos permite pensar que podemos atribuirles legítimamente las creencias particulares relacionadas a estas nociones, sin necesidad de atribuirles una infinidad de creencias generales que las respalden.

Una posible objeción a este razonamiento podría ser que la noción de proto-inferencia que utilizamos para explicar el comportamiento de los simios no alcance a satisfacer la noción de creencia lógica que Davidson demanda como requisito para la atribución de creencias. Esto es así, efectivamente, porque Davidson postula una caracterización sumamente radical del holismo conceptual con el fin de demarcar al ser humano con respecto al resto de las criaturas que se comportan de alguna forma (animales o autómatas).

Por esta razón la postura de Wittgenstein resulta más adecuada para dar cuenta del comportamiento animal, ya que admite versiones deflacionadas de las nociones que intervienen en la atribución de creencias. Aun así, Wittgenstein sólo estaba dispuesto a atribuirle a los animales creencias constituidas por meros estados informativos de su entorno. Sin embargo, desde una perspectiva wittgensteiniana es posible ofrecer una noción un poco más apropiada de creencia que no dependa de habilidades lingüísticas. Dicha noción está basada en la posesión de ciertos conceptos que la criatura en cuestión adquiere a través de sus prácticas, en sus capacidades proto-inferenciales y en los estados informativos que brinda su entorno.

## **Conclusión**

En el presente trabajo argumenté que la postura de Wittgenstein acerca de la atribución de creencias a animales es la más adecuada para explicar el comportamiento de algunas especies animales. En primer lugar, expuse

los argumentos principales de las posturas de Davidson y Wittgenstein. En segundo lugar, brindé evidencia empírica acerca de la posesión de contenidos conceptuales y de las capacidades proto-inferenciales que presentan algunas especies de simios. Finalmente, realicé un análisis crítico del holismo conceptual davidsoniano a partir de la evidencia aportada. Dicho análisis resultó en las siguientes ideas: i. Es posible atribuirles a estos animales las creencias particulares basadas en la posesión de los conceptos de rango y parentesco apelando a un holismo moderado; ii. Es posible establecer el carácter conceptual de estas creencias apelando a la corrección doxástica basada en estados informacionales y razonamiento proto-inferencial. Esto indica que es posible ir más allá de lo planteado por Wittgenstein al atribuirle a estos animales una noción de creencia basada en determinados conceptos, proto-inferencias y estados informacionales.

## Referencias

- Bergman, T., Beehner, C. J., Cheney, D., & Seyfarth, R. (2003). Hierarchical classification by-rank and kinship in baboons. *Science*, 302, 1234-1236.
- Call, J. (2006a). Descartes' two errors: Reason and reflection in the great apes. En S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 219-234). Oxford: Oxford University Press.
- Call, J. (2006b). Inferences by exclusion in the great apes: The effect of age and species. *Animal Cognition*, 9, 393-403.
- Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition*, 4, 207-220.
- Danón, L. (2011). Normatividad doxástica en animales. *Analítica*, 5, 9-28.
- Davidson, D. (1984). Thought and talk. En *Inquiries into truth and interpretation* (pp. 155-170). New York: Oxford University Press.
- Davidson, D. (2001). Rational animals. En *Subjective, intersubjective, objective* (pp. 95-106). New York: Oxford University Press. (Obra original de 1982)
- Glock, H. J. (2000). Animals, thoughts and concepts. *Synthese*, 123(1), 35-64.

- Hampton, R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences*, 98, 5359-5362.
- Kripke, S. (2006). *Wittgenstein a propósito de reglas y lenguaje privado* (J. Rodríguez Marquese, Trad.). Madrid: Tecnos. (Obra original de 1982)
- Malcom, N. (1973). *Thoughtless brutes. Proceedings and Addresses of the American Philosophical Association*, 46, 5-20.
- Premack, D., & Premack, A. J. (1994). Levels of causal understanding in chimpanzees and children. *Cognition*, 50, 347-362.
- Seyfarth, R., & Cheney, D. (2015). The evolution of concepts about agents: Or what do animals recognize when they recognize an agent? En E. Margolis & S. Laurence. (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 57-76). Cambridge: MIT Press.
- Smith, J. D., Schull, J., Strote, J., McGee, K., Egnor, R., & Erb, L. (1995). The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*). *Journal of Experimental Psychology: General*, 124(4), 391-408.
- Wittgenstein, L. (2009a). *Investigaciones filosóficas* (A. García Suárez & C. U. Moulines, Trad.). En I. Reguera (Ed.), *Obra completa* (vol. I). En *Colección Biblioteca de Grandes Pensadores*. Madrid: Gredos. (Obra original de 1953)
- Wittgenstein, L. (2009b). *Zettel*. En I. Reguera (Ed.), *Obra completa* (vol. II). En *Colección Biblioteca de Grandes Pensadores*. Madrid: Gredos. (Obra original de 1967)



# Razones para saber-cómo y saber-que

Santiago A. Vrech\*

## 1. Introducción

**A**l respecto del debate “saber-cómo y saber-que” existen dos posiciones principales que se diferencian respecto a la explicación de lo que hace que una acción sea Inteligente. Estas posiciones son el Intelectualismo y el Anti-Intelectualismo. Ambas pueden rastrear sus orígenes al ya famoso libro de Gilbert Ryle *El concepto de lo mental* (1949/2009). De la manera en que Ryle aborda y caracteriza estas posiciones puede entenderse, por un lado, que el Intelectualismo es la conjunción de dos tesis: una concerniente a lo que es un estado de inteligencia en general y otra respecto a la relación entre este estado y una acción:

**[Mente]** Para todo  $X$ ,  $X$  es un estado de Inteligencia si y sólo si  $X$  involucra un estado interno-psicológico en el cual se considera contenido proposicional.

**[Acción]** Para todo  $P$ ,  $P$  lleva a cabo una acción inteligente si y sólo si  $P$  (1) lleva a cabo una acción y la misma (2) involucra un estado interno-psicológico en cual se consideran contenidos proposicionales.

De la conjunción de [Mente] + [Acción] se obtiene entonces que una persona  $P$  sabe cómo hacer una acción  $A$  inteligentemente si  $P$  al ejecutar  $A$  considera o intuye el contenido proposicional relevante para que  $A$  sea una exhibición de inteligencia y no solamente mero hábito, costumbre o suerte (Bengson & Moffet, 2011, pp. 7-8).

Por otra parte, el anti-intelectualismo explica la acción inteligente sin recurrir a la consideración de proposiciones en la mente de las personas, sino más bien de la siguiente manera:

**[Acción-NoI]** Para todo  $P$ ,  $P$  ejerce una acción  $A$  inteligentemente si y sólo si  $P$  al llevar a cabo  $A$  la misma (1) es el producto de una habilidad que ha sido desarrollada y entrenada, (2) es entrenable [*trainable*] en tanto que

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH) Centro de Investigaciones María Saleme de Burnichon (CIFYH). Córdoba, Argentina. savrech@protonmail.com

puede ser modificada y mejorada y por último (3) es “*multitrack*” en tanto que A puede ejecutarse de diversas o diferentes maneras.

El debate contemporáneo gira en torno a cuál de estas dos explicaciones generales es la correcta; considerando qué razones existen para sostener una u otra posición, qué consecuencias se siguen de cada una al modificar los principios [Mente], [Acción] y [Acción-NoI] de diferentes maneras (Cath, 2019).

En el presente trabajo dirijo mi atención a los autores intelectualistas que sostienen que oraciones como “Anna sabe cómo andar *bien* en bicicleta”, “Pedro sabe [cómo] tocar el piano *excelentemente*” y “Sofía sabe [cómo] jugar al ping-pong *habilidosamente*” no son casos en los cuales la manifestación de la Inteligencia se explica recurriendo a [Acción-NoI] o a alguna versión modificada de tal principio, sino que, en realidad, tales oraciones son gramaticalmente engañosas, pues, después de todo, incluso cuando el verbo “saber” se acompañe o pueda ser acompañado por “cómo”, este tipo de saber adquiere su Inteligencia debido a la consideración de proposiciones en el momento anterior o simultáneo a la realización de la acción. Por cuestiones de espacio no podré desarrollar los argumentos que los diferentes autores presentan para defender esta posición.

La consecuencia directa de sostener esta tesis es que “todo el saber-cómo es saber-que. Desatender este hecho empobrece nuestro entendimiento de la acción humana, al oscurecer la manera en la cual ésta es informada por la inteligencia” (Stanley & Williamson, 2001, p. 444). De la misma manera, Snowdon finaliza su artículo estableciendo que en su trabajo “trat[ó] de argumentar en contra la versión estándar del contraste entre ‘saber-cómo/saber-que’ . . . y hacer plausible la afirmación de que [la separación entre ‘saber-cómo’ y ‘saber-que’] es falsa” (2003, p. 28). La conclusión es, entonces, que si P saber-hacer X Inteligentemente, entonces *en realidad* P sabe o tiene contenido proposicional sobre cómo hacer X. ¿Es esto correcto? En principio no. Veamos por qué.

## 2. ¿Cómo sabés eso? O el espacio lógico de las razones

El primer paso que tenemos que tomar para tener una razón de peso para aseverar que los Intelectualistas están equivocados es percibir que la correcta atribución del verbo “saber” a una persona depende primordialmente de que esta persona pueda justificar su conocimiento. En la litera-

tura en epistemología o en teoría del conocimiento esta tesis se encuentra formulada de maneras diversas, aunque todas ellas comparten un núcleo común. ¿Cuál es este núcleo? Ya lo dije: que, si una persona sabe algo, ella debe poder justificar o brindar razones que nos permitan *a nosotros* atribuirle el saber en cuestión. W. Sellars ejemplifica esta idea de la manera siguiente:

El punto esencial es que, al caracterizar un episodio o estado como uno de saber, no estamos dando una descripción empírica de ese estado o episodio; [más bien lo que hacemos es] situarlo en el espacio lógico de las razones, al justificando y al ser capaz de justificar lo que uno dice. (Sellars, 1956/1997, p. 76)

En su famoso artículo “¿Es el conocimiento una creencia verdadera justificada?” de 1963, Gettier cita explícitamente a tres filósofos (Ayer, Chisholm y Platón) que comparten la idea de que P sabe Q si y sólo si

- 1) P cree que Q
- 2) P tiene evidencia, justificación o “el derecho a estar seguro/a que” Q, y
- 3) Q es verdadero.

Si bien el artículo de Gettier ha hecho surgir problemas con esta caracterización del conocimiento, las diferentes respuestas mantienen el punto (2): yo sé que Aristóteles fue un filósofo griego si yo tengo evidencia o justificación para creer ello. O, en términos sellarsianos, si yo declaro saber ello, entonces me sitúo en el espacio lógico de las razones: en el espacio de justificar y ser capaz de justificar mi creencia. Por último, y solo para demostrar la centralidad de la justificación del conocimiento, en el debate en epistemología entre internalismo y externalismo se debate cuál es la unidad que requiere de justificación: si el hablante o lo que el hablante dice. Pero sea cual sea la respuesta victoriosa, al final del día seguirá siendo central el requerimiento de justificación (Alston, 1986; Ginet, 1998).

### 3. Razones para saber-cómo y saber-que

De lo escrito en la sección anterior podemos inferir entonces que la utilización del verbo “saber” conlleva una lógica normativa: el estar en un espacio lógico de razones o el poder justificar lo que se sabe. Esto tiene

relevancia directa para tener una razón para rechazar la reducción del saber-cómo a saber-que. Con este fin, consideremos dos situaciones en las cuales *diríamos*<sup>1</sup> de alguien que sabe(-que) y sabe(-cómo) y analicemos cómo podríamos nosotros saber de esa persona que sabe lo que alega saber.

En primer lugar, empecemos con un caso de saber teórico o proposicional. Imaginemos que un día nos encontramos con un conocido en un café. Este conocido se llama Johann, es austriaco y está por terminar su doctorado en historia sobre la guerra austro-prusiana. Johann nos cuenta que, según el mito popular, el estado independiente de Liechtenstein en la primera guerra mundial envió a 80 hombres al frente italiano y, tras algunos días, 81 volvieron: no solo no hubo ninguna baja en el batallón, sino que además en el camino de vuelta se hicieron amigo de un austriaco y lo invitaron a vivir a su país. “Mucha gente, nos dice Johann, *cree saber* que este hecho fue así, pero en realidad no lo es. Con mis investigaciones, basada en parte con documentos de la época y en parte con entrevistas a los ciudadanos octogenarios de Liechtenstein, puedo afirmar que (*yo sé*) que en realidad el suceso en cuestión tuvo lugar en 1866 cuando el contingente de 80 hombres tomó posesión en el paso de Prato en Italia para defender la frontera entre Austria y Liechtenstein de los ataques italianos. Al finalizar la guerra y sin haber sufrido ninguna baja, al momento de regresar el batallón trajo consigo a un oficial austriaco herido”. ¿Tendríamos alguna duda en aseverar que Johann *sabe que* en realidad el hecho en cuestión sucedió cuándo y cómo él dice que sucedió? En principio no. Si todavía tuviésemos dudas, podríamos presionar a Johann para que nos diga en qué documentos y entrevistas basa su saber. Y he aquí el punto importante para lo que ahora discutimos. La justificación del saber-que (el saber que Johann alega tener sobre la historia del batallón) *es en términos de verdad y falsedad*, pues en caso de que los documentos que cite Johann sean documentos que, en verdad, no existen, y que los testimonios no sean confiables, nosotros podríamos dudar de la justificación traída a colación y, consecuentemente, calificarla de falsa. Como resultado, finalmente podemos decir que Johann *no sabe* que el hecho sucedió como él lo afirma.

Tenemos, entonces, los siguientes elementos relativos a la justifica-

---

<sup>1</sup> La regla de juego filosófica que adopto en este trabajo es que la manera en la que ordinariamente hablamos es interesante e importante para hacer filosofía. Si no se acepta esta regla o si se argumenta metafilosóficamente en contra de la misma, el argumento que desarrollo no tendrá, en principio, sentido.

ción de una atribución de saber-que: para la justificación del saber se presentan razones, las mismas son evaluadas en términos de verdad o falsedad; si las razones no parecen convincentes, uno puede dudar del saber en cuestión y, ante esta situación, las razones que se dan para la justificación pueden mejorarse.

Ahora, para analizar un caso de saber-cómo, imaginemos el siguiente escenario. Miles Davis es un trompetista muy pero muy talentoso a tal punto que cada vez que establece una fecha para un concierto las entradas se agotan dentro de un día. Nos encontramos en el café todavía con Johann, y en la televisión se encuentra una entrevista a Miles. En medio de la entrevista se produce una escena un tanto particular: el entrevistador le dice a Miles que él sabe también tocar muy bien la trompeta, “incluso mejor que usted”, dice, a modo de desafío, el provocador entrevistador. “Bueno ¿a ver?” responde Miles, y el entrevistador procede a citar la biblia del trompetista (suponiendo que tal cosa exista, esta biblia contendría todo el saber proposicional habido y por haber sobre cómo tocar la trompeta). Un tanto desconcertado, Miles le entrega su trompeta para *ver*<sup>2</sup> si el entrevistador, a pesar de saber mucho sobre trompetas, sabe cómo tocarla. Con la trompeta en sus manos el entrevistador no puede hacerlo, y tras unos minutos incómodos Miles responde “sabe mucho *sobre* cómo tocar la trompeta, pero todavía no me he *mostrado* que *sabe tocar* la trompeta”.

Lo que este escenario ejemplifica es que al momento de justificar el saber-cómo no se puede aceptar como justificación razones proposicionales<sup>3</sup>, sino que por el contrario la justificación del saber-cómo (en algunos casos) es primordialmente mediante la exhibición ostensiva del saber en cuestión. Nosotros sabremos que el entrevistador sabe cómo tocar la trompeta si lo vemos haciéndolo; sabemos que Federer sabe cómo jugar al tenis porque lo vemos jugar magníficamente, sabemos que Bryan Magee sabe cómo liderar entrevistas porque lo vemos haciéndolo. Pero si quitáramos la justificación ostensiva ¿diríamos que el entrevistador sabe tocar la trompeta, que Federer sabe cómo jugar bien al tenis y que Magee sabe cómo realizar buenas entrevistas? Si se responde que no ¿cómo podríamos aseverar que ellos saben lo que decimos o dicen saber? La

<sup>2</sup> Como se verá a continuación, la justificación del saber-cómo tiene una relación particular con la percepción.

<sup>3</sup> Tengo que insistir en que esto es cierto relativo a algunos casos. Hay otros como jugar al ajedrez que están en una frontera difusa: ¿si un ajedrecista cita la biblia del ajedrecista no diríamos que también sabe cómo jugar al ajedrez?

justificación es imprescindible. Pero hace falta remarcar que, a diferencia de la justificación del saber-que, la justificación de saber-cómo es evaluada en términos de corrección e incorrección: evaluamos si alguien realiza la acción inteligente en cuestión correcta o incorrectamente (o el entrevistador toca la trompeta magníficamente como Miles o no la toca) y, en conexión con esto, no puede reforzarse la justificación, aunque, claro, el saber(-cómo) puede hacerse más versátil, complejo y mejor.

Teniendo en mente estas dos diferencias respecto a las características de la justificación en cada tipo de saber en mente, puede verse que la consecuencia intelectualista de establecer que el saber-cómo es una instancia de saber-que es falsa, puesto que ello traería como consecuencia que la justificación del saber-cómo sería de tipo proposicional. De acuerdo con esta imagen, la respuesta a nuestro inquisitivo “¿sabes jugar bien al ping-pong?” sería “Sí, sé que la raqueta se agarra con la mano así, que la pelota solamente puede picar una vez en mi mesa, que son dos servicios cada uno, etc.”. ¿Sería esta justificación correcta para adjudicarle a una persona el saber cómo jugar al ping pong? No parece. Más bien, la manera en la cual justificaría su saber es exhibiendo sus habilidades. La consecuencia Intelectualista respecto a la justificación del saber-cómo llevaría, pues, a una falsa imagen de nuestro uso ordinario del verbo “saber” y el juego normativo-lógico que este conlleva.

## Referencias

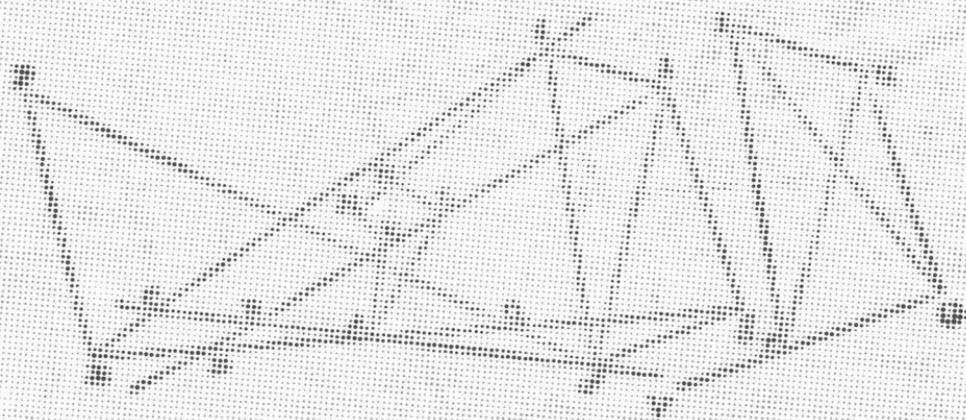
- Alston, W. P. (1986). Internalism and externalism in epistemology. *Philosophical Topics* (14), 179-221.
- Bengson, J., & Moffett, M. (2011). Two conceptions of mind and action: Knowledge how and the philosophical theory of intelligence. En J. Bengson & M. Moffett (Eds.), *Knowing how: Essays on knowledge, mind, and action* (pp. 3-55). Oxford: Oxford University Press.
- Cath, J. (2019). Knowing how. *Analysis*, 79(3), 487-503.
- Dretske, F. (1985). Precis of knowledge and the flow of information. En H. Kornblith (Ed.), *Naturalizing epistemology* (pp. 169-187). Cambridge, MA: MIT Press.
- Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis*, 23(6), 121-123.

- Ginet, C. (1998). The general conditions of knowledge: Justification. En L. Alcoff (Ed.), *Epistemology: The big questions* (pp. 79-89). Blackwell: London.
- Ichikawa, J. J., & Setup, M. (2014). The analysis of knowledge. En E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2014 edition). Recuperado de <https://plato.stanford.edu/archives/spr2014/entries/knowledge-analysis/>
- Ryle, G. (2009). *The concept of mind*. London: Routledge. (Obra original de 1949)
- Sellars, W. (1997). *Empiricism and the philosophy of mind: with an introduction by Richard Rorty and a study guide by Robert Brandom* (R. Brandom, Ed.). Cambridge, MA: Harvard University Press. (Obra original de 1956)
- Snowdon, P. (2003). Knowing how and knowing that: A distinction reconsidered. *Proceedings of the Aristotelian Society*, 104(1), 1-29.
- Stanley, J., & Williamson, T. (2001). Knowing how. *The Journal of Philosophy*, 98(8), 411-444.



---

## 8. Historia y filosofía de las ciencias formales







# ¿Por qué no hay argumentos concluyentes en filosofía?

## El caso de los argumentos por regresión infinita

Gustavo Arroyo\*

### Introducción

Cuenta la leyenda que en el antiguo oriente dos astrónomos discutían cierta vez sobre la manera de explicar el hecho de que la tierra se mantenga inmóvil en el centro del universo. Uno de ellos sugirió que podría deberse a una enorme tortuga que la sostenía sobre su caparazón. El otro preguntó entonces qué era lo que impedía que la tortuga cayera. El primer sabio respondió que tal vez fuera un inmenso elefante. Pero ¿quién sostiene al elefante? El primer sabio respondió que podría ser un enorme reptil. Pero esta hipótesis, como ya podrá suponerse, no convenció a su interlocutor que volvió a plantear la misma pregunta, esta vez referida al enorme reptil.

Este ejemplo sencillo contiene todos los componentes del tipo de argumento que discutiremos aquí. En primer lugar, se trata de un argumento que tiene como objetivo refutar una cierta teoría. Y lo hace mostrando que esa teoría implica una contradicción: El defensor de la teoría da por sentado que la tierra está fija en el centro del universo (eso es precisamente lo que desea explicar). Pero ¿cómo puede saberlo una vez que, para terminarlo, debe recorrer una serie infinita de verificaciones? (debe estar seguro primero de que la tortuga no está cayendo, pero para eso debe estar seguro de que el elefante tampoco lo está, etc.). Por lo tanto, o bien no es un hecho que la tierra esté inmóvil o bien es un hecho, pero no es esa la explicación. Nuestro sabio es confrontado a una dolorosa elección.

Los Argumentos por Regresión Infinita son muy utilizados en los debates filosóficos. Se ha dicho incluso que constituyen un patrón de argumentación distintivo y exclusivo de la filosofía. Como parte de esta mesa sobre Argumentos por Regresión Infinita deseo responder la pregunta de

\* Universidad Nacional de General Sarmiento. Malvinas Argentinas, Buenos Aires, Argentina.

gjarroyo@yahoo.com.ar

por qué estos argumentos raramente constituyen argumentos concluyentes. La distinción entre argumentos concluyentes e inconcluyentes (también podríamos hablar de argumentos “demoledores” y “no-demoledores”) es una distinción firmemente arraigada en el lenguaje pre-filosófico. Por ejemplo, hay argumentos concluyentes a favor de la esfericidad de la tierra y de la infinitud de los números primos, pero, en cambio, no hay argumentos demoledores para la conjetura de Goldbach o para la hipótesis catastrofista en el debate sobre la extinción de los dinosaurios. Como ha señalado Ballantyne (2014), hay diferentes maneras posibles de establecer la distinción entre ambos tipos de argumentos. En este trabajo utilizaré una versión simplificada de la definición propuesta por este autor:

x es un argumento concluyente si y solo si cualquier persona que entienda x y no acepte la conclusión en base a las premisas se comporta irracionalmente. (Ballantyne, 2014, p. 528)

Tengamos presente que, de acuerdo a esta definición, decir que un argumento sea demoledor no significa que sea válido y decir que un argumento sea válido no quiere decir que sea demoledor. Los argumentos a favor de la esfericidad de la tierra son argumentos inductivos, pero las dudas que podrían plantearse para no aceptar la conclusión no parecen ser razonables. Por otro lado, ningún argumento válido con premisas falsas o que puedan ser puestas en duda, podrá ser demoledor. El que se niegue a admitir que la tierra es esférica luego de haber visto una foto satelital del planeta o se niegue a admitir que hay infinitos números primos una vez que ha comprendido la demostración de Euclides puede ser acusado legítimamente de ser irracional (dogmático, intelectualmente perverso, etc.). ¿Por qué los Argumentos por Regresión Infinita no suelen ser argumentos demoledores?

Volvamos a la historia de los dos sabios. Notemos que la regresión no se produce solo de aceptar la hipótesis de que la tierra es sostenida por una enorme tortuga. Debemos asumir dos premisas más que no han sido explicitadas. Lo haré con la ayuda de un mínimo formalismo: La teoría objeto de la refutación (T) ha sido propuesta para explicar el hecho de que una determinada entidad (E) (la tierra) tiene una cierta propiedad (P) (permanecer inmóvil).

- a) Para explicar el hecho de que E posea la propiedad P, la teoría postula (en el explanans) la existencia de una entidad del mismo tipo que E. Para explicar la inmovilidad de un cuerpo (la tierra) la teoría postula la existencia de una tortuga, que también es un cuerpo. Siendo que la tierra puede estar móvil o inmóvil por ser un cuerpo, la tortuga, por ser un cuerpo, también puede estar en cualquiera de esos dos estados. Ahora bien, al derivar la conclusión contradictoria hemos asumido que la tierra no puede estar inmóvil si no está inmóvil también la tortuga. Si la tortuga estuviera cayendo, caería con ella también la tierra. (Formalmente, lo que nuestro sabio asume es que E no puede tener la propiedad P a menos que E1 también la tenga)
- b) Pero el argumento parece concluyente porque tácitamente estamos aceptando adicionalmente un postulado de orden general: la única manera de que un cuerpo pesado permanezca suspendido en el espacio es ser sostenido por otro. Es este postulado el que pone en marcha la serie infinita. Si nuestro sabio dijera: “la tortuga no se sostiene en nada. Se mantiene estática por la acción de fuerzas magnéticas”, podría ser acusado de introducir una hipótesis ad-hoc, pero no de caer en una regresión infinita. (Formalmente, lo que nuestro sabio asume es que hay solo una manera de tener la propiedad P)

Mi argumento de por qué los Argumentos por Regresión Infinita no son demoleedores es que en los casos más sustanciales de argumentos de este tipo algunas de estas dos premisas resultan cuestionables. Recordemos un punto remarcado originariamente por Pierre Duhem (1906/2003) respecto de las teorías científicas: si a partir del conjunto formado por una hipótesis principal y varias hipótesis auxiliares derivamos una consecuencia inaceptable (por ejemplo, una predicción fallida), lo que resulta así refutado no es la hipótesis principal sino todo el conjunto. Formalmente:

$$(H_1, H_2, H_3, \dots, H_n) \rightarrow p$$

$$\neg p$$

---


$$\neg(H_1, H_2, H_3, \dots, H_n)$$

El científico en este caso deberá rechazar al menos una de las hipótesis del conjunto. Pero si a alguna de las hipótesis auxiliares fuera en principio cuestionable, el científico podría responsabilizarla por la conclusión adversa y escapar así de la refutación. Esto mismo se aplica a los Argumentos por Regresión Infinita.

### **Argumentos filosóficos por regresión infinita**

Para mostrarlo discutiré algunos ejemplos paradigmáticos. El primer ejemplo proviene de la filosofía de la psicología. Todos aceptamos que hay una diferencia entre hacer algo voluntariamente y hacerlo de manera involuntaria, por ejemplo, guiñar un ojo de manera voluntaria y guiñarlo por el efecto de una droga o un tic nervioso. Está claro que los actos voluntarios son aquellos que una persona decide llevar a cabo y que los involuntarios son aquellos que no están mediados por una decisión. Existe una inclinación filosófica, sin embargo, a pensar que tomar una decisión es algo que ocurre en nuestra mente, algo que tiene un carácter episódico y es causa del acto físico (por ejemplo, de guiñar el ojo). Podemos denominar Mentalismo a cualquier teoría que intente reconstruir en términos episódicos conceptos psicológicos como “decidir”, “pensar”, “creer”, etc. Gilbert Ryle, que defendía una teoría rival al mentalismo, ofreció un conocido argumento por regresión infinita en contra de la misma. He aquí la reconstrucción simplificada de ese argumento ofrecida por Jay Rosenberg:

Si esto es así [si un acto es voluntario si y solo si es producido por un acto de la voluntad], aún no hemos recibido la respuesta a nuestra pregunta original. Solo podemos entender esta pregunta si sabemos de antemano qué es lo que hace voluntario a un acto de la voluntad. ¿Y qué es lo que hace voluntario a un acto de la voluntad? El único curso posible para nosotros es volver aplicar la teoría. Al hacerlo, sin embargo, descubrimos que necesitamos otro acto de la voluntad que sea voluntario. La pregunta no se resuelve. (Rosenberg, 1984, p. 62)

No hay dudas de que para el mentalista un acto es voluntario si y solo si es precedido por otro acto, este último de naturaleza psíquica. ¿Pero está obligado a sostener que el segundo deberá ser voluntario para que el primero también lo sea? En absoluto. Aquí hay una diferencia sustancial con el ejemplo de los sabios. La tierra no puede estar quieta si aquello

que la sostiene no lo está. Pero no parece absurdo sostener que un acto (físico) pueda ser voluntario si es causado por un acto mental que no es voluntario, que surge o aparece de manera involuntaria en la mente de un sujeto. De hecho, esta es la posición compatibilista respecto del libre albedrío. Para los compatibilistas un acto es libre si y solo si es determinado por causas psicológicas. Deja de ser libre si es determinado por causas “externas”. Rosenberg (y Ryle) sobrestiman la fuerza destructiva de ese argumento pues no ven que el filósofo mentalista no está obligado a aceptar la premisa auxiliar de la que depende el argumento.

Consideremos ahora un argumento sugerido por Wittgenstein en un pasaje de las *Investigaciones Filosóficas*:

¿Cómo sabrá él qué color tiene que elegir cuando oye «rojo»? — Muy sencillo: debe tomar el color cuya figura se le ocurre al oír la palabra. — ¿Pero cómo sabrá qué color es aquel ‘cuya figura se le ocurre’? ¿Se necesita un criterio ulterior para ello? (Wittgenstein, 1954/1988, § 239, p. 217)

Wittgenstein intenta refutar aquí una cierta concepción intelectualista del uso del lenguaje. Según esta concepción disponemos siempre de una razón o criterio cuando decidimos aplicar un término a un cierto objeto. Pero ¿cuál es nuestro criterio, pregunta Wittgenstein, cuando aplicamos el término “rojo” a una muestra de color? Una posible respuesta en línea con la tesis intelectualista podría ser que la muestra de color coincide con la imagen mental asociamos inmediatamente con la palabra “rojo”. Para determinar el color de una muestra postulamos la existencia de otra muestra. ¿Pero cómo sabemos, pregunta Wittgenstein, que la imagen mental que se nos ocurre es de color rojo y no de algún otro color? El autor del argumento parece suponer que la única respuesta legítima es sostener la existencia de una segunda imagen que sirva de criterio para la primera, lo cual encerraría a su oponente en una regresión infinita: si sabemos que la primera imagen es roja porque la comparamos con una segunda imagen, entonces necesitamos una tercera con la cual comparar la segunda y así infinitamente. No podemos aceptar que esa regresión tiene lugar pues implica una consecuencia absurda: si debemos pasar por una secuencia infinita de pasos nunca llegaríamos a una decisión final (justificada) respecto de si la muestra es o no es roja.

Pero debemos cuidarnos de sobreestimar la fuerza de este argumento. Parece implausible sostener que podemos estar justificados en afirmar que la muestra de color es roja si no tenemos una razón para afirmar que la

imagen mental lo es, pero ¿tiene que ser esa razón el hecho de que la hemos comparado con una segunda imagen? ¿No podría ser tal vez una posible justificación el hecho de que los seres humanos poseemos la capacidad innata para producir la imagen mental adecuada una vez que escuchamos la palabra “rojo”? ¿Y no podría ser el conocimiento de que poseemos esa capacidad la mejor justificación para sostener que sabemos que la imagen mental es roja?

## Conclusión

Esta es, de hecho, la estrategia más frecuente para escapar de la regresión. Cuando el escéptico argumente que la exigencia de justificar nuestros juicios conducirá a una regresión infinita (si  $q$  es nuestra justificación para afirmar  $p$  y si  $r$  es nuestra justificación para afirmar  $q$  entonces habrá que preguntarse qué nos justifica en afirmar  $r$ ...), el fundacionalista detendrá la regresión argumentando que existen proposiciones que se auto-justifican. Cuando los seguidores de Fodor (1975) explican la adquisición del lenguaje natural postulando el dominio de un lenguaje previo (un “lenguaje del pensamiento”), disponen de una manera muy simple de evadir la regresión, a saber, sostener que ese segundo lenguaje es innato. Cuando el teísta afirma que el mundo fue creado por Dios y el ateo le pide que explique quién creo a Dios, el primero puede responder (sin contrasentido aparente) que Dios es la única existencia increada.

## Referencias

- Ballantyne, N. (2014). Knockdown arguments. *Erkenntnis*, 79(S3), 525-543.
- Clark, R. (1988). Vicious infinite regress arguments. *Philosophical Perspectives*, 2, 369-380.
- Duhem, P. (2003) *La teoría física, su objeto y estructura* (M. Pons Irazazábal, trad.). México: Herder. (Obra original de 1906)
- Fodor, J. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Passmore, J. (1961). *Philosophical reasoning*. New York: Charles Scribner's sons.
- Rescher, N. (2010). *Infinite regress*. New Jersey: Transaction Publishers.

Rosenberg, J. (1984). *The practice of philosophy*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Wittgenstein, L. (1988). *Investigaciones filosóficas* (A. García Suarez & U. Moulines, trad.). Barcelona: Crítica. (Obra original de 1954)



## Series temporales y regresión al infinito: el argumento de McTaggart

Ulises Dávalos\*

### Introducción

El objetivo de este artículo es reconstruir y defender la validez del argumento de McTaggart a favor de la irrealdad del tiempo. Llegamos a este argumento en particular motivados por el estudio de distintos tipos de regresión al infinito tratando de comprender cómo se utiliza este tipo de argumento en la filosofía. Ahora bien, la explicación que expondremos a continuación está impulsada por una cuestión muy mundana, a saberse, si un argumento de este tipo logra rebatir la idea del tiempo como real. Dada la exigencia de su autor es primordial entender en primer lugar cómo este comprende el tiempo, por lo cual tratemos directamente qué es lo que describe con series temporales para luego plantear si su objetivo está ajustado.

### Estructura del argumento por regresión

Existen, de acuerdo a McTaggart, dos posibles maneras de establecer series temporales. Las denomina series-A y series-B (McTaggart, 1927). Las series A se valen de los predicados monádicos “pasado”, “presente” y “futuro” (Por ejemplo “el atentado a las torres gemelas es pasado”). Las series B utilizan los términos relacionales “antes” y “después”. Podría decirse que la diferencia entre las series A y las series B reside, esencialmente, en que las primeras son dinámicas y las segundas estáticas. Por ejemplo, el atentado a las torres gemelas alguna vez fue en el futuro, pero luego devino presente y ahora es en el pasado. Pero en cambio, siempre fue y será el caso que el atentado a las torres gemelas fue antes de esta ponencia.

Valiéndonos de esta distinción podemos exponer de manera esquemática el argumento de McTaggart. Su argumento contiene tres premisas:

\* Universidad Nacional General Sarmiento. Grupo de lógica y epistemología. Buenos Aires, Argentina.

ulisesdavalos21@gmail.com



Para que el tiempo fuera real debería ser real el cambio.

Para que el cambio se real deberían ser reales las series A

Pero las series A no son reales

Luego, el tiempo no es real.

Se trata indudablemente de un argumento válido. Por lo tanto, la verdad de la conclusión estará garantizada si sus premisas son verdaderas. ¿Son verdaderas las premisas de este argumento? No disponemos aquí del espacio para ocuparnos del argumento en su totalidad. Nos limitaremos por ello a discutir la tercera premisa, que contiene el argumento por regresión al infinito, tema de esta mesa redonda. Pero permítaseme hacer unas breves observaciones acerca de las dos primeras, lo cual permitirá que nos ubiquemos de mejor manera en el argumento. Que el tiempo requiere movimiento es una afirmación plausible. Parece ser apoyada por el hecho de que solo observamos el pasaje del tiempo observando el cambio en las cosas (por ejemplo, el movimiento de las manecillas del reloj). También es plausible la segunda premisa. Las Series B son estáticas y por lo tanto, variación a lo largo de estas series no cuentan como cambio. Para que exista el cambio necesitamos que existan las series A.

McTaggart sostiene que las series A no pueden ser reales porque son en sí mismas contradictorias y una expresión contradictoria (como por ejemplo “el círculo cuadrado”) nunca puede ser aplicada a nada realmente existente. Así, si las series A son contradictorias no pueden ser aplicadas a nada realmente existente. Para demostrarlo, parte del supuesto no controvertido de que los términos “pasado” “presente” y “futuro” que caracteriza a las series A incompatibles entre sí:

Si  $x$  es presente, entonces no es ni pasado ni futuro

Si  $x$  es pasado, entonces no es ni presente ni futuro

Si  $x$  es futuro, entonces no es ni presente ni pasado

Esto es obviamente verdadero, pero además deberá serlo si es que las series A son reales. Si las series A son reales, entonces cuando un evento cambia su posición en una de estas series, esto tiene que ser un cambio

genuino. Y el cambio no sería genuino si fuera posible que un evento conservara las otras posiciones. Pero las series A son contradictorias porque las propiedades A son incompatibles entre sí y al mismo tiempo cada evento debe tenerlas todas:

En el 2005, esta ponencia era en el futuro, por lo tanto tiene la propiedad “futuro”.

Hoy, esta ponencia es presente, por lo tanto tiene la propiedad “presente”.

En el 2025, estará en el pasado, por lo tanto tiene la propiedad “pasado”.

Así, cada evento solo puede tener una de estas propiedades y necesariamente las tiene a todas. Esto es una contradicción. Luego, concluye McTaggart, las series A son contradictorias.

Todavía no ha aparecido el argumento por regresión al infinito. Este argumento aparece como respuesta a una objeción que cualquiera de nosotros haría al argumento que acabamos de esbozar. El argumento que acabamos de esbozar podría parecer ingenuo o falaz. Porque cuando decimos que el presente y el pasado son incompatibles lo que queremos decir es que un evento no puede tener esas propiedades al mismo tiempo. Es como cuando decimos que “redondo” y “cuadrado” son propiedades incompatibles. Lo que queremos decir es que un cuerpo no puede ser las dos cosas al *mismo tiempo*. Pero bien puede ocurrir que un cuerpo tenga esas propiedades de forma sucesiva. Lo que McTaggart mostró (y esta es la crítica) es que un evento tiene las 3 propiedades de manera sucesiva.

Esta ponencia alguna vez *fue* en el futuro

Esta ponencia *es* ahora en el presente

Esta ponencia *será* en el pasado

Por lo tanto, no hay contradicción alguna. Pero esta crítica no puede tener éxito porque, como bien observa McTaggart, al intentar escapar así de la contradicción caemos en una regresión infinita. Preguntémosnos esto: ¿Qué podría significar que esta “ponencia fue futuro” sino que esta ponencia “es futuro en el pasado”? Lo mismo vale para los otros tiempos:

$x$  fue futuro =  $x$  es futuro en el pasado

$x$  es ahora futuro =  $x$  es futuro en el presente

$x$  será futuro =  $x$  es futuro en el futuro

Habrán tiempos más complejos:

Esta ponencia era en el futuro = esta ponencia, en el pasado, es en el futuro.

Esta ponencia es ahora en el presente = esta ponencia en el presente, es en el presente.

Esta ponencia será en el futuro = esta ponencia, en el futuro, es en el futuro.

De manera más general, podemos decir que en lugar de tener 3 propiedades-A simples (pasado, presente y futuro), tenemos 9 propiedades complejas:

Pasado		Pasado
Presente	en el	Presente
Futuro		Futuro

Debemos notar ahora que algunas de esas propiedades son compatibles entre sí (por ejemplo “pasado en el pasado” y “presente en el pasado”; “futuro en el futuro” y “presente en el futuro”), pero algunas de ellas son incompatibles (por ejemplo “presente en el presente” y “pasado en el presente” o “presente en el presente” y “futuro en el presente”).

Habiendo llegado a este punto, McTaggart insiste en que todo evento debe tener todas las propiedades:

En el 2005, esta ponencia era en el futuro en el presente.

Hoy, esta ponencia es en el presente en el presente.

En 2020, esta ponencia será en el futuro en el presente.

En síntesis, esta ponencia tiene pasado en el presente, presente en el presente, y futuro en el presente. Pero acabamos de decir que esas propiedades son incompatibles. Por lo tanto, la serie A sigue siendo contradictoria.

Alguien podría insistir en que este argumento es ingenuo o falaz. Porque esta ponencia por cierto tiene pasado en el presente, presente en el presente y futuro en el presente, ¡pero en diferentes momentos!

Esta ponencia *era* en el futuro, en el presente.

Esta ponencia ahora *es* presente en el presente.

Esta ponencia *será* en el pasado en el presente.

El lector podrá intuir ya que esta respuesta no funcionará pues nos lleva directamente a una regresión infinita. “Esta ponencia era en el futuro en el presente” equivale a “es en el futuro en el presente en el pasado”. En lugar de 9 propiedades tendremos 27:

Pasado		Pasado		Pasado
Presente	en el	Presente	en el	Presente
Futuro		Futuro		Futuro

Algunas de ellas son compatibles y otras incompatibles, pero...

Esta regresión no parece ser virtuosa: en cada nueva fase quien crea en las series A se verá obligado a moverse a la fase siguiente para evitar la contradicción, pero la contradicción indefectiblemente volverá a aparecer.

Así es como McTaggart resuelve su argumento acerca de la irrealidad del tiempo. Pues, como ya hemos visto si aceptamos que el cambio es esencial al tiempo, así como que el cambio solo puede darse en series del tipo A no podremos avanzar más de allí. Como podrá apreciarse las series del tipo A guardan en ellas una contradicción cuyo intento de resolución nos arrincona a un argumento por regresión vicioso. Pues, el tiempo presenta una cantidad limitada de propiedades simples (pasado o presente o futuro) pero contradictorias entre sí, o tiene una cantidad infinita de propiedades complejas que se escalan de manera exponencial. Con este argumento de regresión, McTaggart niega la premisa acerca de que las series A son reales y con ello, está en condiciones de asumir que el tiempo mismo no es real. Esto dentro de su propio marco idealista en el cual justifica una realidad única e indivisible. El tiempo como “real” debe ser negado en or-

den de su argumento supuestamente demoleedor. Puesto que este regreso al infinito compromete a quien pretenda defender la realidad del tiempo a una asunción de infinitas entidades, algo con lo que naturalmente nadie estaría dispuesto a comprometerse. Esa negación casi categorial es lo que comprendemos la fuerza más grande dentro de argumentos de este tipo.

Ahora bien, ¿es esto suficiente para negar la realidad del tiempo? Al inspeccionar nuestra posición, este argumento no parece suficientemente fuerte como para desestimar nuestras concepciones previas acerca del tiempo. De igual manera lo han creído otros filósofos que ante esto, se dieron a la tarea de dar solución al intrincado problema que el autor nos propone. Pues como hemos visto, este argumento parte de premisas verdaderas y se corresponde con una deducción válida, mas creemos que, aunque válido, no genera una discusión inapelable ni tampoco posea un carácter que demuela cualquier propuesta anterior. Es decir, si bien este argumento parece no tener fallas, creemos que la pretensión de McTaggart por sostener una irrealidad del tiempo basado en él no es del todo ajustada.

Por otro lado, nos lleva a interrogarnos acerca de la naturaleza de este tipo de argumentos. Estos nos otorgan cierta sensación de fortaleza, pero demuestran ser bastante poco seductores a nuestras concepciones, mucho menos a aquellas tan primigenias como la existencia del tiempo. Pues bien, si su entereza estructural no garantiza su eficacia para convencer ¿deberíamos revisarlo? En este punto hemos comenzado pensar que un argumento por regresión no es tal vez inapelable como nos quiso hacer creer el autor. Ahora bien, ¿debemos atribuírselo al argumento en si o buscar algún tipo de justificación alterna? Creemos que el argumento no puede reformularse de manera más clara o precisa para llegar al cometido planteado por lo que aquí no nos detendremos demasiado, en su lugar inspeccionaremos otras dos posibles respuestas al problema de este argumento.

En primer lugar, proponemos que simplemente, en la filosofía no hay tal cosa como argumentos finales. Es decir, creemos que parte de cierta esencia de la filosofía versa en poder rebatir aquellos argumentos que parecer solidos explorando alternativas tangenciales. Esta característica tal vez alejaría a la filosofía de una posición univoca de la verdad, pero garantizaría (y de hecho es característico de esta) que sus temas sean revisitados infinitamente a lo largo del tiempo.

Por otro lado, proponemos que la ambición del filósofo fue demasiado grande y que, si bien su argumento no tiene fallas, hay lugares en donde la filosofía no puede argumentar de manera efectiva. Aquí tal vez en un pragmatismo muy cerrado decimos que no importa la perfección del argumento, ideas tan constitutivas como la del tiempo, no pueden ser desarmadas por ninguna estratagema lógica.

En conclusión, creemos que, si bien este argumento es ciertamente consistente, en suma, carece de la fuerza necesaria para sostener lo que propone el autor. Tal vez por ser una regresión que nos evade de lo cotidiano o tal vez porque la regresión en si es precaria para fundamentar, el quid de la cuestión residiría tal vez en una vía alterna de argumentación que no dependa de este método.

## **Referencias**

McTaggart, J. M. E. (1927). *The nature of existence* (vol. 2). Cambridge: Cambridge University Press.



# Lenguajes y arquitecturas

Pío García\*

Javier Blanco\*\*

## Introducción

Hemos propuesto en trabajos anteriores a la efectividad y programabilidad como dimensiones adecuadas para abordar la noción de computación. En este trabajo queremos avanzar en la sugerencia de que dichas dimensiones pueden representar una perspectiva ventajosa para comprender aspectos conceptuales e históricos de la computación.

Como es bien sabido, en la búsqueda de lograr una caracterización no ambigua ni circular de procedimiento efectivo, Alan Turing propuso la noción que hoy consideramos estándar de computación. La descripción de los elementos integrantes de una *máquina de Turing* vendría a dar cuenta del comportamiento de un ser humano actuando mecánicamente. De esta manera, el aspecto mecánico aparecía como constitutivo de lo computacional.

Ahora bien, la discusión sobre lo que constituía un procedimiento efectivo estaba, a principios del siglo XX, circunscripta a la lógica y los fundamentos de la matemática. Pero la tradición de las máquinas que calculan, como la pascalina o los motores de Babbage, y de las máquinas que decodifican, como el telar de Jacquard, se originó y desarrolló de manera bastante independiente. Aunque luego, en el siglo XX, terminarían por confluir en la construcción de las primeras computadoras.

La efectividad y programabilidad permiten, desde nuestra perspectiva, describir las propiedades fundamentales de la tradición de las máquinas de cálculo y aquellas que tenían algún sistema de codificación. La efectividad de estas máquinas estaba asociada con diversas propiedades y características: desde la confiabilidad de los materiales hasta la organización de los sub-sistemas para ganar en automatización de procesos. También, por

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH), Córdoba, Argentina.

\*\* Universidad Nacional de Córdoba (UNC), Facultad de Matemática, Física, Astronomía y Computación (FAMAF), Córdoba, Argentina  
piogarcia@ffyh.unc.edu.ar

supuesto, estaba asociada con una de las promesas por las cuales se construían máquinas: una mayor velocidad en las tareas mecanizadas. Efectividad aquí, entonces, está vinculada con aquellos aspectos que hacen que una máquina sea veloz, autónoma y confiable.

A su vez, un aspecto central de estas máquinas era el grado en el cual se podían codificar comportamientos o las tareas para las cuales eran construidas. El que un sistema sea programable y en qué medida lo era, constituía una propiedad definitoria no sólo para proto-computadoras —en el sentido temporal— como las de Babbage sino para dispositivos automatizados como el telar de Jacquard. Una de las innovaciones características en este tipo de telar era la codificación de patrones en un sistema de tarjetas perforadas.

Dentro de este esquema era esperable, y hasta natural, asociar la efectividad de las máquinas al *hardware* (a aquellos aspectos más directamente vinculados con lo físico) y la programabilidad con el *software* o con los procedimientos de codificación.

En gran medida, la propuesta de Turing vino a «atar» de una manera original la efectividad y la programabilidad a través de las ideas de máquina de Turing y máquina universal. Pero esta propuesta estuvo en un principio circunscripta al contexto lógico y matemático del cual surgió.

Nuestra propuesta es que la discriminación entre estos aspectos no solo puede servir para iluminar aspectos *conceptuales* de la noción de computación sino para entender algunos aspectos de su *historia*. En este trabajo vamos a tomar como ejemplo algunos hitos de la constitución de la noción de arquitectura y de la actividad de programar.

Pero antes de abordar esta cuestión creemos que puede ser útil especificar en qué sentido las propuestas originales de Turing podrían complementarse señalando los alcances de la noción de computación, lo que habitualmente se ha tratado bajo la denominación de la tesis Church-Turing.

## Gandy y las tesis Church-Turing

La cuestión de qué está involucrado en la noción de computación podría ser investigada, asimismo, a partir de distinguir entre perspectivas diferentes que a veces, histórica y conceptualmente, han sido vistas como equivalentes. En un texto de 1980 —*Church's thesis and principles of mechanism*— Gandy distingue entre tres versiones de la llamada tesis

Church-Turing. El primer enunciado es la tesis de Church: “lo que es efectivamente calculable es computable” (Gandy, 1980, p. 123). El segundo enunciado es llamado el teorema T (por Turing): “lo que puede ser calculado por un ser humano abstracto trabajando de una manera rutinaria es computable” (Gandy, 1980, p. 124). Finalmente, el tercer enunciado es el que Gandy llama Tesis M: “lo que puede ser calculado por una máquina es computable”. Lo calculable, lo realizable por un ser humano bajo ciertas restricciones y aquello que puede hacer una máquina son contextos todos ellos relevantes para entender a la computación, pero no equivalentes.

Estas versiones de la tesis Church-Turing suelen plantearse tomando en consideración la idea de procedimiento efectivo, la cual se entiende en términos de una máquina de Turing. A partir de lo reseñado por Gandy y de los recientes comentarios de Copeland y Shagrir (2007, 2019), se podrían hacer las siguientes consideraciones. En primer lugar, en la ciencia computacional contemporánea los algoritmos —o procedimientos efectivos— están asociados con *máquinas* y no con seres humanos (Copeland & Shagrir, 2019, p. 68). Entonces, ¿qué diferencias habría entre el tipo de restricción que encontramos en los seres humanos y en las máquinas?:

el argumento [de] Turing, contiene “pasos cruciales . . . donde él [Turing] apela al hecho de que el cálculo lo está realizando un ser humano”. Por ejemplo, Turing supuso que “un ser humano solo puede escribir un símbolo a la vez”, y Gandy señaló que esta suposición no se puede transferir a una máquina paralela que “imprime un número arbitrario de símbolos simultáneamente”. (Copeland & Shagrir, 2019, p. 69)

En segundo lugar, como es evidente, los algoritmos computacionales actuales están muy lejos de poder ser descriptos de manera *directa* por una máquina de Turing.

La máquina de Turing [no] puede duplicar (a diferencia de simular) *El juego de la vida* de John Conway, donde, a diferencia de una máquina de Turing, *cada celda se actualiza simultáneamente* [cursivas nuestras]. (Copeland & Shagrir, 2019, p. 68)

Siguiendo este razonamiento, Gandy, en su análisis de la idea mecánica de computación incluye el *paralelismo* como una propiedad de computadoras que no toman a los seres humanos como modelo y límite. En este sentido, más que una descripción completa de lo que es un algoritmo, una máquina de Turing representa un paso, crucial, pero en cierto sentido in-

suficiente, para la construcción de una noción adecuada de computación. Evidentemente el paso conceptual importante involucra comprender de qué manera un programa se relaciona con sus datos en diferentes niveles de análisis; lo cual es, en cierto sentido, equivalente a la idea de máquina universal. Sin embargo, a pesar de su importancia conceptual, y como vimos en el ejemplo presentado por Gandy, tampoco parece suficiente el mero señalamiento de este paso: hace falta la *especificación* de la manera en la cual esta idea puede implementarse.

Creemos que la discusión sobre la manera de articular las arquitecturas computacionales y el desarrollo de la actividad de programación en los '40 y '50 del siglo pasado pueden entenderse como parte de esta tarea de especificación.

En concreto, nuestra estrategia para las secciones siguientes será mostrar cómo en un principio los aspectos ligados a la efectividad en la tradición de las máquinas, a saber, velocidad, automatización y confiabilidad, se trasladan casi de manera directa a la construcción y al uso de las primeras computadoras. De esta manera, se entendía que una máquina solo puede ser más efectiva si hay una mejora cuantitativa en el *hardware*. Las primeras computadoras estaban en la senda marcada por la tradición de las máquinas, en el doble aspecto señalado más arriba.

Pero una progresiva comprensión de los problemas involucrados en la computación –a través del trabajo de von Neumann– permitió el surgimiento de un tipo de respuesta, una nueva arquitectura, que se traducía, por supuesto, en un nuevo *hardware*, pero cuya motivación era más bien conceptual. A su vez, problemas típicamente asociados con la efectividad como la mayor velocidad y la automatización comenzaron a ser abordados a partir de la codificación. De esta manera, lo que en la tradición de las máquinas era un aspecto principalmente asociado con la plasticidad o la flexibilidad, pasó a ser una estrategia o un camino para lograr mayor efectividad. Y es justamente esta relación original entre efectividad, de la manera en la cual la hemos propuesto, y programabilidad, que desdibuja el límite entre *software* y *hardware*, lo que, a nuestro juicio, caracteriza a la computación.

Veamos a continuación cómo puede leerse a partir de esta clave de análisis algunos aspectos históricos de la computación.

## Arquitecturas

La historia de las arquitecturas computacionales difícilmente pueda ser contada de manera lineal y clara. Las máquinas del siglo XIX, como el motor diferencial y el analítico de Babbage, realizaron aportes relevantes en cuanto a la organización, sin embargo, su importancia fue reconocida de manera tardía.<sup>1</sup>

Hay otros ejemplos del siglo XX como las máquinas construidas por Konrad Zuse desde 1935, Z1, Z2 y Z3, que pueden ser consideradas como las primeras computadoras electrónicas controladas por programas. Pero luego de ser destruidas en 1944, su relevancia histórica fue reconocida recién muchos años después. Por tanto, en este trabajo vamos a concentrarnos en las arquitecturas computacionales Harvard y von Neumann.

En un trabajo anterior hemos presentado los orígenes de la computación científica a través del trabajo de Howard Aiken y Leslie Comrie (Blanco & García, 2020). Aquí solo hace falta recordar la máquina que representa la arquitectura Harvard: la Mark I. A finales de los 1930, Aiken visita varias empresas con la idea de construir una máquina de cálculo.

Finalmente, IBM se interesa por el proyecto y entre 1939 y 1943 se construye la ASCC, Calculadora Secuencial de Control Automático [*Automatic Sequence Controlled Calculator*], llamada luego, cuando se instala en Harvard, la Mark I.

Esta máquina fue utilizada para cálculos “balísticos” (Priestley, 2011, p. 102). A esta computadora le siguieron una serie de máquinas diseñadas y construidas en Harvard en los ‘40 y ‘50 del siglo pasado (las Mark II, III y IV, de las cuales la Mark IV era completamente electrónica).

La complejidad técnica de esta máquina era importante. Tenía más de 700.000 componentes electromagnéticos y pesaba más de cuatro toneladas. Lo importante, para nosotros, es que en esta máquina había una separación clara entre *datos* e *instrucciones*. No solo había una separación física, sino que los *formatos* en los cuales estaban codificados los datos e instrucciones eran diferentes.

Una cinta separada podría contener números para entrada, pero los formatos de cinta no eran intercambiables. Las instrucciones no se podían ejecutar desde los registros de almacenamiento (Priestley, 2011).

---

<sup>1</sup> Por esta razón presentamos la efectividad y la programabilidad en la tradición de las máquinas de esta manera. Con el aporte de Turing podría leerse en retrospectiva de una manera diferente el trabajo de Babbage.

Esta estructura particular que caracteriza a una arquitectura Harvard estaba motivada por una razón de peso ya que dicha estructura permitía acceder de manera *simultánea* a los datos y al programa. La mayor *velocidad* era la razón que motivaba la apuesta por esta organización.

La otra arquitectura importante de esta época surgió en la Escuela Moore de Ingeniería Eléctrica de la Universidad de Pensilvania. A mediados de los años cuarenta del siglo pasado, J. Presper Eckert y John Mauchly construyeron una máquina automática de cálculo electrónica que llamaron ENIAC. Una de las funciones principales de esta máquina era calcular las tablas balísticas para el Laboratorio de Investigación correspondiente de la Armada (Aspray, 1990). Eckert y Mauchly al hacerse conscientes de las limitaciones de la ENIAC comienzan a discutir una versión mejorada de esta computadora: la EDVAC.

En 1944 había comenzado a participar de algunas reuniones en la Escuela Moore, John von Neumann, quien quería usar recursos de cálculo para su proyecto de la bomba atómica. De manera inmediata se compromete con la discusión del diseño de la EDVAC. Así, von Neumann escribió un reporte en 1945 –*First Draft of a Report on the EDVAC*. Este reporte suele ser considerado como la primera publicación del concepto de “programa almacenado”.

El aspecto definitorio de la arquitectura von Neumann es que la distinción entre *datos* y *programa* es relativizada desde el *hardware* mismo. Esta relativización implica un paso atrás, al menos en un sentido inmediato, en la búsqueda de mayor velocidad, mayor efectividad podríamos decir, para obtener una flexibilidad cualitativamente significativa. Evidentemente, la noción de máquina universal es la que permite describir este paso.<sup>2</sup>

Bajo esta perspectiva aparece una ventaja cualitativa de la arquitectura von Neumann con respecto a la de Harvard. Pero hay un aspecto del trabajo concreto de los investigadores en la Mark I, la arquitectura Harvard, que queremos destacar: la creciente importancia de la programación para solucionar problemas que en principio parecían ser de otra clase o abordables con otras herramientas. Además, como veremos, muchos de los avances tecnológicos relacionados con la programación no apuntaban solo a la *plasticidad* o flexibilidad del sistema, sino a lo que hemos llamado antes la *efectividad*, representada en la mayoría de los casos por la búsqueda

---

<sup>2</sup> Michael Godfrey cita en su introducción al comentario del *First Draft* que Turing en su reporte sobre la ACE cita el informe de von Neumann como un ejemplo de máquina de propósito general.

de mayor velocidad. También aparece aquí como objetivo la búsqueda de una mayor automatización. Todos estos aspectos muestran una creciente influencia de una manera de entender la efectividad y la programabilidad en términos computacionales, esto es en términos de *integración*.

Para mostrar estos aspectos y simplificar la presentación vamos a concentrarnos en el trabajo pionero de Grace Hopper en programación.

## Programación

Cuando Hopper fue enviada por la Armada a trabajar con Aiken en la Mark I, se dio cuenta de que para poder programar esa máquina necesitaba entender la *organización física* de la computadora (Beyer, 2009, p. 47). Hopper tuvo que aprender electrónica y estudiar los planos de circuitos; lo cual no representaba una tarea sencilla dado, como vimos, el tamaño y complejidad del aparato. Es más, si consultamos el manual de operaciones que escribió Hopper, veremos que en aquellos lugares en donde se explica cómo programar la computadora aparecen, habitualmente, diagramas de circuitos.

Hay que señalar que estos límites de la Mark I, en cuanto a su configuración para cómputos específicos, no dependían de su carácter *mecánico*. Algo similar ocurría con los primeros sistemas computacionales *electrónicos*. La computadora ENIAC, de la Universidad de Pensilvania era mucho más rápida por tener un sistema de tubos de vacío. Pero su «lógica» era la misma: codificar la máquina era equivalente a la manipulación del *hardware* (Beyer, 2009, p. 51). A fines de los cuarenta Eckert y Mauchly contratan a seis mujeres del Laboratorio de Investigación Balística de la Armada para trabajar en la ENIAC.<sup>3</sup> Al igual que Hopper, estas mujeres pronto se dieron cuenta que para programar la ENIAC tenían que conocer el *hardware*.<sup>4</sup>

Pero Hopper había desarrollado en poco tiempo, una estrategia de trabajo con la computadora mecánica *Mark I* que privilegiaba la codificación. Lo cual contrastaba con la aproximación que se llevaba adelante con otras computadoras más avanzadas desde el punto de vista del *hardware*.

---

<sup>3</sup> Los otros programadores de la Eniac fueron Kay McNulty, Betty Jennings, Betty Snyder, Marlyn Wescoff, Fran Bilas y Ruth Lichterman.

<sup>4</sup> Este aspecto se reflejaba en una división de trabajo clara: los oficiales eran los *codificadores* y los reclutas eran los *operadores* que cableaban la máquina.

En palabras de la propia Hopper:

Quando visité [la gente que trabajaba en] ENIAC, el tremendo contraste entre Harvard y las actividades de Eckert y Mauchly fue la *programación*. Lo de ENIAC era: conectar piezas y esencialmente construir una computadora *especial* para cada trabajo, y nosotros estábamos acostumbrados al concepto de programación y de controlarlo a través de nuestro programa. (citada en Beyer, 2009, p. 52)

De esta manera, aquí estaba en consideración también, la apuesta por máquinas de propósito general y de propósito especial que caracterizó los inicios de la computación científica.

Ahora bien, ¿cómo había llegado Hopper a esta idea de la relevancia de la «programación»? Ella comenzó a trabajar en la Mark I durante los últimos años de la segunda guerra mundial y los requerimientos del ejército eran apremiantes. Sin embargo, solo Hopper y Bob Campbell eran los encargados de manejar la máquina. Y manejar la máquina significaba, como ya vimos, cambiar la organización de los cables que constituían la computadora. Dada la urgencia y cantidad de demanda comenzaron a construir «rutinas» de código con una aproximación sistemática para ahorrar en trabajo de modificación del cableado físico (Beyer, 2009, p. 53).

Ahora bien, la *Mark I* era demasiado lenta para la demanda durante la guerra. Funcionaba 24/7 pero no era suficiente. Por esta razón, Hopper y Bloch comenzaron en 1944 a optimizar las rutinas del *software* para minimizar la *cantidad de ciclos* necesarios para realizar cada computo. Así, se redujo en un 36 por ciento el tiempo de ejecución de los programas. La búsqueda era de programas más *eficientes*. La contrapartida fue que los programas se hicieron cada vez más complejos y difíciles de entender.

Otra cuestión era el formato del resultado. En un principio la Mark I imprimía un conjunto de números que tenía que ser interpretado por un operario quien, utilizando una máquina de escribir «traducía» el resultado de la computadora. A Hopper se le ocurrió *automatizar* ese proceso con un programa para traducir el resultado.

Así, como muestran estos ejemplos, a los problemas de mayor velocidad y de automatización se los comenzó a atacar con soluciones de codificación. Evidentemente, la importancia de la aproximación de la automatización a través de la codificación podía verse también en el trabajo con otras computadoras de la época. En 1950 Mauchly comenzó a utilizar en la UNIVAC lo que llamó «códigos breves» [*Short codes*], los cuales consistían

en ordenes básicas o “código especial elegido para simplificar el trabajo del programador humano y dejar de lado mucho del trabajo tedioso de codificación detallada en la computadora” (Priestley, 2011, p. 187).

Pero esta manera de trabajar tenía como contrapartida, tal como señalamos más arriba, una complejización del código. La *accesibilidad epistémica*, un problema hasta ahora típicamente asociado con la complejidad de la máquina física,<sup>5</sup> aparece en el contexto de la tarea de codificación. Además, la multiplicidad de códigos asociados con máquinas específicas venía a agravar este problema.

El señalamiento de estos problemas provocó diferentes respuestas. Por ejemplo, a fines de los cincuenta se propone un lenguaje de programación, Algol 58, con la finalidad de colaborar en la estandarización. Uno de los objetivos de este lenguaje era la “expresión y comunicación de los algoritmos” (Priestley, 2011, p. 206). Además, el lenguaje debía ser «trasladable» mecánicamente en programas de máquina. Sin embargo, un aspecto inusual, para la época, era que este lenguaje no hacía referencia a una máquina particular. Las dificultades para implementar Algol 58 llevaron al desarrollo de Algol 60 cuyo objetivo principal, ahora mucho más explícito, era la *comunicación de los algoritmos*.

Algunos hitos del trabajo posterior de Hopper muestran otro aspecto interesante. Cuando Hopper comenzó a escribir subrutinas para la UNIVAC, a partir de 1951, de manera semejante a las implementadas en la Mark I, su trabajo derivó en la construcción del sistema A-0. Desde una perspectiva contemporánea el sistema A-0 es considerado como el primer compilador. Pero, en esa época, no era sencillo conceptualizar qué se estaba haciendo. Priestley (2011) señala que había varias interpretaciones sobre este tipo de programas. En primer lugar, se interpretaban como *extensiones* del código de máquina. En segundo lugar, una interpretación habitual era como una *traducción de un tipo de lenguaje* al lenguaje de máquina. Otra interpretación era verlo como una *traducción o extensión de la máquina misma*.

Esta última idea es la que también había propuesto Turing para entender la clase de algoritmo involucrado aquí. Turing lo llamaba «rutina interpretativa» y lo describía como aquellos programas que “permitían a la computadora convertirse en una máquina que usa una instrucción

---

<sup>5</sup> La tapa de la revista especializada en computación *Communications of the ACM* de 1961 llevaba la imagen de la torre de Babel con los nombres de los lenguajes de programación más populares de la época.

de código diferente a la originariamente requerida para el diseño de la computadora” (Priestley, 2011, p. 190). Esta referencia sirve para poner de manifiesto cómo el problema de la construcción de un tipo de programa que permita comunicarse de manera más simple con el lenguaje de máquina, desde un nivel conceptual, es un tipo de problema similar al que abordaban aquellos que proponían las arquitecturas von Neumann. Y es finalmente el mismo tipo de solución que propuso Turing, aunque desde una perspectiva más general, con su máquina universal.

## Consideraciones finales

En este trabajo hemos intentado mostrar las ventajas que tendría analizar aspectos conceptuales e históricos de la computación a partir de las dimensiones de la efectividad y la programabilidad. Estas dimensiones aparecen configuradas de manera diferente en lo que hemos llamado la tradición de las máquinas y en la tradición de las computadoras.

Sin embargo, las primeras computadoras construidas en la primera mitad del siglo XX muestran una continuidad importante con la tradición de las máquinas. Esto se evidencia tanto en las arquitecturas, ejemplificada en la Mark I de Harvard, como en la práctica de la programación entendida en términos de modificación de cableado físico de las máquinas. Además, esta continuidad permite ver el paulatino desplazamiento de problemas típicamente asociados con la efectividad de las máquinas al terreno de la programación. Este desplazamiento no fue homogéneo ni directo en el caso de las arquitecturas y los lenguajes –como vimos con el caso de la importancia cada vez mayor de la programación en la Mark I. Este aspecto histórico permite destacar con mayor fuerza el creciente desdibujamiento entre los límites de los problemas tradicionalmente asociados con el *hardware* y el *software*. Lo que venía a representar la solución teórica de Turing en 1936, en lo que hemos llamado en otros trabajos una noción relacional de computación, se fue configurando como un horizonte, no siempre claro, de solución de problemas. En este camino de solución de problemas concretos, de velocidad, automatización y confiabilidad se fueron reconfigurando las notas características de lo que significaba efectividad y programabilidad en la tradición de las máquinas.

## Referencias

- Aspray, W. (1990). The stored program concept. *IEEE Spectrum*, 27(9), 51.
- Beyer, K. (2009). *Grace Hopper and the invention of the information age*. Cambridge, MA.: MIT Press.
- Blanco, J., & García, P. (2020). En torno a las nociones de efectividad y programabilidad para comprender la noción de computación. En M. O'Lery, L. Federico, & Y. Ariza (Eds.), *Filosofía e historia de la ciencia en el cono sur. Selección de trabajos del XI Encuentro de la Asociación de Filosofía e Historia de la Ciencia del Cono Sur* (pp. 102-109). São Carlos-Buenos Aires: AFHIC.
- Copeland, B. J., & Shagrir, O. (2007). Physical computation: How general are Gandy's principles for mechanisms? *Minds and Machines*, 17(2), 217-231.
- Copeland B. J., & Shagrir, O. (2019). The Church-Turing thesis: logical limit or breachable barrier? *Communications of the ACM*, 62, 1, 66-74
- Gandy, R. (1980). Church's thesis and principles for mechanisms. En J. Barwise, H. J. Keisler, & K. Kunen (Eds.), *Studies in logic and the foundations of mathematics* (pp. 123-148). Elsevier.
- Priestley, M. (2011). *A science of operations: machines, logic and the invention of programming*. Springer Science & Business Media.



## Definiciones implícitas y estructuralismo matemático *ante rem*

Eduardo N. Giovannini\*.\*

Mayra Huespe\*

### Introducción

El método de las definiciones implícitas juega un papel central en la filosofía contemporánea de la matemática, en particular en el debate sobre el estructuralismo matemático. Esta influyente posición filosófica sostiene que la matemática no se ocupa del estudio de las 'propiedades internas' de sus objetos, sino más bien del modo en que estos objetos se 'relacionan entre sí' (Shapiro, 1997; Parsons, 1990). Para el estructuralismo matemático, las teorías matemáticas describen estructuras abstractas, en oposición a sistemas particulares de objetos y relaciones que ejemplifican estas estructuras. Por ejemplo, la aritmética de Peano describe la estructura de los números naturales, que puede ser interpretada o tener un modelo en diversos sistemas conjuntistas; el análisis matemático describe la estructura de los números reales, la geometría la estructura de diferentes espacios, y así sucesivamente.

El estructuralismo *ante rem* defendido por Shapiro (1997) constituye una de las variantes más elaboradas del estructuralismo matemático. Esta variante postula a las estructuras como entidades existentes en un dominio abstracto. Más precisamente, para el estructuralismo *ante rem* las estructuras abstractas y las posiciones deben ser consideradas como objetos *bona fide* que existen independientemente de los sistemas de objetos que constituyen sus instancias concretas. En otras palabras, esta variante del estructuralismo matemático se caracteriza por postular la independencia ontológica de las estructuras y los 'lugares' o 'posiciones' en las estructuras respecto de los sistemas particulares de objetos que las instancian o ejemplifican. En este contexto, una piedra angular de la epistemología del estructuralismo *ante rem* consiste en la tesis según la cual las

\* Universidad Nacional del Litoral (UNL).

\* CONICET.

mayrahuespe@hotmail.com



definiciones implícitas, *i.e.*, las axiomatizaciones de teorías matemáticas, nos proporcionan un acceso a las estructuras ‘puras’ y a las ‘posiciones’ en las estructuras.

Ahora bien, a pesar del papel central que desempeña el método de las definiciones implícitas en el estructuralismo *ante rem* desarrollado por Shapiro (1997), resulta llamativo que la noción misma de definición implícita es descripta allí de un modo esquemático, e incluso a veces informal. El objetivo del presente trabajo es examinar críticamente el papel desempeñado por el método de las definiciones implícitas en la variante del estructuralismo matemático *ante rem* elaborada por Shapiro (1997). En primer lugar, sostendremos que, en la descripción de dicho método presentada por este autor, es posible distinguir dos nociones diferentes de definición implícita. Esquemáticamente, por un lado, una primera noción según la cual una definición implícita consiste en la especificación por medio de un sistema de axiomas del significado de los términos primitivos de una teoría matemática. Por ejemplo, en el caso de la axiomatización de la geometría euclídea elaborada por Hilbert (1899), los axiomas de dicha teoría proporcionan una definición implícita de los conceptos primitivos ‘punto’, ‘línea’, ‘plano’, ‘estar en’, ‘estar entre’, etc. Por otro lado, una segunda noción de definición implícita, según la cual un conjunto de axiomas define un concepto de objeto matemático de orden superior, o más precisamente, una clase de modelos o estructuras. En el ejemplo antes mencionado, los axiomas de Hilbert (1899) definen el concepto de orden superior de espacio euclídeo tridimensional. Luego, una segunda tesis que defenderemos en este trabajo consiste en sostener que estas dos nociones de definición implícita no sólo difieren en cuanto a su funcionamiento semántico, sino además en lo que respecta a sus criterios de éxito o adecuación.

El trabajo estará estructurado de la siguiente manera. En primer lugar, realizaremos una presentación esquemática de la versión del estructuralismo *ante rem* propuesta por Shapiro. En particular, prestaremos especial atención a la distinción trazada por este autor entre teorías matemáticas “algebraicas” y “no-algebraicas”. En segundo lugar, analizaremos las dos nociones de definición implícita antes mencionadas y examinaremos cómo dichas nociones pueden ser caracterizadas de un modo más preciso en términos modelo-teóricos. En tercer lugar, discutiremos el papel epistemológico asignado por Shapiro a dichas nociones, y en particular, identificaremos una serie de problemas en relación los criterios de ade-

cuación de “unicidad” y “existencia” que deben satisfacer ambas nociones de definición implícita.

### **El estructuralismo matemático *ante rem***

Shapiro (1997) afirma que las estructuras matemáticas están conformadas por lugares o posiciones que mantienen entre sí relaciones de carácter formal. A tales lugares los pueden ocupar diversos tipos objetos que respetan las relaciones formales en cuestión. Estos conjuntos de objetos que instancian las estructuras reciben el nombre de sistemas. Tanto las estructuras como los lugares son ontológicamente independientes de los sistemas y los objetos que las instancian. Esta independencia está estrechamente vinculada con el carácter formal de las relaciones entre las posiciones de las estructuras. Los objetos que componen los sistemas pueden estar relacionados de otros modos, además de formalmente. Por lo cual, tales elementos pueden tener otras propiedades, además de las estructurales, como por ejemplo propiedades físicas. Las posiciones, por el contrario, únicamente tienen propiedades de carácter formal o estructural. Lo único que importa de ellas es cómo se relacionan con los demás elementos de la estructura.

Desde cierta perspectiva, las estructuras pueden ser analizadas como sistemas. Por lo cual, todas las estructuras pueden ejemplificarse a sí mismas. Esta relativización conceptual se entiende mejor a la luz de las perspectivas *places are offices* y *places are objects*. Desde la primera de ellas, las estructuras se analizan en los términos de los sistemas que las ejemplifican y los lugares en base a los objetos que los ocupan. Los supuestos términos singulares que refieren a las posiciones se interpretan como variables ocultas. Desde esta perspectiva, los numerales son interpretados como variables cuyos rangos son todos los objetos de los sistemas que ejemplifican a la estructura de la aritmética. Por el contrario, la perspectiva *places are objects* permite analizar a los lugares como objetos *bona fide*. Los términos singulares recuperan su estatus. De allí que, una estructura dada –analizada desde la perspectiva *places are objects*– puede ejemplificarse a sí misma. Pese a la validez teórica de ambas perspectivas, las tesis ontológicas centrales del estructuralismo *ante rem* son formuladas desde la perspectiva *places are objects*. En sí misma, la validez teórica de las dos perspectivas no es más que un resultado del realismo estructural defendido por el autor.

Para Shapiro, las teorías matemáticas algebraicas son aquellas que es-

tudian diferentes clases de modelos no equivalentes, como por ejemplo la teoría de grupos y la teoría de anillos. Contrariamente, las teorías no algebraicas, como por ejemplo la aritmética, estudian clases de modelos equivalentes:

A field is nonalgebraic if it has a single “intended” interpretation among its possible models or, more precisely, if all of its “intended” models are isomorphic (or at least equivalent). . . . A field is algebraic if it has a broad class of (nonequivalent) models. Group theory, for example, is about all groups. (Shapiro, 1997, p. 50)

Siguiendo esta distinción, las teorías matemáticas no algebraicas estudian ciertos dominios matemáticos definidos y las teorías algébricas, por el contrario, estudian ciertas clases de permutaciones que están presentes en diferentes áreas de la matemática. Por lo cual, bajo los postulados ontológicos del estructuralismo que defiende Shapiro, una teoría no algébrica dada estudia una estructura *ante rem* específica o una clase de estructuras isomórficas y una teoría algébrica, por el contrario, estudia diferentes clases de estructuras no isomórficas.

## **Dos tipos de definición implícita para el estructuralismo matemático**

El método de las definiciones implícitas ocupa un papel destacado en la concepción estructuralista en la filosofía contemporánea de la matemática. La idea central aquí es que un modo efectivo de comprender y adquirir conocimiento de una estructura matemática es a través de una descripción directa de la misma. Ello se logra proporcionando una axiomatización de la teoría matemática que determina –si es exitosa– la clase de todos sus modelos. La tesis de que la matemática consiste en el estudio de las estructuras matemáticas y de las propiedades estructurales de sus objetos es así una motivación fundamental en el llamado método de las definiciones implícitas. Más específicamente, la noción de definición implícita que resulta más habitual en este contexto consiste en sostener que lo que se define por medio de un sistema de axiomas es la clase de los modelos de la teoría matemática.

Esta concepción de las definiciones implícitas resulta además particularmente apropiada para las versiones no-eliminativista del estructuralismo matemático, especialmente para el estructuralismo *ante rem*. En efecto, para esta posición resulta esencial que uno pueda comunicar y re-

ferirse a las estructuras puras y a las posiciones en las estructuras puras, independientemente de sus instanciaciones concretas. Ahora bien, una definición implícita exitosa no define un sistema matemático concreto, sino la clase de los modelos de la teoría axiomática. En consecuencia, en virtud de esta concepción de definición implícita, es posible describir una estructura matemática pura, incluso si ninguna instancia particular de ella es conocida o exhibida. En palabras de Shapiro:

Structures successfully described by implicit definitions are naturally construed as *ante rem* (if they exist at all, of course). . . . In an implicit definition, asking about Julius Caesar is similar to the aforementioned listener who is wondering about the name of the center fielder's mother. The speaker was describing a structure, not a system of particular people. The mathematics book is not describing a system of sets or Platonic objects or people. It describes a structure or a class of structures. (Shapiro, 1997, pp. 130-132)

Ahora bien, es interesante observar que otras descripciones realizadas por Shapiro (1997) de este método definicional sugiere una noción diferente de definición implícita. De acuerdo con esta segunda noción, lo que se define a través de un sistema de axiomas es el significado de los términos primitivos de una teoría matemática. Esta concepción alternativa es sugerida en pasajes como los siguientes:

There is an ambiguity in the phrase "implicit definition". . . . In the present context is a simultaneous characterization of a number of items in terms of their relations to each other. In contemporary philosophy, such definitions are sometimes called "functional definitions". (Shapiro 1997, p. 130)

The other sense of "implicit definition" occurs where one defines several new linguistic items in terms of each other (and terms already in use). Hilbert's axiomatization of geometry is an implicit definition in this sense: "point", "line", and "plane" are defined in terms of their relations to each other. (Shapiro, 1999)

La concepción de las definiciones implícitas según la cual los axiomas de una teoría matemática definen el significado de los términos primitivos tiene sus orígenes en el surgimiento de la moderna concepción de método axiomático, en los trabajos de Hilbert, Dedekind y Peano, entre otros. Un aspecto central de esta concepción, sin embargo, es que no resulta claro cómo funcionan semánticamente este tipo de definiciones. En efecto, a

partir de la aparición de dichos trabajos, existe un consenso en la literatura sobre la axiomática formal de que los sistemas axiomáticos –tomados como definiciones implícitas de los términos primitivos– no deben fijar una única referencia para dichos términos. Es decir, en una teoría axiomática formal los conceptos y relaciones primitivas no están ligados a una referencia fija, sino que puede ser libremente interpretados. La especificación de los términos primitivos a través de un sistema formal de axiomas es así dejada de un modo deliberado semánticamente indeterminada. Este hecho ha provocado que numerosos autores rechacen esta noción de definición implícita, en tanto consideran que no puede ser explicada de un modo adecuado en términos de una semántica basada en la teoría de modelos (cf. Hodges, 2001). Sin embargo, el estructuralismo *ante rem* parece vindicar la noción, aunque sin proporcionar una explicación explícita de cómo los axiomas de una teoría matemática pueden especificar el significado de sus términos primitivos y cuál es su relación con la concepción de las definiciones explícitas como definiciones de clases de modelos o tipos isomórficos de estructuras.

### **Definiciones estructurales: criterios de adecuación**

Del mismo modo que la noción de definición explícita está ligada a estrictos criterios de adecuación (i.e., eliminabilidad y no-creatividad), una tarea central en la teoría de las definiciones implícitas consiste en especificar sus criterios de adecuación. Shapiro identifica dos criterios fundamentales: existencia y unicidad.

El primer requerimiento que debe cumplir una definición implícita para definir exitosamente a una estructura es la unicidad: la definición debe caracterizar *a lo sumo* una estructura (salvo isomorfismo). Shapiro entiende a este primer criterio en términos semánticos: los axiomas que sirven de definiciones deben ser categóricos. Esta maniobra impone una primera limitación sobre las teorías matemáticas axiomatizadas en una lógica de primer orden. En efecto, el teorema de Löwenheim-Skolem muestra que toda teoría contable de primer orden con un modelo infinito, tiene un modelo de cualquier cardinalidad infinita. Este resultado dificulta considerablemente el cumplimiento del requisito de unicidad, debido a que no permite descartar los modelos no estándares de las definiciones implícitas más interesantes. Para sortear este problema, Shapiro propone circunscribirse a los lenguajes de segundo orden (u orden superior).

El segundo requerimiento que propone Shapiro es la existencia: las definiciones implícitas deben caracterizar *al menos* una estructura. Para caracterizar este criterio, Shapiro apela al principio de coherencia que formula en el marco del desarrollo de su propuesta ontológica estructuralista. Según este principio, si es una fórmula coherente en un lenguaje de segundo orden, entonces existe una estructura que satisface a  $\Phi$  (cf. Shapiro, 1997, p. 95). Los problemas en la caracterización de este segundo requisito surgen al tratar de proporcionar una definición rigurosa de coherencia. Una primera opción puede ser apelar a la consistencia deductiva. Es decir, sostener que, si no es posible derivar consecuencias contradictorias de un conjunto de axiomas, entonces esos axiomas describen al menos una estructura. El fundamento riguroso de esta primera opción es el teorema de Gödel: si un conjunto de sentencias de primer orden es deductivamente consistente, entonces tiene un modelo.

Esta maniobra tiene dos problemas fundamentales. En primer lugar, es circular: el teorema de completitud es un resultado en la matemática, específicamente en la teoría de conjuntos. En este sentido, los diversos modelos de axiomatizaciones consistentes se encuentran en la jerarquía conjuntista, que es otra estructura. El segundo problema surge al tratar de articular ambos criterios. El cumplimiento de unicidad y de existencia depende de la lógica subyacente que se emplea en la axiomatización de una teoría matemática. De allí que, ambos requisitos no pueden ser satisfechos simultáneamente. Esquemáticamente, mientras toda teoría consistente de primer orden tiene un modelo (Teorema de Completitud de Gödel), las teorías de primer orden no son categóricas, esto es, tienen modelos no-estándar (Teorema de Löwenheim-Skolem). Por el contrario, mientras las teorías matemáticas de segundo orden o superior son categóricas, existen teorías de segundo orden que son (deductivamente) consistentes, pero no son satisfacibles (Teorema de Incompletitud de Gödel).

Una posible solución a este primer problema es entender a la coherencia como una analogía de satisfacibilidad: una sentencia  $\Phi$  es satisfacible si existe un modelo de  $\Phi$ . En este contexto se manifiesta nuevamente la circularidad de la maniobra de Shapiro. En efecto, si la coherencia se explica a partir de la satisfacibilidad, entonces la locución “existe” se entiende como “es un miembro de la jerarquía conjuntista”, que es otra estructura. Shapiro reconoce la circularidad en su propuesta, pero argumenta que tal circularidad no es viciosa:

There is no getting around this situation. We cannot ground mathematics in any domain or theory that is more secure than mathematics itself. All attempts to do so have failed, and once again, foundationalism is dead (see Shapiro [1991, chapter 2]). The circle that we are stuck with, involving second-order logic and implicit definition, is not vicious and we can live with it. I take “coherence” to be a primitive, intuitive notion, not reduced to something formal, and so I do not venture a rigorous definition. (Shapiro, 1997, p. 135)

La estrategia adoptada por Shapiro para solucionar este último problema consiste en renunciar a dar una caracterización rigurosa del criterio de existencia en términos de la noción conjuntista de satisfacibilidad, y sustituirla por la noción primitiva de coherencia. La coherencia no es definida, sino que es explicada a partir de la noción conjuntista de satisfacibilidad. Más específicamente, la satisfacibilidad es a la coherencia lo que la noción de recursividad es a la computabilidad. En resumen, coherencia es la noción intuitiva que sirve como criterio para la existencia de estructura y satisfacibilidad es su análogo matemático riguroso.

## Discusión

En esta sección desarrollamos una serie de consideraciones críticas en relación a la descripción del método de las definiciones implícitas realizada por Shapiro y a su variante *ante rem* del estructuralismo matemático. Más precisamente, concentramos nuestro análisis en tres aspectos que consideramos problemáticos en dicha propuesta. En primer lugar, problematizamos la posibilidad de aplicar los mismos criterios para los usos diferentes que Shapiro les atribuye a las definiciones implícitas. En segundo lugar, señalamos la dificultad de compatibilizar el requerimiento de unicidad, entendido como categoricidad, con la distinción entre teorías algebraicas y teorías no algebraicas que propone Shapiro (1997, 2005). Por último, mostramos en qué sentido el éxito del rol epistemológico que Shapiro le asigna a las definiciones implícitas depende de un posicionamiento antifundacionalista en filosofía de las matemáticas.

## 1. Definiciones funcionales y estructurales

Un primer aspecto problemático se manifiesta al comparar las definiciones funcionales y las estructurales. Como hemos visto, para Shapiro, la noción de definición implícita puede entenderse en dos sentidos. Desde el primero de ellos, las definiciones describen estructuras o clases de estructuras. El segundo uso, por otro lado, refiere a la caracterización simultánea de un número de ítems en términos de las relaciones que mantienen entre sí.

La comparación de las definiciones estructurales con las funcionales sugiere que la segunda apunta a definir los primitivos de las teorías matemáticas. Sin embargo, las definiciones implícitas en primer sentido son claramente definiciones de clases de estructuras. En el caso de ambas nociones pueden ser defendidas, una pregunta pertinente es: ¿cómo puede explicarse la relación entre ambos tipos de definiciones? Los criterios de adecuación mencionados para las definiciones implícitas –unicidad y existencia– claramente se aplican a la concepción de definición implícita como clases de modelos o estructuras. En este sentido, otra pregunta relevante es ¿pueden aplicarse sin más los mismos criterios para la noción de definición implícita como definición del significado de términos primitivos?

## 2. Definiciones estructurales y teorías algebraicas

Dada la distinción entre teorías algebraicas y no algebraicas, Shapiro nos debe una explicación de cómo el método de las definiciones estructurales puede ser aplicado exitosamente para adquirir conocimiento de teorías matemáticas que son esencialmente algebraicas. En la medida en que se presenta a la categoricidad (unicidad) como una condición necesaria para el conocimiento de estructuras *ante rem*, no resulta claro cómo las definiciones estructurales pueden proporcionar un conocimiento de aquella clase de estructuras. En efecto, un rasgo central de las teorías algebraicas (por ejemplo, la teoría de grupos o la de anillo) es que no son descritas categóricamente, al menos en sus formulaciones habituales. Ello plantea entonces el interrogante de cómo el método de las definiciones implícitas puede resultar una técnica efectiva para adquirir conocimiento de las llamadas teorías matemáticas algebraicas. ¿Significa ello que las teorías esencialmente algebraicas no pueden ser consideradas desde la perspectiva *ante rem*? ¿O se trata más bien de un límite para el método de las definiciones estructurales (y sus criterios de adecuación)?

### 3. Definiciones estructurales, antifundacionalismo y practica matemática

Shapiro intenta fundar nuestro conocimiento de las estructuras matemáticas en nuestra comprensión de la coherencia y la categoricidad de las descripciones (*i.e.*, definiciones implícitas) de estas estructuras. Estas nociones son explicadas, aunque no definidas, en términos conjuntistas. La coherencia y la categoricidad de las descripciones, y por consiguiente la existencia de las estructuras de las mismas, debe ser decidida o establecida dentro de la teoría de conjuntos. Para Shapiro esto no representa un círculo vicioso, puesto que la plausibilidad del conocimiento matemático de estructuras *ante rem* no puede ser “probada” a partir de premisas “extra-matemáticas”:

In mathematics as practiced, set theory (or something equivalent) is taken to be the ultimate court of appeal for existence questions. . . . the thesis that satisfiability is sufficient for existence underlies mathematical practice. . . . Structuralists accept this presupposition and make use of it like everyone else, and we are in no way better (and no worse) of a position to justify it. This presupposition is not vicious, even if it lacks external justification. (Shapiro, 1997, p. 136)

En este contexto se evidencia que, un presupuesto necesario para asegurar el éxito del rol epistemológico que Shapiro le asigna al método de las definiciones, en el marco del estructuralismo *ante rem*, es su posicionamiento antifundacionalista.

#### Perspectiva

A modo de conclusión, en esta sección final proponemos dos ejes articuladores para aplicar en ulteriores investigaciones, tomando como base para su formulación los problemas detectados en el presente trabajo. Un primer eje articulador surge al analizar la función de las estructuras *ante rem* en la propuesta epistemológica del estructuralismo de Shapiro. El conocimiento de la coherencia y categoricidad de una descripción se reduce, en última instancia, a la cuestión de la existencia de conjuntos. En este sentido, la ontología del estructuralismo *ante rem* no parece desarrollar un papel substancial en la parte más fundamental de la epistemología de la matemática de Shapiro. Una de las críticas de MacBride (2008) a la propuesta estructuralista apunta a este aspecto: “Shapiro’s account offers no

distinctively structuralist insight into the epistemology of mathematics” (MacBride, 2008, p. 163). En este marco, el primer eje articulador que proponemos es analizar si las definiciones estructurales, en el marco de la propuesta *ante rem*, logran solucionar el problema del acceso, como Shapiro pretende.

El segundo eje articulador gira en torno a la categoricidad en tanto condición necesaria de adecuación de las definiciones estructurales. Por el teorema de quasi-categoricidad de Zermelo sabemos que, dados dos modelos de ZFC2, o bien son isomórficos o bien uno de ellos es isomórfico a un segmento inicial propio del otro. En este sentido, la quasi-categoricidad sería suficiente para determinar referencia, al menos interna a la teoría de conjuntos. El conjunto vacío,  $\omega$ , el continuo, etc., representan lo mismo en todos los modelos (salvo isomorfismo). En este marco, el segundo eje articulador que proponemos consiste en indagar sobre la caracterización de la condición de unicidad de las definiciones estructurales, teniendo como base los distintos resultados de categoricidad (interna, externa, cuasi-categoricidad, etc.).

## Referencias

- Hilbert, D. (1899). *Grundlagen der Geometrie [Foundations of geometry]*. Leipzig: Teubner.
- Hodges, W. (2001). Model theory. En E. Zalta, *The Stanford encyclopedia of philosophy* (Fall 2018 Edition). <https://plato.stanford.edu/archives/fall2018/entries/model-theory/>
- Macbride, F. (2008). Can «ante rem» structuralism solve the access problem? *The Philosophical Quarterly* (1950-), 58(230), 155-164.
- Parsons, C. (1990). The structuralist view of mathematical objects. *Synthese*, 84, p. 303-346.
- Shapiro, S. (1997). *Philosophy of mathematics: Structure and ontology*. Oxford: Oxford University Press.
- Shapiro, S. (1999). Implicit definition and abstraction [manuscrito no publicado]. *Unpublished notes*.
- Shapiro, S. (2005). Categories, structures, and the Frege–Hilbert controversy: The status of meta-mathematics. *Philosophia Mathematica*,



## Datos y semántica en computación Análisis de la discusión de Fresco y Wolf con Floridi

Xavier Huvelle\*

### Introducción

En los últimos años la discusión sobre la naturaleza de la información ha resurgido en gran parte gracias a Floridi. Pero esta discusión no es nueva y ya fue ampliamente abordada por la Cibernética a principios del Siglo XX, en particular por Wiener (1948), Ashby (1960), Shannon (1948) y Simon (1947). Pero este abordaje e identificación contemplan nociones muy dispares acerca de la información, o por lo menos plantean funciones distintas. Van Benthem y Adriaans (2008) enumeran por lo menos tres grandes tradiciones que se fueron conformando con el tiempo: una noción de información algorítmica y compresión (tipo Kolmogorov), una noción probabilística y cuantitativa de la información (tipo Shannon) y una noción basada en conocimiento y lógica, en la que un agente usa y representa información. Resultaría complicado poder encajar a los Cibernéticos en alguna de estas nociones dado que comparten varios elementos de estas perspectivas. Por ejemplo, plantear una noción de información sin agente y dejar de lado la teoría de la información de Shannon resultaría imposible para Simon. En cambio, para Shannon el estatuto del agente no importa, sino que importa lo que se transmite, no por quién. Hoy en día, estas cuestiones son relevantes y nos permiten observar conflictos en algunos de los supuestos que proponían los Cibernéticos, en particular acerca de la naturaleza de la información. Aunque los Cibernéticos hablan en términos “físicos” lo que más interesa para estos autores es responder la pregunta por cómo se procesan los datos y la información, pero no realmente sobre lo propiamente físico de la información. ¿Es la información una señal eléctrica, un signo lógico u otro elemento? Fresco y Wolf (2014), para responder a esta pregunta, propusieron una noción de computación física no simbólica basada en

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades, Córdoba, Argentina.  
xavier.huvelle@gmail.com

la *información instruccional* de la teoría de la información propuesta por Floridi (2010, 2011). En un artículo de 2018, Fresco y Wolf vuelven sobre esta noción de información instruccional y proponen cambiarla por una *información funcional* de tipo *teleológica* basada en Corning (2007). En este trabajo presentamos esta discusión entre Fresco, Wolf y Floridi respecto de la información instruccional, además presentamos tres críticas basadas en la lógica programada. Estas críticas cuestionan la propuesta de Fresco y Wolf para plantear si la información instruccional (funcional) que proponen está bien caracterizada, si su propuesta puede evitar el *pancomputacionalismo* y si pueden prescindir de un contenido semántico externalista como ellos afirman.

## Discusión

El origen de la concepción actual de información se encuentra atado con la teoría de la información de Shannon y de la Cibernética (Ashby 1960; Wiener 1948). La perspectiva tomada por Shannon y Wiener se basa en el aspecto ingenieril de transmisión de la información y no tanto con los elementos semánticos que la componen. Floridi (2010, 2011) ofrece, por su lado, una noción muy peculiar de información semántica que distingue con una caracterización de los *datos*. Los datos en esta perspectiva son la falta de uniformidad entre por lo menos dos variables no interpretadas que son distintas una de otra y que han sido dejadas para futuras interpretaciones (Floridi, 2011, p. 85). Mientras que la información semántica se trata de una definición formal en la que un objeto  $O$  es una instancia de información semántica, entendida como contenido semántico, si y solo si: (a) consiste en  $n$  data  $D$ ; (b)  $D$  está bien formado; (c)  $D$  es significativamente independiente de su recipiente (Floridi, 2010, p. 21). El punto (a) es lo que se entiende por un dato positivo y (b) significa que el dato es estructurado de acuerdo a algunas reglas (reglas sintácticas, por ejemplo). Cuando (a) y (b) son combinados, forman la forma más simple de información: *Datos estructurados* (Fresco & Wolf, 2014, p. 1472). Mientras que (c) es el *contenido semántico* y requiere que el dato estructurado se adecue con la semántica del sistema o del código. Fresco y Wolf (2014) recuperan las caracterizaciones ofrecidas por Floridi para elaborar una noción de información para la computación digital. La información instruccional es entonces prescriptiva y no descriptiva. Fresco y Wolf dirán que es más bien imperativa y la definen de esta manera: “el contenido semántico es

una instancia de la información instruccional si y solo si su satisfacción resulta en una acción en un contexto dado” (2014, p. 1473).

No hizo falta esperar mucho para que en un artículo de Fresco y Wolf (2018) se pueda leer: “La conclusión principal de este artículo es que, si adoptamos la posición conceptual de Floridi sobre información para formular una postura de procesamiento de la información de la computación, [esta] nos lleva a un callejón sin salida conceptual” (Fresco & Wolf, 2018, p. 2). En otras palabras, hay que abandonar la idea de información instruccional de Floridi. Según Fresco y Wolf (2018, p. 5), podemos interpretar los datos en Floridi en relación con lo que se diferencian: (a) diferencias *de re* que son una falta de uniformidad en el mundo real considerado como “datos puros” o “datos proto-epistémicos”; (b) diferencias *de signo* que son una falta de uniformidad entre por lo menos dos señales que por ejemplo son menos o más capacidad de carga de un teléfono o una señal eléctrica variable en una conversación telefónica; (c) diferencias *de dicto*, que son una falta de uniformidad entre dos símbolos, por ejemplo la diferencia en el alfabeto griego entre  $\alpha$  y  $\beta$ . La forma analizada por ambos autores es la (b), *de signo*, pero para ellos no queda claro si la definición de Floridi para las señales solo se aplica en un mundo físico. Si fuera así, Fresco y Wolf afirman que esta definición no sería válida para toda función matemática o máquinas de Turing. Las máquinas de Turing, por ejemplo, caerían en el dominio de diferencias *de dicto* en el que los datos mostrarían una falta de uniformidad entre símbolos (cuadrados vacíos y una marca en el cuadrado). Pero en el caso de las funciones matemáticas no podemos describir como símbolos a los miembros de un dominio o co-dominio siendo un subconjunto de los reales. Por lo tanto, se podría interpretar a estos datos de diferencias *de signo* para tratar sobre entidades abstractas y físicas. Dos nociones de diferencias son ofrecidas por Fresco y Wolf (2018, p. 6), una *actual* en la que dos señales son actualmente distintas y una *potencial* que existe, pero podría no haber existido. Estos conceptos buscan explicar por ejemplo cómo caracterizaríamos un dato (singular) en la computación de una puerta AND de dos entradas y una salida (Fig. 1). En donde el *input A* procesa una señal (dato singular) de un voltaje entre 0.1 V y 0.5 V (bajo), mientras que el *input B* procesa una señal de un voltaje entre 1.1 V y 1.5 V (alto). En caso que obtenga un voltaje entre 0.6 V y 1 V será considerada como indefinida.



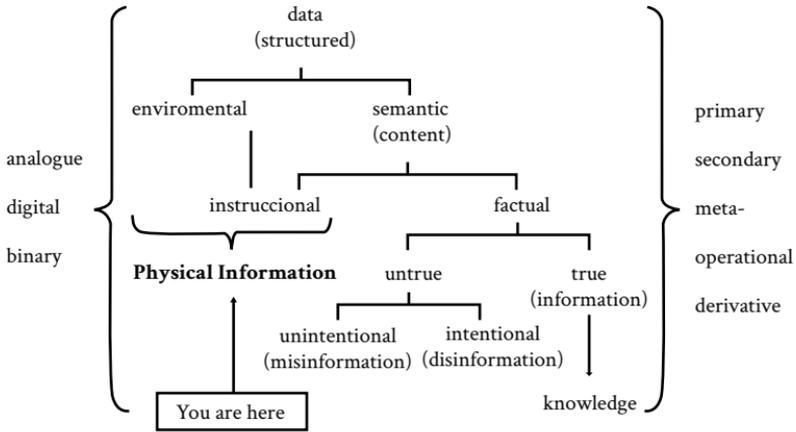
**Fig. 1:** Ejemplo de puerta AND de dos entradas y una salida, donde A es un voltaje bajo entre 0.1 V y 0.5 V, y B es un voltaje alto entre 1.1 V y 1.5 V.

A partir de estas nociones de *actual* y *potencial*, Fresco y Wolf ofrecen cuatro posibles interpretaciones:

- i) El input A y el input B tienen una señal *actual* y el datum es la diferencia entre estas dos señales.
- ii) El input A y el input B tienen cada uno una señal *actual* y cada señal es un datum de la distinción espacial de ambos.
- iii) El input A y el input B son distintos en tanto que uno es *actual* y el otro *potencial*.
- iv) El input A y el input B son ambas señales (alta o baja) que pueden *potencialmente* existir o no. (2018, p. 8)

La interpretación i) es problemática, dado que no permite diferenciar entre el *datum* del *input* A y aquel del *input* B. Por ejemplo, tendríamos tres posibles resultados: uno en el que no hay diferencias, otro en el que el *input* A tiene un voltaje alto y otro en el que el *input* B tiene un voltaje alto. En este sentido nunca podríamos determinar los valores del *output* ya que no podríamos distinguir el voltaje alto o bajo de A y B. ii) es también problemática por dos razones: 1) Si A y B reciben 0,5 V y que la diferencia entre ambas es solo espacial (A y B). ¿Qué diferencia A con 0,5 V de B con 1,1 V? Ninguna. Pero a pesar de esto serán procesados distintamente por la puerta ya que el voltaje implica un cambio en su valor de verdad. Y 2) Una señal ocupando distintas regiones en el espacio no se aplica en una puerta NOT ya que solo tiene una única entrada. iii) es directamente inconsistente ya que no cumple con el diseño que posee la puerta dado que, si uno de los *inputs* posee una señal *actual* y el otro no,<sup>1</sup> entonces no habrá *output* posible. Solo si la señal fluye en A y B tendremos a Q.

<sup>1</sup> Al no ser *actual* es *potencial*, si es *potencial* significa que es posible que tenga un valor X. Las puertas lógicas no permiten posibilidades si una de las entradas es *actual*.



**Fig. 2:** Cuadro de Floridi jerarquizando los tipos de información y ubicando la información instruccional.

Nota: Imagen extraída de *Information: A very short introduction* (fig. 11) de L. Floridi, 2010, Oxford University Press.

En la interpretación iv), la puerta AND procesa datos debido que el input A y el input B son señales (*datum*) que se distinguen por la presencia o ausencia de una señal de alto voltaje o bajo voltaje (diferencia potencial). Un “enabler”<sup>2</sup> asegura la presencia o la ausencia de la corriente por lo que el input A y el input B siempre tendrán un estado. En caso de que el “enabler” tenga una memoria, dicen Fresco y Wolf (p. 9), la puerta booleana hará uso de la memoria para procesar los datos y por lo tanto convirtiendo todos los circuitos booleanos en poseedores de estados.

El punto iv) motiva a Fresco y Wolf a preguntarse si la definición de dato ofrecida por Floridi es realmente no interpretada, además de que la propia lógica booleana describe el funcionamiento físico de las puertas en 0 y 1 que, ya de por sí, son una interpretación. Concordamos con esta idea de que existen interpretaciones en estos dos niveles, complicando la perspectiva de “dato puro”<sup>3</sup> avanzada por Floridi. El ejemplo propuesto por Fresco y Wolf nos muestran la presencia de algún contenido semántico

<sup>2</sup> Un *enabler* es para, Fresco y Wolf (2018, p. 7), una característica que permite hacer un puente entre la separación conceptual o temporal que existen entre una señal potencial y una actual, en donde indica la posibilidad contrafactual de que una señal potencial podría haberse manifestada o de que la señal actual no se habría manifestada.

<sup>3</sup> Por dato “puro” queremos decir “dato no interpretado”.

que posee la señal y que puede interpretarse como dato. La cuestión ahora es cómo establecer el contenido semántico y qué sería “significar”. Para Floridi este proceso se da a partir de la información instruccional.

Información instruccional aparece en Floridi como un subtipo de “contenido semántico” o de “información ambiental” (Fig. 2). Ambos (contenido semántico e información ambiental) se distinguen respecto de si se requiere de un “significado” de información o no. En el caso de información ambiental se busca ofrecer un tipo de información que posee un significado de forma independiente a un productor inteligente. Por ejemplo, las puertas booleanas no involucran semánticas según Floridi (2010, p. 35). Mientras que el contenido semántico depende de que los datos se ajusten al significado (semántico) del sistema, código o lenguaje elegido. De este modo la información es significativa independientemente del receptor. El problema que ven tanto Fresco y Wolf (2018) como Corning (2007) es: ¿cómo podemos caracterizar lo que es significativo o no? En Floridi el proceso no es claro, y si se sale de un entorno de computación simbólica es aún más problemático debido que existen, como se mostró en el ejemplo de los circuitos booleanos, algunas formas de interpretación.

Fresco y Wolf (2018) afirman entonces que hay cuatro mayores problemas con las nociones de Floridi: (a) que la noción de “dato” no se encuentra bien definida para los sistemas físicos; (b) el requisito de forma adecuada<sup>4</sup> [*well-formedness*] se aplica para los sistemas simbólicos, pero no es claro cómo se aplica en sistemas físicos no simbólicos; (c) el requisito de significado implica una teoría de la significación que no es dada por Floridi; y (d) la información instruccional se encuentra definida desde la teoría *General de la Información* propuesta por Floridi.

Una salida interesante es la propuesta por Fresco y Wolf (2018) que se basa en una teleología funcional del receptor/usuario y en donde diferencian de forma jerárquica los conceptos *datum* (dato solo), *signo*, *señal* y *símbolo* siendo el más abstracto. Fresco y Wolf los definen de la siguiente manera:

- a) Datum es un acto sistemático, evento, proceso o estructura (o su ausencia) desde el cuál un receptor puede, pero no todavía, funcionalmente responder dado que es sensible a las variaciones de sus formas.

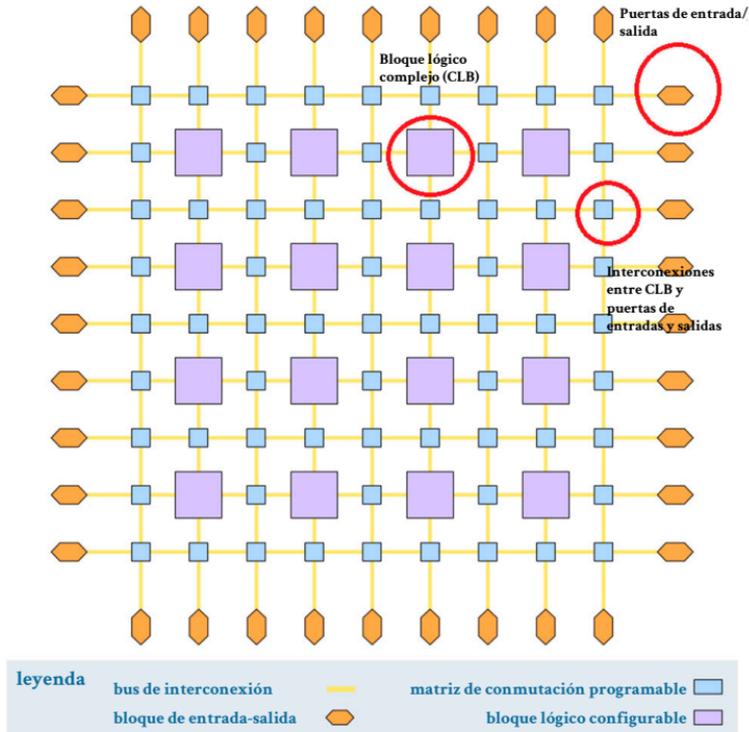
---

<sup>4</sup> Preferimos dejar este punto de lado en este trabajo por falta de extensión. Pero la idea principal de “*well-formedness*” para Floridi es que los datos corresponden a las reglas sintácticas que son manejadas por el sistema, código o lenguaje elegido. Se vincula con el contenido semántico que hemos descrito.

- b) Signo es un datum en el cual el receptor es diseñado en responder sistemáticamente.
- c) Una Señal es un signo que ha sido enviado por un emisor diseñado en emitirlo como un signo para un tipo de receptor particular.
- d) Un símbolo es una señal que es parte del sistema auto-referencial y gobernado por reglas. (2018, p. 16)

Retomando el ejemplo de la puerta AND, un *datum* es una corriente eléctrica que solo puede ser calificada como tal para esta puerta que es este receptor en particular. Si en vez de corriente tenemos agua o rayos ultra violetas, estos, no serán recibidos, el receptor en este sentido debe ser capaz de hacer la diferencia (de forma negativa). En caso del *signo*, la puerta debe responder sistemáticamente a ciertos valores (entre 0.1 V y 0.5 V –bajo–, entre 1.1 V y 1.5 V –alto–, y entre 0.6 V y 1 V –indeterminado–) que son *signos*. La *señal* es el ordenamiento espacio-temporal del output de la puerta AND con otras señales proviniendo de las otras puertas. Este ordenamiento es crucial ya que puede implicar comportamientos o computaciones distintas y es en esta etapa superior que se ejerce la información instruccional (desarrollada a partir de las características citadas aquí). Existe en esta etapa una distinción funcional, dicen Fresco y Wolf, entre los *inputs* operacionales (*outputs* de la puerta) y el *input* de control (información instruccional). Finalmente, los *símbolos*, enfatizan la referencia inter-simbólica dentro de un sistema particular. Los símbolos pueden, en caso de la lógica formal, referirse a otros símbolos sin hablar de cosas que existen “en el mundo” y por lo tanto pueden ser considerados como sistemas “cerrados”. De esta manera, Fresco y Wolf plantean la posibilidad de construir máquinas abstractas de procesamiento de la información simbólica. Por ejemplo, dicen Fresco y Wolf (2018, p. 20), un programa es un conjunto de símbolos y el lenguaje de programación es el sistema auto-referencial que definen los *inputs* y operaciones permitidas en el programa. El código que no ha sido compilado puede, según los autores, manifestarse físicamente como una señal.

La estrategia de Fresco y Wolf es, entonces, usar la descomposición jerárquica de Floridi y el aporte teleológico del usuario/receptor de Corning para el manejo de los datos/signos/señales y símbolos. Para Corning (2007, p. 302) el ambiente contiene información de control latente o potencial pero que este potencial no difiere de sus propiedades físicas. En este sentido, este potencial se actualiza solamente cuando un sistema



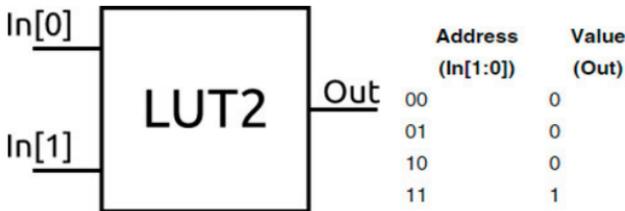
**Fig. 3:** Representación de un chip FPGA de lógica programada (Ventura, 2014). Los círculos (nuestros) representan los niveles donde actúa la programación en un FPGA.

*Nota:* Imagen adaptada de “¿Qué es la lógica programada?” de V. Ventura, <https://polaridad.es/logica-programada-programable-que-es-pld-fpga-hdl-cpld/>

propositivo (cibernético) lo utiliza, por lo que es contexto-dependiente y específico al usuario. La idea de fondo de esta estrategia es interesante para tratar el problema de si hay que diferenciar los datos de una noción semántica o no, pero esta no se encuentra libre de problemas. La lógica programada puede representar un contraejemplo del ejemplo de puerta lógica usado por Fresco y Wolf. En efecto, para algunos aparatos más pequeños se hace uso de un tipo de diseño implementado en chips que reconfigura los circuitos a partir de modificaciones en un programa o a partir de los componentes.

Estos se diferencian de los circuitos clásicos permitiendo cambiar y modificar desde los componentes (*hardware*) o desde un programa (en una memoria de estado sólido) la configuración lógica de los circuitos.

Si tomamos un FPGA (un tipo de chip de lógica programada, ver figura 3) podemos encontrar por lo menos tres problemas. Primero, para Fresco y Wolf el argumento para evitar el *pancomputacionalismo* era que la puerta lógica AND podía solamente procesar un único tipo de *input* por ejemplo el voltaje *bajo* y *alto*. En la concepción de la lógica programada esto no es así, la misma puerta AND puede ser reconfigurada y procesar otros valores que no necesariamente sean voltajes sino cualquier elemento expresado mediante una memoria RAM a nivel simbólico. Segundo, la jerarquización de *datum*, *signo*, *señal* y *símbolo* puede verse modificada por el rol que puede tener la programación del circuito a partir de un programa. En la figura 4 un LUT2 de un FPGA se compone de un *input* de memoria RAM y de un *output*. El *output* posee los valores binarios 0 y 1 que forman una tabla lógica, el LUT2, al emplear dos variables, permite computar todas las tablas posibles dado que se puede programar el *output* deseado. En este sentido, la información instruccional no actúa solo en el ordenamiento de las señales, sino que configura y determina qué señales computar. Por lo tanto, lo que determina la computación es cómo hemos programado el LUT y lo que se considera válido como signo. Además, existen por lo menos dos instancias de ordenamientos posteriores ejercidas por la programación: a) se programa las interconexiones o el ruteo que permite implementar funciones (Fig. 3); y b) los bloques de *inputs* y *outputs* del FPGA son también programables. Por lo tanto, tenemos un ordenamiento estructural de las señales en a) y b). Una vez realizado este ordenamiento, no es necesario el uso de un programa una vez obtenido el *output* del FPGA ya que la señal se encuentra ordenada y cumple la función deseada.

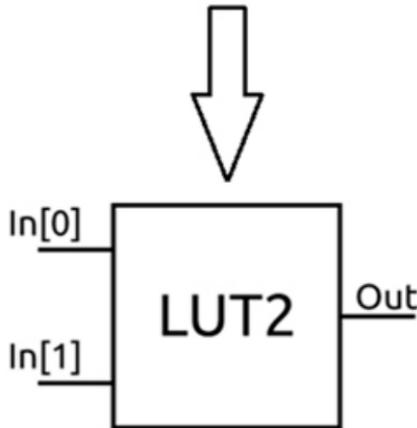


**Fig. 4:** A la izquierda un Look-Up Table (LUT) de dos entradas. A la derecha una tabla de verdades de un LUT2 en donde la columna de izquierda representa los valores ofrecidos por la RAM y a la derecha los valores de verdades binarios 0 y 1.

*Nota:* Imagen tomada de "How does an FPGA work?" de J. Rajewski, 2018, <https://alchitry.com/blogs/tutorials/how-does-an-fpga-work>

Una tercera dificultad, para Fresco y Wolf, es el tratamiento del “significado”, en este caso el significado se encuentra asociado con la computación física y la teleología de la información para la computación. Un *datum* o signo es significativo en tanto que la puerta lo reconoce como signo y permite producir una señal. A esto lo consideran como una semántica interna y no como una semántica externa. Pero ahora que la lógica programada puede determinar lo que consideramos como signo: ¿cuán interna es esta semántica? O mejor expresado: ¿cuán externa es? Si se define el *output* deseado a partir de la programación nos encontramos con dificultades para poder diferenciar ambas semánticas. El *output* del LUT2 es definido a partir de una semántica deseada por el programador para una función determinada. Esta semántica puede ser considerada como “interna” ya que es propia a la computación deseada en el *Bloque Lógico Complejo* (CLB) pero es externa por ser establecida desde el nivel simbólico (Fig. 5). Para Fresco y Wolf es este nivel simbólico que determina la externalidad de la semántica del sistema.

Ahora bien, la segunda crítica es la más problemática y responderla desde Fresco y Wolf es relativamente complejo. Una de las propuestas es que, en realidad, en un circuito no disponemos siempre de *data* y signos, sino que ambos son reemplazados por señales ya que una puerta procesa señales que pueden haber sido producidas por puertas anteriores (2018,



**Fig. 5:** La instrucción establece cuál puerta el LUT2 debe ser.

*Nota:* Imagen adaptada de “How does an FPGA work?” de J. Rajewski, 2018, <https://alchitry.com/blogs/tutorials/how-does-an-fpga-work>

p. 18). Pero si fuera este el caso, ya que poseemos señales en los *inputs*, significa que tenemos información instruccional que ya actuó en este nivel y la noción de dato se entremezcla con la de señal. La terminología es confusa y tenemos entonces dos posibles interpretaciones: (a) una señal como *datum* y signo en donde no actuaría la información instruccional hasta convertir el *output* en señal que luego es ordenado por la información instruccional (que no resuelve nuestra crítica); y (b) en la que la señal es entendida como *datum* y signo pero que esta vez sí actúa la información instruccional. Si (b) es cierta entonces tenemos dos problemas: i) no tiene sentido diferenciar *datum*, signos y señales ya que con señales sería suficiente; y ii) el nivel simbólico no sería independiente de la computación física como Fresco y Wolf desean ya que la información instruccional podría actuar en el nivel de la señal y por lo tanto tener una incidencia física.

La primera crítica parece solucionarse si aceptamos el aspecto propositivo y funcional de la computación, en este sentido solo sería aceptado como datos y signos lo que pueda convertirse en señales a través de la puerta lógica. Pero, como vimos con la crítica anterior, si el signo y el *datum* son considerados como señales donde la información instruccional puede actuar y que solo necesitaríamos señales para computar, entonces perdemos la fortaleza que permitía la definición del signo. La señal podría ser cualquier elemento que pueda convertirse en señales y daría lugar a un pancomputacionalismo. Pero por otro lado podemos si se mantiene la diferencia entre signo y señal disolver la primera crítica ya que distinguimos señal de signo.

## Conclusión

En este trabajo hemos analizado principalmente las propuestas de Floridi y de Fresco y Wolf respecto de la cuestión de la conformación del concepto de dato e implícitamente de información en computación. Hemos descriptos las críticas realizadas por Fresco y Wolf acerca de Floridi, y también hemos planteado el problema conceptual que reviste hablar y usar una noción de dato “puro”. Luego hemos presentado la alternativa ofrecida por Fresco y Wolf basada en la postura funcional y teleológica de Corning. Hemos presentados las críticas realizadas por Fresco y Wolf y su adaptación de la teoría de Corning y Floridi. Finalmente planteamos tres críticas a partir de la lógica programada hacia la postura de Fresco y Wolf en el que tenemos algunas dificultades para poder mantener la

jerarquía propuesta. (a) Respecto de dónde y cómo se ejecuta la información instruccional en el nivel de la señal, (b) Que no consideramos que el argumento en contra del *pancomputacionalismo* puede aplicarse en la lógica programada, (c) Que su postura no puede evitar realmente un tipo de semántica externalista como ellos creen.

## Referencias

- Adriaans, P., & Benthem, J. F. A. K. van (Eds.). (2008). *Philosophy of information*. Amsterdam: Elsevier.
- Ashby, W. R. (1960). *Design for a brain: The origin of adaptive behavior* (2<sup>nd</sup> ed.). Chapman & Hall.
- Corning, P. (2007). Control information theory: The “missing link” in the science of cybernetics. *Systems Research and Behavioral Science*, 24, 297-311.
- Floridi, L. (2010). *Information: A very short introduction*. Oxford: Oxford University Press.
- Floridi, L. (2011). *The philosophy of information*. Oxford: Oxford University Press.
- Fresco, N; Wolf, M. (2014). The instructional information processing account of digital computation. *Synthese*, 191, 1469-1492. doi: 10.1007/s11229-013-0338-5
- Fresco, N; Wolf, M. (2018). Data, signals and information in digital computation. *Logique & Analyse*, 241, 1-23.
- Rajewski, J. (2018). Ilustración del Lut2 (Figura 4). Recuperado de <https://alchitry.com/blogs/tutorials/how-does-an-fpga-work>
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423.
- Simon, H. A., (1947). *Administrative behavior: A study of decision-making processes in administrative organization*. Macmillan Co.
- Ventura, V. (2014). Ilustración del circuito FPGA (Figura 3). Recuperado de <https://polaridad.es/logica-programada-programable-ques-pld-fpga-hdl-cpld/>

Wiener, N. (1948). *Cybernetics: Or control and communication in the animal and the machine*. Cambridge, MA: MIT Press.



## El problema de significación de los símbolos

Mariana Olezza\*

### Introducción

En el presente trabajo se estudiará el problema de significación de los símbolos (PSS), que tiene que ver con cómo los datos adquieren significado. Este problema fue introducido en 1990 por Stephen Harnad a partir de la siguiente pregunta: ¿Cómo la semántica de un sistema de símbolos formal puede ser intrínseca al sistema (en vez de ser parásita a nuestras mentes)?

En primer lugar, se explorará el clásico argumento de la habitación china de Searle, que es una respuesta a las afirmaciones del computacionalismo y la IA fuerte de que a través de la mera manipulación de símbolos una computadora podría llegar a tener significado y una mente. Éste es un experimento mental que enfrenta la analogía entre mente y computadora, y muestra que la mente implica no solo manipulación de símbolos, sino capacidad semántica para darse cuenta del significado de los mismos.

En segundo lugar, se verá el PSS y sus implicancias. Dentro de éstas, nos referiremos a las condiciones Z planteadas por el filósofo Luciano Floridi, que son las condiciones que en teoría un sistema tendría que cumplir para no estar comprometido semánticamente y significar sus símbolos de manera autónoma.

Finalmente, se desarrollarán algunas críticas a la semántica basada en la acción (SbA) de Floridi, que es la solución que él propone para el problema de significación de los símbolos. Una primera crítica viene por el lado del lenguaje privado de Wittgenstein, ya que Floridi plantea un sistema en el que el agente artificial (AA) elabora una semántica propia antes de haberse comunicado con otro AA, y recién en una segunda etapa se comienza a comunicar con otros AA. En la primera etapa el AA tendría una semántica y a la vez estaría aislado, cosa que de acuerdo a la teoría del lenguaje privado no es posible. Otra crítica tiene que ver con el modo de funcionamiento de la asociación entre símbolos de un conjunto de símbolos con los estados del AA (estados relacionados con el significado). Tam-

\* Universidad de Buenos Aires/IIEP-BAIRES.  
marianaolezza@gmail.com

bién se añadirá para complementar la crítica del filósofo Vincent Müller, que plantea que las condiciones Z son muy fuertes, y señala la asociación entre semántica y normatividad.

## La habitación china

El argumento de la habitación china elaborado por Searle (1980), revisado por Preston y Bishop (2002) y Cole (2008) se trata de un escenario hipotético, en el que se ubica a un no hablante de chino dentro de una habitación. Dicha habitación tiene rendijas para enviarle preguntas en chino, y el no hablante de chino busca las respuestas en una tabla y genera la respuesta correcta. Desde afuera pareciera que el “procesador central” entiende chino, pero en realidad no es así. La semántica es externa, “parásita” al observador de afuera. El problema tradicional sugiere que meras manipulaciones sintácticas en una computadora nunca llevarán a generación de significado.

El argumento de Searle es entonces:

- La manipulación sintáctica no es suficiente para que un sistema adquiera significado

- Una computadora solo realiza manipulaciones sintácticas

---

- Una computadora no adquirirá significado.

## El problema de significación de los símbolos

Harnad (1990) introduce el problema de significación de los símbolos (PSS), usando el mencionado argumento de la habitación china. Un sistema formal de símbolos, como el que constituye un agente artificial (AA), un robot, por ejemplo, parece no tener acceso al significado de los datos que puede manipular exitosamente sintácticamente. Es como un ser humano que debe aprender chino como primer idioma consultando un diccionario chino-chino.

Tanto el AA como el ser humano no hablante de chino no podrán tener éxito, ya que un símbolo puede tener sentido, pero sus propiedades

físicas y sintácticas normalmente no proveen de ninguna pista acerca de su valor semántico, siendo el valor semántico relacionado al sintáctico de una forma enteramente arbitraria. Usualmente, los símbolos que constituyen un sistema simbólico no se parecen ni están causalmente conectados a sus correspondientes significados. Se debe explicar entonces qué significa que los datos tengan significado, lo cual es un problema semántico. El PSS es el problema de cómo los datos adquieren significado, uno de los más importantes dentro de la filosofía de la información.

¿Cómo pueden los datos constituyentes de información semántica adquirir significado en primer lugar? El PSS tiene que ver con especificar precisamente cómo un sistema de símbolos formal puede elaborar su propia semántica para los símbolos que manipula y hacerlo desde cero, interactuando con el ambiente y con otros sistemas de símbolos formales. La interpretación de los símbolos debe ser intrínseco al sistema en sí mismo, no puede ser extrínseco, es decir, por un intérprete. (Como en el caso de la habitación china de Searle, del observador que lo ve desde afuera).

Floridi (2014) propone tratar este problema en dos partes: una negativa y una positiva. En la negativa señalará los problemas de las alternativas existentes de resolución a este problema, y en la parte positiva propondrá su propia solución al PSS: la semántica basada en la acción (SbA). Nos concentraremos en las condiciones que se deben cumplir para tratar el PSS por un lado, y por otro en la parte positiva de la SbA.

## **Las condiciones Z**

El tratamiento adecuado del PSS requiere, según Floridi (Floridi, 2014, p. 137) cumplir las así llamadas condiciones “Z”, a saber: que no haya innatismo ni externalismo. Con respecto al innatismo, no se deben suponer recursos semánticos preinstalados en el AA. La segunda restricción que Floridi propone imponer al agente artificial es más fuerte y consiste en no suponer externalismo. Esto quiere decir que no se le deberían proveer recursos semánticos desde afuera, desde una entidad ya competente semánticamente (a través de entrenamiento, por ejemplo). Si una población ya es semánticamente competente, la comunicación y el entrenamiento pueden jugar un rol esencial en el desarrollo de capacidades semánticas<sup>1</sup>

---

<sup>1</sup> Wittgenstein en éstas obras sostiene que el significado de las palabras y el sentido de las proposiciones está en su función, su uso en el lenguaje. Preguntar por el significado de una

(Wittgenstein 1953, 1958) de nuevas generaciones de agentes, pero dicha transmisión semántica *no brinda una explicación de cómo surge ésta en primer lugar*. Si lo que se busca explicar es cómo surgió la semántica, debe adoptarse un enfoque más restrictivo, ya que no tendría que aceptarse el externalismo. Este enfoque elimina la posibilidad de utilizar técnicas computacionales tales como *machine learning* (aprendizaje máquina) para los AA, ya que en éstas se requiere que el programador entrene al sistema (le provea una semántica).

## La semántica basada en la acción (SbA)

### El lenguaje privado

Las condiciones Z que se deben cumplir, según Floridi, imponen un primer momento en el cual el sistema no tiene ninguna semántica (ni vista desde afuera), y la desarrolla por sí mismo. En un *segundo lugar* se comunicará con otros AA. La pregunta que podemos hacer a esto es ¿es esto posible?, ¿puede desarrollar una semántica para sí mismo sin comunicación con otros AA? Nos hace pensar en el lenguaje privado de Wittgenstein, y su imposibilidad de acuerdo a él.

De acuerdo a Floridi, los juegos del lenguaje de Wittgenstein no cumplen con las condiciones Z y no dan cuenta de cómo surge el significado en primer lugar. Él desarrollará una semántica alternativa llamada Semántica Basada en la Acción (SbA), con grandes diferencias con respecto a la semántica de Wittgenstein.

La SbA está basada en unas máquinas que poseen arquitectura de metaprogramación<sup>2</sup> y los significados son los estados internos del AA (por ejemplo “giro a la izquierda 30°”). Los estados internos del AA son excelentes candidatos para el rol de no ser semánticos, pero ser inductores de recursos semánticos. En la SbA el propósito de la acción no tiene influen-

---

palabra o por el sentido de una proposición equivale a preguntar cómo se usa. Además, puesto que dichos usos son muchos y multiformes, el criterio para determinar el uso correcto de una palabra o de una proposición estará determinado por el contexto al cual pertenezca, que siempre será un reflejo de la forma de vida de los hablantes. Dicho contexto recibe el nombre de juego de lenguaje.

<sup>2</sup> Un sistema capaz de metaprogramación opera en dos niveles, que interactúan entre sí. Organiza acciones al nivel Objeto, en donde interactúa con el ambiente. También toma acciones en sus estados internos y sus propias elaboraciones. En este caso opera en el Meta nivel que toma como datos las acciones del nivel Objeto.

cia directa en la generación del significado. Ningún tipo de teleosemántica es presupuesto. No hay criterios semánticos externos manejando el proceso de generación de significado, por lo tanto satisface la condición Z.

Como contraparte, los juegos del lenguaje de Wittgenstein son teleológicos. De acuerdo a la teoría de Wittgenstein, el significado surge de las interacciones sociales y por lo tanto esto representa *criterios externos* y representa habilidades semánticas preestablecidas en los AA involucrados. La asociación entre significados y símbolos es contextual, se basa en negociación entre los hablantes y requiere entrenamiento. Es una teoría basada en convención y no cumple con las condiciones Z.

Podemos remarcar que de acuerdo al SbA, la semántica tiene su origen en los comportamientos individuales de los agentes, *no en la comunidad*.

## Críticas

Una duda que surge al estudiar la SbA de Floridi es ¿cómo se asocian los estados (significados) a los símbolos? Porque en este caso, partiendo de un AA aislado, no se tiene una asociación contextual y basada en entrenamiento. Lo que se desea obtener es que el AA haga surgir la semántica por sí solo, sin ningún entrenamiento externo ni nada por el estilo.

La asociación entre estado y símbolo en la SbA ocurre según Floridi de una manera “arbitraria pero no aleatoria” (Floridi, 2014, p.169), arbitraria para no comprometer semánticamente el sistema (violando las condiciones Z): que no se siga ninguna lógica en la asociación, y por otro lado “no aleatoria”, ya que “máquinas similares generarán resultados similares”.

Sin embargo, esto último no es correcto, ya que si bien es posible la arbitrariedad, por más que sean máquinas similares, al introducir una función aleatoria para la elección de estado-símbolo, resultan todas con diferentes resultados. Es lo mismo que un programa de computadora tenga la misma estructura pero en cierto punto tenga una función aleatoria que arroja un resultado; en ese caso los resultados serían todos aleatorios, diferentes unos de otros; más allá que los distintos equipos sean “similares”.

Esto acarreará problemas más graves en una segunda etapa, ya que Floridi se basa en que los agentes tengan los mismos pares estado-símbolo para lograr una comunicación exitosa entre ellos. Pero, si cada AA asocia a un estado diferentes símbolos, ¿cómo podrían comunicarse de manera efectiva?

Una vez que el nuevo estado es obtenido, M2 [la máquina, el AA] asocia el estado transducido con un símbolo removido de un conjunto de símbolos. El proceso de remover un símbolo y acoplarlo con un estado es discreto, no recursivo y *arbitrario pero no es aleatorio*, de la siguiente manera. M2 hace explícito solo un símbolo para cada entrada que recibe; y no puede remover el mismo símbolo más de una vez. La elección del símbolo es arbitraria, ya que no está semánticamente relacionada a los estados transducidos, pero no es aleatoria, ya que similares tipos de agentes asociarán similares símbolos con similares estados transducidos. (Floridi, 2014, p. 169)

## Otras críticas

Las condiciones Z le parecen muy fuertes a Vincent Müller (2015). En la SbA de Floridi no hay teleosemántica presupuesta (el propósito de la acción no tiene influencia directa en la generación del significado). Luego, en la SbA no hay criterios de semántica extrínseca guiando el proceso de generación de significado. Sin embargo, por otro lado, se dice que “a diferencia del comportamiento heliotrópico de los girasoles”, existe significado. Müller no ve como puede ser que no haya semántica extrínseca guiando el proceso de generación de significado, y que por otro lado exista significado. No se obtienen los objetivos que se intentan conseguir, ni innatamente ni dados externamente, por las restricciones fuertes de las condiciones Z.

Müller apunta a que ya Floridi había criticado otros sistemas como “Kismet”, que tienen algo construido que les permite resaltar ciertos datos con respecto a otros, marcando cierto “éxito”. Otros sistemas, como el de Cangelosi y Riga (2006), usan un usuario competente de lenguaje para dar la marca de éxito. El problema que ve Müller es que cuando una solución es mejor que otra, Floridi dice que la condición Z ha sido violada.

Sin metas, nada es “mejor” que nada, no hay “éxito”. O el sistema realmente es un agente, lo cual implica que tenga metas, dadas innatamente o externamente (con lo cual no se resolvería el PSS), o es simplemente un sistema que interactúa con el ambiente, sin objetivos. Pareciera que tener una semántica implica normatividad, ya que sin ella no tenemos el uso “correcto” o “incorrecto” de los símbolos, de “éxito” o “fracaso”.<sup>3</sup>

---

<sup>3</sup> La reversa no aplica: se puede tener normatividad sin tener semántica (por ejemplo, en un sistema sintáctico, fórmula bien formada, reglas de derivación, etc.).

## Conclusión

En el presente trabajo se ha introducido el PSS a través de la presentación del argumento de la habitación china. Se puede decir que un sistema como un AA no tiene acceso al contenido semántico de los símbolos que simplemente manipula, ya que la relación entre sintaxis y semántica es arbitraria. Se estudiaron las condiciones Z, expuestas por Floridi, necesarias para que un sistema signifique sus símbolos de manera autónoma, y luego se exploró la SbA que él propone.

El estudio de dichas condiciones Z y SbA nos llevan a realizar una serie de críticas. Con respecto a las condiciones Z, como bien señala Müller pueden ser condiciones demasiado “fuertes” de manera tal que cada vez que el sistema favorezca una serie de datos frente a otros, Floridi diga que se violaron las condiciones Z y que el sistema no cumple con el PSS. Con respecto a la SbA, no encaja con la teoría del lenguaje privado de Wittgenstein, ya que requiere que un AA elabore su propia semántica de manera autónoma, significando sus símbolos, y recién en una segunda etapa se comunique con otros AA. Otro problema que tiene la SbA es la forma de funcionamiento, relacionada con la arbitrariedad con la que los AA eligen sus símbolos para no comprometer semánticamente al sistema. Si esto es así, a la hora de comunicarse en la “segunda etapa” los AA tendrán problemas para ponerse de acuerdo en los símbolos a utilizar. Pero por otro lado, si los pares estado-símbolo de los AA fueron elegidos con algún criterio, terminan estando comprometidos semánticamente. De este modo se puede decir que la SbA parece no resolver efectivamente el PSS.

## Referencias

- Cangelosi, A., & Riga T. (2006). An embodied model for sensorimotor grounding and grounding transfer: Experiments with epigenetic robots. *Cognitive Science*, 30(4), 673–689.
- Cole, D. (2008). The chinese room argument. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2008 edition). <http://plato.stanford.edu/archives/fall2008/entries/chinese-room/>
- Floridi, L. (2014). *The philosophy of information*. Oxford: Oxford University Press.

- Floridi, L. (2010). *Information: A very short introduction*. Oxford: Oxford University Press.
- Harnad, S. (2002). Symbol grounding and the origin of language. En M. Scheutz (Ed.), *Computationalism: New directions* (pp. 143–158). Cambridge, MA.: MIT Press.
- Harnad, S. (2000). Minds, machines and Turing: The indistinguishability of indistinguishables. *Journal of Logic, Language, and Information*, 9(4), 425–445.
- Harnad, S. (1990). The symbol grounding problem. *Physica Scripta, D*(42), 335–346.
- Harnad, S. (1987). *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.
- Müller, V. (2015). Which symbol grounding problem are we trying to solve? *Journal of Experimental and Theoretical Artificial Intelligence*, 27, 73–78.
- Preston, J., & Bishop, M. (Eds.). (2002). *Views into the Chinese Room: New essays on Searle and Artificial Intelligence*. Oxford: Clarendon Press.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–457.
- Searle, J. R. (1990). Is the brain a digital computer? *Proceedings and Addresses of the American Philosophical Association*, 64, 21–37.
- Searle, J. R. (1992). *The rediscovery of the mind*. Cambridge, MA; London: MIT Press.
- Wittgenstein, L. (1953). *Philosophische Untersuchungen*. Oxford: Blackwell.
- Wittgenstein, L. (1958). *The blue and brown book*. Oxford: Blackwell.



## Acerca de la definición de “número” Un abordaje wittgensteiniano

Pablo Ruiz Lezcano\*

### Introducción

En el ejercicio de la matemática hay algunos “objetos” en los que tienen poca relevancia el carácter ontológico y, en algunos casos, el epistemológico. Los números, el punto, la recta y el par ordenado son algunos ejemplos. Lo que importa, en general, es la relación entre ellos y su estructura. Los problemas que ha planteado la matemática a la filosofía son variados y complejos, uno de ellos es analizar la naturaleza de los conceptos mencionados. A lo largo de la historia de la matemática (y de la filosofía) algunos de ellos han sido elucidados gracias al avance de nuevas herramientas conceptuales. Por ejemplo, Euclides (trad. 2015) define a la recta como “aquella que yace por igual respecto de los puntos que están en ella” (p. 53), mientras que actualmente se la define de diferentes formas, una de ellas es:

$$(x, y) \in \mathbb{R} \text{ tal que } ax + by = c, \text{ donde } a, b \text{ y } c \in \mathbb{R}$$

El acto de definir se puede describir como el ejercicio de enunciar las condiciones que un objeto ( $x$ ) debe cumplir para ser aceptado o no como miembro de un conjunto. Las definiciones son respuestas a la pregunta *¿qué es  $x$ ?*, *¿cuál es la naturaleza de  $x$ ?* o *¿cuál es la esencia de  $x$ ?* Las ciencias, así como la filosofía, intentan dar respuesta a preguntas de este tipo. La matemática no es ajena a dicha cuestión, pues gracias a ella surge una de las características relevantes de la disciplina: el acuerdo entre miembros. Sin embargo, esta situación no ha tenido lugar en relación a el concepto de número. Pues bien, ¿puede haber una definición que abarque a los distintos conjuntos numéricos, cuya introducción en un curso de matemática es heterogénea: axiomática para los naturales, algebraica para los racionales e imaginarios, topológica para los reales? ¿Qué concepto puede considerar

\* Universidad Nacional de General Sarmiento (UNGS). Los Polvorines, Malvinas Argentinas, Buenos Aires, Argentina.  
pf.ruiz.lezcano@gmail.com

la discretitud de los enteros, la densidad de los racionales y la completitud de los reales? En otras palabras, ¿por qué a  $-2$ ;  $\frac{3}{5}$ ;  $\log 7$ ;  $\sqrt{3}$  o  $-6 + 3i$  los llamamos números?

El presente artículo tiene un doble objetivo. En primer lugar, analizar algunas definiciones o elucidaciones de número que fueron dadas a lo largo de la historia de la matemática; y, en segundo lugar, señalar que la herramienta conceptual *semejanza de familia* que ofrece el filósofo Wittgenstein es adecuada para pensar la naturaleza de dicho concepto.

### **Análisis de algunas definiciones de número**

Una posible forma de abordar la pregunta *qué es x* es indicando un conjunto de condiciones necesarias y suficientes para la aplicación del concepto definido [*definiendum*]. Esta tradición, llamada *esencialismo*, se origina en Sócrates y es el modelo que guió la búsqueda de definiciones en toda la historia de la filosofía hasta mediado del siglo XX (Arroyo, 2019). Estos intentos definicionales están presentes en la actividad matemática y se ilustra fácilmente en conceptos como el de número primo o circunferencia. Por ejemplo, el número primo es aquel entero que es divisible por sí mismo y por la unidad, solamente. En este caso, “número primo” es el *definiendum*<sup>1</sup> y las condiciones necesarias y suficientes son “ser entero”, “divisible por sí mismo” y “divisible por la unidad”, que forman parte del *definiens*.<sup>2</sup> La misma se puede representar como:

$$(x)(Px \leftrightarrow (Ex \wedge Dx \wedge Ux))$$

donde:

$Px$  = *x es primo*;

$Ex$  = *x es entero*

$Dx$  = *es divisible por sí mismo y*;

$Ux$  = *x es divisible por la unidad*.

El argumento que sostiene este enfoque se puede visualizar en varios diálogos platónicos. Sócrates intenta buscar lo común en todos los objetos que llevan el mismo nombre. Por lo tanto, el hecho de que designemos a un elemento con un único nombre es la evidencia de que debe haber un rasgo (o conjuntos de rasgos) compartidos por todas las cosas que caen

<sup>1</sup> Objeto que se define.

<sup>2</sup> Parte de la definición en la que se anuncia las condiciones.

bajo el mismo concepto (Arroyo, 2019). En consecuencia, para esta tradición, la definición es señalar aquellos rasgos comunes que comparten los miembros de una clase.

Distintas corrientes de la filosofía de la matemática o algunos filósofos/ matemáticos de renombre han desarrollado, dentro del paradigma señalado, una ardua tarea en la búsqueda de una definición o elucidación de número. A continuación, presentaremos algunas de ellas.

Euclides, en su famoso libro llamado *Los elementos*, afirma que “un número es una pluralidad compuesta de unidades” (trad. 2015, p. 284). Dicha definición es una de las muchas de la antigüedad. Por ejemplo, Nicómaco combina varias definiciones entonces en boga, al afirmar que es “una pluralidad definida” (igual que Eudoxo) o un “conjunto de unidades”. Teón, por su parte, dijo que un número es “una colección de unidades”; mientras que Jámblico que la colección de unidades es aplicada a la unidad, es decir, al número.

Ya en la modernidad, Kant (1781/2014) sostiene que los números están ligados a nuestra intuición del tiempo, siendo este, igual que el espacio, una de las formas de sensibilidad pura (“estructuras” a priori). El número es, por lo tanto, una pluralidad conocida de manera determinada mediante el acto de contar, añadiendo sucesivamente una unidad a una unidad; es el resultado de desplegar en el tiempo el concepto puro de cantidad.

A partir de la mitad del siglo XIX, cuando comienza la carrera de la axiomatización de la aritmética con Dedekind y Peano, se busca un sistema de postulados capaz de deducir la aritmética de los números a la aritmética de los números naturales. En la obra *Los principios de la aritmética expuestos según un nuevo método*, Peano propone un sistema empleando una lógica inspirada en Russell y como primitivos al “0” (cero),<sup>3</sup> “número natural”<sup>4</sup> y “sucesor”.<sup>5</sup> Es a partir de esta base que Frege comienza la tarea de la fundamentación de la Matemática, ya que para él los axiomas que había postulado Peano no eran realmente primeros principios, en otras palabras, no eran evidentes y, por lo tanto, necesitaban ser probados. Por otro lado, Frege rechaza la intuición como fundamento del conocimiento. Es así que el matemático y lógico alemán funda el proyecto logicista, con el propósito de reducir la aritmética a la lógica y erradicar a la intuición como fun-

<sup>3</sup> La categoría de *cero* es la de *constante individual*.

<sup>4</sup> Se aplica a una determinada clase de individuos

<sup>5</sup> Es un término para una operación que aplicada a un individuo origina otro al cual llamamos sucesor de él. Por ejemplo, “1” se define como el “sucesor de 0”. De esta forma define a todos los números naturales a partir del “1”.

damento del conocimiento. En *Los fundamentos de la aritmética*, luego de criticar las definiciones de Mill, Newton, Hankel, Scröder, Schloemilch, Hilbert y, en particular, Kant, propone una definición lógica.

Según Frege (1884/1973), Mill considera que el número es una propiedad como el color o la dureza; en cambio, para Scröder es una copia de la realidad: “es sacado de ella, por medio de la representación de las unidades mediante unos” (Frege, 1884/1972, p. 47). Por otro lado, Newton propone entender al número en un sentido amplio<sup>6</sup>: como una relación abstracta de cada una de las magnitudes con otra de la misma clase, que se toma como unidad. Asimismo, el psicologismo (Schloemilch) afirma que los números son productos de la abstracción, esto es, representaciones mentales de objetos concretos.<sup>7</sup> Mientras que el formalismo (Hilbert) considera que son signos sin contenido expreso y que no precisa definición alguna. No contento con ambas definiciones y luego de haber argumentado en contra de las mismas, Frege sostiene que los números son objetos lógicos que caen bajo determinados conceptos. Por ejemplo: una clase  $C$  tiene el número 2 si existen objetos,  $x$  e  $y$  que están en  $C$ , tal que  $x \neq y$ ; y si  $z$  está en  $C$  entonces o bien  $z = x$  o bien  $z = y$ .

Las definiciones mencionadas son insatisfactorias, pues: o bien son de naturaleza circular, o bien sus condiciones no son suficientes o necesarias. La primera característica se evidencia en las propuestas de Euclides, Nicómaco, Eudoxo, Teón y Jámblico puesto que apelan a la idea de unidad, igual que Scröder; mientras que Kant, a la noción de cantidad, y Newton, a la de magnitud. En cambio, las definiciones de Mill, Schloemilch y Hilbert tienen un *definiens* cuyas características no son necesarias o suficientes. Mientras que Frege, fiel a su propósito logicista, solo define a los números naturales. Consideramos que si el resto de los conjuntos numéricos se definen a partir de la noción de número natural, entonces el concepto número pierde sentido o no debería existir para el resto de los conjuntos, de la misma manera que la palabra mamífero es utilizada (entre otras) para definir perro. De esta forma, a cada conjunto se lo deberían llamar como “los enteros”, “los racionales”, “los reales”, etc. Sin embargo, esto no ocurre, dado que a tales conjuntos se los llama como “los números enteros”, “los números racionales”, “los números reales”, etc.

En consecuencia, no se ha ofrecido hasta ahora una definición aceptable de “número”, la razón tal vez sea que esa definición no es posible.

<sup>6</sup> Incluye a los racionales e irracionales.

<sup>7</sup> Como si fuera una imagen.

El hecho de que aún no haya sido encontrada la definición no implica, claro está, su imposibilidad; pero cuando se han realizados varios intentos para dar con este problema y todos han sido insatisfactorios, tal vez pueda tomarse como un indicio de que la pregunta por la definición de número está mal planteada.

### **Semejanza de familia: abordaje wittgensteiniano del concepto de número**

Entre los párrafos 65 a 68 de *Investigaciones filosóficas* (IF en adelante), Wittgenstein (1954/2008) cuestiona el presupuesto que se esconde en las definiciones esencialistas. En conformidad a su postura sobre la Filosofía y sobre la naturaleza de los problemas filosóficos, el filósofo critica el impulso a la búsqueda de un elemento (o conjuntos de elementos) en común; o sea, a la noción de esencia:

en vez de indicar algo que sea común a todo lo que llamamos lenguaje, digo que no hay nada en absoluto común a estos fenómenos por lo cual empleamos la misma palabra para todos—sino que están *emparentados* entre sí de muchas maneras diferentes. Y a causa de este parentesco, o de estos parentescos, los llamamos a todos “lenguaje”. (Wittgenstein, 1954/2008, p. 87 [§65])

A diferencia de lo que afirmaba en el *Tractatus logico-philosophicus*, Wittgenstein sostiene aquí que no hay un rasgo común (o conjunto de rasgos) a todos los fenómenos que llamamos lenguaje. El concepto “lenguaje” puede ser construido de forma análoga a lo que Wittgenstein hace con el concepto de “juego”:

Considera, por ejemplo, los procesos que llamamos “juegos”. Me refiero a juegos de tablero, juegos de cartas, juegos de pelota, juegos de lucha, etc. ¿Qué hay común a todos ellos? —No digas: “Tiene que haber algo común a ellos o no los llamaríamos ‘juegos’”—sino *mira* si hay algo común a todos ellos. —Pues si los miras no verás por cierto algo común a *todos*, sino que verás semejanzas, parentescos y por cierto toda una serie de ellos. Como se ha dicho: ¡no pienses, sino mira! [...] Podemos ver como los parecidos surgen y desaparecen. [...] Y el resultado de este examen reza así: vemos una complicada red de parecido que se superponen y entrecruzan. (Wittgenstein, 1954/2008, p. 87 [§66])

De esta forma, Wittgenstein piensa que no hay una característica co-

mún a todo lo que llamamos “lenguaje” o “juego”. Pero esto no implica afirmar que los distintos significados de “juego” son independientes unos de otros, ya que los objetos que caen bajo la clase que denominamos “juegos” es por una relación de *semejanzas*: un conjunto de rasgos que no están presentes en todos los miembros, sino que se superponen y entrecruzan de la misma manera que se superponen y entrecruzan los rasgos físicos y psicológicos entre los miembros de una familia. A estas relaciones, Wittgenstein (1954/2008) las denomina *semejanzas (o parecidos) de familia*. Wittgenstein sostiene que la razón por la cual a dichos objetos los llamamos “juego” o “lenguaje” es la existencia de un “buen número” de características asociadas entre los miembros de esa clase (Arroyo, 2019). Hay sin duda algo en los fenómenos que llamamos “juego” o “lenguaje” que permite su reconocimiento, este es un “rango de aplicación que se va ampliando en función no de un elemento que esté presente en todos los casos, sino porque se van dando relaciones de semejanza entre los objetos de modo tal que nos permiten seguir aplicando el mismo término” (Bassols, 2014, p. 19), y son los hablantes, en base a consideraciones de orden práctico, quienes fijan hasta donde se da dicha semejanza, fijando una flexibilidad sin que sea arbitraria. Podría decirse que Wittgenstein inicia una nueva tradición en las teorías de las definiciones: el *antiesencialismo*.

El concepto de *semejanza de familia* es pertinente para comprender la noción de número. Es más, Wittgenstein (1954/2008) sostiene que estos conforman una familia: están unidos entre sí por relaciones de semejanzas, en donde hay propiedades que se entrecruzan, se superponen y, en algunos casos, no están presentes. ¿Cuáles podrían ser esas propiedades? Un libro clásico de matemática, *Calculus* de Spivak (1967/2014), nos puede ayudar sobre esta cuestión. Define a los números reales como aquellos que cumplen una serie de propiedades básicas, estas son:

- 1) Propiedad asociativa de la suma.
- 2) Elemento neutro de la suma.
- 3) Existencia del opuesto (o inverso aditivo).
- 4) Propiedad conmutativa de la suma.
- 5) Propiedad asociativa de la suma.
- 6) Elemento neutro de la multiplicación.
- 7) Existencia del inverso multiplicativo (excepto el cero).

- 8) Propiedad conmutativa de la multiplicación.
- 9) Propiedad distributiva.
- 10) Ley de tricotomía.
- 11) Si  $a$  y  $b$  pertenecen a  $P$ ,<sup>8</sup> entonces  $a + b$  pertenecen a  $P$ .
- 12) Si  $a$  y  $b$  pertenecen a  $P$ , entonces  $a.b$  pertenecen a  $P$ .

Si consideramos los otros conjuntos numéricos “clásicos” (Naturales, enteros, racionales y complejos) observamos que algunas de estas propiedades no están presentes en los mismos, tal como se muestra en el siguiente cuadro:

<i>Propiedades</i>	Naturales	Enteros	Racionales	Reales	Complejos
Asociatividad de la suma.	x	x	x	x	x
Elemento neutro de la suma.		x	x	x	x
Existencia del opuesto (o inverso aditivo).		x	x	x	x
Conmutativa de la suma.	x	x	x	x	x
Asociatividad de la multiplicación.	x	x	x	x	x
Elemento neutro de la multiplicación.	x	x	x	x	x
Existencia del inverso multiplicativo (excepto el cero).			x	x	x
Conmutatividad de la multiplicación.	x	x	x	x	x
Propiedad distributiva.	x	x	x	x	x
Ley de tricotomía.	x	x	x	x	
Si $a$ y $b$ pertenecen a $P$ , entonces $a + b$ pertenecen a $P$ .	x	x	x	x	
Si $a$ y $b$ pertenecen a $P$ , entonces $a.b$ pertenecen a $P$ .	x	x	x	x	

---

<sup>8</sup> Conjunto de los números positivos.

Notemos que las propiedades mencionadas no son condiciones necesarias y suficientes para caracterizar a los naturales, a los enteros y mucho menos a los complejos; es más, el cuadro se puede complejizar si tomamos otros “rasgos”, como la discretitud, la densidad o la completitud. Los naturales, por ejemplo, no tienen un inverso aditivo; los enteros, en cambio, sí, pero no cuentan con un inverso multiplicativo. Los racionales, por su parte, cumplen todas las propiedades y encima es un conjunto denso en  $\mathbb{R}$  cosa que no sucede con los enteros. Los reales cumplen las propiedades mencionadas y es completo: “rasgo” que no tiene los racionales. A su vez, los complejos son un conjunto completo, pero no satisfacen las propiedades de orden. Entonces, ¿basta que cumplan las propiedades de asociatividad y/o conmutatividad de la suma y/o multiplicación? La respuesta es “no”: existen objetos matemáticos que satisfacen algunas de las mencionadas y no son números; ejemplo: sean  $A$  y  $B$  dos matrices de igual tamaño, entonces  $A + B = B + A$ . Asimismo, existen números que no la cumplen: los cuaterniones, por ejemplo.

Gracias a la representación vectorial de los números complejos, el matemático Hamilton decidió crear un “estructura análoga” para representar vectores en el espacio, dando origen a los números cuaterniones en 1853. Un ejemplo es

$$3 + 2i + 7j - 6k,$$

donde  $i$ ,  $j$  y  $k$  tienen un papel semejante al que juega  $i$  en los números complejos, asumiendo que se debe cumplir:

$$i^2 = j^2 = k^2 = i.j.k = -1;$$

$$i.j = -j.i = k;$$

$$j.k = -k.j = i;$$

$$k.i = -i.k = j$$

Hasta el momento, los números señalados (naturales, enteros, racionales, reales y complejos) cumplen la propiedad conmutativa de la multiplicación. Sin embargo, los cuaterniones no la obedecen. Por ejemplo, sean los números

$$3 + 2i + 7j - 6k$$

y

$$4 + 6i + 8j - 9k$$

Por lo tanto, si realizamos la distributiva y la asociatividad y conmutatividad de la suma, queda:

$$(3 + 2i + 7j - 6k).(4 + 6i + 8j + 9k) = -2 + 137i - 2j - 23k$$

y

$$(4 + 6i + 8j + 9k).(3 + 2i + 7j - 6k) = -2 - 85i + 106j + 29k$$

Así se concluye:

$$(4 + 6i + 8j + 9k).(3 + 2i + 7j - 6k) \neq (3 + 2i + 7j - 6k).(4 + 6i + 8j + 9k)$$

Hamilton debió ir en contra de la tendencia establecida por los algebristas, de la misma manera que los fundadores de las geometrías no euclídeas, Bolyai y Lobachevsky, al no considerar el quinto postulado de la geometría euclidiana. Al renunciar la propiedad conmutativa de la multiplicación, Hamilton da lugar a las nuevas álgebras, aquellas donde esta propiedad no es necesaria. Una década antes, John Thomas Graves llevó a cabo un estudio (aunque no lo publicó) sobre los octoniones (un álgebra 8-dimensional), donde el producto no satisfacía ni la propiedad conmutativa ni la propiedad asociativa (Sánchez Muñoz, 2011).

Los últimos párrafos son ejemplos de un aspecto que Wittgenstein resalta en *IF*:

Y extendemos nuestro concepto de número como cuando al hilar trenzamos una madeja hilo a hilo. Y la robustez de la madeja no reside en que una fibra recorra toda su longitud, sino en que se superpongan muchas fibras. (Wittgenstein, 1954/2008, p. 89 [§67])

En esta cita el autor, fiel a su estilo metafórico, compara la cohesión entre los miembros de una clase con la manera en que se consigue la robustez de una madeja de hilo. La razón por la cual un objeto es incluido a una clase puede ser porque se *parece* a los objetos previamente incluidos. Asimismo, los hilos que se superponen corresponden a las propiedades que conectan entre sí a los individuos haciéndolos miembros de una clase. La robustez de la madeja corresponde a la racionalidad que guía la inclusión o exclusión de un individuo en una clase. De la misma manera que una madeja no pierde su robustez al perder un hilo, nuestra decisión no

deja de ser racional por el hecho de que no exista un conjunto preestablecido de propiedades. Al incluir un miembro en una clase, ese trae consigo nuevas propiedades o, incluso, desaparecen (Arroyo, 2019). El concepto de número es, en consecuencia, un concepto abierto y no puede ser considerado como la suma lógica de los casos particulares:

“Perfecto; así pues, el concepto de número se explica para ti como la suma lógica de estos conceptos individuales emparentados entre sí: número cardinal, número racional, número real, etc., y del mismo modo el concepto de juego sería la suma lógica de los correspondientes conceptos parciales” —No tiene por qué ser así. Pues *puedo* darles límites rígidos al concepto de “número” así, esto es, usando la palabra *número* como designación de un concepto rígidamente delimitado, pero también puedo usarla de modo que la extensión del concepto *no* esté cerrada por un límite. Y así es como empleamos de hecho la palabra “juego”. ¿Pues de qué modo está cerrado el concepto de juego? ¿Qué es aún un juego y qué no lo es ya? ¿Puedes indicar el límite? No. Puede *trazar* uno: pues no hay aún ninguno trazado. (Pero eso nunca te ha incomodado cuando has aplicado la palabra “juego”). (Wittgenstein, 1954/2008, p. 89 [§68])

## Conclusión

En primer lugar, consideramos que el concepto de número, igual que el de “juego”, “lenguaje” e incluso el de “arte” (cf. Weitz, 1956), no puede ser definido en términos esencialistas, pero conforman una familia. En segundo lugar, si bien hay una imposibilidad de definirlos bajo una tradición socrática, esto no implica que no se pueda reconocerlos e incluso enseñarlos, la relación de semejanza que hay entre ellos es el motivo por el cual se puede dar el reconocimiento; así, llamamos  $\aleph_0$  o  $4 + 6i + 6j + 9k$  “número” porque entra en el juego del lenguaje de los matemáticos (Bassols, 2014). A su vez, el concepto de número no puede ser considerado cerrado ni como la suma lógica de todos los casos particulares y, por otro lado, la relación de semejanza entre ellos se torna más evidente si consideramos a otros conjuntos numéricos, tales como los  $p$ -ádicos o los hipercomplejos u otras estructuras. Por último, las definiciones de número que han sido dadas en las primeras páginas son, de una forma u otra, de carácter ontológico, mientras que el análisis que empleamos por medio de la herramienta *semejanza de familia* es de tipo algebraico; esta diferencia se dio porque no debemos pensar la noción número ignorando a la comunidad que trabaja con ellos. Los pasajes señalados de *IF* nos transmite la imposibilidad de

concebir al “lenguaje” y al “juego” como actividades aisladas y separada del resto de actividades no-lingüísticas (Rivera, 2006). También nos dice que las afirmaciones filosóficas que se construyen alrededor de un concepto son incomprendimientos del lenguaje e ignoran su carácter contextual: en la búsqueda de esencias inevitablemente se olvida la diversidad y los casos particulares. Por tal motivo es que consideramos los libros de matemática para comprender los rasgos de los números. Asimismo, si un concepto no puede ser definido en términos tradicionales no es algo problemático para Wittgenstein, el significado de un término, según él, no depende de un conjunto de propiedades sino del “uso” que le damos; es decir, la función que juega la palabra en un “juego de lenguaje” (Wittgenstein, 1954/2008, p. 309 [§432]). Teniendo en cuenta la categoría wittgensteiniana de *semejanza de familia* y su enfoque pragmático en *IF* y un enfoque histórico y fenomenológico para comprender la actividad matemática y sus objetos (Zalamea, 2009), entonces, se concluye que el concepto de “número” es abierto. Es, en palabras de Wittgenstein, “un concepto de bordes borrosos” (1954/2008, p. 91 [§71]). No hay un conjunto de propiedades preestablecidas para definir número, a diferencia de otros conceptos matemáticos; es más, al incluir un miembro este “trae” consigo nuevas propiedades o desaparecen otras. Este miembro se comprende a partir de la actividad matemática y cobra “vida” con su “uso”. La inercia de los matemáticos (la necesidad de resolver problemas) es la que permite que el concepto de número se “amplíe”, dando lugar, de esta forma, a “nuevos números”.

## Referencias

- Arroyo, G. (2019). *Teoría de las definiciones: una introducción crítica*. Los Polvorines: Universidad Nacional de General Sarmiento.
- Bassols A. T. (2014). La filosofía de las matemáticas del segundo Wittgenstein. *Praxis filosófica* (39), 11-40.
- Euclides. (2015). *Elementos* (M. L. Puertas Castaños, trad.). Madrid. Gredos. (Obra original ca. 300 a.C.)
- Frege G. (1973). *Los fundamentos de la aritmética* (U. Moulines, trad.). Barcelona: Laia. (Obra original de 1884)
- Kant I. (2014) *Crítica de la razón pura* (2.<sup>a</sup> ed., M. Caimi, trad.). Buenos

- Aires: Colihue clásica. (Obra original de 1781)
- Rivera, S. (2006). *Ludwig Wittgenstein. Entre paradojas y aporías*. Buenos Aires: Prometeo Libros.
- Sánchez Muñoz, J. M. (2011). Hamilton y el descubrimiento de los cuaterniones. *Pensamiento Matemático*, 1(2), 1-27.
- Spivak, M. (2014). *Calculus* (4.º ed., J. M. Oller Sala, trad.). Barcelona: Reverté. (Obra original de 1967)
- Weitz, M. (1956). The role of theory in aesthetics. *The Journal of Aesthetics and Art Criticism*, 15(1), 27-35.
- Wittgenstein, L. (2008) *Investigaciones filosóficas*. (4.ª ed., A. García Suarez & U. Moulines, trads.). Crítica, Barcelona. (Obra original de 1954)
- Zalamea, F. (2009). *Filosofía sintética de las matemáticas contemporáneas*. Bogotá: Editorial Universidad Nacional de Colombia.



# La aplicación de los argumentos de regresión al infinito

Omar Vásquez Dávila\*.\*

## 1. Introducción

Con el fin de precisar cuál es la aplicación de los argumentos de regresión al infinito (también aparecerán como ARI), analizo por qué se suele afirmar conjuntamente que: a) los argumentos por regresión al infinito son aplicables únicamente en el ámbito filosófico; y b) los argumentos por regresión al infinito son un tipo de prueba por reducción al absurdo. Ambas afirmaciones son esgrimidas por Clark (1988) y, con algunos matices, por Post (1992). Sin embargo, si consideramos conjuntamente que, las pruebas por reducción al absurdo (también aparecerán como ARA) no se aplican meramente en el contexto de las discusiones filosóficas, y que la formalización de ambos tipos de argumentos es básicamente la misma, podríamos tener dudas para aceptar la conjunción entre a) y b). Más precisamente, si los argumentos de regresión al infinito tienen la misma forma que los de reducción al absurdo, y lo que caracteriza a un argumento es su forma, ¿por qué restringir los argumentos de regresión al infinito únicamente al ámbito de la filosofía?<sup>1</sup> ¿No se pueden dar regresos en matemática, biología, ciencias de la computación, etc.? Así pues, estas consideraciones impulsan una evaluación acerca de cuál es el rol y el alcance de los argumentos de regresión al infinito. *Prima facie*, podríamos estar en la siguiente disyuntiva: que el uso de estos argumentos no es exclusivo del ámbito filosófico, o que no son un tipo de prueba por reducción al absurdo. Esta segunda opción puede significar que los ARI tienen distinta forma que los ARA o que generar una regresión al infinito no

<sup>1</sup> Entiendo aquí por ámbitos de la filosofía: la ontología, la epistemología, la moral, etc. En éstos hay ejemplos emblemáticos del uso de los argumentos de regresión al infinito (cf. más adelante).

\* Instituto de Ciencias (ICI), Universidad Nacional de General Sarmiento. Malvinas Argentinas, Buenos Aires, Argentina.

\* Buenos Aires Logic Group, Sociedad de Análisis Filosófico (SADAF). CABA, Buenos Aires, Argentina.  
hildebrando2@hotmail.com

es lo mismo que obtener un absurdo. Veremos que ambas opciones son incorrectas. Por ende, concluiré que el uso de los argumentos de regresión al infinito no se restringe a ámbitos filosóficos.

Este trabajo tiene la siguiente estructura. En la segunda sección, explícito el tipo de análisis que hago aquí de los ARI: éste consiste en caracterizar el tipo de tesis que han sido objeto de la regresión y en precisar cuáles son las posibles consecuencias de una contradicción (o de un absurdo). Para ello, presento algunos casos típicos de argumentación por regresión al infinito y, luego, caracterizo a las contradicciones en función de su rol inferencial. En la tercera sección presento algunos ejemplos paradigmáticos de argumentos de regresión al infinito. En la cuarta sección analizo la forma estándar de estos argumentos. Finalmente, a partir de la caracterización del rol inferencial de las contradicciones, en la quinta sección sugiero que se puede extender el ámbito de aplicación de los ARI.

## **2. Los argumentos de regresión al infinito: un posible análisis**

Este análisis de los ARI está motivado por la idea de (Dummett, 1973/1981) según la cual el uso de una determinada forma de oración tiene dos aspectos: las condiciones bajo las cuales se hace una aseveración, y las consecuencias de ésta.<sup>2</sup> En sus palabras:

Crudely expressed, there are always two aspects of the use of a given form of sentence: the conditions under which an utterance of that sentence is appropriate, which include, in the case of an assertoric sentence, what counts as an acceptable ground for asserting it; and the consequences of an utterance of it, which comprise both what the speaker commits himself to by the utterance and the appropriate response on the part of the hearer, including, in the case of assertion, what he is entitled to infer from it if he accepts it. (Dummett, 1973/1981, p. 396)

Así, para determinar las características de las tesis que son objeto de una regresión, analizaremos las condiciones de su aserción. Esto, a su vez, nos permitirá precisar el ámbito de aplicación de los argumentos de regresión al infinito. Y, para analizar la eficacia de dichos argumentos, analizaremos las consecuencias de dichas tesis, lo cual tiene que ver, principalmente, con qué se entiende por obtener una contradicción y las consecuencias de esta obtención.

---

<sup>2</sup> Esta referencia a Dummett es sólo metodológica. Los objetivos de Dummett en el trabajo citado son distintos a los de este trabajo.

### 3. Algunos ejemplos

Los argumentos de regresión al infinito se han propuesto tradicionalmente en favor de posturas epistémicas acerca de la justificación tales como el fundacionalismo, el escepticismo, el coherentismo, etc. (cf. Annis, 1978). Por ejemplo, si una creencia está justificada en la medida en que se infiere de otra creencia justificada, debería haber una secuencia infinita de justificaciones;<sup>3</sup> pero debido a que se asume que no puede haber una secuencia de ese tipo, no es cierto que toda creencia se justifica inferencialmente a partir de otra creencia justificada. Dicho de otro modo, debe haber creencias justificadas cuya justificación no se basa en otras creencias justificadas, por ejemplo, creencias autoevidentes, inmediatas, etc. Así, la justificación no sería necesariamente inferencial. Esto nos lleva a aceptar creencias básicas, inmediatas o fundacionales. Esta es en esencia la postura fundacionalista de la justificación.

Otra aplicación paradigmática de los argumentos de regresión es en contra de la existencia de las relaciones. Brevemente, si se afirmara que hay una relación  $R$  entre un hecho (u objeto)  $A$  y un hecho (u objeto)  $B$ , debería haber una relación  $R_1$  entre  $R$  y  $A$  (y también entre  $R$  y  $B$ ). Pero entonces, hay también una relación  $R_2$  entre  $R_1$  y  $R$  (y también entre  $R_1$  y  $A$ ), y así *ad infinitum*. Pero si la aceptación de la existencia de relaciones nos compromete con esta secuencia infinita de relaciones, debemos negar la existencia de relaciones. Esta es una versión simplificada del argumento clásico de Bradley en contra de la existencia de las relaciones. Asimismo, encontramos argumentos de regresión al infinito en otras disciplinas: en ética a favor de la existencia de un único fin último de la acción racional, en metafísica a favor de la existencia de un motor inmóvil, etc.

### 4. La forma de los ARI

Según algunos autores (Clark, 1988; Post, 1992) la diversidad de aplicación de estos argumentos, siempre dentro del ámbito de lo filosófico, se debe a que “tienen la forma de una reducción al absurdo de una conjunción de supuestos”, (Post, 1992). Las distintas posturas respecto de la justificación, por ejemplo, surgen de la revisión de alguno de los supuestos que generan la regresión. Esto evidencia que la regresión no es, necesar-

<sup>3</sup> Si la relación de justificación es asimétrica y transitiva (véase sección 4).

amente, un problema en sí mismo, sino que es síntoma de que se puede estar incurriendo en algún tipo de vicio teórico. En general, podríamos decir que la forma de un argumento de regresión al infinito es la siguiente (cf. Black, 1988):

1.  $\forall x (Ax \rightarrow (\exists y) (Ay \& xRy))$  (e.g., toda creencia justificada está justificada por otra creencia justificada)
2.  $\exists x Ax$  (e.g., una instancia sería: hay creencias justificadas)
3. R es irreflexiva (i.e., una creencia justificada no se justifica a sí misma)
4. R es transitiva (i.e., si una creencia justificada  $x$  justifica una creencia  $y$ , e  $y$  justifica  $z$ ,  $x$  justifica  $z$ )
5. No hay una secuencia infinita cada uno de cuyos elementos tenga A y a la vez mantenga R con su predecesor. (supuesto finitista de la justificación)
6.  $\perp$  (de 1-5), pues se sigue de 1 a 4 que hay una secuencia infinita tal que cada uno de sus elementos tiene A y a la vez mantiene R con su predecesor. Y esto contradice 5.

Esta formalización evidencia tres cosas. Primero, a diferencia de lo que pasa en una reducción al absurdo, uno de los supuestos consiste en una afirmación de existencia (el supuesto 2). En segundo lugar, se asume que la relación R tiene ciertas características (supuestos 3 y 4). En un argumento por reducción al absurdo no es necesario asumir ni siquiera la existencia de R. Finalmente, lo que considero más controversial –y que llamativamente no ha sido discutido en la literatura sobre el tema– es que la relación R es la versión diádica de la propiedad (o predicado monádico) A. En principio, se puede pensar que esto no representa un problema, pues la propiedad de *ser alta* puede convertirse en *ser más alta que*, la propiedad de *ser pesado*, en *ser más pesado que*, etc. Sin embargo, es discutible, primero, que toda propiedad monádica pueda convertirse en una relación; y segundo, que la propiedad de *ser alta* (o *ser pesado*) sea transferible, como sucede, por ejemplo, con la propiedad de *ser una creencia justificada*, o con la propiedad de *estar en movimiento*. De esta manera, tenemos que en los argumentos por regresión al infinito hay un supuesto más: el supuesto de que A representa una propiedad transferible. Ninguno de estos supuestos es necesario en una reducción al absurdo. Todos ellos tienen una carga

ontológica que podría explicar por qué se ha restringido la aplicación de los argumentos de regresión al infinito al ámbito de lo filosófico.

## **5. El rol inferencial de las contradicciones**

Clark sostiene que la aplicación de los ARI se restringe al ámbito de lo filosófico. Según este autor:

It is not necessary but it is natural that arguments of this sort tend to be philosophical arguments. Dealing in infinite totalities is already a fairly exotic sort of enterprise. (It is not, e.g., the common stuff of politics or the law.) Dealing in infinite sequences without the possibility of establishing a relevant base case for an inductive argument, is a pretty restricted sort of task. (It is not, e.g., the typical mathematical sort of task.) Drawing a universal negative conclusion, where there can be no theory-independent fact of the matter, is not ordinarily a matter of common science, or common sense. (Clark, 1986, pp. 378-379)

Si bien podemos discutir que sólo en filosofía lidiamos con totalidades infinitas, o que en filosofía cuando se hace esto es imposible establecer un caso base, prefiero, al menos en este trabajo, analizar el último punto mencionado por Clark. Según éste, obtener una conclusión universal negativa no es algo común en ciencia o algo de sentido común. En lugar de discutir si esto es común o no, lo cual considero poco fructífero, me concentraré en precisar qué se puede concluir de un regreso al infinito en tanto argumento que tiene la forma de una reducción al absurdo. Para esto, sugiero que nos concentremos en un componente común, y a la vez esencial, tanto a los ARI como a los ARA, a saber, la obtención de una contradicción. Veremos que, a partir de esto, podremos arrojar ciertas dudas a esta restricción que se hace de los ARI al ámbito de lo filosófico.

Si entendemos la noción de absurdo técnicamente como una contradicción,<sup>4</sup> y ésta se obtiene en los dos tipos de argumentos que estamos analizando –i.e., los de regresión al infinito y los de reducción al absurdo–, entonces, es necesario considerar cuestiones tales como qué es una contradicción, cuán inaceptable es (¿nos lleva ésta necesariamente a revisar los supuestos?), cuál es el rol inferencial de las contradicciones, etc. Un análisis de la última cuestión puede ser iluminadora al respecto. De acuerdo con Routley y Routley:

---

<sup>4</sup> Qué es una contradicción no es un asunto carente de controversia (cf. Grimm, 2006).

Theories of negation differ, very obviously, in the roles they allow or assign to, contradictions. Contradictions may be allowed no inferential role (they imply nothing, except perhaps themselves), a total inferential role (they imply everything), or some limited inferential role (they imply some things, such as their contradictory components, but not others). (Routley & Routley, 1985, p. 205)

Estos distintos roles de las contradicciones dependen de cuál sea nuestra caracterización de la negación. Como se sabe, el segundo rol se corresponde con la negación clásica y la intuicionista; el tercer rol se corresponde con la negación de la *logic of paradox* (LP) de Priest (1979) (entre otros sistemas); y el primero, con la concepción de la negación como cancelación.<sup>5</sup> Según esta última,  $\sim A$  cancela el contenido de  $A$  (y viceversa), de modo que  $A \wedge \sim A$ , es vacía de contenido; y de algo que no tiene contenido, no se puede inferir ni  $A$  ni  $\sim A$  ni nada.<sup>6,7</sup> Para nuestros propósitos es importante advertir que tanto en la primera como en la tercera caracterización de la negación, el principio *ex contradictione quodlibet* ( $A \wedge \sim A$ )  $\vdash$   $B$  no es válido.

La razón que se suele proporcionar para enfatizar que los argumentos de regresión al infinito no son concluyentes es que la contradicción obtenida nos obliga a revisar los supuestos, y dado que hay más de un supuesto revisable, distintas tesis son defendibles a partir de la obtención de una regresión. Por ejemplo, luego de obtener una contradicción, el fundacionalista niega el paso 1 de la formalización presentada en la sección anterior, i.e., niega que toda creencia justificada se justifique a partir de otra creencia justificada. Pero también se puede negar el paso 2 para evitar la contradicción; y esto es lo que haría un escéptico, en efecto, negaría que hay creencias justificadas. De modo que la obtención de la regresión es funcional tanto al fundacionalista como al escéptico.

Sin embargo, este impulso hacia la revisión de los principios está motivada por la idea de que una contradicción genera trivialización. Como hemos visto, el rol inferencial de las contradicciones es variado. Si aceptá-

---

<sup>5</sup> No me detendré aquí en la discusión acerca de cuál es la negación con más virtudes teóricas. Simplemente intento señalar que cuál es el rol de las contradicciones es una cuestión abierta. Para una discusión más detallada y técnica de estos distintos tipos de negación (cf. Wansing & Skurt, 2018).

<sup>6</sup> En una lectura más moderada de la negación por cancelación ( $A \wedge \sim A$ ) se puede implicar a sí misma, pero a nada más.

<sup>7</sup> Wansing y Skurt (2018) proponen que la ineffectividad inferencial de las contradicciones no supone necesariamente la concepción de la negación como cancelación.

ramos la ineffectividad inferencial de las contradicciones, ¿nos sentiríamos obligados a revisar los supuestos? ¿qué otras actitudes podemos tomar ante la obtención de una contradicción? Una respuesta es que la obtención de una contradicción nos puede llevar temporalmente a suspender el juicio.<sup>8</sup> Los procedimientos en ciencias de la computación apoyan esta posibilidad. Según, Gerd Wagner: “From an information processing point of view, however, it seems to be more natural to discard contradictory information as neither true nor false, so it cannot be used in the derivation of further information” (Wagner, 1991, p. 542).

No parece difícil aceptar que la suspensión de juicio no es algo tan grave como la trivialización. Estas consideraciones parecen indicar que los argumentos de regresión (y los de reducción al absurdo) no son concluyentes debido a que nos llevan a la suspensión de juicio. Esta postura puede resultar más atractiva que aquella que propone que la obtención de una contradicción nos obliga a revisar los supuestos debido a que nos lleva a reiniciar la investigación desde cero (i.e., a partir de un nuevo conjunto de supuestos), o bien debido a que evita que el mismo argumento nos sirva para sostener tesis incompatibles entre sí, como el fundacionalismo y el escepticismo de la justificación. A su vez, hemos ejemplificado este análisis basado en un caso de las ciencias de la computación, ámbito en el cual no es difícil aceptar que, primero, lidiamos con secuencias infinitas y, además, como hemos visto, la obtención de una contradicción puede tener distintas consecuencias.

## **6. Conclusión**

Básicamente he tratado de motivar la idea de que los ARI pueden aplicarse a ámbitos que no sean propiamente filosóficos sin renunciar a la idea de que dichos argumentos son un tipo de argumentos de reducción al absurdo. A su vez, hemos caracterizado las diferencias que hay entre unos y otros a partir de los compromisos ontológicos que encontramos en los ARI y no en los ARA. Estos compromisos explicarían por qué todos los ARI son un tipo de ARA pero no todos los ARA un tipo de ARI. Al analizar las condiciones de aseveración de las tesis que generan regresiones al infinito hemos visto que los supuestos que tienen afirmaciones de existencia nos permiten diferenciar este tipo de argumentos de los argu-

---

<sup>8</sup> Sobre si esta actitud se condice con la actitud escéptica es algo que no discutiré aquí.

mentos por reducción al absurdo. Por otro lado, la caracterización del rol inferencial de las contradicciones nos permitió proponer una explicación, alternativa a la usual, de por qué los argumentos de regresión no son concluyentes.

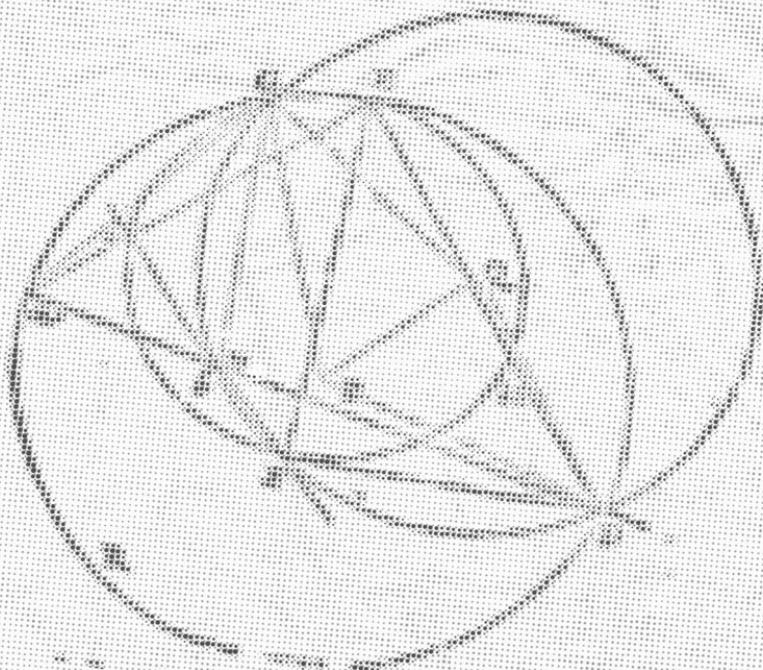
## Referencias

- Annis, D. B., (1978). A contextualistic theory of epistemic justification. *American Philosophical Quarterly*, 15, 213-219.
- Black, O., (1988). Infinite regresses of justification. *International Philosophical Quarterly*, 28(4), 421-437.
- Clark, R., (1988). Vicious infinite regress arguments. *Philosophical Perspectives*, 2, 369-380.
- Dummett, M., (1981). *Frege: Philosophy of language* (2<sup>nd</sup> ed.). Cambridge, MA: Harvard University Press. (Obra original de 1973)
- Grimm, P., (2006). What is a contradiction? En G. Priest & J.C. Beall (Eds.), *The law of non-contradiction* (pp. 49-72). Oxford: Oxford University Press.
- Post, J. F., (1992). Infinite regress argument. En J. Dancey & E. Sosa (Eds.), *A companion to epistemology* (pp. 209-212). Oxford, Blackwell.
- Priest, G. (1979). The logic of paradox. *Journal of Philosophical Logic*, 8(1), 219-241.
- Routley, R., & Routley, V. (1985). Negation and contradiction. *Revista Colombiana de Matemáticas*, 19 (1-2), 201-230.
- Wagner, G. (1991). Ex contradictione nihil sequitur. En J. Mylopoulos & R. Reiter (Eds.), *IJCAI'91: Proceedings of the 12th International joint conference on artificial intelligence* (vol. 1, pp. 538-543). Sydney: Morgan Kaufmann Publishers Inc.
- Wansing, H., & Skurt, D. (2018). Negation as cancellation, connexive logic, and qLPm. *The Australasian Journal of Logic*, 15(2), 476-488. doi: 10.26686/ajl.v15i2.4869



---

## 9. Lógica y filosofía de la lógica







# Hacer lógica y saber historia de la lógica

## Hacia una reconciliación metodológica

Miguel Álvarez Lisboa\*

### Introducción

Dicen que Wittgenstein se vanagloriaba de nunca haber leído a Aristóteles. Esta anécdota ilustra bien uno de los rasgos metodológicos más sobresalientes de la primera filosofía analítica: su resistencia a las preocupaciones históricas.

Esta actitud fue originalmente una reacción a la tendencia contraria, que por aquellos años ganaba popularidad en las universidades europeas. Tal doctrina proponía que los problemas modernos están indisolublemente atados a sus versiones originales, y todavía estas últimas son inseparables del contexto, las ideas e incluso el lenguaje que utilizaron los filósofos que primero los pensaron. Por lo tanto, toda filosofía debe adoptar metodologías filológicas y exegéticas si quiere proponerse como auténtica o valiosa. Para los filósofos analíticos de vieja escuela esto es una exageración. En el mejor de los casos, la historia provee el vocabulario básico y los problemas originales, pero cada filósofo está en posición de formular las preguntas como le parezcan más interesantes y desde el léxico que le resulte más apropiado (Beaney, 2013; Glock, 2008).

Hoy en día, un gran número de filósofos “analíticos” todavía están dispuestos a aceptar su compromiso con una versión ligeramente más moderada de la misma actitud. Si hubiera que enunciarla en un *dictum*, acaso el siguiente sea apropiado: el conocimiento de la historia de la filosofía, disciplina respetable e interesante por mérito propio, no es indispensable para que el filósofo haga bien su trabajo. Voy a llamar a esta actitud *ahistoricismo metodológico* (AM).

De todas las disciplinas que conforman el baluarte de la tradición dicha analítica, se ha pretendido que sea la lógica aquella que ilustra de mejor manera la fertilidad del AM. Podemos explicar esto sobre la base de un hecho evidente: la introducción de los métodos de la matemática contribuyó

\* IIF-SADAF-CONICET/Buenos Aires, Argentina.  
miguel.alvarez@um.uchile.cl

muy positivamente al auge de la lógica a finales del siglo XIX y explica su rápido desarrollo en el XX. Este fenómeno suele ser racionalizado por la filosofía contemporánea como un despertar del sopor del aristotelismo y la escolástica. Esto relaciona a la lógica con el AM de al menos dos maneras distintas:

1. Los trabajos que permitieron el desarrollo de la lógica Matemática ignoraron en muy buena medida la tradición lógica anterior.
2. La matemática en sí misma es también fuertemente ahistórica.

Por el momento no interesa si esta imagen del “despertar” es apropiada o no; tan sólo quiero llamar la atención sobre lo convincente, o, a lo menos, verosímil, de esta justificación para confiar en el AM en la lógica.

Mi propósito en este artículo es debilitar precisamente dicha confianza. Lo que haré será presentar tres razones para creer que la historia de la lógica le provee al lógico<sup>1</sup> herramientas de trabajo que de otra manera le serían difíciles de conseguir. En este sentido, no busco defender que el conocimiento de la historia de la lógica sea *indispensable* para el trabajo del lógico, sino tan sólo mostrar cómo él puede serle útil en un sentido no trivial.

## La naturalización del presente

Una de las consecuencias de la adscripción al AM es un fenómeno llamado *naturalización del presente*. Un científico trabajando bajo un paradigma<sup>2</sup> comprende sus problemas y los de sus colegas contemporáneos como los únicos problemas científicos, de los cuales los problemas antiguos (que no conoce en profundidad) fueron versiones rudimentarias o imperfectamente formuladas. Mira toda la historia de la ciencia como un reflejo de su presente: en el pasado se formularon los problemas y en el futuro se resolverán.

---

<sup>1</sup> Aunque simpatizo y promuevo el uso de formas gramaticales inclusivas en el discurso académico, aquí uso las expresiones “el lógico” y “los lógicos” como género mixto porque “la lógica” y “las lógicas” son equívocas en este contexto. En concreto, “la lógica” de igual manera refiere a la profesional que cultiva la disciplina y dicha disciplina en cuanto tal, por lo que me parece que usar “la lógica” en el primer sentido dificulta innecesariamente la lectura.

<sup>2</sup> A lo largo de mi escrito utilizo una serie de conceptos kuhneanos que *no* deben entenderse como un compromiso fuerte con su filosofía sino como un *façon de parler*.

La lógica Matemática, en tanto ciencia formal, está supeditada a las mismas formas de desenvolvimiento y desarrollo de las demás ciencias. Por lo tanto, estas mismas consideraciones aplican para ella.

Un lógico profesional trabaja en las áreas delimitadas por su tiempo y se aboca a buscar soluciones a problemas ortodoxos. Algunos de estos intentos de solución serán igualmente ortodoxos, pero otros no. Cuando este último es el caso, cierta resistencia es esperable por parte de la comunidad; a veces, simplemente, por influencia de ese presente naturalizado.

Durante la primera mitad del siglo XX, el paradigma en lógica se consolidó en torno a lo que hoy se conoce como la “Lógica Clásica”. Dicha matriz disciplinaria reúne las doctrinas de la filosofía analítica de Cambridge en torno a las nociones de significado y validez, las de la escuela de Leópolis-Varsovia en torno a las de modelo y verdad, además de varios resultados matemáticos. Consideremos ahora el argumento del “cambio de tema” de Quine. El contexto de este conocido alegato del filósofo es el capítulo 6 de su libro *Philosophy of logic*, lugar donde discute las lógicas que llama “*deviant*” [desviadas]. Él escribe:

To turn to a popular extravaganza, what if someone were to reject the law of non-contradiction and so accept an occasional sentence and its negation both as true? . . . My view of this dialogue is that neither party knows what he is talking about. They think they are talking about negation,” ‘not’; but surely the notation ceased to be recognizable as negation when they took to regarding some conjunctions of the form “ as true, and stopped regarding such sentences as implying all others. Here, evidently, is the deviant logician’s predicament: when he tries to deny the doctrine he only changes the subject. (Quine, 1970/1986, p. 81)

El uso de la expresión “*popular extravaganza*” [extravagancia popular] es sugerente. Uno podría preguntarse qué tan extravagante es el dudar de que las contradicciones no impliquen cualquier cosa, desde un punto de vista histórico. En el sistema lógico aristotélico, por ejemplo, la necesaria conexión de la conclusión con las premisas por medio de la comunión de términos hace formalmente imposible que un argumento tenga por conclusión “cualquier cosa”. Los estoicos, por su parte, conocieron el contenido formal del *ex falso*, pero lo tuvieron por paradójico (D’Ottaviano & Gomes, 2017). Para los autores escolásticos el tema también es motivo de debate. La doctrina es aceptada por quienes defienden que validez es preservación de verdad, pero a regañadientes, pues tiene consecuencias

indeseables desde un punto de vista teológico. Por otra parte, escuelas que no reconocen el criterio de validez como preservación de verdad están abiertamente en contra de la doctrina.

El (hoy) famoso argumento de *Pseudo-Scotus*, la primera y más simple prueba formal del Principio de Explosión, fue conocido y discutido por los lógicos escolásticos, pero es un hecho que su impacto no fue profundo para la lógica inmediatamente posterior. No en vano el autor del texto original, *Universam logicam quaestiones*, fue olvidado. El texto propiamente tal fue redescubierto en el siglo XVII e incluido erróneamente en la obra completa de Duns Scotus. Así, el aclamado *Ex falso sequitur quodlibet* apenas figura en la literatura lógica anterior a Frege y Peano.

Por lo tanto, cuando Quine acusa a los revisionistas de cambiar el tema, realmente cabe discutir si es legítimo atribuir el Principio de Explosión al “significado” de la negación, cuando tal doctrina fue abiertamente considerada polémica durante siglos y por no pocas escuelas lógicas que tuvieron perfecta conciencia de sus fundamentos y sus consecuencias. El alegato de Quine sólo tiene sentido si la historia de la lógica es mirada desde una óptica simplificadora: si la lógica clásica no es un nuevo paradigma sino *el* paradigma, gestado lentamente a lo largo de los siglos, entonces los lógicos de la antigüedad sólo no tenían conciencia del significado completo o perfecto de la negación. Esto, como ya hemos dicho, es falso.

Lo que está a la base de esta idea prejuiciosa es una falsa imagen de la historia de la lógica a la que llamo: *El mito fundacional de la lógica clásica*. Él consiste en la narración siguiente: la lógica fue inventada por Aristóteles. El sistema, aunque primitivo e imperfecto en sus métodos, sentó las doctrinas principales y permitió obtener algunos resultados clave. En los siglos posteriores los lógicos se dedicaron sólo a releer la obra del estagirita y a inventar métodos mnemotécnicos para optimizar su aplicación a casos concretos. Finalmente, los lógicos matemáticos del siglo XIX consiguieron superar el sistema original de Aristóteles, corrigiendo sus errores y sobrepasando sus límites conceptuales. Se llegó entonces a la versión más perfecta y en alguna medida definitiva del sistema lógico, que es la moderna lógica clásica.

Es difícil decir cuál es el origen del mito fundacional. Los Kneale, por ejemplo, responsabilizan a Kant y sus afirmaciones apresuradas sobre la perfección del *Órganon* aristotélico por su propagación original y la perseverancia en su creencia (Kneale, M. & Kneale, W., 1962, pp. 354). En cualquier caso, me parece que la emergencia de este discurso es perfecta-

mente natural y compatible con la naturalización del presente. Es “normal” que el lógico “normal” vea a los lógicos antiguos como investigadores que trabajan los mismos temas que él y bajo los mismos presupuestos, sólo que con métodos imperfectos y versiones rudimentarias de sus teorías.

Por lo tanto, al menos una manera efectiva, si no la única, que tiene el lógico “desviado” para defenderse contra críticas como la de Quine es apelar a conceptos históricos. Esto siguiendo dos estrategias: o bien a) mostrar que su idea va en la línea de algún programa de investigación olvidado, o bien b) defender que la visión hegemónica es reciente y se funda en concepciones de corta data que no proveen soluciones satisfactorias a todos los problemas.

Hay antecedentes notables del uso de ambas estrategias en la literatura reciente. Tomemos por primer caso la Dialógica. Se trata de un programa de investigación muy reciente (apenas sesenta años) que se basa en ideas muy diferentes a las que maneja el resto de la comunidad. Por supuesto, desde un punto de vista matemático no hay objeción que quepa al intento de desarrollar nuevos métodos para resolver viejos problemas. Sin embargo, ellas no son tan convincentes para los filósofos. Los lógicos dialógicos han sabido defender su programa de investigación ante posibles alegatos de cambio de tema racionalizando su genealogía en retrospectiva. Por ejemplo, en el proemio del libro introductorio *How to play dialogues* los autores escriben:

The expression Dialogical Logic refers to a research tradition *that can be traced back to at least Greek antiquity*, [énfasis agregado] when logic was conceived as the systematic study of dialogues in which two parties exchange arguments over a central claim. In its modern form, dialogical logic uses concepts of game theory to design dialogue games that provide semantics for a wide range of logical systems. (Redmond & Fontaine, 2011, p. xv)

También en el artículo “Dialogical logic” de E. Krabbe para el *Handbook of philosophical logic* encontramos un giro parecido: el autor, antes de reconstruir la historia reciente de la Dialógica, agrega toda una sección titulada “The agonistic roots of logic” que comienza así:

The word ‘agonistic’ in the title of this section derives from the Greek *agon* (=game, contest). In what follows I hope to show that originally logic was concerned with a kind of discussion game or verbal contest and that, therefore, *contemporary dialogue logic has its roots in the oldest kind of logic* [énfasis agregado]. (Krabbe, 2006, p. 666)

El movimiento es atrevido: no sólo la dialógica es antigua, sino que es tan antigua como toda lógica, y más aún, rescata el espíritu de la lógica *más antigua*. Ni esta observación ni la de Redmond y Fontaine tienen utilidad teórica ni aportan ventajas técnicas de algún tipo: ellas tan sólo sirven como carta de presentación filosófica.

Un ejemplo de la segunda estrategia lo encontramos en la lógica relevantista<sup>3</sup>. En su obra fundacional, *Entailment: the logic of relevance and necessity*, Belnap y Anderson justifican su sistema diciendo:

*For more than two millennia logicians have taught that a necessary condition for the validity of an inference from A to B is that A be relevant to B [énfasis agregado]. Virtually every logic book up to the present century has a chapter on fallacies of relevance, and many contemporary elementary texts have followed the same plan. (Anderson & Belnap, 1975, p. 17)*

Los autores evitan la estrategia de atacar directamente a la lógica clásica por haber ignorado la característica de relevancia; antes bien, lo que ellos hacen es poner a la historia de *su* lado, desafiando (aunque sea en una sola oración) la plausibilidad del Mito Fundacional.

## Historia de los errores

Al comienzo de su libro *Intuition pumps and other tools for thinking* (2006), el filósofo Daniel Dennett caracteriza a la historia de la filosofía más o menos como la historia de personas muy inteligentes que cometen errores muy tentadores. Por esta razón, dice él, es que los filósofos se esmeran tanto en enseñar a sus estudiantes la historia de su disciplina: porque quien no conoce la historia de dichos errores está condenado a cometerlos de nuevo.

Los científicos no suelen tener la misma idea de la historia de su disciplina. Pero esto es en gran medida porque en la ciencia, y sobre todo en la ciencia empírica, los resultados positivos se acumulan: el científico puede avanzar con la confianza de que todo lo acertado que hicieron sus colegas del pasado ha sido recuperado de alguna u otra manera en *sus* propios trabajos y teorías.

---

<sup>3</sup> Uso la palabra relevantista para traducir *relevantist* y relevante para traducir *relevant* para no desviarme del uso común en lenguaje española, pero es sensible notar que lo que los lógicos de esta tradición entienden por *relevance* se corresponde con lo que en español se dice “pertinencia”.

Esto aplica también para la lógica: resultados como la prueba de la validez de los silogismos de primera figura, la demostración de Pseudo-Scotus o los Teoremas de Incompletitud de Gödel están “a la vista” tanto como el investigador los necesite y se los presupone en trabajos posteriores sin necesidad de una lección de historia. Pero estos son los resultados positivos de la ciencia: por regla general, los resultados negativos, errores o intentos fallidos no suelen ser recordados con la misma fuerza.

Como en la historia de toda ciencia, hay en el pasado de la lógica muchas historias de programas de investigación fallidos. Dado que la lógica *no* es una ciencia empírica, sus resultados positivos no son de la misma naturaleza que los resultados de la física o la química. En esto el lógico es más como la filósofa: el peligro de los errores tentadores está siempre al acecho.

Un ejemplo particularmente bueno de esto es la lógica de la negación. Tal como atestigua Lawrence Horn en *A Natural history of negation* (1989/2001), la literatura en torno a esta conectiva lógica es muy antigua, amplia y profunda. Para la investigadora novel esto implica necesariamente el peligro de volver a incurrir en un error pasado, si el panorama de los intentos y errores no se ha considerado al menos de manera general.

Miremos una de estas propuestas específicas. A lo largo de los siglos ha habido numerosos intentos de reducir la negación a una relación de “incompatibilidad” pretendida como primitiva. En la literatura contemporánea más reciente la escuela australiana ha alcanzado relativa notoriedad por su propuesta modal de la estrategia (Berto & Restall, 2019), pero de ninguna manera ellos la han inventado. De hecho, se la puede rastrear al menos hasta *El Sofista* de Platón, por lo que es una idea filosófica de larga data.

El problema con esta doctrina de la negación es el siguiente (Horn, 1989/2001): si el criterio de otredad o *incompatibilidad* ha de ser tomado como primitivo, entonces él describe una relación que se ajusta o no al Principio de No Contradicción; si lo hace, entonces hay circularidad (pues lo incompatible sólo se comprende por apelación a la negación, que es la noción que pretende definir); si no, entonces es inadecuado (pues incluso el dialicismo reconoce que hay negaciones que respetan de manera *a priori* el Principio de No Contradicción). “Yet, as untenable analyses go, the Otherness thesis has proved surprisingly resilient” (Horn, 1989/2001, p. 51).

No hay ningún problema con intentar salvar una doctrina filosófica, sobre todo si ella se muestra intuitiva o tiene casos notables de aplicación; pero si hay un desconocimiento de los intentos anteriores, es más que probable que se repita un error flagrante. La discusión de las propuestas en Horn (1989/2001) es ilustrativa y sugerente: tanto los problemas como los intentos de solución suelen parecerse entre sí.

## Bibliografía e inspiración

A diferencia de otras disciplinas científico-filosóficas, que crecieron lentamente con el pasar de los siglos y sufrieron cambios sustantivos de forma y contenido en ese tiempo, la lógica maduró muy rápidamente. Sin embargo, el prejuicio clásico tiende a ignorar este hecho. El *Órganon* de Aristóteles es un caso: según el mito fundacional, estos libros constituyen el nacimiento de la Lógica como disciplina filosófica y proto-ciencia. Sin embargo, tal como atestiguó tempranamente Łukasiewicz, la obra del Estagirita presenta rasgos de una sistematicidad que fue en gran parte subestimada por los lógicos matemáticos de nuestro tiempo.

Es cierto que los lógicos de cada siglo no siempre trabajaron los mismos temas o vieron con igual importancia unos problemas u otros. Esto es así porque, en buena medida, dichos trabajos se desarrollaban bajo paradigmas diferentes. Pero ya hemos dicho que la forma de progreso en la lógica no es acumulativa, por lo que admite un desarrollo, por así decirlo, en paralelo antes que serial.

Si esta visión de la lógica es correcta, entonces los textos de los lógicos del pasado, tanto los más lejanos en el tiempo como los más recientes, no son sólo antecedentes o curiosidades relativas a nuestra ciencia sino textos con los que se puede dialogar y extraer información y comprensión de nuestros problemas modernos.

Mi sugerencia metodológica en esta tercera sección es que uno puede inspirarse de los textos de lógicos del pasado. Y uso la palabra “inspiración” en el siguiente sentido: cuando uno toma ideas o sugerencias de otras fuentes, para componer en la mente de uno una idea original que comparte con la fuente sólo rasgos lejanos o líneas generales.

Hay ejemplos de inspiración en este sentido en la lógica contemporánea, por supuesto. El mejor debe ser el de las lógicas conexas [*connexive logics*]. Estos sistemas de lógica fueron propuestos originalmente como “an attempt to formalize the species of implication recommended

by Chryssipus: And those who introduce the notion of connection say that a conditional is sound when the contradictory of its consequent is incompatible with its antecedent” (McCall, citado en Wansing, 2016, párr. 7). En la literatura sobre el tema se suele hablar de estos dos esquemas:

$$(TA) \sim(\sim A \rightarrow A)$$

$$(TB) (A \rightarrow B) \rightarrow \sim(A \rightarrow \sim B)$$

como las tesis de Aristóteles y Boecio respectivamente, porque ambos lógicos parecen haber sugerido en su obra tales pensamientos. Pero dírimir si (TA) es exactamente algo que dijo Aristóteles, o si puede decirse que Aristóteles hubiera asentido ante una formalización de sus ideas más cercana a la lógica conexiva que a la lógica clásica ya no es una discusión lógica propiamente tal sino exegética; ella pertenece a la historia de la lógica. Lo importante aquí es que los precursores de estos sistemas de lógica no-clásica se *inspiraron* en textos antiguos para desarrollar sus propios trabajos. Una vez que los formalismos han sido presentados, ellos pueden desembarazarse felizmente de sus fuentes; la lógica conexiva se convierte en un objeto de estudio por mérito propio, que tendrá o no sus aciertos lógicos y filosóficos dependiendo de su desarrollo moderno. Esto es a lo que me refiero con *inspiración*.

Se ha dicho en ocasiones que Wittgenstein previó, sobre todo en algunos pasajes de sus observaciones sobre los fundamentos de la matemática, el desarrollo de las lógicas paraconsistentes e incluso del dialecismo. Ha habido cierto crédito reconocido al respecto: sin ir más lejos, el neologismo *dialethism* fue acuñado precisamente teniendo a la vista uno de estos pasajes. Una lectura pesimista sería afirmar que en realidad Wittgenstein ni previó ni sugirió ninguna de estas ideas novedosas. Jean-Yves Girard llega a comparar a Wittgenstein con Nostradamus, en el sentido de que en su obra todo puede hallarse, pero sólo *en retrospectiva*.<sup>4</sup> Me parece que esto es injusto: ciertamente los apuntes de Wittgenstein, así como también los escritos de los antiguos muchas veces resultan oscuros para el ojo moderno; pero es precisamente esta oscuridad la que permite la inspiración en el sentido arriba esbozado. Después de todo, los científicos que promovieron las más grandes revoluciones no fueron los mejor instruidos en la ciencia de su tiempo sino precisamente los que la comprendían

<sup>4</sup> La frase exacta de Girard, que me ha costado un poco traducir, dice: “Except perhaps Wittgenstein, an inspired guy, in the work of which everything can be found, but –like Nostradamus– only *afterwards*” (Girard, 2011, p. 181).

medianamente o mal, y que por esto fueron capaces de pensar de maneras heterodoxas.

## Conclusión

Contra la seducción que todavía podría ejercer en los lógicos modernos el AM, en este artículo he presentado tres argumentos para reconciliar a la Lógica con su Historia. El primero es que las apelaciones históricas pueden servir como defensa contra el conservadurismo. Probar la antigüedad de una idea o cuestionar la longevidad de la ortodoxia es una estrategia eficiente para debilitar ciertas críticas reaccionarias. El segundo argumento es que en lógica es importante tener a la vista los proyectos fallidos, pues es fácil caer en errores tentadores si no se los conoce. Finalmente, el tercero de mis argumentos mostró que los textos de los antiguos son una valiosa fuente de inspiración. Por estas tres razones es que sostengo que saber Historia de la Lógica de hecho resulta útil para el lógico, en el sentido de que se provee de herramientas que de otra forma le serían difíciles de conseguir.

## Referencias

- Anderson, A., & Belnap, D. (1975). *Entailment: the logic of relevance and necessity*. New Jersey: Princeton University Press.
- Beaney, M. (2013). The historiography of analytic philosophy. En M. Beaney (Ed.), *The Oxford handbook of the history of analytic philosophy* (pp. 30-60). doi: 10.1093/oxfordhb/9780199238842.013.0029
- Berto, F., & Restall, G. (2019). Negation on the Australian plan. *Journal of Philosophical Logic*, 48, 1119-1144.
- Dennett, D. (2006). *Intuition pumps and other tools for thinking*. Londres: W. W. Norton and Company.
- D'Ottaviano, I. M. L., & Gomes, E. L. (2017). *Para além das Colunas de Hércules, uma história da paraconsistência: de Heráclito a Newton da Costa*. Campinas: Editora UNICAMP.
- Girard, J. Y. (2011) *The blind spot*. Bad Langensalza: European Mathematical Society.

- Glock, H. J. (2008). Analytic philosophy and history: A mismatch? *Mind*, 117(468), 867-897. doi:10.1093/mind/fzn055
- Horn, L. R. (2001). *A natural history of negation*. Leland: CSLI Publications. (Obra original de 1989)
- Kneale, M, & Kneale, W. (1962). *The development of logic*. Londres: Oxford University Press.
- Krabbe, E. C. W. (2006). Dialogue logic. En D. M. Gabbay & J. Woods (Eds.), *Handbook of the history of logic 7: Logic and the modalities in the twentieth Century* (pp. 665-704). doi:10.1016/s1874-5857(06)80035-x
- Quine, W. O. (1986). *Philosophy of logic*. Cambridge: Harvard University Press. (Obra original de 1970)
- Wansing, H., (2016). Connexive logic. En E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2016 edition). Recuperado de <https://plato.stanford.edu/archives/spr2016/entries/logic-connexive/>



## Una consideración del juicio reflexionante kantiano desde la lógica stit

Inés Crespo\*

Sebastián Ferrando\*\*

Luis A. Urtubey\*\*

### Introducción

Según Kant en la *Crítica del juicio*, el juicio de gusto, un caso del juicio estético, como “Este hombre es bello” (1790/1992, Libro primero, §25) es universal, aunque subjetivo. Los juicios de gusto y los de lo agradable son estéticos, basados en la experiencia de un individuo singular, ya que se formulan a partir de respuestas subjetivas: “[...] el sujeto se siente a sí mismo tal como es afectado por la representación” (Kant, 1790/1992, Libro primero, §1).

No obstante, para Kant no tiene sentido decir que algo es bello para mí: “pues no debe llamarlo *bello* si le place meramente a él” (Kant, 1790/1992, Libro primero, §7). En el juicio de gusto, se “hablará [...] de lo bello, como si la belleza fuese una cualidad del objeto, y el juicio fuese lógico (como si constituyera este un conocimiento del mismo, a través de conceptos del objeto)” (Kant, 1790/1992, Libro primero, §6).

Hablamos en tal caso como si [*als ob*] el juicio fuera lógico y basado en conceptos. Se trata sin embargo de otro tipo de necesidad:

la necesidad que es concebida en un juicio estético, sólo podrá llamársela ejemplar, es decir, [tratase de] la necesidad del asentimiento de todos a un juicio que es considerado como ejemplo de una regla universal que no puede ser aducida [*die man nicht angeben kann*]. (Kant, 1790/1992, Libro primero, §18)

No opera aquí el juicio determinante, sino el juicio reflexionante, función de la facultad de juzgar (de pensar lo particular bajo lo universal), en la que, dado lo particular, se trata de hallar lo general. El juicio de gusto es

\* New York University / París, Francia

\*\* Universidad Nacional de Córdoba, Córdoba, Argentina  
luis.urtubey@gmail.com



una suerte de caso límite, allí se ejerce el juicio *meramente* reflexionante: nuestro sentimiento reclama su propia conformidad, de manera que se exige que los otros tengan una respuesta como la nuestra. Podemos hablar de esta *necesidad ejemplar* como de un tipo de validez intersubjetiva que Kant introduce como un punto medio entre la objetividad y la subjetividad.

En este trabajo, realizaremos una consideración formal del juicio reflexionante kantiano, en especial de la normatividad del juicio de gusto, desde la perspectiva de la lógica stit, siguiendo una adaptación al caso del modelo que presentan Marra y Klein (2015). Las semánticas stit (Belnap, Perloff, & Xu, 2001; Horty, 2001, 2019) se apoyan, por un lado, en la idea de indeterminación temporal (apertura hacia el futuro) y, por otro, en la noción de “actuar” o acción libre [*agency*]. La lógica stit hace uso de sistemas modales de tiempo ramificado (de tipo *branching-time* o *branching space-times*) y la modalidad stit, por “*seeing to it that*”, que da cuenta de las acciones de los agentes como elecciones sobre historias, que se ramifican temporalmente. El operador modal stit tiene por objeto representar el “actuar” de un individuo al “hacer o provocar que” (en inglés, *bringing it about* o *seeing to it that*) se obtenga algún estado de cosas.

En nuestro análisis, conforme al juicio reflexionante de Kant, un agente requiere que las acciones de los otros se ajusten a un cierto principio o regla general de la cual resulta que esta experiencia particular sea tal. Dado un agente **a** y un principio normativo **N**, nuestra propuesta es obtener la formalización o modelización dentro de la lógica stit, en términos de la reformulación siguiente: “**a** requiere que la experiencia de **b** se atenga a **N** o siga el objetivo dado por **N**”.

## El caso del juicio reflexionante

La facultad de juzgar es nuestra capacidad de pensar lo particular bajo lo universal. Se distinguen dos operaciones. En el juicio determinante se subsume un particular bajo una regla, principio o ley ya dados (por ejemplo, conceptos del entendimiento o categorías). En el juicio reflexionante sólo está dado lo particular, y la operación consiste en hallar lo general.

La facultad de juzgar opera bajo el principio transcendental según el cual la naturaleza está organizada en conformidad con nuestras facultades (cf. Kant, 1790/1992, Introducción, V), principio que sostiene que, para todo lo que existe naturalmente, pueden encontrarse conceptos empíricamente determinados.

En el caso del juicio de gusto, como “Este hombre es bello”, el juicio es *meramente reflexionante*, nuestro sentimiento reclama su propia conformidad, de manera que se exige que los otros tengan una respuesta como la nuestra. Hallamos un universal para el particular, tomando nuestra respuesta frente al objeto evaluado como universalmente válida. Exigimos a otros que acuerden con nuestra respuesta afectiva porque ésta depende de las mismas capacidades que nos permiten formular juicios lógicos.

En el juicio lógico, como “Este hombre es argentino”, la facultad de juzgar tiene un rol clasificatorio: subsumimos el particular bajo el universal al considerar en el objeto las propiedades que tiene en común con otros objetos. En cambio:

En todos los juicios a través de los que declaramos a algo bello, no permitimos que nadie sea de otra opinión; y sin embargo, no fundamentamos nuestro juicio en conceptos, sino sólo nuestro sentimiento, que ponemos por fundamento, pues, no como un sentimiento privado, sino como uno común. (Kant, 1790/1992, Libro primero, §22)

En el caso del juicio de gusto, la condición de necesidad a la que aspira es la idea de un sentido común, o *sensus communis*:

sólo bajo el supuesto de que haya un sentido común (por tal no entendemos, empero, un sentido externo, sino el efecto [que proviene] del libre juego de nuestras fuerzas cognoscitivas), sólo bajo la suposición, digo, de un tal sentido común, puede ser emitido el juicio de gusto. (Kant, 1790/1992, Libro primero, §20)

En la tradición aristotélica, el *sensus communis* era la facultad mental que reconoce que las representaciones provenientes de los sentidos proceden de un mismo objeto. En la tradición cartesiana en la que se inscribe Kant, tal reconocimiento supone el concepto de un objeto, concepto que se sitúa en el mallado de las categorías de la razón pura según Kant.<sup>1</sup> Dice Kant:

Por *sensus communis* hay que entender [...] la idea de un sentido común a todos, esto es, de una facultad de juzgar que en su reflexión tiene en cuenta, en pensamiento (a priori), el modo representacional de cada uno de los demás, para atener su juicio, por así decirlo, a la entera razón humana y huir así a la ilusión que, nacida de condiciones subjetivas privadas

---

<sup>1</sup> Cf. una nota al pie editorial en la “Analítica de lo bello” en (Kant, 1790/2000, p. 372).

que pueden fácilmente ser tenidas por objetivas, tendría una desventajosa influencia sobre el juicio. (Kant, 1790/1992, Libro primero, §40)

Por lo tanto, visto que el juicio de gusto no se basa en conceptos sino sobre la base de nuestro sentimiento, reclama su propia conformidad, bajo el amparo del *sensus communis* que permite imputarle a otros la exigencia de nuestra propia respuesta afectiva; un cambio de opinión no se da simplemente a partir de elementos discursivos: no puede demostrarse (que algo es bello), como puede demostrarse un teorema.

El juicio de gusto se diferencia del juicio lógico en que el último subsume una representación bajo conceptos del objeto y el primero no la subsume bajo concepto alguno, pues de otra manera la aprobación universal necesaria podría ser forzada mediante pruebas. (Kant, 1790/1992, Libro primero, §35)

En lo que sigue, lo que procuramos es ubicar el juicio reflexionante de Kant, entendido como acción de juzgar, en el marco de la lógica de la acción. Debemos aclarar que no se trata de una interpretación de la obra kantiana, sino más bien de una aplicación inspirada en sus ideas. Entenderemos que el juicio reflexionante —desde la perspectiva asumida— constituye una “acción de juzgar” de los agentes en una comunidad con ciertas características especiales (*sensus communis*). El juicio *meramente* reflexivo culmina en las acciones (doxásticas) de un agente, que se ajustan a ciertas *consideraciones normativas* (que serán representadas por “N”), no completamente determinables, en las cuales se encuadra una *experiencia particular* (que viene a plasmar la necesidad ejemplar en la que reposa la universalidad del juicio de gusto).

## Elementos de la lógica stit

Marra y Klein (2015) elaboran un modelo que incorpora la estructura de árbol temporal de lógica stit como lógica temporal subyacente al modelo deóntico. Un árbol es en este contexto un conjunto finito  $W$  junto a un orden-para-el-árbol  $<$ . Si  $x$  se da antes que  $y$  en el árbol, es decir,  $x < y$ , se dice que  $x$  es un *predecesor* de  $y$ , siendo  $y$  un *sucesor* de  $x$ . La *raíz*  $r$  de un árbol  $\langle W, < \rangle$  es el elemento  $r$  que se halla debajo de todos los demás elementos, esto es,  $r < x$  para todo  $x \neq r \in W$ . Finalmente, una *historia* es una *rama* maximal de un árbol, esto es, una secuencia que consta de sucesores

inmediatos  $r < x_1, \dots, < x_n < l$ , en la cual  $r$  es la raíz del árbol y  $l$  es una hoja. Tenemos así la siguiente definición:

**Definición (Modelo de árbol).** Un modelo-árbol es un 4-tuplo  $T = \langle W, w_0, <, V \rangle$  en el cual:

- $W$  es un conjunto de estados o mundos con  $w_0 \in W$ ,
- $<$  es un orden-del-árbol sobre  $W$  cuya raíz es  $w_0$ ,
- $V: At \rightarrow \wp(W)$  es una valuación de las oraciones atómicas.

El modelo deóntico de Marra y Klein (2015) incorpora además normas, a partir de la extensión a un lenguaje normativo. Siendo más precisos, incorpora obligaciones que se presentan luego de aceptar una norma. Para un árbol dado  $T$  y cada norma posible  $\phi \in \mathcal{L}_{\text{norm}}$ , se puede identificar el sub-árbol  $O_\phi$  (denominado “conjunto de obligaciones”) de  $T$  de historias posibles compatibles con la satisfacción de  $\phi$ . El sub-árbol  $O_\phi$  es la representación formal de las acciones admisibles, esto es, aquellas con las cuales el agente se compromete si intenta seguir la norma.<sup>2</sup>

## El modelo NIA: normas, intenciones y acciones

Procuremos en esta sección ubicar entonces el juicio reflexionante kantiano, entendiéndolo como “acción de juzgar”, en el marco formal dado por la lógica de la acción.<sup>3</sup> Particularmente, entre las lógicas de la acción, consideraremos una formalización en el contexto dado por la lógica stit, que venimos de presentar brevemente. Para evitar malentendidos, debemos aclarar que no se trata de que pretendamos alcanzar una interpretación de la obra kantiana, sino más bien de una aplicación o de un desarrollo en el contexto de la lógica de la acción inspirada en sus ideas. Vamos a entender entonces que el juicio reflexionante —desde la perspectiva asumida— constituye una “acción de juzgar” de los agentes en una comunidad,

<sup>2</sup> Cf. Marra y Klein (2015, p. 272).

<sup>3</sup> Seguiremos en cuanto a esta incorporación de la lógica de los imperativos la teoría elaborada por Veltman (2011), a la que también apelan Marra y Klein (2015). Nos apoyamos aquí en buena medida en el trabajo que Marra y Klein (2015) presentan para un marco más estrictamente deóntico o normativo. Nuestro enfoque parte de un debilitamiento de ciertos aspectos de este marco normativo más estricto.

con ciertas características especiales. Asimismo, asumiremos que el juicio *meramente* reflexionante culmina o se halla al final de las “acciones doxásticas” de un agente,<sup>4</sup> que se ajustan a ciertas consideraciones normativas (que representaremos a través de **N**), no completamente determinables o especificables, en las cuales se encuadra una experiencia particular (considerada como *ejemplar*). De este modo, resulta plausible sostener, que para un agente **b** y asumiendo un *componente normativo N* con las características descritas con anterioridad, es posible obtener una formalización o modelización del proceso indicado en el marco de la lógica stit, que antes consideramos, en términos de un principio así formulado:

**Principio (Acciones doxásticas bajo el juicio meramente reflexionante).** Las acciones (doxásticas) de **b** que llevan a un juicio al tipo del juicio *meramente* reflexionante son acciones que se atienen a **N** o que siguen “líneas” compatibles con **N**.

Al considerar este principio o esta caracterización en el marco de la lógica stit, nos damos pronto con una limitación para su expresión formal en este contexto, que resulta del hecho de que stit sólo permite incorporar una perspectiva “externa” de la acción, dejando de lado el otro componente que entra en juego en este caso, a saber, el elemento intencional (interno) y dinámico. Está claro que a través de stit alcanzamos una caracterización de la acción por su resultado, o sea, lo que produce, pero el aspecto intencional, que queda fuera, parece ineludible en la consideración del juicio reflexionante que estamos planteando, dado que pretende incorporar el componente subjetivo, que viene impuesto por el carácter *ejemplar* de la experiencia de un sujeto. La alternativa que presentaremos en esta sección se apoya en la posibilidad de incorporar de un modo general este elemento subjetivo, el cual está dado, en términos más explícitos, por la sujeción de las acciones doxásticas al componente normativo **N**, mediante la incorporación de imperativos en el marco dado por stit. Seguiremos en cuanto a esta incorporación de la lógica de los imperativos la teoría elaborada por Veltman (2011), a la que también apelan Marra y Klein (2015).

---

<sup>4</sup>Hablamos de este tipo de acciones en tanto son conducentes a una forma de opinión y no estrictamente hablando a un conocimiento. Encontramos en McHugh (2013) una expresión de este concepto, que es referida del siguiente modo: “It seems that we do indeed exercise doxastic agency. We often, though by no means always, form, maintain and revise our beliefs . . . through conscious mental activity” (p. 134). Si todo marcha bien, agrega el autor, esto conduce al acto de juzgar.

Brevemente introduciremos algunos conceptos previos. El lenguaje normativo  $\mathcal{L}_{\text{norm}}$  que estaremos usando es el definido en Marra y Klein (2015):

**Definición (Lenguaje normativo  $\mathcal{L}_{\text{norm}}$ )**  $At = \{p_1 \dots p_n\}$  es un conjunto de proposiciones atómicas.

Así se define la sintaxis de las fórmulas del lenguaje:

**Definición (Sintaxis de  $\mathcal{L}_{\text{norm}}$ )**  $\Phi := p^{\exists} | p^{\forall} | \neg p^{\exists} | \neg p^{\forall} | \phi \wedge \phi | \phi \vee \phi$

En esta definición,  $p^{\exists}$  y  $\neg p^{\exists}$  refieren a las normas denominadas “*one-time goal norms*”, esto es, en las cuales se requiere que  $p$  (o respectivamente  $\neg p$ ) sea verdadera una vez; mientras que  $\neg p^{\forall}$  y  $\neg p^{\forall}$  refieren a las denominadas “*standing goal norms*”, o sea, que requieren que  $p$  (respectivamente  $\neg p$ ) sea siempre verdadera. Esta diferencia es importante, en particular cuando se trata de su cumplimiento: las normas del primer tipo expiran una vez que se hayan cumplido (por ejemplo: escribir este artículo), mientras que las del segundo tipo se mantienen siempre activas y no pueden descargarse (como sería el caso de: respetar a nuestros padres).

Siguiendo el trabajo de Marra y Klein (2015), introducimos ahora una versión simplificada de su modelo NIA, de normas, intenciones y acciones.

**Definición (modelo pre-NIA).** Adaptado de Marra y Klein (2015). Un modelo pre-NIA es un 6-tuplo  $M = \langle W, w_0, <, V, O_N, F \rangle$ , en el cual

- $\langle W, w_0, <, V \rangle$  es un modelo-árbol (como en stit),
- $O_N \subseteq W$  es el conjunto de obligaciones relativas a  $N$ ,
- $F : W \rightarrow \wp(\wp(At^{\{\exists, \forall\}} \times \{\text{verdadero, falso}\}))$  es la función “*planning*” que asigna un plan a cada mundo posible.

La interpretación que se pretende sería en los siguientes términos:  $w_0$  es el momento actual en el tiempo (la situación actual). El agente decide aceptar o rechazar una nueva norma y forma su lista (en inglés, su *to-do list*) en el momento  $t$ .<sup>5</sup> Las listas (*to-do lists*) y planes en todos los nodos

---

<sup>5</sup> Cf. Veltman (2011) para más detalles sobre estas nociones. Brevemente: “A to-do list is a set  $D \subseteq At^{\{\exists, \forall\}} \times \{\text{true, false}\}$ . A plan  $P$  is a set of consistent to-do lists such that for all  $D, D'$

subsecuentes (del árbol) permiten mantener un registro de los compromisos que ya se han cumplido y cuáles aún permanecen sin cumplir.

Asimismo, el modelo incorpora *obligaciones* que se desprenden de la aceptación del *componente normativo* asociado a cada situación. Para un árbol dado  $T$  y cada elemento  $\phi \in \mathcal{L}_{\text{norm}}$  se identifica un subárbol  $O_\phi$  de  $T$  (el *conjunto de obligaciones*) conformado por las historias posibles, que son compatibles con la satisfacción de  $\phi$ . El subárbol  $O_\phi$  es la representación formal de las acciones admisibles antes descriptas.

Resulta obvio, además, que se deberán establecer algunos requerimientos sobre la coherencia de los planes individuales y así mismo sobre la relación entre diferentes planes. Antes de expresarlos, conviene agregar alguna notación más (cf. Marra y Klein, 2015).

**Notación (Fuente)** Si  $v$  es el predecesor inmediato de  $w$  y  $D \in F(w)$ , entonces se dice que  $D' \in F(v)$  es una fuente de  $D$  si  $D \subseteq D'$  y todo ítem en  $D' - D$  es de la forma  $\langle p^3, \text{verdadero} \rangle$ , para  $w \in V(p)$  o  $\langle p^3, \text{falso} \rangle$ , para  $w \notin V(p)$ .

Entonces, podemos aproximar una idea de coherencia respecto de la representación de obligaciones.

Respecto a esto último, recordemos que un *plan*, denotado por “ $D$ ”, se conceptualiza como un conjunto de tuplos de la forma  $\langle \phi, \text{verdadero/falso} \rangle$ , con  $\phi \in A_i^{\{3, V\}}$ . Esta expresión formal sirve para expresar la idea de que cuando una norma está presente en una situación (un mundo  $w$ ) y es aceptada por un agente (esto es, resulta verdadera en este mundo  $w$ ), la norma en cuestión debe satisfacerse. El agente debería obrar en el futuro de acuerdo con esa norma, ya que la está aceptando. Ciertamente, bajo un supuesto de *racionalidad*, de una cierta racionalidad. La noción de *fuentes* introducida sobre la base del *plan* servirá en la siguiente definición (cláusulas 4 y 5), para regular los compromisos que puede manejar un agente a través de la ramificación temporal del modelo. Obsérvese, que sólo podrían quedar en el camino normas del tipo *one-time goal*.

**Definición (Coherencia de F).** Sea  $M$  un modelo NIA con la función *planing*  $F$ . Se dice que  $F$  es coherente sii:

---

$\in P$  it holds that  $D \not\subseteq D'$ . [...] A to-do list is the set of basic commitments the agent aims to jointly realize” (Marra & Klein 2015, p. 275). Intuitivamente, al considerar lo que un agente se compromete a realizar, una lista de este tipo denominado “to-do list”, sirve entonces para representar y “seguir el rastro” de lo que el agente intenta hacer.

1. *Éxito*: Si  $w$  es una hoja en NIA, entonces ningún  $D \in F(w)$  contiene compromisos de la forma " $\langle p^3, \text{verdadero/falso} \rangle$ ".
2. *No redundancia*: si  $w \in W$  con  $w \in V(p)$ , entonces ningún  $D \in F(w)$  contiene compromisos de la forma " $\langle p^3, \text{verdadero} \rangle$ ". De modo similar para  $w \notin V(p)$  y " $\langle p^3, \text{falso} \rangle$ ".
3. *Cumplimiento de las normas establecidas*: Si  $w \in W$ , con  $w \in V(p)$ , entonces ningún  $D \in F(w)$  contiene " $\langle p^y, \text{falso} \rangle$ ". De igual modo para  $w \notin V(p)$  y " $\langle p^y, \text{verdadero} \rangle$ ".
4. *Conservadurismo*: Sea  $v$  el predecesor inmediato de  $w$ , entonces para todo  $D \in F(w)$ , hay una fuente  $D'$  de  $D$  en  $F(v)$ .
5. *Libre elección*: Para todo  $D \in F(w)$  hay algún sucesor inmediato  $v$  de  $w$  y algún  $D' \in F(v)$  tal que  $D$  es una fuente de  $D'$ .

Explicemos un poco esta definición. La cláusula del *Éxito* expresa el hecho de que el árbol en el modelo  $M$  constituye el horizonte temporal del agente y cada norma del tipo *one-time norm* debe satisfacerse en este horizonte. La condición de *no-redundancia* establece que las normas del tipo *one-time* se descargan una vez que se satisfacen o se cumplen. La tercera condición, *cumplimiento de las normas establecidas*, indica que el plan  $F(w)$  no puede contener listas (*to-do lists*) incompatibles con  $w$ . El carácter *conservador* significa que no pueden introducirse nuevos compromisos en el camino. El agente puede, de este modo, tomar nuevos compromisos sólo en el estado inicial  $w_0$ . La última condición, *libre elección*, establece que todas las listas de  $F(w)$  son compatibles con algún futuro estado de cosas. Tomadas en su conjunto estas condiciones establecen la coherencia para los diferentes  $F(w)$ .

Definimos ahora un modelo NIA, según Marra y Klein (2015):

**Definición (Modelo-NIA).** Un modelo-NIA es un modelo pre-NIA en el cual la función *planning*  $F$  es coherente.

Descriptos de este modo los rasgos principales del modelo, indicaremos brevemente en la conclusión las líneas centrales de su aplicación al juicio reflexionante, que redondea nuestra propuesta. Dada la breve extensión del trabajo, podremos sólo esbozar esta aplicación en forma general.

## Conclusión

Iniciamos este trabajo señalando nuestra idea consistía en alcanzar una formalización o modelización, inspirada por el juicio reflexionante kantiano, dentro de la lógica *stuit*, en términos de una reformulación de su contenido en los siguientes términos: “**a** requiere que la experiencia de **b** se atenga a **N**”. Seguimos así —en términos generales— la interpretación normativa del juicio de gusto kantiano en cuanto se ejerce el juicio *meramente* reflexivo, aunque existen interpretaciones filosóficas alternativas, como la recientemente elaborada en Harbin (2020).

Desde una perspectiva normativa, el juicio de gusto es prescriptivo, tiene un carácter imperativo al enunciarse, ya que **a** procura que un agente **b** considere modificar su agenda, que sus preferencias se alineen con las del agente **a** que expresa el juicio, respecto del objeto de la evaluación, lo que puede entenderse como una situación dada. Esto sería así, pese a la indeterminación del contexto normativo **N**. Al enunciar el juicio de gusto, **a** pretende afectar el curso de acción de **b**, que éste se atenga a **N**, por ejemplo, a un principio de gusto no del todo determinable (pero garantizado por el *sensus communis*), a sabiendas de que se trata de una conformidad contingente. La necesidad ejemplar procura extender como una norma aquello que se presenta en un ejemplo singular. El carácter prescriptivo del juicio de gusto reside en la enunciación y no en la necesidad de la norma **N** en cuestión. El efecto imperativo del juicio se sostiene pese a la contingencia de lo que se supone como norma **N**. Entendemos entonces, que cuando un agente **a** expresa un juicio de esta naturaleza induce o dispara en su interlocutor, un agente **b**, u otros agentes en una comunidad, un tipo de *acción* que consideramos como *doxástica* y que produce cierto resultado. Se les exige a los interlocutores de **a** que adecuen sus propias agendas a la que experimenta **a**, dada la situación de la enunciación. Destacamos también que esta acción debe incluir o estar acompañada por un componente intencional: un objeto o un aspecto de la situación es lo que se evalúa, aquello acerca de lo cual se pretenden modificar las agendas de otros. El elemento normativo del juicio de gusto puede identificarse en el modelo en las obligaciones que resultan de la aceptación de una norma imprecisa, que se vehiculiza a través del juicio propiamente dicho, a través de la acción de juzgar. Es interesante hacer notar que el carácter impreciso de la norma en cuestión, que antes señalamos, no impediría la aparición de obligaciones más precisas, que incluso alcanzan un status

cuasi-conceptual que pretende regir el juicio de gusto (acaso el tabú acerca del canibalismo o la coprofagia sean obligaciones precisas pero contingentes, basales para el *sensus communis* pero no necesarias desde un punto de vista lógico). Esto puede destacarse como un rasgo favorable de la modelización.

Por estas razones, entendemos que la adaptación de un modelo del ámbito de la lógica deóntica, como el modelo NIA antes expuesto, proporciona una primera aproximación adecuada para la representación formal del proceso que dispara la expresión de un juicio de gusto. Los componentes centrales del juicio reflexionante, que hemos tenido en cuenta bajo nuestra conceptualización, son reconocibles también en los elementos del modelo propuesto. La modelización misma procura esclarecer principalmente —en estos términos— la noción indicada inicialmente como *la experiencia del agente b*, relativa a una situación, de la cual *a* requiere su adecuación a cierta normatividad cuya base es subjetiva. Si bien se deben realizar diversos ajustes y especificaciones al caso concreto del juicio reflexionante, entendemos que la aplicación de este enfoque es promisoria y alienta a su desarrollo en ulteriores trabajos. Asimismo, resultará interesante considerar su funcionamiento confrontado con otras interpretaciones filosóficas del juicio de gusto kantiano, como la antes desarrollada por Harbin (2020).

## Referencias

- Belnap, N., Perloff, M., & Xu, M. (2001). *Facing the future: Agents and choices in our indeterminist world*. Oxford: Oxford University Press.
- Harbin, R. K. (2020). Universality without normativity: Interpreting the demand of kantian judgements of taste. *Dialogue: Canadian Philosophical Review/Revue Canadienne de Philosophie*, 1-24. doi:10.1017/S0012217319000398
- Horty, J. (2001). *Agency and deontic logic*. Oxford: Oxford University Press.
- Horty, J. (2019). Epistemic oughts in stit semantics. *Ergo*, 6(4), 71-120.
- Kant, I. (1992). *Crítica de la facultad de juzgar* (P. Oyarzún, trad.). Caracas: Monte Ávila Editores. (Obra original de 1790)

- Kant, I. (2000). *Critique of the power of judgment*. (P. Guyer & E. Matthews, trads.) Cambridge: Cambridge University Press. (Obra original de 1790)
- Marra, A. & Klein, D. (2015). Logic and ethics: An integrated model for norms, intentions and actions. En W. van der Hoek, W. Holliday & W. Wang (Eds.), *Logic, rationality, and interaction: LORI 2015, Taipei, Taiwan, October 28–31, 2015, Proceedings* (pp. 268-281). *Lecture notes in computer science*, 9394. Berlin; Heidelberg: Springer.
- McHugh, C. (2013). Epistemic responsibility and doxastic agency. *Philosophical Issues*, 23, 132-157.
- Veltman, F. (2011). *Or else, what? Imperatives on the borderline of semantics and pragmatics* [Presentación en PDF]. Institute for Logic, Language and Computation, Amsterdam. <https://staff.fnwi.uva.nl/u.endriss/teaching/lolaco/2011/slides/veltman.pdf>



## Acerca de estados de estados de cosas verdaderos en otros mundos, pero que nunca serán verdaderos en este mundo

Manuel Dahlquist\*

### Algunos rasgos importantes de la lógica de Buridan

Jean Buridan (1300-1361) es probablemente el mejor lógico que dio la Edad Media. Fue el principal representante de la denominada Escuela de París (Klima, 2016; Read, 2012; Sundholm 2012). Buridan, como lógico nominalista –y los lógicos medievales en general– suscribieron la idea de que las proposiciones son objetos espacio-temporales. Es imprescindible para una correcta comprensión de los aspectos semánticos de la lógica de Buridan, tener en cuenta que las “proposiciones nacen y dejan de existir como cualquier otro objeto en el mundo” (Epstein, 1992, p. 153). De hecho, al menos desde Ockham se aceptaba comúnmente que el número de objetos del mundo está lejos de ser fijo, ya que los objetos comienzan a existir y dejan de existir, y las proposiciones son un tipo particular de objeto; un objeto capaz de representar otros objetos. Para Buridan, una proposición es una entidad lingüística específica, un *token*, i.e., una oración que se piensa, se pronuncia o se escribe en un momento y lugar específicos.<sup>1</sup> Como resultado, los objetos de la lógica medieval son menos “abstractos” que los de los sistemas de lógica contemporáneos. Esta concepción de las proposiciones genera que la lógica de Buridan se desarrolle en el marco de lo que ahora llamamos semántica basada en *tokens* (Klima 2004; Zupko 2018).

Otro rasgo importante de la lógica medieval al que no es ajeno Buridan es que la consecuencia lógica es el tema central de la teoría lógica. Hubo que esperar a las producciones de Tarski y Carnap para encontrar

---

<sup>1</sup> Los medievales consideraron tres tipos de lenguaje: el mental, el hablado y el escrito; todos ellos se plasman en un tiempo determinado, en un lugar determinado (cf. Dahlquist, 2018, secc. 5.2.).

\* Universidad Nacional del Litoral / Instituto de Humanidades y Ciencias Sociales del Litoral (CONICET) / Universidad Autónoma de Entre Ríos.  
manuel.dahlquist@gmail.com

una línea de trabajo comparable a los desarrollos del siglo XIV.<sup>2</sup> La lógica del siglo XIV comparte dos características con la lógica contemporánea:

- (1) Su tema principal es la noción de consecuencia;
- (2) La noción de consecuencia se define en términos semánticos, es decir, apelando al concepto de verdad, junto con algún concepto modal.

La caracterización general de la consecuencia en la lógica medieval es similar a lo que ahora llamamos el concepto *intuitivo* o *pre-teórico* o *común* de consecuencia lógica: hay una buena consecuencia si y solo si es imposible que la conclusión sea falsa cuando las premisas son verdaderas. Con matices, esta es la base para la caracterización de la consecuencia lógica en el siglo XIV (cf. Dutilh Novaes, 2008, p. 472). Veamos para dejarlo claro dos ejemplos:

Una inferencia sólida es aquella en la cual el opuesto del consecuente resulta repugnante al antecedente. (Paul of Venice, 1472/1984, p. 167)

En toda buena inferencia simple, el antecedente no puede ser verdad, sin que lo sea el consecuente. (Burley, trad. 2000, p. 3)

### **“Toda oración es afirmativa; por lo tanto, ninguna oración es negativa”**

Buridan planteó críticas a la definición tradicional de consecuencia, agregando el tiempo como un componente esencial para definirla correctamente<sup>3</sup>. Buridan nos da dos ejemplos de inferencias que no cumplen con los requisitos de la noción temporal de consecuencia, aunque deberían considerarse intuitivamente válidas. En otras palabras: después de agregar el componente temporal a la noción de consecuencia, todavía hay casos

---

<sup>2</sup> Ninguna escuela de lógica previa había ofrecido una oferta tan completa enfoque a la noción de consecuencia –aunque Aristóteles habla de implicación entre proposiciones–, ni ninguna otra escuela había sido tan prolífica en el desarrollo de una explicación sistemática y rigurosa del tema.

<sup>3</sup> Dicho toscamente: Buridan plantea la necesidad de que antecedente y consecuente sean evaluados en el mismo tiempo. “Todos los hombres son mortales; por lo tanto, los argentinos son mortales” puede ser una consecuencia inválida si el antecedente se evalúa cuando todavía hay hombres sobre la tierra y el consecuente en un momento de la historia donde los argentinos no existan.

más allá del alcance de la noción temporal de consecuencia lógica que, sin embargo, representan inferencias válidas.

El ejemplo que nos ocupa es uno de ellos y aparece en el capítulo 8 de su *Sophismata*, que es un capítulo dedicado a los *Insolubles*; dice: “Toda oración es afirmativa; por lo tanto, ninguna oración es negativa” (Buridan, trad. 2001, p. 952).

Esta inferencia no preserva la verdad de las premisas en la conclusión, ya que, si la premisa es verdadera, la conclusión es falsa, pues dice que no hay oraciones negativas, habiendo al menos una, a saber, la misma conclusión.

“No hay oraciones negativas”, es, como habrá notado el lector, auto-referencial. Buridan fue consciente de ello y nos dice que va a tratar con “proposiciones que son autorreferenciales [*de propositionibus habentibus reflexionem supra seipsas*] debido a la significación de sus términos” (trad. 2001, p. 952). Luego agrega que el capítulo “contiene proposiciones llamadas insolubles” (Buridan, trad. 2001, p. 952).

Deben hacerse algunas aclaraciones para comprender correctamente cómo estos temas –sofismas, insolubles y autorreferencia– interactuaron entre sí en los tiempos de Buridan. Primero: las oraciones autorreferenciales eran un subconjunto de los Insolubles, que son un subconjunto de los sofismas:

Como es evidente en el trabajo de Buridan, en el siglo XIV, los sofismas habían adquirido una gran importancia filosófica y su alcance se había restringido, en general, a una consideración de cuestiones lógicas y semánticas. Buridan ignora las imposibilidades, como ejercicios meramente formales, y las insolubles se retienen solo como sofismas de un tipo especial. (Scott, 1966, p. 5)

Segundo, lo que se llamó sofismo en la Edad Media no debe equipararse con ningún tipo de truco retórico o sofisma:

Los argumentos a favor y en contra de la misma oración reflejan desconcierto, pero el objetivo no era invitar al engaño. Por el contrario, las habilidades obtenidas residen en la desambiguación de las expresiones latinas, en las formulaciones exactas de sus condiciones de verdad y en el reconocimiento de las conexiones inferenciales entre ellos. En resumen, en lugar de sofistería, los sofismas llevaron al tratamiento del latín como un lenguaje preciso y lógico. (Yrjönsuuri & Coppock, 2016, p. 265)

Los sofismas designados simplemente como sofismas se caracterizaron más ampliamente como proposiciones que parecen seguir reglas bien establecidas, que de alguna manera son inaceptables o presentar problemas especiales. Se utilizaron principalmente, pero no exclusivamente, para la prueba y aplicación de reglas lógicas (Scott, 1966, p. 5).

En segundo lugar, la insolubilia generalmente se asocia con paradojas semánticas, particularmente la paradoja conocida como Mentiroso, pero no todas las insolubles son formas de Mentiroso (Scott, 1966, p. 50). “Debes saber que los sofismas llaman ‘insolubles’ a algunas cosas no porque no se puedan resolver de ninguna manera, sino porque se resuelven [solo] con dificultad” (Ockham, trad. 2002, p. 2). Los insolubles no solo abordan la paradoja del mentiroso (Yrjönsuuri, 2011, p. 550); en general, involucran oraciones que se refieren a sí mismas de una manera u otra (a veces en virtud del contexto en el que aparecen), pero su referencia no siempre está vinculada a afirmar la verdad (falsedad) de la misma. Cuando hablamos de autorreferencia en general, pensamos en oraciones que “dicen algo sobre sí mismas”, pero también en oraciones que reflejan el significado de sus propios términos. Entonces, de acuerdo con esto, “Esta oración es falsa” y “Ninguna oración es negativa” son autorreferenciales. Una diferencia importante entre “Esta oración es falsa” y “Ninguna oración es negativa” es que en la primera la autorreferencia está vinculada a lo falso (o verdadero) y esto le da a la oración “Esta proposición es falsa” un carácter semántico; en el segundo, la autorreferencia se basa en preguntas más bien sintácticas: “Ninguna proposición es negativa” es autorreferencial porque habla de oraciones que tienen la partícula “no”, y usa la partícula “no” para negar que haya oraciones que tienen la partícula “no”.

### **La solución de Buridan: un enfoque modal**

Ahora bien; como siempre hicieron los medievales, luego de presentar un insoluble, el autor presenta las razones a favor de considerarlo una buena inferencia y las razones en contra. Buridan hace esto y elige defender que “Toda proposición es afirmativa; por lo tanto, ninguna proposición es negativa” se trata de una inferencia válida. Claro que Buridan es consciente de que la inferencia no preserva la verdad (y este es el argumento más fuerte para considerarla una mala inferencia):

Nuevamente, una consecuencia no es válida si el antecedente puede ser verdadero sin la verdad del consecuente. Pero este es el caso aquí, ya que el antecedente puede ser verdadero y el consecuente no puede ser verdadero, está claro que el antecedente puede ser verdadero sin la verdad del consecuente. (Buridan, trad. 2001, p. 953)

En favor de la validez de la inferencia Buridan dará tres conclusiones: dirá primero que la inferencia se trata de un entimema y que, agregando la premisa (necesaria) que falta, “Ninguna afirmativa es negativa” se puede derivar “Ninguna oración es negativa” en el modo (válido) *Celarent*. Luego, y también a su favor, dirá que la inferencia cumple con la regla de contraposición, regla válida para todas las inferencias en el corpus medieval. Por último, dará la razón que más nos interesa en el contexto de este trabajo.

La tercera conclusión es que es posible alguna proposición que no puede ser cierta. Prueba: un posible antecedente nunca implica un consecuente imposible, como supongo en [el libro I de *Primeros analíticos*]. Pero “toda proposición es afirmativa”, lo cual es posible, implica “ninguna proposición es negativa”; por lo tanto, es posible y, sin embargo, no puede ser cierto. *Y, por lo tanto, está claro que no se dice que una proposición sea posible porque puede ser verdadera, ni imposible porque no puede ser verdadera, pero se dice que es posible porque las cosas pueden ser como significan [énfasis agregado]*, tomando estas palabras en el sentido correcto de acuerdo con la determinación del capítulo 2, e imposible porque las cosas no pueden ser así, etc. (Buridan, trad. 2001, pp. 954-955)

El énfasis en la cita es donde Buridan deja en claro la distinción conceptual entre posibles proposiciones y posible-verdadero. Que una proposición sea posible depende de si las cosas pueden ser como la proposición dice que son. Sin embargo, esto no es lo mismo que decir que la proposición es verdadera (o no puede ser verdadera). En otras palabras, una oración es posible cuando describe una posible situación, y posiblemente verdadera “cuando es cierta y en una posible situación” (Buridan, trad. 1985, p. 29). El ejemplo arquetípico de una posible oración es, por supuesto, “ninguna proposición es negativa”.

Una oración solo es posible-verdadera si puede ser parte de la posible situación que describe. El ejemplo estándar que Buridan usa para distinguir lo posible y lo posiblemente verdadero es la oración “Ninguna oración es negativa”: describe una posible situación, porque seguramente es posible

que todas las declaraciones hayan sido afirmativas, pero la oración en sí misma nunca puede ser parte de tal situación, ya que no es de forma afirmativa. (Scott, 1966, p. 29)

En otras palabras: “Ninguna propuesta es negativa” puede ser parte de una posible descripción del mundo (una en la que no hay proposiciones negativas, porque Dios las ha aniquilado) pero no puede, en sí misma, formar parte de ese mundo, ya que, en el momento preciso en que aparece, el mundo deja de tener solo oraciones afirmativas porque, ella misma, es una oración negativa.

### Posible, posible-verdadero y meramente posible

La lógica modal contemporánea (en sus aspectos semánticos) nace de las ideas de Leibniz y se estructura a mediados del siglo XX en los trabajos de Prior, Kanger, Hintikka y Kripke. Informalmente, la idea es la siguiente: Una proposición es *necesaria* si es verdadera en todos los mundos posibles, *posible* si es verdadera en al menos uno. La idea es que diferentes cosas pueden ser verdaderas en diferentes mundos posibles, pero aquello que sea verdadero en todos los mundos posibles es necesario, mientras que aquello que es verdadero en al menos un mundo posible, es posible (Chellas, 1980). En términos más cercanos a la visión de la semántica relacional, diríamos:

La fórmula  $\Box A$  es verdadera en un mundo  $w$  en un modelo si y solo si  $A$  es verdadera en cada mundo  $w^*$  accesible desde  $w$  en el modelo; de manera similar,  $\Diamond A$  es verdadera en  $w$  si y solo si  $A$  es verdadera en algún mundo  $w^*$  accesible desde  $w$ . (Williamson, 2013, p. 82)

Buridan, como vimos en la cita anterior, acuerda con esta idea y la presenta casi en los mismos términos. Pero entiende que “Ninguna oración es negativa” escapa de esta condición. Sin embargo es capaz de formar parte de la descripción de estados de cosas, razón que filósofos como Wittgenstein o Carnap juzgaron como suficiente para ser posibles. Para entender la inferencia “Toda oración es afirmativa; por lo tanto, ninguna oración es negativa” como válida es necesario, según Buridan, atender a la distinción entre posible y posible-verdadero. De ella se desprenden nuevas definiciones de posibilidad

Arthur Prior era un ávido lector de lógica medieval. En 1969 publicó un artículo –“The possible-true and the possible”– dedicado a la distinción de Buridan. Los objetivos son presentar la distinción con mayor precisión y mostrar que la semántica más compleja a la que nos lleva la distinción también debería generar una lógica modal más compleja. Prior presenta para presentar la ubicación de Buridan un escenario que, en lugar de mundos posibles, presenta hojas; en estas hojas aparecen las proposiciones que no son más que marcas en las hojas. Las hojas que forman un conjunto de hojas se designan con las letras  $X$ ,  $Y$ . En su presentación, Prior distinguirá entre un lenguaje de objetos  $L$  y un metalenguaje  $M$ . Esto es importante: las oraciones del lenguaje de objetos se usan para describir la apariencia física de ciertas marcas en hojas de papel, incluidas las de las oraciones latinas que están escritas en las hojas. El lenguaje objeto es el primero que consiste en una cierta clase limitada de oraciones latinas que presentaremos de inmediato. El metalenguaje  $M$  “es el inglés en el que se escriben las secciones de este artículo que siguen. Estoy preparado para que me digan, y de hecho creo, que no es un inglés estándar” (Prior, 1969, p. 483). Nuestra presentación de las ideas de Prior será esquemática y de acuerdo con las presentaciones contemporáneas, en aras de hacerlas breves, claras y útiles a nuestro argumento.

**LENGUAJE:** es un lenguaje de marcas o *tokens*, que consta de tres categorías disjuntas: términos, cuantificadores y cópulas. Es el lenguaje que se necesita para expresar las inferencias inmediatas típicas del famoso *Cuadrado de la oposición*, por lo que lo llamaremos LENGUAJE C.O.

**Lenguaje C.O.** = {Términos: {*propositio*; *affirmativa*; *negativa*}; Signos de cantidad: {*omnis*; *quaedam*; *nulla*}; Cópulas: {*est*; *non est*}

**Fórmula bien formada:** una línea de marcas se llama oración si y solo si consiste en un signo de cantidad seguido de un término seguido por la cópula *est* (o por *non est*, si el signo de cantidad es *quaedam*) seguido de un término. Así, las oraciones son líneas de marcas de estas cuatro formas

**FBF** = {*omnis A est B*; *quaedam A est B*; *nulla A est B*; *quaedam A non est B*}

Con los elementos anteriores ahora podemos dividir las oraciones en las que son verdaderas en sus hojas y las que no lo son:

Si las oraciones son verdaderas o no en sus hojas está determinado por las formas de las marcas en sus hojas. *En este sentido, las oraciones son “sobre”*

*las formas de las marcas* [énfasis añadido]. Cada término está asociado con un grupo particular de formas, que se puede decir que connota, aunque esto significa no más que eso, la presencia en una hoja de marcas de ciertas formas determinará, en formas que detallaremos en breve, si las oraciones que contienen ciertos términos deben contarse como “verdaderas en sus hojas”. (Prior, 1969, p. 483)

La cita anterior está de acuerdo con las características de una semántica de *tokens*, la semántica de Buridan como vimos arriba. Ahora se puede distinguir entre las oraciones posibles y las posibles-verdaderas; la oración “*Quaedam oratio est negativa*” es posible, ya que puede ser verdadera en su hoja, además de ser verdadera en otras hojas. Por el contrario, la oración “*Nulla oratio est negativa*” no puede ser bajo ninguna circunstancia verdadera en su propia hoja, ya que su propia existencia (su apariencia en la hoja) la hace falsa.

Prior nos da las condiciones de oraciones posibles y posibles-verdaderas, que transformaremos en definiciones. Para que las diferencias puedan contrastarse y no haya confusión, simbolizaré el *posible de Buridan* con  $\diamond$ , la definición de *posible de modelos relacionales contemporáneos* como  $\diamond$ , y *posible-verdadero* como  $\spadesuit$ .

**Definición posible- $\diamond$ :**  $(M, w) \models \diamond\varphi$  si hay algún  $w'$ , tal que,  $wRw'$  y  $(M, w') \models \varphi$

Esta es una definición contemporánea de posibilidad, expresada en términos de la semántica de mundos posibles.

**Definición Posible- $\diamond$ :**  $(M, w) \models \diamond\varphi$  sii (i)  $w$  es el mundo actual; y (ii) existe algún  $w'$  ( $w \neq w'$ ) y  $(M, w') \models \varphi$ .

Esta es la definición de posible de Buridan. Exige que  $\varphi$  pertenezca –sea verdadera– en el mundo  $w$  (que en  $w$  sea el caso de que  $\varphi$ ) y además que  $\varphi$  sea verdadera en otro mundo (estado)  $w'$  diferente a  $w$ .

**Definición Posible- $\spadesuit$ :**  $(M, w) \models \spadesuit\varphi$  si hay alguna  $w'$ , tal que en  $(M, w') \models \varphi$  y (i)  $w$  es el mundo actual; (ii) ( $w \neq w'$ ); (iii)  $(M, w) \not\models \spadesuit\varphi$

Esta es la definición de Posible-verdadero, que simbolizamos con el símbolo  $\spadesuit$ . Exige que  $\varphi$  sea la descripción de un mundo diferente al mundo actual, pero que, a su vez no sea verdadera en el mundo actual. Por cuestiones de espacio dejaré de lado las especulaciones sobre los conceptos

de posible indicados por los diamantes. El que indica la pica ♠, es el que nos interesa, pues queremos ponerlo en relación con el resto del aparato lógico para mostrar cómo puede incluirse en las discusiones modales –lógicas, metafísicas y ontológicas– contemporáneas.

### **Objetos imposibles (a manera de conclusión)**

Como vimos arriba, la semántica de *tokens* de Buridan asume lo dispone a asumir las oraciones como objetos. Entre estos objetos que son las oraciones hay un tipo de ellos –las posibles– que *nunca serán actuales*. Este descubrimiento (digo descubrimiento pues no se habla de ello antes del siglo XIV) es de suma actualidad y se inserta sin problemas en las discusiones entre actualistas y posibilistas, donde las distinciones modales son el alfa y el omega de las discusiones, como bien lo muestran trabajos como los de Plantinga (1974).

Buridan nos presenta un tipo de objeto que no sólo no conforma la realidad actual, como Superman, Santa Claus o el Yasí Yateré, sino que nos presenta un objeto que, a diferencia de los nombrados, jamás pertenecerá al mundo actual. No por sus condiciones ontológicas sino por sus condiciones lógicas, que a su vez, determinan su ontología. Esto los hace únicos y fascinantes. “No hay oraciones negativas” es un objeto que no puede existir con verdad en el mundo actual y la actualidad es el concepto básico de las discusiones acerca de objetos posibles y reales:

La posibilidad de un objeto debe entenderse junto con otra noción, a saber, la actualidad. Las dos nociones están relacionadas al menos de la siguiente manera: (i) cada objeto actual es un objeto posible. Una relación más controvertida entre las dos nociones se expresa de la siguiente manera: (ii) no todos objeto posible es un objeto actual, es decir, algunos objetos posibles son objetos no-actuales. (Yagisawa, 2020, secc. 1)

Del siglo XIV nos llega una voz poderosa que acerca al debate objetos-imposiblemente-actuales (por razones lógicas) a la discusión metafísica. Desarrollaremos esta posición en trabajos ulteriores.

## Referencias

- Buridan, J. (1966). *Sophisms on meaning and truth* (T. Scott, trad.). New York: Meredith Publishing Company.
- Buridan, J. (1982). *Chapter eight of Buridan's sophismata* (G. E. Hughes, trad.). Cambridge University Press.
- Buridan, J. (1985). *Jean Buridan's logic* (P. King, trad.). Dordrecht: Reidel Publishing Company.
- Buridan, J. (2001). *Summulae de dialectica* (G. Klima, trad.). New Haven; London: Yale University Press.
- Burley, W. (2000). *On the purity of the art of logic: The shorter and longer treatises* (P.V. Spade, trad.). New Haven, CT; London: Yale University Press.
- Chellas, B. (1980). *Modal logic: An introduction*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511621192
- Epstein, R. (1992). A theory of truth based on a medieval solution to the liar paradox. *History and Philosophy of Logic*, 13(2), 149-177.
- Dutilh Novaes, C. (2008). 14<sup>th</sup> century logic after Ockham. En D. Gabbay & J. Woods (Eds.) *Mediaeval and Renaissance logic* (pp. 433-504). *The handbook of the history of logic* (vol. 2). Amsterdam: Elsevier.
- Dahlquist, M. (2018). *Una introducción a la teoría lógica de la Edad Media*. College Publications.
- Klima, G. (2004). Consequences of a closed, token-based semantics: the case of John Buridan. *History and Philosophy of Logic*, 25(2), 95-110.
- Klima, G. (2016). Consequence. En C. Dutilh Novaes, & S. Read, (Eds.) *The Cambridge companion to medieval logic* (pp. 316-341). Cambridge: Cambridge University Press.
- Ockham, W. (2002). *Summa logicae*, III. 3.46 (P.V. Spade, trad.). En P. Boehner, G. Gál & S. Brown, (Eds.), *Guillelmi de Ockham Opera philosophica et theologica, OPh I*; St. Bonaventure, NY: The Franciscan Institute.

- Paul of Venice. (1984). *Logica parva*, (R. Perreiah, trad.). München; Wien: Philosophia Verlag. (Obra original de 1472)
- Plantinga, A. (1974). *The nature of necessity*. Oxford: Oxford University Press.
- Prior, A. (1969). The possibly-true and the possible. *Mind*, New Series, 78, 312.
- Read S. (2001). Self-reference and validity revisited. En Yrjönsuuri M. (Ed.), *Medieval formal logic: Obligations, insolubles and consequences* (pp. 183-196). *The new Synthese historical library* (vol. 49). Dordrecht: Springer.
- Read, S. (2012). The medieval theory of consequence. *Synthese*, 187, 899-912. doi:10.1007/s11229-011-9908-6
- Scott, T. (1966). Introduction. En J. Buridan, *Sophisms on meaning and truth* (T.Scott, trad). New York: Meredith Publishing Company.
- Sundholm, G. (2012). "Inference versus consequence" revisited: inference, consequence, conditional, implication. *Synthese*, 187, 943-956. doi:10.1007/s11229-011-9901-0
- Yagisawa, T. (2020). Possible objects. En E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2020 ed.). <https://plato.stanford.edu/archives/sum2020/entries/possible-objects>
- Yrjönsuuri, M. (2011). Insolubles. En H. Lagerlund (Ed.), *Encyclopedia of Medieval philosophy* (pp. 550-551). Dordrecht: Springer.
- Yrjönsuuri, M., & Coppock, E. (2016). Sophisms and insoluble. En C. Dutilh Novaes & S. Read (Eds.), *The Cambridge companion to medieval logic* (pp. 265-289). Cambridge: Cambridge University Press.
- Zupko, J. (2018). John Buridan. En E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2018 ed.). <https://plato.stanford.edu/archives/fall2018/entries/buridan>
- Williamson, T. (2013). *Modal logic as metaphysics*. Oxford University Press.



## Pluralismo lógico y normatividad

Alba Massolo\*

### Introducción

Las tesis pluralistas sobre la lógica sostienen que hay más de un sistema lógico correcto, bueno, verdadero o legítimo. A su vez, según la tesis que defiende el carácter normativo de la lógica, esta disciplina funciona como una guía prescriptiva para el razonamiento. Pero entonces, si hay más de una lógica correcta y al mismo tiempo la lógica constituye una guía normativa para el razonamiento, ¿cuál de todas esas lógicas deberíamos emplear como guía normativa para razonar? De esta manera, se pone de manifiesto que existe un conflicto entre el pluralismo lógico y la normatividad. Este conflicto se agudiza aún más cuando el pluralismo no sólo da lugar a la diversidad sino también a la rivalidad entre lógicas. En este caso, la lógica no sólo ofrecería más de un estándar normativo adecuado sobre cómo razonar, sino que estos estándares, además, se contradecirían entre sí. De este modo, esta aparente tensión lleva a cuestionar si es posible conciliar el pluralismo lógico con la normatividad. Este problema es denominado en la literatura filosófica actual el problema del colapso para el pluralismo lógico.

Partiendo de una concepción externalista sobre la normatividad de la lógica, el objetivo de este trabajo consiste en argumentar a favor de una conciliación entre el pluralismo lógico y la normatividad. Nuestra propuesta está basada en la metodología del equilibrio reflexivo de Nelson Goodman (1954/1983). Adoptando esta metodología, es posible sostener que las leyes de la lógica se establecen a partir de un proceso de equilibrio reflexivo entre las prácticas argumentativas sociales, por un lado, y los intentos de formular principios reguladores para esas prácticas, por otro. Como no existe una única manera de realizar este proceso de ajuste mutuo entre las leyes de la lógica y las prácticas inferenciales de los agentes, esto deriva en la existencia de una pluralidad de sistemas lógicos correctos que regulan diferentes prácticas inferenciales sin entrar en conflictos en-

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH).  
Córdoba, Argentina.  
albamassolo@unc.edu.ar

tre sí. De esta manera, nuestra hipótesis de trabajo sostiene que es posible plantear una conciliación entre el pluralismo lógico y el estatus normativo de la lógica si entendemos que la lógica es el resultado de un proceso de equilibrio reflexivo y caracterizamos la normatividad de la lógica desde una perspectiva externalista.

La estructura de este trabajo es la siguiente: en la segunda sección, ofrecemos una descripción del pluralismo lógico y del problema del colapso. En la tercera sección, establecemos una distinción entre dos perspectivas sobre el carácter normativo de la lógica, a saber, la perspectiva internalista y la perspectiva externalista. Argumentamos a favor de la perspectiva externalista sobre la normatividad. En la cuarta sección, presentamos una caracterización de la lógica a partir de la tesis del equilibrio reflexivo y proponemos una versión pluralista sobre la lógica. Finalmente, en la quinta sección mostramos que la perspectiva externalista sobre la normatividad es compatible con la versión de pluralismo lógico elaborada a partir de la tesis del equilibrio reflexivo. Además, mostramos de qué manera esta versión de pluralismo lógico resulta inmune al problema del colapso.

## El pluralismo lógico y el problema del colapso

El pluralismo es una tesis filosófica sobre la lógica que sostiene que existe más de una lógica correcta. Si bien es posible encontrar una férrea defensa de esta tesis en los escritos de Carnap (1934/1937) (Caret & Kouri Kissel, 2020), la versión de pluralismo lógico propuesta por Beall y Restall (2006) ha sido la gran responsable de avivar el debate actual en torno al pluralismo. La tesis de Beall y Restall se centra en la caracterización de la relación de consecuencia lógica a partir de la teoría de modelos. Estos autores sostienen que es posible realizar una caracterización general de la noción de consecuencia lógica a partir de la (por ellos) denominada tesis generalizada de Tarski (GTT):

(GTT): un argumento es válido<sub>x</sub> si y sólo si, en todo caso<sub>x</sub> donde las premisas son verdaderas, la conclusión también lo es.<sup>1</sup> (Beall & Restall, 2006, p. 29)

Para Beall y Restall, además, la noción de consecuencia lógica posee

---

<sup>1</sup> La traducción al español de esta cita, así como la de las otras dos citas que se presentan en este artículo, es mía.

tres rasgos centrales: es una relación necesaria, formal y normativa. De esta manera, el argumento de estos autores esgrime que al haber al menos dos maneras diferentes de instanciar caso  $x$ , ya sea como modelos tarskianos, situaciones o modelos de Kripke, y a su vez, cada una de estas instancias da lugar a relaciones de consecuencia que respetan los rasgos centrales de necesidad, formalidad y normatividad, se sigue entonces que hay más de una instancia admisible de la noción de consecuencia lógica. Además, al no haber motivos o razones para preferir una de esas instancias por sobre las otras, o para determinar cuál de esas instancias es la adecuada; debemos aceptar que existe más de un sistema lógico correcto.

Uno de los principales problemas que debe afrontar esta versión de pluralismo es el denominado problema del colapso, término que fue acuñado por Colin Caret (2017) para englobar una serie de críticas que advierten las dificultades que el carácter normativo de la lógica supone para el pluralismo (Priest, 2006; Read, 2006; Keefe, 2014; Barrio, Pailos, & Szmuc, 2018). Según el problema del colapso, la pluralidad de lógicas difícilmente se puede conciliar con el carácter normativo de la lógica. Esto es así puesto que, si la lógica fuese normativa para el razonamiento, entonces debería poder determinar, para cada inferencia, si una conclusión se sigue (o no se sigue) de un conjunto de premisas. Consideremos, por ejemplo, la inferencia que permite concluir  $B$  a partir del conjunto de premisas  $\{A, \neg A\}$ . Entre las diferentes relaciones de consecuencia lógica que se han aceptado como correctas, una de esas relaciones, la relación de consecuencia clásica  $\vdash_C$ , valida esa inferencia; mientras que existe otra relación de consecuencia diferente, la relación de consecuencia intuicionista  $\vdash_{int}$ , que no valida esa inferencia. No obstante, al ser la lógica una disciplina normativa se espera que aporte una única respuesta acerca de la validez de la inferencia en cuestión. De esta manera, según el planteo de Caret (2017), el pluralismo lógico colapsa en un monismo. En el caso de nuestro ejemplo, tenemos dos relaciones de consecuencia lógica: una que hace válida esa inferencia y otra que la hace inválida. ¿Qué se debería hacer entonces frente a esta inferencia? ¿Deberíamos aceptar  $B$ ? ¿O deberíamos rechazar  $B$ ? Una de las respuestas que se ha dado (Read, 2006), es que deberíamos aceptar  $B$ . Esto es así porque, por un lado,  $\vdash_C$  asegura que  $B$  es verdadera pero, por otro lado,  $\vdash_{int}$  no asegura que  $B$  sea falsa. Por esto,  $\vdash_C$  da una respuesta sobre  $B$  que  $\vdash_{int}$  no da. Así, frente a cualquier conflicto de este tipo, siempre se va a seguir como guía normativa a aquel sistema lógico que ofrezca una respuesta con respecto a aceptar o rechazar

una conclusión. De esta manera, si se acepta que la lógica es una disciplina normativa, el pluralismo de Beall y Restall parece estar destinado a colapsar en un monismo.

¿Es acaso el problema del colapso una debilidad inherente al pluralismo lógico? La tesis pluralista de Beall y Restall parece sucumbir inevitablemente a este problema. Se ha señalado que la mayor dificultad para esta tesis pluralista reside en el hecho de otorgar a la relación de consecuencia lógica un carácter neutral con respecto al tema (en inglés, *topic-neutral*) (Kouri Kissel & Shapiro, 2020). En este sentido, esta tesis sostiene que si un argumento es válido en al menos un sistema lógico correcto, entonces cualquier argumento que tenga la misma forma, independientemente de su contenido, también será válido. Esta parece ser la principal dificultad del pluralismo lógico de Beall y Restall para ser consistente con el carácter normativo de la lógica. Más allá de esta versión pluralista en particular, se han señalado algunas características que comparten las tesis pluralistas que colapsan ante la normatividad, a saber, (i) sostienen que hay más de una relación de consecuencia lógica correcta para un mismo lenguaje; (ii) la consecuencia lógica que definen tiene alcance global; (iii) aceptan la existencia de relaciones de consecuencia rivales y (iv) sostienen que la relación de consecuencia lógica es normativa (Stei, 2020). La inconsistencia se genera a partir de aceptar, por un lado, la aplicabilidad universal de la lógica y, por otro lado, la normatividad de la lógica para el razonamiento. De esta manera, el problema del colapso resulta inherente a las tesis pluralistas que defienden el carácter neutral de la lógica y su universalidad.

Se han formulado dos propuestas alternativas para enfrentar este problema. Por un lado, rechazar la tesis de la normatividad de la lógica para el razonamiento y, por otro lado, proponer una caracterización de pluralismo lógico que no sea neutral con respecto al tema ni pretenda tener un alcance universal. En la primera alternativa, se ha argumentado que la lógica no constituye una guía normativa para el razonamiento (Blake-Turner & Russell, 2018; Russell, 2020). De esta manera, se rechaza la característica (iv) descrita en el párrafo anterior y la normatividad deja de ser un problema para el pluralismo lógico. En la segunda alternativa, se han planteado diferentes tesis pluralistas locales sobre la lógica que rechazan la característica (ii) del párrafo anterior (Shapiro, 2006, 2014; Cook, 2010; Caret, 2017; Kouri Kissel & Shapiro, 2020). Como se pondrá de manifiesto en las próximas secciones, nuestra propuesta está vinculada a esta segunda alternativa en tanto proponemos un pluralismo lógico de alcance local.

## ¿En qué sentido es la lógica una disciplina normativa?

Con respecto al estatus normativo de la lógica, uno de los núcleos centrales de esta discusión consiste en tratar de dilucidar qué significa afirmar que la lógica es una disciplina normativa. En líneas generales, podría hacerse una distinción entre dos maneras de entender la normatividad de la lógica. En primer lugar, como un fenómeno propio de la esfera privada, vinculado a los procesos individuales de pensamiento y razonamiento. Y, en segundo lugar, como un fenómeno propio de la esfera pública, vinculado a las prácticas discursivas públicas de argumentación (Skelac, 2017). Mientras que la primera perspectiva lleva a adoptar una concepción internalista acerca de la normatividad, la segunda perspectiva lleva a plantear una concepción externalista.

Por un lado, desde una perspectiva internalista, la discusión en torno a la normatividad de la lógica está centrada en determinar en qué sentido puede afirmarse que la lógica es normativa para el razonamiento humano. El razonamiento se caracteriza como una actividad individual. Así, la lógica es normativa en tanto ofrece una guía directiva de primera persona para razonar (Steinberger, 2019). Por otro lado, desde una perspectiva externalista, se ha sostenido que la lógica tiene principalmente un rol normativo en las prácticas dialógicas de interacción racional entre agentes. Los procesos de razonamiento individuales serían el resultado de un proceso de internalización de esas prácticas sociales (Dutilh-Novaes, 2015). Desde esta perspectiva externalista, la discusión está centrada en determinar en qué sentido puede afirmarse que la lógica es normativa para las prácticas dialógicas de intercambios racionales que involucran discusiones críticas e intentos de salvar diferencias de opinión. En este sentido, el rol normativo de la lógica consistiría en fijar estándares o criterios para participar en esos intercambios argumentativos (Steinberger, 2019).

A partir de esta diferenciación, es posible elaborar un argumento a favor de la perspectiva externalista sobre la normatividad. El carácter normativo de la lógica reside fundamentalmente en ofrecer estándares objetivos para regular los intercambios argumentativos que tienen lugar en los diferentes contextos dialógicos de interacción humana. Estos estándares objetivos permiten regular estos intercambios a partir de establecer prácticas inferenciales que pueden ser aceptadas y prácticas inferenciales que deben ser rechazadas. A fin de completar nuestro argumento, ofrecemos

en la siguiente sección una caracterización de la naturaleza de la lógica que, como argumentaremos en la quinta sección de este artículo, es compatible con esta perspectiva externalista sobre la normatividad.

### **Los sistemas lógicos como el resultado de un proceso de equilibrio reflexivo**

La idea de equilibrio reflexivo fue introducida por Nelson Goodman (1954/1983) en el contexto de la discusión acerca de la justificación de las reglas de inferencia deductivas e inductivas (Brun, 2017). En la tesis del equilibrio reflexivo hay dos ideas centrales, a saber, ajuste y acuerdo:

Los principios de la inferencia deductiva están justificados por su conformidad con una práctica deductiva aceptada. Su validez depende de acuerdos con inferencias particulares que realmente hacemos y aprobamos. (Goodman, 1954/1983, p. 63)

Si bien este proceso puede parecer circular, Goodman aclara que se trata de un círculo virtuoso:

El punto es que las reglas y las inferencias particulares están justificadas al estar en un acuerdo entre sí. Una regla se cambia si lleva a una inferencia que no estamos dispuestos a aceptar y una inferencia se rechaza si viola una regla que no estamos dispuestos a cambiar. El proceso de justificación es el proceso delicado de hacer un ajuste mutuo entre reglas e inferencias aceptadas. En este acuerdo alcanzado subyace la única justificación necesaria para ambas. (Goodman, 1954/1983, p. 64)

Posteriormente, la tesis del equilibrio reflexivo fue utilizada en la filosofía de la lógica en el marco de la justificación del carácter normativo de la lógica (Resnik, 1985) y en el marco de la justificación de los principios lógicos (Peregrin & Svoboda, 2017). A partir de establecer algunas diferenciaciones con estas dos propuestas,<sup>2</sup> nuestro planteo consiste en

---

<sup>2</sup> Por una cuestión de espacio no podemos explayarnos demasiado con respecto a las diferencias que mantenemos con respecto a las propuestas de Resnik (1985) y de Peregrin y Svoboda (2017). No obstante, no queremos dejar de señalar, aunque sea brevemente, que nuestra propuesta, a diferencia de la Resnik no está basada en intuiciones, sino en prácticas argumentativas. Asimismo, mientras que Peregrin y Svoboda entienden que el proceso de equilibrio reflexivo permite explicitar los principios lógicos que preexisten bajo alguna forma de proto-lógica en las reglas constitutivas del lenguaje, nuestra propuesta no adhiere a la distinción entre reglas constitutivas y reglas regulativas ni tampoco acepta la existencia previa de una proto-lógica.

sostener que la lógica es el resultado de un proceso de equilibrio reflexivo entre, por un lado, las prácticas inferenciales sociales y, por otro lado, los principios o leyes que regulan esas prácticas. Este proceso de ajuste y acuerdo mutuo configura y transforma tanto las prácticas de los agentes como las reglas que se aceptan. Ninguno de estos dos extremos del proceso es independiente del otro. La lógica emerge en este proceso a partir de los intercambios argumentativos que resultan indispensables para el desarrollo de la vida en sociedad. Vale señalar que estos mecanismos de ajuste y acuerdo pueden realizarse de diferentes maneras y llevar así a resultados diversos. Además, pueden existir prácticas inferenciales específicas que requieran ajustes especiales. De esta manera, la diversidad de procesos de ajuste y acuerdo puede derivar en la configuración de diferentes sistemas de principios lógicos. Es así como se obtiene una pluralidad de lógicas.

El pluralismo lógico, por lo tanto, responde a dos orígenes que no son mutuamente excluyentes, por un lado, el proceso de equilibrio reflexivo no produce necesariamente resultados únicos y así diferentes procesos de equilibrio reflexivo pueden dar lugar a diferentes sistemas lógicos. Este es el caso, por ejemplo, de la práctica matemática. Mientras que algunas teorías matemáticas solo admiten demostraciones constructivas de sus teoremas, otras teorías matemáticas admiten demostraciones no constructivas. Esto trae como consecuencia la existencia de al menos dos sistemas de principios lógicos diferentes que regulan la práctica de la demostración matemática. Por otro lado, las prácticas argumentativas se desarrollan en diferentes contextos. Algunos de estos contextos poseen particularidades que influyen en los procesos de equilibrio reflexivo, dando lugar así a diferentes sistemas lógicos que responden a la regulación de cada uno de estos contextos específicos de argumentación. Este es el caso, por ejemplo, de las particularidades inherentes a la argumentación en matemática, en ciencias, en el ámbito jurídico-legal, en el debate cívico-político, entre otros contextos que podrían mencionarse.

### **Una manera de conciliar el pluralismo lógico con el estatus normativo de la lógica**

Según la perspectiva externalista sobre la normatividad defendida en la tercera sección, la lógica es normativa en tanto ofrece reglas que permiten regular los intercambios argumentativos que tienen lugar en los diferentes contextos de interacción social. En este sentido, el rol normativo de

la lógica consiste en ofrecer estándares objetivos que permiten regular esas prácticas inferenciales realizadas entre múltiples agentes. Asimismo, la caracterización de la lógica a partir de la tesis del equilibrio reflexivo sostiene que la lógica es el resultado de un proceso de acuerdo y ajuste mutuo entre las prácticas argumentativas sociales, por un lado, y las leyes o principios que regulan esas prácticas, por otro. En este proceso de equilibrio reflexivo se produce un intercambio de ida y vuelta en el cuál tanto las prácticas argumentativas como las leyes reguladoras sufren diferentes modificaciones. Podría afirmarse, entonces, que la fuerza normativa que las leyes de la lógica poseen sobre las prácticas argumentativas reside en el hecho de que esas leyes se obtuvieron a partir de un proceso de equilibrio reflexivo en la búsqueda de una regulación para esas prácticas. Durante el desarrollo de este proceso, las leyes permitieron que se pudieran configurar esas prácticas inferenciales y, a su vez, las prácticas inferenciales permitieron la configuración de esas leyes.

Esta caracterización de la lógica a partir de la tesis del equilibrio reflexivo permite defender una versión pluralista sobre la lógica a partir de sostener que, por un lado, el proceso de equilibrio reflexivo no produce resultados únicos. Así, este proceso puede dar lugar a diferentes sistemas de principios lógicos que regulan diferentes tipos de prácticas inferenciales. Además, por otro lado, este proceso tiene lugar en diferentes contextos donde se desarrollan estas prácticas argumentativas. La particularidad de algunos de estos contextos puede imponer requerimientos especiales tanto a las prácticas como a las leyes, dando lugar así a resultados de los procesos de equilibrio reflexivo que son exclusivos para cada contexto específico de argumentación. Sobre este punto, creemos necesario enfatizar que esta versión de pluralismo lógico se elabora sobre la base de prácticas argumentativas. En una línea similar, Caret (2019) ha señalado algunas dificultades de defender tesis pluralistas sobre la base de intuiciones y, a partir de esta crítica, ha propuesto un pluralismo lógico basado en prácticas. Si bien este autor elabora un argumento a favor del pluralismo desde la evidencia aportada por la práctica matemática, nuestra propuesta no se limita a la práctica matemática. Por el contrario, la variante de pluralismo lógico que intentamos defender en este artículo toma como punto de partida todas las prácticas argumentativas sociales que son, o intentan ser, reguladas por algún sistema de principios lógicos.

¿Cómo puede disolverse, entonces, el problema del colapso para esta variante de pluralismo lógico? Como mencionamos anteriormente,

la fuerza normativa de la lógica reside en su capacidad para regular las diferentes prácticas argumentativas sociales. La lógica es normativa en tanto ofrece reglas que permiten diferenciar, dentro de esas prácticas, acciones admisibles de acciones inadmisibles. No obstante, estas reglas no son guías directivas individuales. Se trata de reglas a partir de las cuales se configuraron esas prácticas inferenciales y cuyo fin reside en permitir algunos intercambios y rechazar otros para el conjunto total de agentes que participan en dichas prácticas. Cada tipo de práctica argumentativa posee su propio sistema de principios reguladores. La fuerza normativa de estos principios está limitada al tipo de práctica, o contexto argumentativo, que pretende regular. Así, existen diferentes guías normativas para diferentes prácticas argumentativas. De esta manera, cada práctica, o cada contexto específico de argumentación, establece sus propias prescripciones normativas. En caso de existir un desacuerdo o una rivalidad en el interior de una práctica inferencial específica, este conflicto tenderá a resolverse en el proceso de ajuste y acuerdo mutuo que se establece entre las interacciones inferenciales admitidas y las reglas aceptadas. Por lo tanto, diferentes sistemas lógicos con diferentes prescripciones normativas pueden convivir con la existencia de diferentes prácticas inferenciales y diferentes contextos específicos de argumentación evitando caer en el problema del colapso para el pluralismo lógico.

## Referencias

- Barrio, E. A., Pailos, F., & Szmuc, D. (2018). Substructural logics, pluralism and collapse. *Synthese*. <https://doi.org/10.1007/s11229-018-01963-3>
- Beal, J. C. & Restall, G. (2006). *Logical pluralism*. Oxford: Oxford University Press.
- Blake-Turner, C., & Russell, G. (2018). Logical pluralism without the normativity. *Synthese*. <https://doi.org/10.1007/s11229-018-01939-3>
- Brun, G. (2017). Conceptual re-engineering: from explication to reflective equilibrium. *Synthese* 197, 925–954. <https://doi.org/10.1007/s11229-017-1596-4>

- Caret, C. (2017). The collapse of logical pluralism has been greatly exaggerated. *Erkenntnis*, 82, 739-760.
- Caret, C. (2019). Why logical pluralism? *Synthese*. <https://doi.org/10.1007/s11229-019-02132-w>
- Caret, C., & Kouri Kissel, T. (2020). Pluralistic perspectives on logic: An introduction. *Synthese*. <https://doi.org/10.1007/s11229-019-02525-x>
- Carnap, R. (1937). *The logical syntax of language* (A. Smeaton, trad.). New York: Harcourt, Brace and Company. (Obra original de 1934)
- Cook, R. (2010). Let a thousand flowers bloom: A tour of logical pluralism. *Philosophy Compass*, 5, 492-504.
- Dutilh-Novaes, C. (2015). A dialogical, multi-agent account of the normativity of logic. *Dialectica*, 69, 587-609.
- Goodman, N. (1983). *Fact, fiction and forecast* (2.<sup>a</sup> edición). Cambridge, MA: Harvard University Press. (Obra original de 1954)
- Keefe, R. (2014). What logical pluralism cannot be. *Synthese*, 191, 1375-1390. <https://doi.org/10.1007/s11229-013-0333-x>
- Kouri Kissel, T. (2018). Logical pluralism from a pragmatic perspective. *Australasian Journal of Philosophy*, 96, 578-591.
- Kouri Kissel, T., & Shapiro, S. (2020). Logical pluralism and normativity. *Inquiry*, 63(3-4), 389-410. <https://doi.org/10.1080/0020174X.2017.1357495>
- Peregrin, J., & Svaboda, V. (2017). *Reflective equilibrium and the principles of logical analysis*. New York: Routledge.
- Priest, G. (2006). *Doubt truth to be a liar*. Oxford: Oxford University Press.
- Read, S. (2006). Monism: The one true logic. En D. Devidi & T. Kenyon (Eds.), *A logical approach to philosophy* (pp. 193-209). Dordrecht: Springer.
- Resnik, M. (1985). Logic: normative or descriptive? The ethics of belief or a branch of Psychology? *Philosophy of Science*, 52, 221-238.

- Russell, G. (2020). Logic isn't normative. *Inquiry*, 63(3-4), 371-388. <https://doi.org/10.1080/0020174X.2017.1372305>
- Shapiro, S. (2006). *Vagueness in context*. Oxford: Oxford University Press.
- Shapiro, S. (2014). *Varieties of logics*. Oxford: Oxford University Press.
- Skelac, I. (2017). What we talk about when we talk about logic as normative for reasoning. *Philosophies*, 2, 1-8.
- Stei, E. (2020). Rivalry, normativity, and the collapse of logical pluralism. *Inquiry*, 63(3-4), 411-432. <https://doi.org/10.1080/0020174X.2017.1327370>
- Steinberger, F. (2019). Consequence and normative guidance. *Philosophy and Phenomenological Research*, 98, 306-328.



# Aplicaciones de minería de argumentos y modelos de esquemas de argumentos

Pablo Torres\*

Diego Letzen\*

## 1. Introducción

En el ámbito del estudio de la argumentación existe la idea –rastreada hasta los orígenes de la lógica en los *Tópicos* de Aristóteles– de que es posible tipificar la variedad de casos de argumentos que pueden encontrarse en el uso cotidiano de los mismos o en dominios más específicos (como el científico), mediante el uso de esquemas de argumentos. La posibilidad de tipificar los argumentos supone tres cosas:

- que es posible encontrar rasgos identificables en argumentos particulares que permiten reconocer una estructura común;
- que mediante ese procedimiento es posible cubrir una gran cantidad de los casos de argumentos que es posible observar en el uso argumentativo dentro del lenguaje natural;
- que la identificación de esos esquemas se apoya en conceptos que permiten eventualmente construir una teoría sobre los tipos de argumentos y por extensión sobre la argumentación en general y consiguientemente determinar enfoques metodológicos para el reconocimiento de los casos que caen bajo la descripción.

Expuesto de esta forma, la construcción y sistematización de estos esquemas puede tener como objetivo ulterior la postulación de una teoría argumentativa. En este trabajo comparamos dos posturas alternativas sobre los esquemas argumentativos con respecto a su aplicación a un mismo corpus en lenguaje natural y atendiendo a las posibilidades de su aplicación computacional en minería de argumentos.

\* Universidad Nacional de Córdoba (UNC), Facultad de Filosofía y Humanidades (FFyH) Centro de Investigaciones María Saleme de Burnichon (CIFYH). Córdoba, Argentina. dletzen@gmail.com

La motivación de este trabajo puede entenderse en estos términos: ¿cuáles son los criterios sobre los que sentar una interpretación acerca de lo que hace mejor a un esquema argumentativo? Este problema está en el fundamento de cualquier teoría argumentativa que apele en algún punto a una tipificación y construcción esquemática de lo que es un buen argumento. El problema particular que tratamos, es el de los criterios de carácter teórico y filosófico que nos permita elegir entre dos esquemas contrapuestos de análisis, más precisamente el esquema de Walton (Walton, Reed, & Macagno, 2008) y la tabla de Wagemans (2016).

A fines del trabajo, estas posturas serán interpretadas aquí como teorías sobre los esquemas argumentativos. Si bien podemos obtener varias conclusiones a partir de estos resultados, logramos establecer específicamente que:

- hay un alto grado de coincidencia empírica de base en la tipificación entre ambas teorías lo que hace suponer un apoyo de base a la teoría más general de tipos de argumentos
- desde el punto de vista metodológico, de la aplicación computacional y humana, parece más adecuado reforzar las teorías con elementos propios de la teoría argumentativa y la filosofía de la lógica, dado que los estadísticos cuantitativos como el  $\kappa$  de cohen parecen ser insuficientes por sí solos para poder decidir entre un enfoque y otro.

El trabajo se divide en tres secciones: una primera sección donde presentamos el marco de la discusión sobre esquemas de argumentos; una segunda sección donde desarrollamos de manera resumida las características principales de los esquemas de argumentos planteados –el de Walton y el de Wagemans– y abrimos el debate sobre los criterios con los cuales juzgar cuál es mejor que el otro y finalmente una sección con algunas conclusiones que podemos extraer acerca de este debate.

## **2. Modelos de clasificación de tipos de argumentos**

En las últimas décadas varias clasificaciones han sido propuestas especialmente desde la filosofía para tipificar y ordenar según diversos criterios las regularidades observadas en argumentos. Una de estas teóricas que ha sentado al base de todas las discusiones actuales es la propuesta de Perelman y Olbrechts-Tyteca de 1969 (que se basa en dos categorías generales

orientadas a encontrar asociaciones y disociaciones entre conceptos y definir subclases sobre estas categorías, lo que da lugar a una tipología básica de 4 grandes tipos de argumentos).

Los esquemas de argumentación representan formas de argumento que se usan ampliamente en la vida cotidiana. argumentación conversacional, y en otros contextos como el jurídico y el científico argumentación. Pero en su mayor parte, estos argumentos no están modelados adecuadamente por formas deductivas de razonamiento del tipo familiar en la lógica clásica o como inferencias estadísticas basadas en los conceptos bayesianos estándares de probabilidad. Ellos representan la estructura premisa-conclusión de un argumento, y son derrotables. Las condiciones de derrotabilidad se muestran como un conjunto de preguntas críticas (podrían entenderse como “instrumentos dialécticos”) para ayudar a comenzar el procedimiento de probar la fuerza y la aceptabilidad de un argumento pesando otros argumentos a favor y en contra. Una definición más concisa la da el propio Walton para quien:

. . . los esquemas de argumentos pueden describirse como estructuras abstractas que representan los tipos de argumentos más genéricos, que constituyen los bloques de construcción en aquellos que aparecen en el razonamiento cotidiano. Los esquemas de argumentación son instrumentos de argumentación que involucran la actividad de evaluar críticamente un punto de vista y las razones dadas en su apoyo. (Walton, Reed, & Macagno, 2008, p. 64)

El desafío fundamental que una teoría de los esquemas de argumentación es que necesita encontrar un sistema de clasificación útil y sólido. Por eso la importancia de los esquemas, tal y como habíamos adelantado: deben ser utilizables, fácilmente identificables y, al mismo tiempo, deben permitir que el usuario detecte el patrón de argumento más específico que puede ajustarse al texto o que puede emplearse para producir un argumento adecuado a las circunstancias y el propósito.

Dicho esto, es importante agregar exponer brevemente el otro costado del asunto: los esquemas de argumentos que vamos a presentar en este trabajo se postulan como herramientas de investigación dentro de un campo aún más amplio, el de la minería de argumentos. Vale la pena que expliquemos aquí brevemente de qué se trata dicha disciplina, para poder entender la naturaleza de los experimentos con corpora de argumentos, por un lado y los alcances e importancia de los esquemas de argumentos dentro del campo, por otro.

La minería de argumentos es el área de investigación cuyo objetivo es extraer argumentos del lenguaje natural y sus relaciones textuales, con el objetivo final de proporcionar datos estructurados procesables mediante programas de computadora, de acuerdo a modelos computacionales de argumentos. Involucra varias áreas de investigación del campo mayor de la Inteligencia Artificial (AI), como el procesamiento de lenguajes naturales (NLP), la representación del conocimiento y el razonamiento (KRR o KR) y las investigaciones en el campo de la interacción humano-computadora (HCI). Según Cabrio y Vilata (2018), la interacción entre estos campos puede entenderse de la siguiente forma:

NLP proporciona, entre otras cosas, los métodos para procesar el texto en lenguaje natural, para identificar los argumentos y sus componentes (básicamente premisas y conclusiones) y para predecir las relaciones entre los argumentos. KRR o KR contribuye a la ‘minería’ con las capacidades de razonamiento sobre los argumentos y relaciones recuperadas de modo que, por ejemplo, las falacias e inconsistencias puedan llegar a ser identificadas automáticamente en dichos textos. La HCI guía el diseño de buenas herramientas digitales de apoyo basadas en argumentos. (pp. 5248-5249)

Para lograr que esta actividad de “minería” automatizada funcione de manera confiable es necesario contar con estándares aceptables de análisis de argumentos que permitan guiar el trabajo de extracción y clasificación, además de poder evaluar la adecuación de los resultados con tales parámetros. Se podría decir que estos estándares deberían buscar correspondencia en la evaluación y análisis que un agente con cierta pericia podría realizar de manera artesanal sobre el mismo corpus de argumentos.

Teniendo en cuenta el contexto de producción e investigación en minería de argumentos, debemos entender las particularidades de los esquemas que presentaremos en la siguiente sección. Si bien es nuestra visión que los mismos pueden sentar las bases de una teoría argumentativa, no son propiamente una teoría acerca de los argumentos: son tipificaciones pensadas como herramientas para la extracción automática de argumentos desde el texto escrito, principalmente digital. Si bien el tema de la minería de argumentos es hartamente interesante de por sí, al igual que el impacto de las afirmaciones en él desplegadas, la idea aquí es abstraerlos de su contexto y pensar en el uso de los mismos más allá de las aplicaciones computacionales, es decir, considerar qué impacto tienen en una gente humano.

### **3. Los esquemas de Walton y Wagemans**

En esta sección haremos una breve presentación de los aspectos esenciales de los esquemas de Walton y Wagemans. La reconstrucción de estas propuestas es breve y sintética, y aquel que quiera adentrarse en el tema puede que las encuentre un tanto insatisfactorias. Pero en aras de la brevedad elegimos resaltar aquellos elementos relevantes para la presente discusión.

Ambas teorías serán reconstruidas a partir de los experimentos de anotación de argumentos desarrollados alrededor del corpus de texto del debate presidencial estadounidense del año 2016. Aquí utilizamos la versión del corpus referida en el estudio de Jacky Visser et al. (2020). Este se obtuvo a través de comentarios en Twitter y las publicaciones dentro de subreddits creado con esta finalidad (similar a un foro) abierto a todo el público.<sup>1</sup> El US2016 ha sido anotado sobre la base de la teoría de anclaje de inferencia (IATE). Esta se basa en ideas del análisis del discurso/conversación, la teoría del acto del habla, y estudios de argumentación, como una forma de explicar cómo es que el razonamiento proposicional involucrado en la argumentación está anclado en el discurso (ya sea escrito o hablado).

En cada experimento se busca determinar el grado de coincidencia entre dos personas que utilizan los esquemas para analizar los mismos argumentos, más allá del azar. Se entrena a los evaluadores o “anotadores” para que puedan reconocer y clasificar cada argumento dentro de los límites de cada teoría. Luego se elige como métrica para medir el resultado determinar el coeficiente  $\kappa$  de Cohen, estadístico que sirve para medir el grado de acuerdo entre dos personas o individuos dada las mismas condiciones teniendo en cuenta el factor azaroso (es decir, considerando que podrían estar de acuerdo por mera casualidad, sin mediar ningún razonamiento). Este coeficiente es que mide el éxito relativo o el fracaso de los experimentos en función de dicha medida de acuerdo o coincidencia parcial.

#### **3.1 El esquema de Walton**

En la comunidad computacional, ha sido el enfoque de Walton de esquemas de argumentos, el que contó tradicionalmente con una mayor po-

---

<sup>1</sup> Reddit es una plataforma de discusión en línea (<http://www.reddit.com>) que cuenta con entre 10 y 18 millones de usuarios por mes. La comunidad de usuarios está organizada en áreas de interés, llamadas “subreddits”, dedicadas a una gran variedad de temas.

pularidad. Este enfoque puede encontrarse sistematizado en el libro de Walton, Reed y Macagno del año 2008. La clasificación de Walton comprende una gran variedad de esquemas, descritos en detalle, pero con la flexibilidad para permitir ajustes con el fin de adaptar un esquema a un dominio de aplicación específica. Se basa en una taxonomía de tipos y subtipos, que recoge de manera principal denominaciones tradicionales en el campo de la filosofía. Uno de los problemas de este enfoque es organizar y manejar un número bastante alto de patrones o esquemas de argumentos, que según la clasificación pueden ser más de 60, reduciendo los elementos pragmáticos o eruditos para su diferenciación y priorizando la diferenciación entre relaciones de tipo semántico y ontológico, frente a aquellos componentes más estructurales (lógicos) que puedan ser sistematizados para un análisis, evaluación y producción de manera regular y estandarizada.

Es por esto que frente a los modelos de Walton se han venido proponiendo otras opciones de análisis de esquemas de argumentos que permitan superar estas limitaciones, o al menos aporten un enfoque alternativo para la organización de las estructuras consideradas. Nos interesa destacar aquí una en particular: la anotación con la tabla periódica de argumentos de Wagemans.

### 3.1.1 Características de los esquemas de Walton

Las características de los modelos de Walton pueden variar ligeramente de formulación en función del corpus o tipos de argumentos o trabajos al que se lo esté aplicando. Aquí vamos a exponer las características del esquema utilizado en el análisis del corpus del debate presidencial en EE. UU. en 2016 que recibe el nombre de US2016G1tv y es un subcorpus del corpus US2016 (Visser et al., 2020). Luego de haber procesado los argumentos y armado el corpus, dos personas, los evaluadores de este experimento, se encargaron de clasificar a cada uno de estos argumentos en función de la tabla de Walton. Luego de haber procesado los argumentos y armado el corpus, dos personas, los evaluadores de este experimento, se encargaron de clasificar a cada uno de estos argumentos en función de la tabla de Walton. A grandes rasgos, los argumentos fueron analizados asumiendo los siguientes conceptos:

- **Locuciones:** el texto original está segmentado en locuciones. Las locuciones se modelan como L-nodos, un subtipo de I-nodo.

- **Transiciones:** las relaciones funcionales del discurso se representan como transiciones que conectan las locuciones segmentadas. Las transiciones, o TA-nodos, son un tipo de S-nodo que conecta los L-nodos.
- **Conexiones ilocucionarias:** la intención comunicativa encapsulada en una locución es anotado por medio de conexiones ilocucionarias. En términos de AIF, las conexiones ilocucionarias son YA-nodos, un subtipo de S-nodo.
- **Proposiciones:** la mayoría de las conexiones ilocucionarias conducen a la reconstrucción del contenido proposicional de la locución asociada. Las proposiciones se modelan como I-nodos.
- **Inferencia, conflicto y reformulación:** generalmente conectando una proposición a otra, las relaciones argumentativas de inferencia, conflicto y reformulación indican respectivamente defensa justificatoria, incompatibilidad refutatoria y reformulación revisionista. Las relaciones proposicionales se modelan como subtipos de S-nodos: como nodos RA, CA y MA.

Los dos evaluadores o “anotadores expertos” capacitados en análisis de argumentación y con conocimiento previo de la tipología de esquemas de argumentos de Walton clasificaron cada uno el 55% de los RA-nodos en el cuerpo de US2016G1tv de acuerdo con la tipología de Walton. Para facilitar el proceso, los anotadores hicieron uso de un árbol de decisión de clasificación.

No se utilizaron todos los argumentos del corpus, apenas se consideró una muestra del 10.2% del corpus. El experimento fue considerado con un éxito relativo, dado que el resultado en un  $\kappa$  de Cohen de 0.61 lo que a priori determinaba un acuerdo sustancial entre los anotadores (la métrica esperada para interpretar el resultado como positivo era un valor de  $\kappa$  mayor a 0.6).

### 3.2 La tabla de Wagemans

La tabla periódica de argumentos es una clasificación de argumento propuesta por Wagemans como una alternativa teóricamente sólida y prácticamente útil para la multitud tradicional de descripciones incompletas, informales y a veces incluso tipos de argumentos inconsistentes.

El marco de referencia de la tabla consta de tres características dis-

tintivas de argumentos, cuya superposición produce una tipología de argumentos factorial que puede ser utilizado con el propósito de analizar, evaluar y generar argumentos del lenguaje natural:

- La primera distinción es entre argumentos de primer orden y argumentos de segundo orden. El enfoque supone que las premisas y conclusiones de los argumentos pueden reconstruirse en términos de proposiciones categóricas que consisten en un término sujeto (a, b, etc.) y un predicado término (X, Y, etc.).
- La segunda distinción es aquella entre argumentos predicados y argumentos sujetos. Si el tema de la proposición expresada en la premisa es idéntico al de la conclusión, el mecanismo subyacente del argumento se basa en una relación entre los predicados (diferentes). Si el predicado de la proposición expresado en la premisa es idéntico a la de la conclusión, el mecanismo subyacente del argumento se basa en una relación entre los (diferentes) sujetos.
- Finalmente, los argumentos se caracterizan sobre la base de la combinación específica de distintos tipos de proposiciones que se instan- cian dentro de ellos a partir del uso de la tabla.

Estas tres clasificaciones se combinan en una caracterización completa del tipo de argumento. Los prefijos 1 y 2 indican argumentos de primer orden y de segundo orden. Los infijos pre y sub indican argumentos predicados y argumentos de sujeto. Finalmente, combinaciones de P, V y F como sufijo distinguen las diversas combinaciones de proposiciones de política, valor y hecho, respectivamente

### 3.2.1 Características de la tabla de Wagemans

De manera análoga al esquema de Walton, el esquema de Wagemans funciona diseccionando los elementos de los argumentos:

- **Argumentos de primer orden y de segundo orden:** un RA-nodo se clasifica como de primer orden si conecta dos I-nodos cada uno con un par sujeto-predicado. Se clasifica un RA-nodo de segundo orden si su premisa es un L-nodo (es decir, una locución, a menudo como resultado de un informe), o si la premisa está aplicando un predicado a la proposición completa en la conclusión.

- **Predicados y argumentos de sujeto:** un RA-nodo se clasifica como argumento predicado si los I-nodos involucrados comparten el mismo término sujeto al cual se aplican diferentes predicados, y como argumento sujeto si sucede de manera viceversa.

- **Proposiciones de hecho, de valor y políticas:** un I-nodo se clasifica como una proposición de hecho si puede verificarse mediante observación empírica, como una propuesta de valor si contiene alguna evaluación (ya sea ética, estética, legal o lógica), y como una propuesta de política si expresa un acto o política a realizar.

En cuanto a los resultados, las pautas de anotación se validan mediante el cálculo del inter-anotador en acuerdo para las tres clasificaciones parciales, así como para los esquemas agregados finales. Para la clasificación de argumentos de primer orden y de segundo orden, una muestra aleatoria del 10.0% fue etiquetada por ambos anotadores, resultando en un  $\kappa$  de Cohen de 0.658. También en una muestra del 10.0%, la clasificación de argumentos predicado/sujeto da como resultado un  $\kappa$  de Cohen de 0.851.

La clasificación de los I-nodos como hecho / valor / política arroja un  $\kappa$  de Cohen de 0.778 en una muestra del 13.4%. El acuerdo entre anotadores para el esquema de argumento agregado basado en una muestra del 10.4% da como resultado un  $\kappa$  de Cohen de 0.689. Esto significa que las anotaciones parciales y finales caen dentro del rango de sustancial de casi perfecto acuerdo.

#### **4. La preferencia por un esquema**

En la sección anterior expusimos de forma sintética y resumida cómo dos esquemas argumentativos que plantean diferentes tipificaciones de los argumentos entran en acción directa con corpora extraídos del lenguaje natural. Sin embargo, y sin entrar en detalle en torno a los tecnicismos propios de cada teoría, uno bien podría preguntarse cuáles son, desde un punto de vista filosófico, las conveniencias de un modelo sobre otro. Claramente apelar a una función estadística que mide el error es una respuesta insatisfactoria (de hecho, las críticas epistemológicas y teóricas a la utilización de dicho estadístico y aquello que dice capturar lo hacen objeto de polémica).

Pero aceptemos sin más los resultados de cada uno de estos experimentos, y digamos que ambos son positivos, en tanto que superan la me-

dia esperada de 0.6. ¿Esta es realmente una buena razón para adscribir a una teoría por sobre otra? Pareciera ser que no: los números por sí no son resultados realmente significativos hasta que los interpretemos de alguna forma. Es decir, asumir que un modelo o teoría es mejor que otra porque tiene implica en un experimento una mejora de un 0.05 dentro de un coeficiente no pareciera decir nada por sí mismo.

Nuestro interés por estos modelos radica justamente en los resultados que arrojan al ser utilizados por seres humanos. No tanto el grado de acuerdo que pueden tener entre uno y otro –esa es una preocupación del científico de datos. La idea es que el sistema que clasifique sea fácil de usar y entender, pero que además permita automatizar la mayor parte de tareas de clasificación.

Creemos que un punto central en la comparación entre esquemas es su comprensibilidad. ¿Puede el esquema A clasificar una mayor cantidad de argumentos que el esquema B? ¿Qué nos permite captar uno y otro?

Según Visser et al. (2018) cada modelo presenta particularidades ausentes en el otro. En el caso del modelo de Walton “la mayoría de las inferencias predeterminadas en US2016G1tvWALTON son también clasificadas como tal en US2016G1tvWAGEMANS, pero no viceversa” (p. 322). Las inferencias predeterminadas o *default inferences* son aquellos argumentos que no encajan en ninguna tipología del esquema. Este parece ser un punto en el que Walton aventaja a Wagemans: al considerar argumentos con más de una premisa, pareciera ser que abarca una mayor cantidad de casos. ¿Puede ser esta una razón para preferir un esquema de tipo Walton?

A pesar de lo anterior, el esquema de Wagemans presenta una ventaja en relación al del Walton. Proporciona el valor adicional a las anotaciones parciales. “La clasificación de los tipos de proposiciones, por ejemplo, tiene un valor intrínseco, de las 798 proposiciones en el corpus, la mayoría de 376 se clasifica como valor, seguido de 298 proposiciones de hecho y 108 clasificaciones como política” (Visser et al., 2018, p. 323). Este esquema permite captar muchas instancias argumentativas que no entran en el esquema clásico de premisas y conclusión y que, sin embargo, uno admite claramente como inferencia o argumento en determinadas situaciones.

Pero, ¿podríamos tal vez considerar algún factor adicional además de la comprensibilidad y la precisión en arreglo al acuerdo? Creemos que hay una razón ulterior para preferir una sobre otra. Como explicamos en el comienzo, todos estos esquemas están planteados dentro del campo más

amplio de la minería de información y NLP. La creación de estos corpora (y la digitalización y procesamiento por parte de modelos computacionales) no debería ser el único fin a tener en cuenta. O mejor dicho: no debe ser dada por hecho. Dijimos que este corpus se crea con arreglo de la IAT, pero en cuanto sobrepasa el campo de la argumentación, es terreno legítimo de los lógicos y los teóricos de la argumentación. El problema de dar con un esquema válido de argumentación para procesar un corpus hereda todos los problemas filosóficos de cualquier tipo de esquema de argumentación.

En este sentido, el resultado de este corpus no será mejor que lo que nuestra mejor teoría pueda predecir o, dicho de otro modo, no mostrará nada nuevo bajo el sol. Es difícil especular aquí con lo que nos podremos encontrar luego de que los modelos comiencen ser entrenados con corpora de datos cada vez más y más grandes, pero es de esperar que si estos problemas no se resuelven o al menos plantean explícitamente en la construcción de las bases de datos, los modelos no aportarán respuestas sino que sumarán interrogantes al asunto.

Se vuelve necesario traer la discusión nuevamente al ámbito práctico y replantearnos *qué es lo que hacen los anotadores cuando reconocen un argumento*. Claramente algo en la teoría, algo en los esquemas funciona de manera tal que genera un consenso entre estos anotadores, lo que es sin duda prometedor. Pero no podemos saltar sin más a decir que ambas teorías son igual de buenas sólo porque un estadístico lo plantea nos permite cuantificar algo tan abstracto como el “acuerdo en una clasificación”. No, al menos en el ámbito de la argumentación, donde problemas como el seguimiento de una regla lógica o la incapacidad para detectar ciertas estructuras argumentativas podrían dar lugar a la existencia de sesgos en el procesamiento de las muestras.

## 5. Conclusiones

En nuestro trabajo analizamos estas dos teorías que se presentan como alternativas para procesar argumentos dentro de corpora escritos de lenguaje natural. La opinión general al respecto de estos dos métodos es que ambos son aceptables en similar medida a la hora de abordar la esquematización de argumentos para su procesamiento. Esta opinión, como vimos, se basa principalmente en el acuerdo entre los anotadores en ambos experimentos: dada que la misma es comparable y sustancial (en un  $\kappa$  de

Cohen de entre 0.65 y 0.75) nos instan a asumir que ambos enfoques pueden ser hasta equivalentes y son igual de buenos a la hora de caracterizar argumentos.

En esta revisión consideramos ponderar por una parte las razones teóricas sólidas, más allá de los resultados empíricos, para tener un enfoque por encima de otro, pero, por otra parte, estudiar la carga teórica detrás de cada uno de los modelos y la incidencia que estos elementos puedan tener en el análisis empírico de argumentos. Llegamos a la conclusión de que es muy difícil sostener una paridad entre las teorías solo en base al acuerdo entre anotadores, que es lo que mide el  $\kappa$  de Cohen. Sin duda es prometedora este dato y dice mucho al respecto de la receptividad en los anotadores y la claridad del esquema. Pero si la idea subyacente de los experimentadores es llegar en algún momento a dirimir la discusión solo a partir de este tipo de herramientas estadísticas, entonces no vemos una salida clara.

La ponderación teórica de un esquema por encima de otro no podrá ser hecha pura y exclusivamente sobre la base de estadísticos y experiencias empíricas. Serán necesarias otras categorías, propias de la argumentación y la lógica, inclusive la epistemología de la lógica, que entren en juego y permitan sentar una base de análisis robusta, que garantice criterios que vayan más allá de las aplicaciones computacionales, y echen raíz en el ámbito del lenguaje cotidiano.

## Referencias

- Cabrio, E., & Villata, S. (2018). Five Years of argument mining: a data-driven analysis. En *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence* (pp. 5427-5433). Recuperado de: <https://www.ijcai.org/Proceedings/2018/0766.pdf>
- Visser, J., Lawrence, J., Wagemans, J., & Reed, C. (2018). Revisiting computational models of argument schemes: Classification, annotation, comparison. En S. Modgil, K. Budzynska, J. Lawrence, & K. Budzynska (Eds.), *Computational models of argument: Proceedings of COMMA 2018* (pp. 313-324). *Frontiers in artificial intelligence and applications* (vol. 305). IOS Press. <https://doi.org/10.3233/978-1-61499-906-5-313>
- Perelman, C., & Olbrechts-Tyteca, L. (1969). *La nouvelle rhétorique. Traité de l'argumentation*. Paris: Presses Universitaires de France. (Obra original de 1958)

- Visser, J., Konat, B., Duthie, R., Koszowy, M., Budzynska, K., & Reed, C. (2020). Argumentation in the 2016 US presidential elections: Annotated corpora of television debates and social media reactions. *Lang Resources & Evaluation* 54, 123–154. Recuperado de: <http://corpora.aifdb.org/US2016G1tv>
- Wagemans, J. H. M. (2016). Constructing a priodic table of arguments. En *Argumentation, objectivity, and bias. Proceedings of the 11th International conference of the Ontario society for the study of argumentation (OSSA)* (pp. 1-12). Windsor, ON. Recuperado de: <http://scholar.uwindsor.ca/ossaarchive/OSSA11/papersandcommentaries/106>
- Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. Cambridge: Cambridge University Press.





ISBN 978-950-33-1597-2



**ciffyh**

Centro de Investigaciones  
María Salmerón de Burnichon  
Facultad de Filosofía y Humanidades UNC

..  
Área de

**Publicaciones**

**ffyh**

Facultad de Filosofía  
y Humanidades UNC



Universidad  
Nacional  
de Córdoba