

Reducción y emergencia en el contexto de la filosofía de la mente: fiscalismo y brecha explicativa

*Reduction and emergence in philosophy of mind:
physicalism and the explanatory gap*

J. Arturo Escobedo*; Miguel Ángel Sebastián**

*Universidad Nacional Autónoma de México, México
j.a.escobedo.06@gmail.com

**Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México, México
msebastian@gmail.com

Resumen

Nuestras experiencias conscientes parecen eludir cualquier explicación en términos científicos, presentándose como el caso paradigmático de fallo en la reducción y como principal motivación para postular relaciones de emergencia. Así, diversos autores han señalado que las propiedades fenoménicas han de ser *algo más que* meras colecciones de propiedades físicas lo que nos lleva a rechazar el fiscalismo. Como consecuencia, las propiedades fenoménicas o bien han de ser añadidas a la lista de propiedades fundamentales, o bien que emergen de estas. Alternativamente, el fiscalismo ha de poder explicar cómo ese fallo en la explicación puede ser compatible con su postura.

El objetivo de este artículo es clarificar los conceptos de reducción y emergencia en el contexto de este debate en filosofía de la mente, distinguiendo claramente diversos tipos de relaciones epistemológicas y sus consecuencias ontológicas. Esto nos permitirá entender mejor los compromisos de las distintas posiciones y los desafíos a los que deben hacer frente.

Palabras clave: explicación reductiva, implicación a priori, consciencia fenoménica, emergencia débil, emergencia fuerte.



Received: 03/01/2023. Final version: 08/01/2024

eISSN 0719-4242 – © 2024 Instituto de Filosofía, Universidad de Valparaíso

This article is distributed under the terms of the

Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 Internacional License



CC BY-NC-ND

Abstract

Our conscious experiences seem to elude any explanation in scientific terms. They present paradigmatic case of reduction failure and they are the main motivation for postulating relations of emergence. Thus, various authors have pointed out that phenomenal properties must be something more than mere collections of physical properties. This leads to the rejection of physicalism. As a consequence, phenomenal properties either have to be added to the list of fundamental properties, or emerge from them. Alternatively, the physicalist must be able to explain how that failure in the explanation can be compatible with their position.

The aim of this paper is to clarify the concepts of reduction and emergence in the context of this debate in philosophy of mind, clearly distinguishing different types of epistemological relationships and their ontological consequences. This will allow us to have a better understanding of the commitments and the challenges different positions face.

Keywords: reductive explanation, *a priori* entailment, phenomenal consciousness, weak emergence, strong emergence.

1. Introducción

Philip W. Anderson defiende, en su artículo “More is Different”, la idea de que los sistemas complejos no se pueden explicar solamente en términos de las propiedades de sus partes individuales. Como consecuencia sostiene que se requieren explicaciones que tomen en cuenta los sistemas tomados como un todo, más allá de las actividades de las partes individuales. Estas interacciones dan lugar a comportamientos nuevos y sorprendentes que no se pueden predecir a partir del estudio de las partes individuales. Es en este sentido que toma los sistemas complejos como siendo “más que la suma de sus partes”. El entendimiento de estas propiedades emergentes requeriría entonces un enfoque diferente al utilizado para estudiar las partes individuales de un sistema. Diversos autores (Bedau, 1997; 2008; Humphreys, 1997; 2015; 2016; Crutchfield *et al.*, 1986) han considerado que este concepto de ‘emergencia’ es importante en diversos campos de estudio como la física, la biología, la química, la sociología y otras disciplinas que estudian sistemas complejos. Sin embargo, es en el estudio de la mente y en particular de la consciencia donde la idea de “emergencia” encuentra su más clara aplicación.

Tradicionalmente, el concepto de *emergencia* se presenta por contraposición al de *reducción*, a menudo presentado bajo el eslogan “...nada más que...” [...*nothing more than*...]. Ambos conceptos buscan capturar las posibles relaciones entre diferentes entidades, como objetos, eventos o propiedades, así como entre teorías, explicaciones o conceptos. En la filosofía de la mente estos conceptos resultan especialmente relevantes para entender el lugar que ocupa nuestra experiencia consciente en la naturaleza y el desafío que parece presentar a la tesis del fisicalismo. A *grosso modo*, el fisicalismo es la tesis de que cuanto encontramos en nuestro

mundo no es *nada más que* entidades físicas; —las mesas, los coches, las instituciones, etc.— siendo la reducción una forma de entender la relación entre las entidades físicas y el resto de entidades que existen en el mundo.

Formalmente, el fisicalismo se compromete al menos con la *superveniencia metafísica* de cuanto hay en nuestro mundo en las propiedades físicas. Esta idea se puede entender de manera intuitiva, y abstrayendo de un formalismo más detallado, como la tesis de que cualquier mundo metafísicamente posible que resulte de simplemente copiar todas las propiedades físicas de nuestro mundo y nada más, resultaría ser un mundo idéntico al nuestro.¹ De esta manera, si en el mundo real hubiera entidades no físicas —fantasmas o almas, digamos— un mundo que tuviera las mismas propiedades físicas que nuestro mundo no sería idéntico al nuestro pues carecería de fantasmas y almas. El fisicalismo sería entonces falso. Dejando de lado creencias teológicas y fijándonos en el tipo de entidades ordinarias que encontramos, no parece haber buenas razones para pensar que haya entidades que no sobrevengan en propiedades físicas y en este sentido sean “algo más que” colecciones de partículas del tipo que la física nos invita a considerar. No obstante, las intuiciones parecen cambiar cuando nos adentramos al interior de la mente: ¿son nuestras experiencias conscientes, sensaciones, emociones y sentimientos meras entidades físicas? El objetivo de este artículo es explicar cómo se han entendido los conceptos de reducción y emergencia en las discusiones en filosofía de la mente. Esto, para entender mejor los compromisos de las distintas respuestas que se puedan dar a la pregunta anterior, y los desafíos a los que las diversas posiciones deben hacer frente.

Los conceptos de emergencia y reducción se pueden utilizar para capturar relaciones epistémicas, típicamente entre entidades representacionales como los conceptos y las teorías, u ontológicas. La discusión en filosofía de la mente permite ahondar y entender estas distinciones pues combinan ambas formas de entender los conceptos de reducción y emergencia. Para dejar estas distinciones claras, en la primera sección vamos a presentar el concepto de explicación reductiva y su relación con la implicación *a priori*, que suele ser utilizada en filosofía de la mente a la hora de evaluar si las propiedades fenoménicas se reducen a las propiedades físicas. En la segunda sección se expone el argumento del conocimiento que da soporte a la idea de que haya una brecha explicativa entre verdades físicas y verdades fenoménicas. Además, se distinguen el tipo de emergencia débil (meramente epistémica) que resulta de aceptar la intuición del argumento, de la emergencia, aún más débil, que surge de la complejidad, y se discute la relación entre emergencia débil y fuerte (ontológica). Finalmente, la tercera sección repasa brevemente los problemas que la fisicalista enfrenta para rechazar la consecuencia metafísica del argumento; esto es, para aceptar emergencia débil sin aceptar la fuerte.

¹ Ver Sebastián (2022a, Ch.2) para una discusión más detallada de la caracterización del fisicalismo en español.

2. El reduccionismo en el contexto de la filosofía de la mente: implicación *a priori* y el modelo funcional de reducción

Aunque resulta tentador entender la idea de la reducción en términos de identidad, de modo que A se reduce a B si $A = B$, el fisicalismo no se compromete a tal cosa. Probablemente por buenas razones. Uno puede pensar que la propiedad de ser una mesa es una propiedad que en ningún sentido es susceptible de requerir algo más allá de las propiedades físicas y dudar que haya alguna propiedad propiamente física que todas las mesas necesariamente compartan. Es por esto por lo que resulta más adecuado capturarlo mínimamente en términos de superveniencia —y pensar que ha de existir algún tipo de conexión *a priori* que permita comprender la entidad a reducir en términos de aquella a la que se reduce. Esta conexión es lo que se conoce como implicación *a priori*.

La implicación *a priori* se entiende como un condicional material que se justifica *a priori*: P implica *a priori* Q cuando se conoce *a priori* el condicional material $P \rightarrow Q$. Aquí P puede estar por algún enunciado que describe las propiedades físicas a nivel microscópico y Q puede estar por algún enunciado que describe las propiedades a nivel macroscópico. Si tomamos Q como un enunciado sobre algún suceso que involucre las propiedades macroscópicas del agua (como que el punto de ebullición del agua es de 100°C) y P como un enunciado sobre un suceso físico típicamente a nivel microscópico (sobre la presión y temperatura de las moléculas de H_2O), P implica *a priori* a Q cuando la justificación de ese condicional se da de forma independiente a la experiencia, desde la comodidad del sofá. La explicación reductiva de P a Q se suele entender como una explicación funcional en términos del rol causal que tienen las propiedades: P explica reductivamente a Q cuando P explica por qué Q tiene las propiedades causales que tiene. La presión y temperatura de las moléculas de H_2O explican por qué el agua tiene un punto de ebullición de 100°C .

La conexión entre la explicación reductiva y la implicación *a priori* ha sido defendida por autores como Levine (1993), Chalmers (1996) y Chalmers y Jackson (2001), quienes afirman que una reducción exitosa debe ser explicativa en un sentido completo. Que la explicación sea completa simplemente quiere decir que la explicación debe proveer de información suficiente para volver inteligible el fenómeno a explicar a partir de aquello que lo explica. La exigencia de una implicación *a priori* en este tipo de reducción se debe a la asunción de que existe una conexión íntima entre el análisis conceptual y la explicación reductiva: si no hay implicación *a priori* entonces no hay explicación reductiva. La exigencia de la implicación *a priori* tiene como objetivo establecer una conexión necesaria entre la entidad a reducir y la entidad que reduce de manera que no capturemos meras correlaciones que serían compatibles con posiciones no fisicalistas donde, a pesar de que el tipo de entidades a explicar y aquellas que las explican correlacionan con el mundo actual, tendríamos mundos metafísicamente posibles en que existe la entidad física pero no la mental.

Chalmers y Jackson (2001) caracterizan la implicación *a priori* en la explicación reductiva de acuerdo con lo que llaman “habilidad condicional” y “condicional aplicado”. La habilidad condicional es la habilidad que tiene una persona para determinar la extensión de un concepto. Para tener habilidad condicional la persona debe cumplir con los siguientes requisitos:

1. Poseer “un concepto y tener capacidades racionales intactas”.
2. Poseer “suficiente información empírica acerca del mundo actual”. (Chalmers y Jackson, 2001, p. 323)

Si la persona cumple con ambos requisitos, entonces esa persona puede identificar la extensión del concepto *a priori*, es decir, puede determinar qué propiedades son instanciadas en el mundo desde la comodidad del sofá. Para que la persona pueda evaluar la extensión del concepto debe poseer la información suficiente sobre cómo es (o podría ser) el mundo actual. Para lograr dicha evaluación la persona analiza el concepto desde una hipótesis que es epistémicamente posible: “intuitivamente, una forma específica en que el mundo actual podría resultar ser...” (Chalmers y Jackson, 2001, p. 324). Bajo esa hipótesis, la persona considera toda la información que es relevante para determinar la extensión del concepto y sobre esa información es que la persona es capaz de determinar qué propiedades son instanciadas en el mundo y cuáles no. Un aspecto importante del primero de los requisitos es que la persona debe poseer capacidades de evaluación de la información no limitadas, en el sentido de que debe poseer la suficiente capacidad racional para evaluar información hipotética desde alguna hipótesis epistémicamente posible.² Lo que estamos haciendo es abstraer de limitaciones cognitivas a la hora de determinar la extensión de un concepto de manera que permitimos que la persona sea capaz de considerar toda la información relevante para determinar la extensión del concepto.

Aunque parece que el segundo de los requisitos contradice el requerimiento de la justificación *a priori*, ese no es el caso, ya que la información empírica que se requiere es parte del antecedente del condicional en la implicación, pero no juega un papel en la justificación de dicho condicional. Retomando el ejemplo de la reducción del agua a H_2O , si la persona tiene toda la información empírica relevante sobre la distribución, comportamiento y características del H_2O , entonces la persona puede saber que el agua = H_2O : “Dada la suficiente información empírica en el antecedente, hay buenas razones para creer que el condicional resultante será *a priori*” (Chalmers y Jackson, 2001, p. 325). El condicional que resulta de la justificación *a priori* es el condicional aplicado. Tanto la habilidad condicional como el condicional aplicado son requeridos para tener una explicación reductiva exitosa.

Un paso importante en el proceso de la justificación *a priori* es el de proveer de información empírica suficiente al antecedente. El tipo de información que se requiere suele ser caracterizado desde el modelo funcional de reducción. El modelo funcional de reducción

² Aunque este requisito es una idealización, ya que ninguna persona tiene ese tipo de capacidad racional, basta con que sea posible en principio para entender la noción de implicación *a priori*.

requiere de dos pasos: primero, la caracterización funcional del concepto a ser reducido, esto es, la caracterización de la entidad de acuerdo con el rol causal que dicha entidad tiene; segundo, la identificación de una entidad que es caracterizada de acuerdo con sus propiedades estructurales (físicas) y que cumple con el rol causal asociado con el concepto a reducir. La reducción se efectúa una vez que se cumplen con ambos pasos. El aspecto clave de la reducción bajo el modelo funcional es la caracterización funcional de la entidad a reducir, ya que es a partir de esa caracterización que se puede explicar completamente por qué una entidad se comporta de la forma en que lo hace o por qué exhibe las propiedades que exhibe. Esto se logra cuando se identifica la entidad física a nivel microscópico que cumple con la descripción funcional, en eso consiste la explicación reductiva. Como escribimos arriba, la caracterización funcional se entiende como una caracterización en términos del rol causal que se asocia con la entidad a ser reducida. El rol causal aquí simplemente son las diferentes relaciones causales que la entidad tiene dentro de un sistema o con otras entidades en el mundo: “en general, el rol causal de una propiedad funcional está dado por el *conjunto de conexiones nómicas* que son satisfechas por la propiedad o estado en cuestión” (Marras, 2005, p. 345). Una vez que se da la caracterización funcional, se puede realizar la reducción al identificar a una entidad, generalmente de nivel inferior, que cumple con el rol causal de la entidad a reducir.

Volviendo al ejemplo de la reducción del agua a H_2O , bajo el modelo funcional, tenemos la siguiente caracterización: primero se caracteriza el agua de acuerdo a un rol funcional, esto es, de acuerdo a las relaciones causales que se asocian al agua, como la de hervir a temperaturas de $100^{\circ}C$ o congelarse a $0^{\circ}C$; segundo, tenemos la identificación del H_2O como la estructura molecular que cumple con el rol causal asociado al agua, esto es, que el H_2O hierve a temperaturas de $100^{\circ}C$ o se congela a temperaturas de $0^{\circ}C$. La caracterización funcional provee la información suficiente para poder justificar *a priori* el condicional en la reducción: si el vaso tiene H_2O entonces el vaso tiene agua. Una vez que la persona tiene toda la información empírica sobre el rol causal del H_2O , es capaz de justificar *a priori* su conocimiento de la implicación del concepto de ‘ H_2O ’ al concepto de ‘agua’.

Así, en el marco de la filosofía de la mente, se suele asumir que el fisicalismo se compromete con una implicación *a priori* entre las verdades de los enunciados físicos y las verdades de los enunciados mentales. En tal caso, una vez que conocemos todas las verdades físicas sobre el mundo —de todas las cosas o colecciones de cosas y propiedades físicas que instancian— tenemos toda la evidencia que necesitamos para saber todas las verdades mentales. De esta manera si, abstrayendo de limitaciones cognitivas, conociéramos todas las verdades físicas podríamos deducir desde la comodidad de nuestro sofá, *a priori*, que en este momento uno de los autores del texto tiene un ligero dolor de cabeza. En este sentido, las verdades mentales se reducirían a las verdades físicas. En tal caso, si hay una explicación reductiva de la mente en términos físicos, entonces el condicional $P \rightarrow Q$ —donde P se refiere a alguna descripción sobre un hecho físico y Q se refiere a alguna descripción sobre un hecho mental— se justificaría *a priori* como hemos visto. Veámoslo con el ejemplo del dolor.

Para dar una explicación reductiva de que *A* tiene un ligero dolor de cabeza, lo primero que necesitamos es una caracterización funcional del dolor, como que el dolor es aquel estado que surge cuando hay daño al cuerpo, que produce una sensación incómoda, que a su vez causa que la persona busque mitigar el dolor, etc. En segundo lugar, identificaríamos el mecanismo al nivel al que queremos reducirlo que cumple con esa caracterización funcional. Imaginemos que queremos dar una explicación en términos neuronales —asumiendo que ese nivel a su vez puede ser explicado reductivamente en términos del tipo de propiedades que postula la física a nivel fundamental. Para ello buscaríamos el mecanismo a nivel neuronal que cumple tal rol funcional. Supongamos que es el mecanismo de disparo de fibras-C. ¿Hemos ofrecido así una explicación reductiva del dolor en el sentido que sea una explicación que haga inteligible el fenómeno a explicar, el dolor, a partir de aquello que lo explica, el mecanismo de disparo de fibras-C? Si la respuesta fuese negativa entonces no habría explicación reductiva y el fisicalismo, en caso de estar comprometido con la existencia de una explicación reductiva sería falso. De hecho parece intuitivo que la explicación ofrecida no resulta satisfactoria y que hay un aspecto del dolor que aún se escapa de tal explicación; a saber la manera en que el dolor se siente. Esto nos lleva al terreno de la consciencia fenoménica.

3. La emergencia de las entidades mentales: consciencia fenoménica y brecha explicativa

Nuestros estados mentales a menudo son *fenoménicamente conscientes*, un término técnico para una noción de sentido común que refiere a las experiencias subjetivas que tenemos los seres humanos y seguramente otros animales. Las experiencias (fenoménicamente) conscientes son aquellas tales que hay algo que se siente desde el punto de vista de la criatura (Nagel, 1974). Ejemplos de experiencias conscientes son aquellas que tenemos al oler el café, ver algo rojo, sentir dolor o tener un orgasmo. A la manera en que se siente tener una experiencia le llamamos el carácter fenoménico de la experiencia. El carácter fenoménico es un aspecto que está presente en nuestra vida mental cotidiana, por lo que solemos tener una relación íntima con esas experiencias en el sentido de que es algo que conocemos de forma directa, en contraste a nuestro conocimiento del mundo exterior: conozco cómo se siente ver el rojo o sentir dolor, aunque desconozca la física detrás de dicho color o sensación. Como hemos visto, si la explicación reductiva es verdadera para el caso de tales estados fenoménicos, entonces hay implicación *a priori* entre las verdades físicas y las verdades fenoménicas. Una falla en la implicación *a priori* tiene como consecuencia una falla en la explicación reductiva.

Nuestras experiencias conscientes son el ejemplo paradigmático de ausencia de implicación *a priori* y parece que las verdades fenoménicas —por ejemplo, el tipo de experiencia que alguien más tiene— no pueden ser derivadas de verdades físicas. De esta manera la consciencia parece escapar a la reducción que busca la fisicalista pues parece que no podemos entender el carácter fenoménico de la experiencia por mucho que nos cuenten acerca de las propiedades físicas a las cuales supuestamente se reduciría. David Lewis presenta esta idea de manera directa:

Dicen que la experiencia es la mejor profesora, y que el salón de clases no sustituye a la Vida Real. Hay algo de verdad en ello.

Si quieres saber cómo es qué es una experiencia nueva y diferente, puedes aprenderlo saliendo y realmente teniendo la experiencia. No puedes aprenderlo por mucho que te cuenten acerca de ella sin importar que intensa sea la clase. (Lewis, 1990, p. 29)

Probablemente la mejor articulación de la intuición sea la presentada por el filósofo Frank Jackson (1982). Jackson plantea la situación hipotética de una mujer llamada Mary que no tiene la capacidad de tener experiencias de colores. Bajo una versión de la historia, Mary nace sin la capacidad de percibir colores, lo único que percibe es un mundo en blanco y negro: su experiencia visual es similar a la que se tiene cuando se ve la película antigua. Dada su condición, Mary se vuelve una experta en la visión en color y conoce todos los procesos fisiológicos que corresponden a la visión en color. Un día, Mary recupera la capacidad de percibir colores y ve una rosa, lo que hace que ella intuitivamente exclame: “¡Oh, así es cómo se siente ver rojo!”. Aunque Mary ya conocía todos los hechos fisiológicos pertinentes a la visión en color, parece que aprende algo nuevo al ver la rosa por primera vez, a saber, cómo se siente ver el color rojo o cómo se siente tener esa experiencia. Si aceptamos la intuición que el experimento mental presenta, lo que estamos haciendo es, de hecho, rechazar que haya una implicación *a priori* entre las verdades físicas, que Mary conoce y las verdades fenoménicas, que Mary no podía deducir desde el interior de su habitación. A pesar de que Mary posee la habilidad condicional falla en establecer un condicional aplicado entre las verdades físicas y las verdades fenoménicas. ¿Qué efecto tiene para el debate sobre la verdad del fisicalismo y el entendimiento de la reducción y la emergencia? Para tratar de responder a esta pregunta debemos refinar nuestras nociones de reducción y emergencia.

3.1 Explicación reductiva: emergencia super-débil y débil

El concepto de reducción busca capturar las relaciones que se pueden dar entre diferentes tipos de entidades, como teorías, conceptos, eventos, propiedades, etc. Podemos clasificar la reducción en dos categorías dependiendo del tipo de entidades que flanquean la relación. Por un lado, podemos hablar de *reducción epistemológica*, cuando la relación aplica a entidades representacionales como las teorías, las proposiciones o los conceptos. La explicación reductiva que hemos estado considerando y que requiere que haya una implicación *a priori* entre el *explanandum* y el *explanans* concierne este tipo de reducción. Podemos entonces decir que si no hay implicación *a priori* el *explanandum* no se reduce epistemológicamente al *explanans*. Aun así, aunque la explicación reductiva falle alguien podría insistir en que, en algún sentido, por articular, el tipo de entidades a las que el *explanandum* refiere no son *nada más que* aquellas a las que el *explanans* refiere. En tal caso hablamos de *emergencia débil* o *emergencia epistemológica*. Por otro lado, se puede hablar de *reducción ontológica* y *emergencia fuerte* cuando la relación aplica a entidades no representacionales del mundo, como objetos, eventos, propiedades, etc. Articulemos estas distinciones por medio de un ejemplo.

Comencemos por regresar al caso de los sistemas complejos de los que nos habla Anderson. Resulta obvio que hay sistemas cuya complejidad supera cualquier capacidad humana de entendimiento. Por ejemplo, el sistema de ecuaciones diferenciales que determina la posición y velocidad de las partículas de un sistema en un tiempo futuro a partir de su posición actual resulta prácticamente irresoluble analíticamente cuando el número de partículas y la complejidad de las ecuaciones excede de un determinado rango. El estudio de los sistemas complejos busca formas de lidiar con esa complejidad de manera alternativa a buscar soluciones analíticas a las ecuaciones, estudiando el comportamiento en términos cualitativos en vez de cuantitativos.³ Por ejemplo, podemos predecir de manera altamente eficiente la evolución de un sistema complejo si abstraemos de los cambios instantáneos y en su lugar nos fijamos meramente en las estabilidades; esto es en las posiciones donde tienen a encontrarse las partículas tras un periodo de tiempo; aquellas posiciones donde la velocidad es cero. Aun así, tales aproximaciones tienen sus límites, ya que hay sistemas complejos relativamente sencillos y perfectamente deterministas en los que surge el caos; esto es comportamientos aperiódicos que exhiben una dependencia sensible a las condiciones iniciales. En estos casos, cambios tan pequeños como queramos en propiedades iniciales como puedan ser la velocidad y la posición de las partículas dan lugar a diferencias radicales en el comportamiento de las partículas. En estas condiciones, si hay un límite al detalle en qué podemos medir, digamos, la posición de las partículas, o en la capacidad de cálculo, es imposible predecir el comportamiento de sistemas del sistema.

Casos como el anterior suelen ser considerados como ejemplos de emergencia débil donde los distintos comportamientos “emergen” de esos cambios infinitesimales. Podemos entender un sentido en que lo hacen ya que, dada la sensibilidad al estado inicial, bajo restricciones de cómputo o medición, no es posible deducir el comportamiento a partir de las ecuaciones dinámicas, perfectamente continuas, que rigen el movimiento de las partículas y en este sentido no es posible predecir el comportamiento del sistema, no se hace inteligible a partir de la de los elementos a los que apelamos para explicarlo; a saber, la posición y velocidad de las partículas que componen el sistema. Llamemos a este tipo de emergencia ‘emergencia super-débil’. Le llamamos emergencia *super-débil* porque meramente nos habla de la sensible dependencia de una propiedad a los mínimos cambios de aquella en términos de la cual la explicamos. De esta manera, no podemos deducir el comportamiento del sistema dada nuestra limitada capacidad de cómputo (cualquier sistema que empleemos estará igualmente limitado) o de medición de las propiedades físicas. Imaginemos que Mary tiene capacidades de razonamiento y cálculo ideales; es decir, no tiene limitaciones relacionadas con su capacidad de cómputo y además tiene perfecto conocimiento de todas las propiedades físicas. Conoce exactamente la posición y velocidad de las partículas.⁴ Bajo esta circunstancia, no habría ignorancia que

³ Véase Strogatz (2001) para una excelente introducción. El libro presenta también los detalles mencionados sobre el caos.

⁴ En mecánica cuántica no es posible conocer ambas con precisión arbitraria (Heisenberg, 1927). Sin embargo, el objetivo en esta sección es señalar meramente la diferencia entre lo que llamamos emergencia super-débil y

Mary tuviera acerca de regímenes caóticos. Hay de hecho una implicación entre la posición y velocidad de las partículas y el comportamiento del sistema —una implicación que está dada, precisamente, por las ecuaciones dinámicas que rigen el sistema. Por lo tanto, no hay emergencia débil. Ni tampoco fuerte pues no hay ningún sentido metafísicamente sustancial en que el comportamiento del sistema —la propiedad de alto nivel de las que nos habla el *expalanadum*— sea algo más allá de la velocidad y posición de las partículas que lo componen —las propiedades de las de bajo nivel a las que refiere el *explanans*—: una vez que fijamos las propiedades físicas, las de alto nivel quedan también fijadas.

Contrastemos este caso con el que nos propone Jackson. Si aceptamos la intuición del experimento mental, el conocimiento exhaustivo de las propiedades físicas que tiene Mary junto a su capacidad ideal de razonamiento y de cálculo no le bastan para poder deducir cómo se siente ver rojo. Aceptar la intuición implica rechazar la implicación *a priori* y la reducción epistemológica. Las teorías físicas parecen resultar insuficientes para dar cuenta de las propiedades fenoménicas, de la manera en que se siente tener una experiencia.⁵ Una forma de ver la emergencia débil en el caso de Mary es recordando que la explicación reductiva debe proveer de un análisis funcional de la entidad a reducir, consciencia fenoménica en este caso. El problema radica en que aun cuando se da dicho análisis parece que aún podemos preguntarnos ¿por qué el dolor se siente de la forma en que se siente? El análisis funcional no captura completamente lo que es tener experiencias fenoménicas y es ese aspecto el que es explotado con la intuición de Mary.

Ahora bien, el fisicalismo no es una tesis epistemológica —una tesis acerca de lo que podemos conocer—, sino es una tesis metafísica que nos habla acerca del tipo de entidades que existen en nuestro mundo y su relación entre ellas. En particular afirma que cuanto existe en nuestro mundo es físico, esto es que depende exclusivamente del tipo de entidades que postula la física.⁶ Queda pendiente ver qué relación hay entre la falla en la reducción epistemológica y la reducción metafísica y análogamente entre la emergencia débil y la fuerte.

débil en sistemas complejos. Con este objetivo podemos abstraer de complicaciones derivadas de la mecánica cuántica y asumir que el sistema se rige por la mecánica clásica.

⁵ No es infrecuente encontrar en la literatura el término ‘qualia’ para referirse a las propiedades fenoménicas. Sin embargo, hay ciertas presuposiciones antifisicalistas asociadas al uso de este término que pueden empañar la claridad de la exposición, por lo que hemos preferido evitar su uso en este artículo.

⁶ Esta caracterización es problemática como señaló Hempel (1969), pues plausiblemente o bien es falsa —si tomamos nuestra teoría contemporánea, es muy plausible que tal teoría sea falsa y, por lo tanto, que varias de las entidades que postule no existan— o bien es trivial —si apelamos a una probable teoría física futura o a alguna teoría física ideal no estaremos en posición de saber qué tipo de propiedades postulará esa teoría— quedando la tesis fisicalista sin clarificar. Podemos no obstante entender que una teoría futura sería un ajuste, más que un cambio radical sobre nuestra teoría actual y en este sentido podemos entender el debate entre la fisicalista y la antifisicalista (ver Sebastián, 2022a). La discusión en la siguiente sección presenta el debate abstrayendo del problema de entender qué es una propiedad física.

3.2 Emergencia débil y emergencia fuerte

Nuestra experiencia consciente parece escapar a cualquier explicación reductiva. Esto es el caso con independencia de cuál sea la naturaleza de las propiedades fundamentales que postulemos —propiedades físicas en el caso del fisicalismo. Podemos asumir, en aras de la simplicidad en la exposición, una ontología suficientemente simplista, donde los hechos consisten en cosas (o n-tuplas de cosas) que ejemplifican propiedades. El mundo se describe exhaustivamente diciendo qué cosas (o n-tuplas de cosas) ejemplifican qué conjunto de propiedades fundamentales (Lewis, 1983; 1986). Podemos pensar en el fisicalismo como la tesis de que todas las cosas y propiedades fundamentales son físicas, y que todos los hechos del mundo son físicos o necesitados por hechos físicos, es decir, que todo mundo que sea una copia de las cosas y propiedades fundamentales físicas será un mundo que sea una copia del resto de cosas y propiedades. La idea detrás de la tesis de la implicación *a priori* es que al conocer todas las verdades fundamentales (físicas) —es decir, de todas las cosas (o n-tuplas de cosas) y las propiedades fundamentales (físicas) que tienen—, uno tiene toda la evidencia que requiere para resolver cualquier incertidumbre respecto al mundo real. En este sentido, el conocimiento de las propiedades fundamentales junto con las capacidades de razonamiento ideales lo colocan a uno en posición de conocer cualquier hecho que se da en el mundo. *Ex hypothesi*, Mary tiene tal conocimiento y capacidades ideales y por la tanto sabe cómo es exactamente el mundo que habita. Si concedemos que Mary aprende algo al ver rojo, parecería que tienen que haber hechos fundamentales que no conoce. De esta manera y con el objetivo de entender la relación entre reducción epistemológica y metafísica al seno de la filosofía de la mente y más concretamente en la discusión sobre la naturaleza de nuestra experiencia consciente, resulta útil presentar la idea que subyace al argumento de Jackson en términos de la tensión entre tres tesis que resultan ampliamente plausibles (Sebastián, MN) pero que no parece que puedan ser conjuntamente verdaderas:

P3. Los hechos fenoménicos no son fundamentales.

P4. Conocer todos los hechos fundamentales y tener capacidades de razonamiento ideales permiten conocer todos los hechos del mundo actual.

P5. Mary conoce todos los hechos fundamentales y tiene capacidades de razonamiento ideales, pero ella no conoce lo que se siente ver rojo.

P2 sostiene que es posible una explicación reductiva de todas las propiedades que se dan en nuestro mundo a partir de las propiedades fundamentales y abstrayendo de limitaciones cognitivas. Hace por tanto explícita la relación entre explicación reductiva y la implicación *a priori*. P3 recoge la intuición de que Mary aprende algo al ser liberada. Si aceptamos P3, parece que, o bien los hechos fenoménicos son fundamentales y rechazamos P1, o bien hay hechos que no pueden ser explicados reductivamente a partir de los hechos fundamentales y rechazamos P2. P1 es rechazado por la dualista que afirma que el conjunto de propiedades fundamentales del mundo que habitamos debe incluir además de las propiedades físicas propiedades fenoménicas. Rechazar P2 supone aceptar la superveniencia débil. En este

caso, el desafío para quien cree que el mundo queda exhaustivamente descrito una vez que enumeramos todas las cosas, las propiedades fundamentales, pero que las propiedades fenoménicas no son fundamentales, es explicar cómo Mary puede tener conocimiento de tales propiedades fundamentales y aun así aprender algo.

Una posibilidad es postular la existencia de emergencia fuerte. La emergencia fuerte de los hechos fenoménicos se da por la imposibilidad de establecer una conexión necesaria con los hechos físicos fundamentales: hay una relación de dependencia, pero es una relación contingente, en otras palabras, los hechos fenoménicos dependen de los hechos físicos, pero no hay implicación necesaria y por lo tanto *a priori*. Así, hay nuevas propiedades, las propiedades fenoménicas que son propiedades que van más allá de las propiedades físicas. Quienes creen que hay propiedades fuertemente emergentes, consideran que una vez que fijamos las propiedades fundamentales aún necesitamos fijar las leyes que contingentemente ligan las propiedades emergentes y las propiedades fundamentales (Chalmers, 1996). En este sentido, puede haber mundos que sean iguales al nuestro al nivel fundamental que solo incluye las propiedades físicas, pero que difiere del nuestro en las propiedades que de ellas emergen como las propiedades fenoménicas. La existencia de emergencias fuertes supone entonces la negación del fisicalismo como lo hemos presentado y la defensa de un dualismo de propiedades (Chalmers, 1996).

4. Fisicalismo y el argumento del conocimiento: el desafío de explicar que sólo haya emergencia débil

El argumento del conocimiento pretende refutar la tesis del fisicalismo al mostrar que hay una brecha explicativa entre los hechos físicos y los hechos fenoménicos y postular que la mejor explicación para esa brecha es la existencia de propiedades que no son físicas. La dualista entonces postula que las propiedades fenoménicas deben ser incluidas dentro del conjunto de propiedades fundamentales que constituyen el mundo y ser añadidas a las físicas. Una alternativa es conceder que al nivel más fundamental solo hay propiedades físicas y admitir la existencia solo de emergencia débil.

Ahora bien, el fisicalismo es una tesis metafísica y no meramente epistemológica. Aceptar la brecha explicativa al aceptar la imposibilidad de explicar el carácter fenoménico de la experiencia en términos físicos y con ello la emergencia débil no nos compromete inmediatamente con el rechazo del fisicalismo. La fisicalista que quiera aceptar la intuición del argumento del conocimiento rechaza el paso de la ausencia de implicación *a priori* (emergencia débil) a la postulación de propiedades fenoménicas al nivel fundamental o a la emergencia fuerte de propiedades. Ahora bien, el desafío que ha de enfrentar la fisicalista es el de explicar cómo es posible que haya una emergencia débil sin abrazar el dualismo o la emergencia fuerte. Si aceptamos que Mary aprende algo debemos negar entonces que lo que aprende implique un hecho que no conocía. Hay diversas estrategias que la fisicalista puede tomar, pero todas ellas presentan claros desafíos.

La primera estrategia consiste en defender que lo que Mary aprende es una nueva habilidad. Esta habilidad no implica nuevo conocimiento de hechos, sino saber cómo hacer algo (Lewis, 1990; Nemirov, 1990). Sin embargo, no queda claro si hay habilidades que son necesarias y suficientes para el nuevo conocimiento de Mary, y muchos han dudado de que ese saber hacer pueda caracterizar con precisión la ganancia epistémica intuitiva de Mary. Además, hay buenas razones para creer que el saber cómo implica conocimiento de hechos (Stanley y Williamson, 2001).

La segunda estrategia consiste en argumentar que Mary adquiere un nuevo concepto y, de esa manera, aprende un hecho que ya conocía de una manera nueva (Balog, 2009; 2012; Block, 2006; Levin, 2006; Papineau, 2006; Perry, 2001). La idea es familiar. Cuando alguien descubre que Héspero es Fósforo aprende algo nuevo, aunque el hecho de que Venus es Venus fuera algo que ya sabía. El problema es que la significatividad cognitiva de pensamientos que involucran términos correferenciales involucra el desconocimiento de alguna u otra contingencia —por ejemplo, que el primer cuerpo celeste al atardecer es el último que se ve al amanecer, o hecho concernientes a la introducción de nombres. Pero Mary conoce todos los hechos. Uno puede mantener que hay hechos como los fenoménicos que sólo pueden ser conocidos por familiaridad (*acquaintance*) al estar en el estado, pero eso involucra que el conocimiento de Mary acerca del mundo no es completo —ver Howell (2007; 2013) para una discusión detallada.

Finalmente, la fisicalista puede defender que el tipo de información que ofrece la experiencia es más fino que el tipo de información que ofrece una teoría científica (Sebastián, 2022b). Si una teoría científica busca decirnos cómo es exactamente el mundo que habitamos, nuestra experiencia consciente nos permite localizarnos al interior de ese mundo. Esto explicaría cómo puede haber aspectos de la experiencia consciente —a saber, la información que la experiencia porta— que Mary no puede deducir de su conocimiento científico completo (Sebastián, manuscrito). El desafío es mostrar cómo un estado puede portar este tipo de información de manera compatible con la verdad del fisicalismo.

5. Conclusiones

El objetivo de este artículo ha sido presentar con detalle la manera en que se conciben los conceptos de reducción y emergencia en los debates dentro del ámbito de la filosofía de la mente, particularmente en relación con el fisicalismo y la explicación de la naturaleza de la consciencia fenoménica. El entendimiento adecuado del papel que desempeñan estos conceptos en dicho debate resulta fundamental para lograr una comprensión sólida de las diferentes posiciones en juego y del propio debate en sí.

Hemos visto que la forma tradicional en que la reducción ha sido entendida en los debates en filosofía de la mente es en términos de explicación reductiva, caracterizada por una implicación *a priori*. Aunque pocas personas ponen en duda que haya una perfecta correlación

entre nuestra experiencia consciente y ciertas propiedades neuronales, hay razones para poner en duda que las experiencias conscientes puedan ser de hecho reducidas a tales propiedades neuronales.

Como el argumento del conocimiento pretende poner de manifiesto, parece que el carácter fenoménico de la experiencia, la manera en que se siente tener la experiencia, escapa a ser explicada en términos de otra cosa. La aceptación de esa intuición nos compromete o bien con que las propiedades fenoménicas formen parte del mobiliario básico con que está construido nuestro mundo, o bien con que emerjan débilmente sobre cualesquiera que sean las propiedades fundamentales. Además, hemos visto cómo bajo esa emergencia débil surgen problemas más profundos que aquellos que se derivan de la posible complejidad que haya en el mundo.

El fisicalismo resultaría falso si aceptamos que las propiedades fenoménicas son algo más que las propiedades de las cuales parecen depender y sobrevienen fuertemente sobre ellas. Pero aceptar la emergencia débil sin comprometerse con la emergencia fuerte requiere una explicación que sea a su vez compatible con la verdad del fisicalismo. Hemos visto brevemente algunas de las estrategias que la fisicalista tiene a su disposición para hacerlo, pero también que hay trabajo aún por desarrollar para mostrar que alguna de esas estrategias pueda de hecho ser exitosa.⁷

Referencias bibliográficas

- Anderson, P. W. (1972). More is different. Broken symmetry and the nature of the hierarchical structure of science. *Science*, 177, 393-396. <https://doi.org/10.1126/science.177.4047.393>
- Balog, K. (2009). Phenomenal concepts. En B. McLaughlin, A. Beckermann & S. Walter (Eds.), *The Oxford handbook of philosophy of mind* (pp. 292-312). Oxford University Press.
- Balog, K. (2012). In defense of the phenomenal concept strategy. *Philosophy and Phenomenological Research*, 84, 1-23. <https://doi.org/10.1111/j.1933-1592.2011.00541.x>
- Bedau, M. (1997). Weak emergence. En J. Tomberlin (Ed.), *Philosophical perspectives: Mind, causation, and world* (pp. 375-399). Blackwell.
- Bedau, M. (2008). Is weak emergence just in the mind? *Minds and Machines*, 18, 443-459. <https://doi.org/10.1007/s11023-008-9122-6>
- Block, N. (2006). Max Black's objection to mind-body identity. *Oxford Studies in Metaphysics*, 2, 3-78.
- Chalmers, D. (1996). *The conscious mind*. Oxford University Press.

⁷ Agradecemos a Axel Tovar, a Aldo Filomeno y a un árbitro de esta revista por sus comentarios. Apoyo financiero para esta investigación fue proporcionado por el UNAM-PAPIIT IN402423 y el Programa de Apoyos para la Superación del Personal Académico de la UNAM (PASPA de la DGAPA).

- Chalmers, D., & Jackson, F. (2001). Conceptual Analysis and Reductive Explanation. *The Philosophical Review*, 110(3), 315-360. <https://doi.org/10.1215/00318108-110-3-315>
- Crutchfield, J. P., Doyné Farmer, J., Packard, N. H., & Shaw, R. S. (1986). Chaos. *Scientific American*, 255, 46-57.
- Hempel, C. (1969). Reduction: ontological and linguistic facets. En P. S. Sidney & M. White (Eds.), *Philosophy, Science, and Method: Essays in honor of Ernest Nagel*. St. Martin's Press.
- Howell, R. (2007). The knowledge argument and objectivity. *Philosophical Studies*, 135(2), 145-177. <https://doi.org/10.1007/s11098-007-9077-0>
- Howell, R. (2013). *Consciousness and the limits of objectivity*. Oxford University Press.
- Humphreys, P. (1997). How Properties Emerge. *Philosophy of Science*, 64(1), 1-17. <https://doi.org/10.1086/392533>
- Humphreys, P. (2015). More is different... Sometimes: Ising models, emergence and undecidability. En B. Falkenburg & M. Morrison (Eds.), *Why more is different: philosophical issues in condensed matter physics and complex systems* (pp. 137-152). Springer.
- Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43(3-4), 172-198. <https://doi.org/10.1007/BF01397280>
- Humphreys, P. (2016). *Emergence*. Oxford University Press.
- Jackson, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly*, 32(127), 127-136. <https://doi.org/10.2307/2960077>
- Levin, J. (2006). What is a phenomenal concept? En T. Alter & S. Walter (Eds.), *Phenomenal concepts and phenomenal knowledge: new essays in consciousness and physicalism* (pp. 87-110). Oxford University Press.
- Levine, J. (1993). On leaving out what it's like. En M. Davies & G. Humphreys (Eds.), *Consciousness: psychological and philosophical essays* (pp. 121-136). Blackwell.
- Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, 61(4), 343-377. <https://doi.org/10.1080/00048408312341131>
- Lewis, D. (1990). What experience teaches. En W. G. Lycan (Ed.), *Mind and cognition* (pp. 29-57). Blackwell.
- Lewis, D. (1986). *On the plurality of worlds*. Blackwell.
- Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, 83(4), 435-450. <https://doi.org/10.2307/2183914>
- Nemirow, L. (1990). Physicalism and the cognitive role of acquaintance. En W. G. Lycan (Ed.), *Mind and cognition* (pp. 490-499). Blackwell.
- Papineau, D. (2006). Phenomenal and perceptual concepts. En T. Alter & S. Walter (Eds.), *Phenomenal concepts and phenomenal knowledge: new essays in consciousness and physicalism* (pp. 111-144). Oxford University Press.

- Perry, J. (2001). *Knowledge, possibility and consciousness*. MIT Press.
- Sebastián, M. Á. (2022a). *El problema de la consciencia. Una introducción crítica a la discusión filosófica actual*. Ediciones Cátedra.
- Sebastián, M. Á. (2024). First-Person Perspective in Experience: Perspectival De Se Representation as an Explanation of the Delimitation Problem. *Erkenntnis*, 89(3), 947-969. <https://doi.org/10.1007/s10670-022-00564-4>
- Sebastián, M. Á. (manuscrito). Perspectival information and the knowledge argument. Reconsidering the first-person point of view.
- Stanley, J., & Williamson, T. (2001). Knowing How. *Journal of Philosophy*, 98(8), 411-444. <https://doi.org/10.2307/2678403>
- Strogatz, S. (2001). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Westview Press.