



Ideas y Valores

ISSN: 0120-0062

revideva_fchbog@unal.edu.co

Universidad Nacional de Colombia

Colombia

Kim, Jaegwon

El fisicalismo no reduccionista y su problema con la causalidad mental

Ideas y Valores, vol. LXIII, núm. 155, agosto-, 2014, pp. 235-259

Universidad Nacional de Colombia

Bogotá, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=80931523011>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

TRADUCCIÓN

NOTA DEL TRADUCTOR

El texto que se traduce es “The Nonreductivist’s Troubles with Mental Causation”. *Mental Causation*. Eds. John Heil y Alfred Mele. Oxford: Oxford University Press, 1993. 189-210. También se hizo uso de la versión publicada en Jaegwon Kim, *Supervenience and mind. Selected philosophical essays*. New York: Cambridge University Press, 1993. 336-357. Realicé esta traducción gracias al apoyo del proyecto DIB código 16012, apoyo que recibí al participar en el grupo de investigación Ética, Comportamiento y Evolución.

Doy mis sinceros agradecimientos al profesor Alejandro Rosas por su invaluable ayuda y a Pedro Morales por la revisión de estilo. Finalmente, agradezco al profesor Jaegwon Kim la autorización otorgada para la traducción del artículo al español y su consiguiente publicación.

JUAN DIEGO MORALES
Candidato a Doctor en filosofía
Universidad Nacional de Colombia
jdmoraleso@unal.edu.co

EL FISICALISMO NO REDUCCIONISTA Y SU PROBLEMA CON LA CAUSALIDAD MENTAL

JAEGWON KIM

Brown University - United States of America

1. ¿Un mundo bifurcado o estratificado?

El dualismo mente-cuerpo, en el estilo clásico cartesiano, concibe dos dominios de particulares (“substancias”) no coincidentes que, en términos generales, tienen el mismo estatus ontológico. Se considera que los elementos mentales comparten cierta propiedad definitoria (“pensamiento” o “consciencia”, de acuerdo con Descartes) que excluye la propiedad que define los elementos del lado físico (“extensión”, de acuerdo con Descartes), y que hay una familia de propiedades distintas asociadas con cada dominio, propiedades mentales en un lado y propiedades físicas en el otro, en términos de las cuales los particulares de cada dominio pueden ser caracterizados exhaustivamente. En consecuencia, estamos en presencia de una imagen *bifurcada* de la realidad: el mundo consiste en dos esferas metafísicamente independientes que existen una junto a la otra.

Pero no todo aquel que acepta una imagen como esta piensa que los dominios se encuentran completamente desconectados. Aunque existen notables excepciones, como Leibniz y Malebranche, muchos dualistas sustanciales, incluyendo por supuesto a Descartes, han sostenido que, a pesar de su separación e independencia, los dominios se encuentran conectados *causalmente*: los eventos mentales pueden ser, y a menudo son, causas y efectos de eventos físicos; y los cambios en una mente pueden ser causas o efectos de los cambios corporales. Esto significa que ambos tipos de eventos pueden ocurrir *como eslabones de la misma cadena causal*: si se escoge un evento físico y se traza su precedencia o posterioridad causal, es posible encontrar eventos mentales, y similarmente si se inicia con un evento mental. Se sigue entonces que bajo el dualismo causal cartesiano *no puede existir una teoría física completa de los fenómenos físicos*, pues permite ocurrencias físicas que no pueden ser explicadas causalmente al invocar solo antecedentes físicos más leyes físicas. Cualquier teoría comprehensiva del mundo físico debe, en el interaccionismo cartesiano, incluir referencias a agentes causales no físicos y leyes que gobiernan su conducta. Podemos decir que *el interaccionismo cartesiano viola el cierre causal*

del dominio físico. Por supuesto, también viola el cierre causal del dominio mental; el cartesianismo implica que ninguna teoría científica podría esperar lograr una cobertura completa, a menos que abarque tanto el reino físico como el mental –a menos, esto es, que tengamos una *teoría unificada* de ambos fenómenos, el mental y el físico–.

La imagen ontológica que ha dominado el pensamiento contemporáneo en la filosofía de la mente es notablemente diferente de la imagen cartesiana. El modelo cartesiano de un mundo bifurcado ha sido reemplazado por el de un mundo estratificado, un mundo de “niveles” u “órdenes” de entidades y sus propiedades características. Generalmente se piensa que existe un nivel inferior, consistente en lo que sea que la microfísica nos vaya a decir acerca de cuáles son las partículas físicas más básicas a partir de las cuales toda la materia está compuesta (electrones, neutrones, *quarks* u otras). Asimismo estos objetos, cualesquiera que sean, son caracterizados mediante ciertas propiedades y relaciones fundamentales (masa, espín, sabor u otras). A medida que ascendemos hacia los niveles superiores encontramos estructuras compuestas de entidades que pertenecen a los niveles inferiores; más aún, se piensa que las entidades de cada nivel poseen un conjunto de propiedades distintivas. Así, en cierto nivel encontraremos agregados de moléculas de H_2O con propiedades como transparencia, cierta densidad y viscosidad característica, poder para disolver el azúcar y la sal, etc. En niveles aún superiores encontraremos células y organismos con sus propiedades “vitales” y, más arriba, organismos con conciencia e intencionalidad. Más allá, encontraremos grupos sociales de organismos, y quizás grupos compuestos de tales grupos.¹ Algunas veces se habla en términos de “niveles de descripción”, “niveles de análisis” o “niveles lingüísticos”; a menudo, el modelo estratificado se encuentra implícito en tal forma de hablar.

De esta forma, el mundo descrito por la nueva imagen consiste en una colección ordenada de niveles, en donde cada uno de estos consta de dos componentes: un conjunto de *entidades* que constituyen el dominio de los particulares para cada nivel, y un conjunto de *propiedades* definidas sobre tal dominio. Lo que proporciona esta estructura ordenada es la relación mereológica *ser parte de*: las entidades que pertenecen a una capa dada están compuestas mereológicamente de entidades que pertenecen a los niveles inferiores. Como ya se ha dicho, esta imagen de varios niveles usualmente asume la existencia de un nivel inferior, una capa de entidades que no tienen partes físicamente significativas.

1 Para una presentación útil e informativa de esta imagen estratificada, véase Oppenheim y Putnam (1958).

Hasta aquí la caracterización del modelo estratificado deja una pregunta importante sin resolver: ¿cómo se relacionan las propiedades características de las entidades de un nivel dado con aquellas que caracterizan las entidades de los niveles adyacentes? Dado que las entidades de los distintos niveles están ordenadas por la relación parte-todo, ¿las propiedades de los diferentes niveles también están ordenadas por alguna relación distinta y significativa?²

Esta es la cuestión crucial que responde a aquello que ha definido las diferentes y actualmente controvertidas posiciones sobre ciertas cuestiones metafísicas y metodológicas, entre las que se incluye, especialmente, el problema mente-cuerpo. La respuesta clásica positivista es que las propiedades distintivas de las entidades de un nivel dado son *reducibles a*, o *explicables reductivamente en términos de*, las propiedades y relaciones que caracterizan a las entidades de los niveles inferiores. En esto consiste el *reduccionismo*. Pero en las últimas décadas el reduccionismo ha pasado por tiempos difíciles y ha sido eclipsado por sus mayores rivales, el *eliminativismo* y el *antirreducciónismo*. Estas posiciones concuerdan en la afirmación de que, por lo general, las propiedades de niveles superiores no son reducibles a propiedades de niveles inferiores y, no obstante, difieren acerca del estatus de las primeras. El antirreducciónismo sostiene que las propiedades superiores son propiedades genuinas de los objetos y eventos de este mundo, y que constituyen una parte imposible de eliminar de su verdadera ontología. El eliminativismo, por su parte, sostiene que son adiciones inútiles que deben ser suprimidas de la imagen correcta de la realidad. De esta forma, la ruptura entre el antirreducciónismo y el eliminativismo se entiende a partir de la posibilidad de reducción: el primero, a diferencia del último, rechaza esta posibilidad como criterio de legitimidad de las propiedades superiores, y sostiene que tales propiedades pueden formar un dominio autónomo, un dominio para una “ciencia especial” independiente que no se reduce a las ciencias que tratan fenómenos de los niveles inferiores. El *emergentismo*, una posición influyente durante la primera mitad de este siglo, fue la primera articulación sistemática de esta aproximación no reduccionista.

A primera vista, el modelo estratificado parece mantener la promesa de constituir una forma elegante de evitar la violación del cierre causal de lo físico: quizás las interacciones causales se confinan dentro de cada nivel, de tal forma que se respete el cierre y la autonomía de

² Cuando se describe el modelo estratificado en términos de “niveles de descripción”, o “niveles lingüísticos”, se mantiene una pregunta correspondiente acerca de cómo se relacionan los aparatos descriptivos (predicados, conceptos, enunciados, etc.) de los distintos niveles.

los procesos causales en el nivel físico fundamental. En particular, en la versión no reduccionista del modelo estratificado es posible ver las cadenas causales en un nivel dado, al igual que las propiedades distintivas de tal nivel, formando un reino autónomo y autocontenido, inmune a las intrusiones causales de niveles cercanos. Es más, esta imagen no excluye la asignación de cierto estatus especial a la causalidad física: a pesar de la autonomía causal-nomológica de cada nivel, las relaciones causales de los niveles superiores pueden, en cierto sentido, depender de, o supervenir sobre, los procesos causales-nomológicos que ocurren en los niveles inferiores (McLaughlin 109-136; Fodor 1989 59-79) –tal y como, desde la perspectiva no reduccionista, se piensa que la irreductibilidad de las propiedades de los niveles superiores es consistente con su supervenencia sobre las propiedades y relaciones de los niveles inferiores–.

Centremos nuestra atención en el problema mente-cuerpo. *El fisicalismo no reduccionista*, una posición que merecidamente puede ser llamada “la concepción heredada” de la actualidad, es el antirreducciónismo aplicado al caso mente-cuerpo. Consta de las dos tesis características del antirreducciónismo: su ontología es el monismo físico, idea según la cual las entidades físicas y sus agregados mereológicos son todo lo que existe; pero su “ideología” es antirreducciónista y dualista, pues consiste en la afirmación de que las propiedades psicológicas son irreduciblemente distintas a las propiedades físicas y biológicas subyacentes. Su dualismo se refleja en la creencia de que las propiedades psicológicas, aunque físicamente irreducibles, son propiedades genuinas, tan reales como las propiedades físico-biológicas. Veamos un corolario acerca de la psicología: esta es una ciencia especial independiente de las ciencias físicas y biológicas en cuanto a su metodología y los conceptos y explicaciones que genera.³

Como hemos visto, el interaccionismo cartesiano implica la violación del cierre causal físico, lo que ha constituido una de las causas de su caída. Voy a argumentar que el fisicalismo no reduccionista y su principal acompañante, el emergentismo, son vulnerables a dificultades similares; en particular, se mostrará que el cierre causal físico en gran medida sigue siendo un problema dentro de la ontología estratificada

3 También encontramos la posición de Donald Davidson en su artículo “Mental Events”, reimpreso en Davidson (1980 [cf. Davidson 1995]), que acepta el fisicalismo no reduccionista tal y como se ha caracterizado aquí, esto es, como un monismo ontológico fisicalista (usualmente formulado como una tesis acerca de los *eventos individuales*) y un dualismo de propiedades; mientras rechaza el corolario acerca del estatus científico de la psicología. La perspectiva de Davidson sobre la psicología se entrevé a partir del título de uno de sus artículos acerca del tema, “La psicología como filosofía”, en *Ensayos sobre acciones y sucesos*.

del antirreducciónismo. El fisicalismo no reduccionista, como el cartesianismo, se funda sobre las rocas de la causalidad mental.

2. El fisicalismo no reduccionista y la “realización física”

La tesis ontológica básica del fisicalismo no reduccionista confiere cierto tipo de supremacía a lo físico: toda existencia concreta es física –no existen particulares que no sean físicos, esto es, almas, sustancias mentales cartesianas, “principios vitales” o “entelequias”–. En términos de propiedades, la primacía física viene a significar esto: todas las propiedades mentales son instanciadas en particulares físicos. Así, aunque puede haber, y presumiblemente hay, objetos y eventos que solo tienen propiedades físicas, no puede haber objetos ni eventos que solo tengan propiedades mentales; lo mental debe ser instanciado en sistemas físicos.

No parece haber una explicación general aceptada de lo que significa exactamente decir que algo es “físico”. Como mínimo, quizás una entidad física debe tener una localización determinada en el espacio y el tiempo; pero esto puede ser insuficiente. Quizás una entidad es física justo en el caso de que tenga una propiedad física u otra. Pero ¿qué hace que una propiedad sea física? Tal vez la mejor solución que podríamos dar es la que Hellman y Thompson han ofrecido: explicar “físico” en términos de la física teórica actual (*cf.* 551-564). Esta estrategia puede ser extendida a las ciencias de los niveles superiores, la química y la biología. Cuando llegamos a las propiedades psicológicas, sin embargo, la cuestión acerca de cuáles deberían ser consideradas nuestras teorías científicas de referencia es en sí una cuestión filosófica sin resolver, que se encuentra envuelta fundamentalmente en los debates sobre el reduccionismo y la causalidad mental. En cuanto a las propiedades mentales, entonces, debemos mirar hacia la psicología popular y su característico lenguaje intencional de creencias, deseos y demás, y hacia sus análogos intencionales en la psicología sistemática. Nada en la discusión que sigue dependerá de la definición general y precisa de “físico” y “mental”.

En cualquier caso, muchos fisicalistas, incluyendo a la mayoría de los fisicalistas antirreducciónistas, no están dispuestos a quedar solo con la primacía ontológica de lo físico en el sentido descrito, así como muchos dualistas sustanciales no se satisfacen únicamente con afirmar que los dos dominios están separados y son independientes. La mayoría de los fisicalistas no reduccionistas desean ir más allá de la afirmación de que las propiedades mentales son instanciadas por sistemas físicos; ellos quieren defender una tesis acerca de *la primacía, o el carácter básico, de las propiedades físicas en relación con las propiedades mentales*. La idea central es que, a pesar de su irreductibilidad,

en un sentido robusto las propiedades mentales son *dependientes de* o están *determinadas por* las propiedades físico-biológicas. Esto significa que el dualismo de propiedades del fisicalismo no reduccionista es un dualismo atenuado: un dualismo con relaciones de dependencia entre los dos dominios, tal y como el dualismo cartesiano es un dualismo con relaciones causales entre estos. Para muchos fisicalistas no reduccionistas, en consecuencia, la irreductibilidad no es la última palabra sobre la relación mente-cuerpo. La afirmación de irreductibilidad es una tesis negativa; los fisicalistas no reduccionistas desean una explicación positiva del vínculo entre los dos conjuntos de propiedades. Aquí las palabras de conexión son “dependencia” y “determinación”. Hellman y Thompson, quienes han propuesto una forma elegante de fisicalismo no reduccionista, describen su proyecto de la siguiente forma:

No obstante, recientemente ha habido una creciente conciencia de que el reduccionismo es una afirmación irrazonablemente fuerte. Junto con esto se ha reconocido que el reduccionismo se distingue de la tesis puramente ontológica acerca de los tipos de entidades a partir de los cuales está constituido el mundo Aunque una tesis puramente ontológica es un componente necesario para el fisicalismo, es insuficiente al no recurrir al poder de la ley física Buscamos desarrollar principios de *determinación física* que expliquen con bastante precisión la intuición fisicalista subyacente de que los hechos físicos determinan todos los hechos. (551-552)

Hellman y Thompson afirman que los *hechos* físicos determinan todos los *hechos*; presumiblemente, esto sucede solo porque el tipo de *propiedades* (incluyendo las relaciones) *físicas* que tienen los objetos determina el tipo de *propiedades mentales* que estos poseen.

Pero ¿cómo capturamos esta relación de determinación, o dependencia, de tal forma que escape de la amenaza del reduccionismo? Uno de los proyectos principales del fisicalismo no reduccionista en las últimas dos décadas ha sido responder esta cuestión, en el que se han destacado dos ideas: “la supervenencia” y “la realización física”. Hellman y Thompson desarrollan una explicación de la relación de determinación que es muy similar a lo que en la actualidad es comúnmente conocido como “supervenencia global”: una vez se fija el carácter físico de un mundo, su carácter entero es por lo tanto fijado.⁴ La idea de que los estados mentales son “realizados físicamente” (o “instanciados” o “implementados”) ganó popularidad a finales

4 Para un análisis de las relaciones de supervenencia, véase Kim (1990a).

de la década de 1960,⁵ principalmente mediante un argumento (“el argumento de la realización múltiple”) que ayudó a rechazar el fisicalismo reduccionista e instalar al antirreducciónismo como una de las posiciones más influyentes de la actualidad. El presunto problema del reduccionismo consistía en la “multiplicidad” de la realización, es decir, en que cualquier estado mental es realizable en una variedad de estructuras físicas completamente diversas, lo que hace que se pierda la esperanza de su identificación reductiva con una sola propiedad física. Los dos enfoques, uno basado en la “supervenencia” y el otro en la “realización física”, no son mutuamente excluyentes: como veremos, a partir de una lectura razonable de los términos en cuestión, es plausible sostener que la afirmación de que los estados mentales son realizados físicamente implica la afirmación de que supervenien sobre lo físico. El que la implicación conversa se mantenga es una cuestión que depende, entre otras cosas, de la fuerza de la relación de supervenencia que se tome en cuenta.

Muchos de los fisicalistas antirreduccionistas se sirven tanto de la supervenencia como de la realización, y algunos lo hacen de forma explícita; por ejemplo, LePore y Loewer caracterizan al fisicalismo no reduccionista mediante tres principios, siendo uno de ellos la “supervenencia global”, la tesis de que “[s]i dos mundos nomológicamente posibles son exactamente iguales en cuanto a los hechos físicos fundamentales... entonces son exactamente iguales con respecto a todos los demás hechos” (1989 177-178). Sin embargo, ellos también sostienen que “[l]a relación entre las propiedades psicológicas y neurofisiológicas consiste en que las últimas *realizan* a las primeras”. ¿Qué significa que una propiedad “realice” a otra? LePore y Loewer explican:

¿En qué consiste exactamente el que una propiedad de un evento *realice* a otra? La concepción común es que el hecho de que *e* sea *P* realiza a *f* si *e* es *P* y *e* es *F*, y existe algún tipo de conexión fuerte entre *P* y *F*. Proponemos entender esta conexión como una conexión necesaria *explicativa*. La existencia de una conexión explicativa entre dos propiedades es más fuerte que la afirmación de que *P* → *F* es necesaria físicamente, puesto que no toda conexión físicamente necesaria es explicativa. (1989 179)

5 Hasta donde sé, Hilary Putnam introdujo por primera vez la idea de “realización física” a principios de los años sesenta para describir la relación entre los estados “lógicos” y “estructurales” de las máquinas de computación, y lo extendió por analogía al caso de la mente y lo físico; véase Putnam (1960). Creo que la idea realmente se empezó a imponer cuando fue usada para formular el influyente “argumento de la realización múltiple” en el artículo seminal de Putnam (1967) “Psychological Predicates” (reimpreso luego bajo el título “The Nature of Mental States”).

Es claro que si todas las propiedades mentales son realizadas por propiedades físicas en este sentido, la supervenencia global de lo mental sobre lo físico es asegurada. De hecho, su Tesis de la realización física, como podríamos llamarla, implica una forma de supervenencia más fuerte, la “supervenencia fuerte” de las propiedades mentales sobre las físicas.⁶

Dado el propósito de mis argumentos en este artículo, deseo enfocarme por dos razones en aquellas versiones del fisicalismo no reduccionista que hacen uso de la “realización física” para explicar la relación entre propiedades psicológicas y físicas: primero, existe una variedad de relaciones de supervenencia no equivalentes, lo que hace difícil formular un argumento razonablemente uniforme y perspicuo acerca de las versiones que se basan en la supervenencia; y, segundo, muchos filósofos se han convertido al fisicalismo no reduccionista gracias al “argumento de la realización múltiple”, al que ya nos hemos referido brevemente. En consecuencia, para muchos fisicalistas antirreducciónistas, la Tesis de la realización física es uno de sus compromisos iniciales y básicos. A menudo es una estrategia expositiva correcta formular un argumento de una manera clara y perspicua, incluso si se trata de hacer uso de premisas relativamente fuertes, para luego pasar a ocuparse acerca de cómo el argumento podría cualificarse y afinarse para acomodar supuestos más débiles. Bien puede ser que existan versiones del fisicalismo no reduccionista para las cuales no aplican mis consideraciones, al menos no directamente;⁷ creo, sin embargo, que son relevantes para muchas de las versiones más populares e influyentes.

Volvamos al concepto de realización. LePore y Loewer, como hemos visto, hablan de “realización” de propiedades de *eventos*. Sin embargo, no es necesario limitar la relación a los eventos, y asumiremos que, sin cambios sustantivos, la relación de realización aplica a las propiedades de los objetos (“sustancias”) de la misma forma que a las de los eventos. En cualquier caso, de acuerdo con LePore y Loewer, *F* realiza a *M* justo en el caso de que: 1) *F* → *M* se mantenga con necesidad nomológica (LePore y Loewer hablan de “necesidad física”), y 2) *F* “explique” a *M*. Creo que LePore y Loewer están en lo correcto al sugerir estos dos *tipos* de condiciones. No obstante, es necesario desarrollar ciertos puntos específicos. Considérese primero el requerimiento

6 Véase Kim (1984a).

7 En particular, el “monismo anómalo” de Davidson, *sin* una tesis de supervenencia sería, en gran parte, inmune a mi argumento. Sin embargo, existen otras dificultades para esta posición; en particular, parece que no hay esperanza para una explicación de los poderes causales de las propiedades mentales en el monismo anómalo. Véase Kim (1989a).

explicativo: aunque creo que la idea que se encuentra detrás de esta condición es correcta, deberíamos buscar una relación objetiva metafísica entre F y M , no una relación esencialmente epistémica como la de explicación; esto es, deberíamos ver la relación explicativa entre las dos propiedades como estando soportada por una relación metafísica de realización. Aquí estoy tomando una actitud realista acerca de la explicación: si F explica a M , es porque se mantiene cierta relación metafísica y objetiva entre F y M . En las explicaciones causales, la relación requerida es obviamente la relación causal. En el caso de la realización, los conceptos clave, quiero sugerir, son los de *mecanismo causal* y *microestructura*. Cuando se dice que F “realiza” a M en el sistema s , F debe especificar una propiedad microestructural de s que provee un mecanismo causal para la implementación de M en s . Además, en los casos interesantes –de hecho, si vamos a hablar con sentido de la “implementación” de M –, F será un miembro de la familia de las propiedades físicas que forman una red de estados microestructurales que se conectan nomológicamente, y que proveen un mecanismo microcausal de sistemas apropiados como s para las conexiones nomológicas de un sistema mayor de propiedades mentales, dentro de las cuales M es un elemento. Estos microestados formarán una base explicativa para las propiedades superiores y su relaciones nómicas; pero la relación de realización como tal debe ser distinguida de la relación explicativa. Así, la diferencia que tengo con LePore y Loewer, en cuanto a su condición 2), es bastante reducida: estoy de acuerdo con que algo como su condición explicativa en general se debería sostener para la relación de realización; no obstante, no debería ser considerada como constitutiva de ella.

¿Qué pasa con la condición 1), acerca de que $F \rightarrow M$ sea nomológicamente necesario? Creo que esta condición es aceptable con la siguiente salvedad: en cada sistema s en el que M es realizado físicamente, debe haber un conjunto determinado, finito o infinito, de propiedades físicas F_1, F_2, \dots , cada una realizando a M en el sentido ya explicado, de tal forma que podamos considerar la propiedad disyuntiva $F_1 \vee F_2 \vee \dots$, como *una coextensión nómica de M*. Ya que en el presente artículo no haré uso de este requerimiento, no ofreceré un argumento a su favor.⁸

3. El fisicalismo no reduccionista como una forma de emergentismo

El fisicalismo no reduccionista que tengo en mente consta de las siguientes tesis:

8 Acerca de las razones para este requerimiento, véase Kim (1993a 309-335).

1. (*Monismo físico*) Todos los particulares concretos son físicos.
2. (*Antirreducciónismo*) Las propiedades mentales no son reducibles a las propiedades físicas.
3. (*Tesis de la Realización Física*) Todas las propiedades mentales son realizadas físicamente; esto es, siempre que un organismo o sistema instancie una propiedad mental *M*, este posee alguna propiedad física *F* que realiza a *M* en organismos de su tipo. Debemos adicionar una tesis que se encuentra implícita en las tres anteriores y que usualmente se sostiene como un supuesto verdadero:
4. (*Realismo mental*) Las propiedades mentales son propiedades reales de objetos y eventos, no son meras ayudas útiles para realizar predicciones o maneras ficticias de hablar.

Creo que estos cuatro principios básicos llevan al fisicalismo no reduccionista muy cerca del “emergentismo” –tan cerca, de hecho, que esta clase de fisicalismo no reduccionista se entiende mejor como una forma de emergentismo–. Explicaré brevemente por qué esto es así.⁹

En general los emergentistas aceptaron una ontología puramente materialista de objetos y eventos físicos concretos. Por ejemplo, Samuel Alexander, uno de los principales teóricos de la escuela emergentista, argumenta que no existen eventos mentales más allá de los procesos neuronales:

De esta forma nos damos cuenta, en parte por experiencia, en parte por reflexión, de que un proceso con la cualidad distintiva mental o consciente se encuentra en el mismo lugar y momento en el que se ubica un proceso neuronal, esto es, un proceso complejo y altamente diferenciado de nuestro cuerpo viviente. Estamos forzados, por lo tanto, a ir más allá de la mera correlación entre lo mental y un proceso neuronal, e identificarlos. Sólo hay un proceso que, teniendo una complejidad específica, posee la cualidad de la conciencia Debe ser entonces aceptado como un hecho empírico que un proceso neuronal de cierto nivel de desarrollo posee la cualidad de la conciencia y es, por lo tanto, un proceso mental; y, alternativamente, un proceso mental es *también* uno vital de cierto orden. (5-6)

Esto es, prácticamente palabra por palabra, lo que afirma el “fisicalismo de instancias” o la “teoría de la identidad de instancias”, una forma de fisicalismo no reduccionista. No es una sorpresa entonces que Alexander llame a su posición una versión de “la doctrina de la identidad de la mente y el cuerpo”. Tanto la estructura “estratificada”

.....

9 Para un análisis detallado, véase Kim (1992).

de la ontología emergentista como su carácter fisicalista fundamental son evidentes en el siguiente pasaje de C. Lloyd Morgan, otro líder del movimiento:

En la lección anterior se propuso la noción de una pirámide con niveles ascendentes. Cerca de su base encontramos un enjambre de átomos con una estructura relacional y una cualidad que podemos llamar atomicidad. Sobre este nivel los átomos se combinan para formar nuevas unidades que tienen la cualidad distintiva de ser moléculas; más arriba, en una de las líneas de avance existen cristales, digamos, donde átomos y moléculas son agrupados en nuevas relaciones, cuya expresión es la forma cristalina; en otra línea de avance existen organismos con un tipo diferente de relaciones naturales que les da la cualidad de la vitalidad; aún más arriba, un nuevo tipo de relación natural superviene, el cual podemos expresar con la palabra “mentalidad”. Si el vitalismo y el animismo implican la inserción de alguna entelequia, entonces serán excluidos. (35)

Los átomos y sus agregados mereológicos agotan toda la existencia concreta; ninguna “entelequia” o entidad que no sea realmente física será “insertada” en algún punto de la jerarquía de niveles de existencia, aunque nuevas propiedades emergen para caracterizar las estructuras de entidades básicas que se hacen más complejas. En esta imagen no hay lugar para una existencia concreta que no se pueda descomponer completamente en átomos y demás particulares físicos básicos.

La doctrina emergentista de que las propiedades “emergentes” son irreducibles a las “condiciones de base” a partir de las cuales emergen es familiar; para la mayoría de nosotros, la afirmación de la irreductibilidad es constitutiva de la cosmovisión metafísica emergentista. Aunque la idea emergentista de reducción o explicación reductiva difiere del modelo de reducción implícito en los argumentos antirreducciónistas actuales,¹⁰ la importancia filosófica del rechazo de la reductibilidad entre los dos niveles de propiedades es la misma: las propiedades de nivel superior, al ser irreducibles, realmente constituyen nuevas adiciones a la ontología del mundo. Alexander, por ejemplo, sostiene:

A partir de ciertas condiciones fisiológicas, la naturaleza ha estructurado una nueva cualidad mental que, en consecuencia, no es en sí misma fisiológica aunque vive, se mueve y tiene su ser en las condiciones fisiológicas. Es gracias a esto que puede haber, y hay, una ciencia

10 Para más detalles, véase Kim (1992).

independiente de la psicología Ninguna constelación fisiológica nos explica por qué debería existir la mente. (8)

Esta idea de propiedades superiores irreducibles se encuentra como base de algunas de las versiones recientes del emergentismo, como la que promueve el destacado neurofisiólogo Roger Sperry, quien escribe:

Primero, la conciencia ... es interpretada como una propiedad dinámica que emerge de la actividad cerebral. Como tal, la experiencia consciente se encuentra inseparablemente ligada al proceso cerebral material con todas sus restricciones estructurales y fisiológicas. Al mismo tiempo, las propiedades conscientes de la actividad cerebral son entendidas como siendo algo distinto y especial por derecho propio (533)

Entre otras de las implicaciones de la perspectiva en cuestión para la investigación cerebral, es la conclusión de que una explicación completa del proceso cerebral en el nivel consciente no será posible únicamente en términos de los datos bioquímicos y fisiológicos. (*id.* 535)

De esta forma, tanto para el emergentismo como para el fisicalismo no reduccionista la doctrina de las propiedades irreducibles de nivel superior es una pieza central en sus respectivas posiciones, y sus defensores la toman como lo que hace que sus perspectivas sean distintas e interesantes. Como adiciones netas al mundo, las propiedades emergentes de nivel superior no pueden ser reducidas o explicadas reductivamente; como rasgos nuevos irreducibles del mundo, forman un dominio autónomo y, en palabras de Alexander, hacen que sea posible “una ciencia independiente de la psicología”. Esto es exactamente lo que los actuales fisicalistas antirreducciónistas han estado defendiendo desde hace más de dos décadas: las “ciencias especiales” son autónomas e independientes de las ciencias físicas y biológicas subyacentes.¹¹

Pasemos ahora al tercer principio básico del fisicalismo no reduccionista, la Tesis de la Realización Física. Esta envuelve la afirmación

11 A pesar de todo esto, existe una aparente diferencia entre el emergentismo y el fisicalismo no reduccionista que concierne a la relación entre las propiedades que pertenecen a los niveles adyacentes. Como se recordará, LePore y Loewer sostienen que la base (“de realización”) física debe *explicar* la propiedad mental que realiza. Sin embargo, los emergentistas negarán que las “condiciones de base” puedan de alguna forma constituir una base explicativa de la propiedad que emerge a partir de ellas. Creo que gran parte de la diferencia se puede remontar a las distintas concepciones de explicación y reducción envueltas; para más detalles, véase Kim (1992). Esta diferencia, sea real o solo aparente, no afectará la aplicabilidad del argumento principal de este artículo tanto al emergentismo como al fisicalismo no reduccionista.

según la cual para que una propiedad mental sea instanciada en un sistema, tal sistema debe instanciar una propiedad física adecuada y, además, que cada vez que un sistema instancie esta propiedad física, instanciará igualmente la propiedad mental de una forma necesaria. Los eventos y estados mentales requieren bases físicas, y cuando las bases físicas requeridas se hacen presentes, estos deben ocurrir. Una tesis completamente paralela hizo parte de la doctrina emergentista: la emergencia de las propiedades de nivel superior requieren “condiciones de base” apropiadas, y cuando estas condiciones de base están presentes, estas deben emerger necesariamente. Tanto para el fisicalista antirreducciónista como para el emergentista las bases físicas son *por sí mismas* suficientes para la aparición de las propiedades de nivel superior. Como afirma Morgan, para la emergencia de las propiedades superiores no se requiere “ninguna inserción de entelequias” u otros agentes que no sean físicos:

Puesto que es bastante seguro que hablar de una cualidad emergente de la vida tiene un sabor vitalista, uno adicionalmente debe decir aquí, con el debido énfasis, que si el vitalismo connota algo de la naturaleza de una entelequia o *élan* –cualquier inserción dentro de la evolución fisicoquímica de una influencia extraña que debe ser invocada para explicar el fenómeno de la vida–, entonces, en la medida en que esto es implicado, el concepto de evolución emergente debe rechazarlo explícitamente. (12)

Además Morgan va a insistir en la necesidad de las bases físicas para todo fenómeno de nivel superior: “[a]sí, para la evolución emergente, los eventos conscientes en el nivel *c* (de la mente) envuelven eventos fisiológicos específicos en el nivel *b* (de la vida), y estos envuelven eventos fisicoquímicos específicos en el nivel *a* (de la materia). No puede haber algún *c* sin *b*, ni un *b* sin *a*. Ninguna mente sin vida; ni vida sin ‘una base física’” (*id.* 15)

No hay duda, creo, de que el emergentismo y el fisicalismo no reducciónista concuerdan ampliamente en estos tres principios cruciales, y que es justo e iluminador ver al fisicalismo no reducciónista como una forma de emergentismo. No es casualidad que los fisicalistas antirreducciónistas a veces hablen de las propiedades de nivel superior como “emergentes”.¹²

En cuanto a la cuarta tesis del fisicalismo no reducciónista, esto es, el Realismo Mental, podemos decir que es claro que las citas que

12 Por ejemplo, Hellman y Thompson dicen que “se permite lo que puede ser llamado ‘emergencia’ de fenómenos de órdenes superiores sin apartarse de la ontología física” (555).

hemos visto de los emergentistas están atravesadas por un realismo acerca de lo mental. Para el emergentista, la evolución emergente es un hecho histórico de importancia capital; a través del proceso de la evolución emergente, el mundo ha alcanzado su estado presente –más complejo, rico y lleno–. La mayoría de fisicalistas que rechazan el reduccionismo también rechazan el eliminativismo de lo mental; esta actitud realista se encuentra implícita en lo que hemos denominado la Tesis de la Realización Física (en la cual se asume el realismo acerca de lo físico). Se hace más explícita en la afirmación, aceptada por los emergentistas y muchos antirreducciónistas, de que la psicología es una ciencia especial legítima (a menos que uno adopte un antirrealismo acerca de la ciencia en general); como tal, debe investigar un dominio de fenómenos reales y sistematizarlos al descubrir las leyes y conexiones causales que los gobiernan.

4. Realismo y causalidad mental

Pero ¿en qué consiste exactamente el compromiso acerca de la realidad de las propiedades mentales? ¿Cuál es el significado de decir de algo que es “real”? Alexander proporciona una respuesta adecuada en un párrafo maravilloso en el que se rechaza secamente el epifenomenalismo:

[El epifenomenalismo] supone la existencia de algo en la naturaleza que no tiene ninguna consecuencia, ningún propósito al cual servir, una especie de *nobleza* que depende del trabajo de sus inferiores pero que se guarda por las apariencias y que igualmente podría, sin duda, llegar a ser abolida con el tiempo. (8)

Este es el que podemos denominar “el dictamen de Alexander”: *ser real es tener poderes causales*. Creo que este principio, aplicado a los existentes concretos y sus propiedades, será aceptado por la mayoría de los fisicalistas antirreducciónistas.

El énfasis en el rol causal de las propiedades emergentes es dominante en la literatura emergentista. Aquí hay un par de citas reveladoras:

Justo como las propiedades holistas del organismo tienen efectos causales que determinan el curso y destino de sus células y moléculas constituyentes, de la misma forma las propiedades conscientes de la actividad cerebral se entienden como teniendo efectos causales análogos en la función cerebral que controla el subconjunto de eventos en el patrón de flujo de la excitación neuronal. En este sentido holista, la propuesta presente puede ser entendida como colocando a la mente sobre la materia, pero no como un agente descorporizado o supernatural. (Sperry 533)

Pero cuando un nuevo tipo de relación es superveniente (digamos en el nivel de la vida), es en virtud de su presencia que los eventos físicos envueltos desarrollan un curso de forma diferente –distinto del que hubiera sido si la vida hubiera estado ausente–. (Morgan 16)

Lo llamativo de estos párrafos es la referencia a la “causalidad descendente”: Morgan y Sperry parecen estar sosteniendo que la mentalidad, una vez emerge a partir de procesos físico-biológicos, adquiere una vida causal propia y empieza *a ejercer influencia causal “descendente” al afectar lo que ocurre en los procesos físico-biológicos subyacentes*. El que la idea de tal causalidad tenga sentido es una de las cuestiones principales que deseo discutir en el resto de este artículo. Pero señalemos primero lo que el fiscalista no reduccionista tiene que decir acerca de la causalidad mental.

No hay duda de que el fiscalista no reduccionista típico tiene un fuerte compromiso con la realidad de la causalidad mental. Como hemos visto, nuestro antirreducciónista no es un eliminativista: ¿para qué molestarse con propiedades mentales a menos que se piense que son buenas para el trabajo causal y que pueden jugar un rol en las explicaciones causales? Fodor pone el punto de esta forma:

No estoy realmente convencido de que importe mucho si lo mental es físico; aun menos que importe mucho si podemos probarlo. Mientras tanto, si no es literalmente cierto que mi deseo es responsable causalmente de mi llegada, y que mi comezón es responsable causalmente de que me rasque, y que mi creencia es responsable causalmente de lo que digo ... si nada de esto es literalmente cierto, entonces prácticamente todo lo que creo acerca de cualquier cosa es falsa y es el fin del mundo. (Fodor 1989 77)

¡Uno difícilmente podría declarar con más sentimiento sus anhelos por la causalidad mental! Los fiscalistas antirreducciónistas en general consideran la causalidad mental con seriedad; han gastado mucha energía e ingenio tratando de demostrar que tienen derecho a ella.¹³ No creo que puedan tener lo que quieren, y en lo que queda de este artículo me dedicaré principalmente a mostrarlo. Mi argumento, si es correcto, también mostrará que Fodor se equivoca en sentir que puede satisfacer sus deseos sin preocuparse acerca del problema mente-cuerpo –acerca de “si lo mental es físico”–.

Este es el plan para lo que queda del artículo: mostraré que el emergentismo y el fiscalismo no reduccionista se comprometen con la causalidad descendente –esto es, la causalidad de la mente-a-lo-

13 Véase, por ejemplo, LePore y Loewer (1987 y 1989); Fodor (1989); Davidson (1993).

físico para el caso mente-cuerpo-. Estas no deberían ser noticias para el emergentista: en algún sentido, la causalidad descendente constituye gran parte de su programa. El argumento será de interés, incluso para el emergentista, pues mostrará claramente que la causalidad descendente se encuentra *implicada* por los postulados básicos del emergentismo y el fiscalismo no reduccionista. Pasaré entonces a argüir que la idea de causalidad descendente es bastante problemática, y quizás incoherente, dados los compromisos básicos fiscalistas.

5. El fiscalismo no reduccionista se compromete con la causalidad descendente

Es fácil ver cómo la causalidad descendente se sigue de los principios básicos del emergentismo y el fiscalismo no reduccionista. Primero, como hemos observado, el emergentista y el fiscalista antirreducciónista son realistas con respecto a la mente, y el Realismo Mental, vía el dictamen de Alexander, implica poderes causales para las propiedades mentales. De hecho, sea que acepten o no el dictamen de Alexander, la mayoría de ellos deseará poderes causales para las propiedades mentales. Ahora bien, según ambas posiciones, las propiedades mentales son adiciones netas e irreducibles del mundo. Esto debe significar, bajo el dictamen de Alexander, que las propiedades mentales traen consigo *nuevos poderes causales, poderes que las propiedades físico-biológicas subyacentes no pueden proporcionar*, pues, a menos que la mentalidad tenga contribuciones causales que son realmente nuevas, la afirmación de que es un fenómeno distinto e irreducible más allá del fenómeno físico-biológico será hueca y vacía. Ser real, como Alexander ha sostenido, es tener poderes causales; *por lo tanto, para ser real, nuevo e irreducible, se debe tener poderes causales nuevos e irreducibles*.

Esto encaja bien con la ya aludida tesis de la autonomía de la ciencia psicológica: como una ciencia empírica, la psicología debe producir explicaciones causales de los fenómenos de su dominio; y como una ciencia irreducible y autónoma, tales explicaciones deben ser igualmente irreducibles, al representar conexiones causales del mundo que no son capturadas por las ciencias subyacentes. En consecuencia, la tesis de la autonomía tiene sentido únicamente si las relaciones causales en las que los eventos mentales están envueltos son nuevas e irreducibles –esto es, en el caso de que las propiedades mentales estén dotadas con poderes causales completamente nuevos e irreducibles a aquellos de las propiedades físico-biológicas subyacentes–.

Si *M* es una propiedad mental, en consecuencia, *M* debe tener poderes causales nuevos. Esto debe significar, vamos a suponer, que *M* manifiesta sus poderes al ser eficaz causalmente con respecto a otra

propiedad, digamos N ; esto es, que una instancia dada de M causa que en tal ocasión se instancie N . Asumiremos aquí una concepción nomológica general de la causalidad, aproximadamente en el siguiente sentido: una instancia de M causa una instancia de N , solo en el caso de que exista una ley causal apropiada que invoque la instanciación de M como una condición suficiente para la instanciación de N . Existen tres casos que se deben distinguir: i) la propiedad N que es causada por M es una propiedad mental, ii) N es una propiedad física y iii) N es una propiedad de nivel superior en relación con M . El caso i) es causalidad mente-a-mente (“causalidad del mismo nivel”); el caso ii) es causalidad de lo mental-a-lo-físico (esto es, “causalidad descendente”); y el caso iii) es la posibilidad de que existan propiedades –quizá propiedades sociales– que emergan a partir de, o sean realizadas por, las propiedades mentales (“causalidad ascendente”).

Mi argumento mostrará que el caso i) es posible únicamente si el caso ii) lo es; a saber, que la causalidad mente-a-mente presupone la causalidad de lo mental-a-lo-físico. Quedará claro que el mismo argumento muestra que el caso iii) presupone el caso ii) y, por lo tanto, el caso i).¹⁴ Supongamos entonces que M es causalmente eficaz con respecto a alguna propiedad mental M^* y, en particular, que una instancia dada de M causa una instancia de M^* . Pero M^* , *qua* propiedad mental, es realizada físicamente; digamos que F^* es su base de realización física. Ahora, parece que tenemos dos respuestas distintas e independientes a la pregunta ¿por qué esta instancia de M^* se hizo presente? *Ex hypothesi*, está ahí porque una instancia de M la causó; es por esto que se hizo presente. Pero hay otra respuesta: está ahí porque F^* realiza físicamente a M^* , y F^* ha sido instanciada en esta ocasión. Creo que estas dos historias acerca de la presencia de M^* en este caso constituyen una tensión que debe ser resuelta.¹⁵

¿Es plausible suponer que la presencia conjunta de M y F^* sea responsable de la instanciación de M^* ? No, puesto que esto contradice la afirmación de que M^* es realizada físicamente por F^* . Como hemos visto, esta idea implica que F^* por sí sola es suficiente para producir M^* , sea que se obtenga o no alguna otra condición, anterior, posterior o simultánea (a menos que esté de alguna forma conectada con la ocurrencia de F^* retornaremos a esta posibilidad más adelante). Además la suposición es también inconsistente con nuestra presun-

14 Nota del traductor: Existe una ligera equivocación en la numeración anterior, ya que debería decirse: “Quedará claro que el mismo argumento muestra que el caso iii) presupone el caso i) y, por lo tanto, el caso ii)”.

15 Esta situación es esencialmente idéntica a la que enfrentamos cuando ofrecemos dos causas distintas e independientes para uno y el mismo evento, al afirmarse cada una como suficiente. Véase Kim (1989b).

ción inicial de que una instancia dada de M constituye una condición suficiente para la instanciación de M^* . Tampoco es plausible suponer que la ocurrencia de M^* en este caso se encuentra de alguna forma *sobre determinada*, en el sentido de tener dos orígenes distintos e independientes en M y F^* , pues esto también entra en conflicto con la presunción de que M^* es una propiedad que requiere una base de realización física para ser instanciada y que esta instancia de M^* se encuentra allí porque es realizada por F^* . En la ausencia de F^* , debemos suponer que M^* no podría haber ocurrido –a menos que una base de realización alternativa se hubiera hecho presente–. En cualquier caso, cada instancia de M^* debe tener alguna base física que sea por sí misma suficiente para su instanciación; y esta idea amenaza con anular la afirmación de que M es la causa de esta instancia de M^* .

Creo que la única historia coherente que podemos contar aquí es la de suponer que la instancia de M causa que M^* sea instanciada *al causar que se instancie F^* , la base de realización física de M^** . Esto es, por supuesto, causalidad descendente de M a F^* , un ejemplo de la causalidad de lo mental-a-lo-físico. Creo que el argumento funciona igualmente al remplazar “realización física” por “emergencia”.¹⁶ El punto central de mi argumento se encuentra encapsulado en el siguiente principio, que creo que será aceptado por la mayoría de fisiocalistas antirreducciónistas:

(*Principio causal de la realización*) Si una instancia dada de s ocurre al ser realizada por Q , entonces cualquier causa de esta instancia de s debe ser una causa de esta instancia de Q (y, por su puesto, cualquier causa de esta instancia de Q es una causa de esta instancia de s).

Muchos emergentistas, o quizá todos, aceptarán un principio correspondiente para la emergencia. En cualquier caso, pienso que aplicamos este principio constantemente en nuestra vida diaria: por ejemplo, tratamos el dolor interviniendo en los procesos corporales y nos comunicamos creando vibraciones en el aire o haciendo marcas sobre papel. (La causalidad directa mente-a-mente entre distintos individuos generalmente se considera desacreditada y poco científica: va con nombres como “percepción extrasensorial”, “telepatía” y “lectura de mentes”).

Ahora bien, ¿no podríamos evitar este compromiso con la causalidad descendente al aprovechar el hecho de que M , en cuanto que propiedad mental, tiene su propia base física de realización, digamos F ? ¿Por qué no decir entonces que el hecho de que M causa a F^* es simplemente esto: M es realizada físicamente por F , y F causa a F^* ? Se obtiene entonces la relación causal más básica entre las dos

16 Véase, Kim (1992)

propiedades físicas F y F^* , y el hecho de que M causa a M^* es fundamentado últimamente en la relación causal entre sus respectivas bases físicas de realización. Creo que esta es una imagen muy atractiva,¹⁷ pero que nuestros antirreducciónistas no pueden acoger. Pues la imagen reduce los poderes causales de M a los de su base de realización F : todo el trabajo causal está siendo realizado por F y la causalidad de M a F^* , o de M^* , resulta ser derivativa de los poderes causales de F . Así, M no tiene poderes causales que vayan más allá de los de F , lo que está en contra del dictamen de Alexander y la presunción de que M es una propiedad irreducible. En la siguiente sección volveré a este punto.

Lo que muestran estas consideraciones es que dentro del mundo estratificado del fisicalismo no reduccionista y del emergentismo la causalidad “dentro de un mismo nivel” puede ocurrir solo si es posible la causalidad “a través de distintos niveles”. No será posible aislar y confinar las cadenas causales dentro de los niveles; se encontrarán inevitables fugas de influencia causal de un nivel a otro.

6. ¿Cuál es el error de la causalidad descendente?

Entonces, ¿tiene sentido la causalidad descendente dentro del esquema del fisicalismo no reduccionista? Creo que existen problemas serios. Como veremos, la tensión surge del intento de combinar la “determinación ascendente” con la “causalidad descendente”. El fisicalista antirreducciónista desea ambas cosas: la mentalidad está determinada por y es dependiente de lo físico, y aun las mentes tienen poderes causales nuevos que pueden ser ejercidos, si mi argumento es correcto, solo al afectar causalmente los procesos físico-biológicos en formas novedosas.

Supóngase entonces que la propiedad mental M es eficaz causalmente con respecto a la propiedad física F^* y en particular que una instancia dada de M causa una instancia particular de F^* . Dada la Tesis de la Realización Física, esta instancia de M ocurre porque es realizada por una propiedad física, digamos F . Puesto que F es una base de realización de M , es suficiente para M , y se sigue como una cuestión nomológica que F es suficiente para F^* . Ahora, la cuestión que debe enfrentarse es esta: ¿cuál es la razón para no tomar a F como la causa de F^* , evitando a M y tratándola como un epifenómeno?¹⁸

17 Similar al modelo de “causalidad superveniente” que he sugerido en artículos anteriores, como por ejemplo “Epiphenomenal and Supervenient Causation” (1984b). Peter Menzies lo ha denominado “reducción causal”, en *Against Causal Reductionism* (1988). Véase también LePore y Loewer (1989), y Gabriel Segal y Elliott Sober (1991).

18 Para ser precisos deberíamos poner esto en términos de las instancias de estas propiedades en lugar de las propiedades mismas. En lo que sigue, se tomarán libertades de este tipo para evitar verbosidad.

Creo que esta solución epifenomenalista con respecto a M no puede ser dejada a un lado tan fácilmente. Estamos buscando una explicación causal acerca de por qué F^* es instanciada en este momento. Vemos que M fue instanciada y podemos invocar una ley que conecte las instancias de M con las de F^* . Pero también vemos que F se instanció en el mismo momento y que hay una ley apropiada que conecta las instancias de F con las de F^* . En consecuencia, la situación es esta: parece que la afirmación de que F es una causa directa de F^* es al menos tan fuerte como la de que M lo es (además, sin tener a M como un eslabón intermedio). ¿Hay realmente alguna razón para invocar a M como una causa de F^* ? La cuestión no es si F debería ser considerada la causa de F^* ; en cualquier explicación debería serlo. Más bien, la cuestión es acerca de si en esta situación M debería tener un rol causal distinto. Creo que hay razones persuasivas para rechazar esta idea.

Inicialmente, está el buen principio antiguo de la simplicidad: nos podemos conformar con F como la causa de F^* , de tal forma que ¿por qué molestarse con M ? Nótese que dada la simultaneidad de las instancias respectivas de M y F , no es posible pensar la instancia de M como un eslabón intermedio en la cadena causal que va de F a F^* . Además, si insistimos en M como una causa de F^* , nos encontraremos con otra seria dificultad, a saber, “el problema de la exclusión causal-explicativa”,¹⁹ pues estaríamos aceptando que un solo evento tiene dos causas que son suficientes, distintas y simultáneas. Esto hace que la situación se vea como un fenómeno de sobredeterminación causal, lo que es absurdo. Ahora, *ex hypothesis*, no es posible considerar que M y F forman una causa conjuntamente suficiente, siendo cada una necesaria pero insuficiente. Dada la supuesta irreductibilidad de M , no podemos considerarla como idéntica a F o como una de sus partes. Este es entonces el problema de la exclusión: dado que F es una causa física suficiente para F^* , ¿cómo es que M también podría ser una causa suficiente de F^* ? ¿Cuál es el trabajo causal que queda para M u otra propiedad mental? La afirmación de que M es una causa de F^* será debilitada especialmente si, como podríamos esperar de la investigación neurobiológica en la vida real, se encuentra una cadena causal continua, un mecanismo que conecte a F con F^* . Es claro que el problema de la exclusión no puede ser resuelto dentro del marco del fiscalismo no reduccionista.

Quiero sugerir que todas estas consideraciones señalan que la imagen natural del mundo fiscalista estratificado es la siguiente: todas las relaciones causales se implementan en el nivel físico, y las relaciones causales que se atribuyan a los procesos de nivel superior

19 Para más detalles, véase Kim (1989a y 1990b).

son derivadas de y fundamentadas en los procesos nómicos básicos del nivel físico.²⁰ Esto va tal vez un poco, pero no mucho, más allá de lo que implica directamente la tesis de la supervenencia que la mayoría de los fisicalistas antirreduccionistas aceptan: si, como la tesis de la supervenencia afirma, todos los hechos están determinados por los hechos físicos, entonces todas las relaciones causales que envuelven eventos mentales deben estar determinadas por los hechos físicos (incluyendo presumiblemente hechos acerca de la causalidad física).

Consideremos, pues, una forma un tanto cruda de afirmar esta idea:

(*Principio de Herencia Causal*) Si M es instanciada en cierta ocasión al ser realizada por la propiedad F , entonces los poderes causales de esta instancia de M son idénticos a (quizás un subconjunto de) los poderes causales de F .

En otras palabras, los estados superiores heredan sus poderes causales de los estados subyacentes que los realizan. Los antirreduccionistas deben rechazar este principio; dirán que los poderes causales de nivel superior son “determinados por”, aunque no son idénticos (o reducibles) a, los poderes causales de nivel inferior. Lo que nuestras consideraciones hacen claro es que si “determinados pero no idénticos” significa que los poderes causales superiores son realmente nuevos, entonces los antirreduccionistas se encuentran atrapados en una maraña de dificultades aparentemente insuperables. Desafío a los antirreduccionistas que rechazan este principio a formular un principio alternativo sobre, justamente, cómo los poderes causales de una propiedad realizada están conectados con los de su base de realización; o a explicar, si tal conexión no es imaginada, el significado de la realización.

Las implicaciones del Principio de Herencia Causal son devastadoras para el fisicalismo no reduccionista: si los poderes causales de M son idénticos a los de sus bases de realización, en efecto M no tendrá ninguna contribución causal nueva y la afirmación de que M es una propiedad nueva e irreducible se pone en entredicho. Como he sugerido, si M es tratada como un epifenómeno que emana de su base de realización física sin que tenga ninguna función causal propia, el siguiente paso del argumento que ordena el dictamen de Alexander es que M debe ser “abolida”. Todo esto parece ser una inescapable lección que se sigue del dictamen de Alexander.

El escepticismo acerca de la causalidad descendente se fortalece cuando vemos que se infringe el cierre causal del dominio físico, tal y como sucede en el interaccionismo cartesiano. Lo que es peor, cuando vemos que la base de la realización de M , esto es, F , se encuentra como

20 Existe un indicio fuerte de que Fodor, por ejemplo, acepta esta clase de principio en su “Making Mind Matter More” (1989).

una causa completa de *F** –lo que siempre será el caso cuando se invoque una causa mental–, donde incluso la violación no se encuentra tan bien motivada como en el interaccionismo cartesiano.

La mayoría de los emergentistas no tendrán problema con la falla del cierre causal físico; aunque quizás tengan que enredarse de alguna forma con sus doctrinas para asegurar la consistencia general de su posición, es probable que no derramen una lágrima por el destino de este principio. Para muchos emergentistas esta es precisamente la consecuencia esperada de su posición. A pesar de esto, dudo que los fisicalistas antirreducciónistas contemporáneos puedan darse el lujo de ser tan displicentes con el problema del cierre causal: dejar de lado este principio es aceptar que ni siquiera en principio puede existir una teoría física completa de los fenómenos físicos; que la física teórica, por cuanto aspira a ser una teoría completa, debe dejar de ser física pura e invocar poderes causales irreduciblemente no-físicos –principios vitales, entelequias, energías psíquicas, *élan vital*, o lo que sea–. Si esto es lo que están dispuestos a aceptar, ¿por qué llamarse a sí mismos “fisicalistas”? Su teoría básica del mundo tendrá que ser mixta, una teoría físico-mental similar a la del interaccionismo cartesiano. Y todo esto puede comprometer la imagen estratificada del mundo; incluso es probable que requiera un serio replanteamiento.

En esta coyuntura parece altamente plausible que la única solución a los problemas de la exclusión y del cierre causal físico se encuentre en alguna forma de reducciónismo que nos permita deshacernos de, o al menos moderar, la afirmación de que las propiedades mentales son distintas a sus propiedades físicas subyacentes. Esta afirmación es la que nos ha obligado a postular poderes causales nuevos y distintos para la mente, y nos ha llevado así a la difícil situación actual. En efecto, identificar los poderes causales de la mente con los de sus bases físicas subyacentes es negar su estatus ontológico distinto y considerarla reducida.

Pero si se está familiarizado con la filosofía de la mente y su sabiduría popular, en este momento debería saltar una pregunta a la mente: dado el fenómeno de la “realizabilidad múltiple”, ¿la Tesis de la Realización Física no excluye por sí misma toda forma de reducciónismo? ¿“El argumento de la realización múltiple” no ha refutado al reducciónismo de una vez por todas?²¹ La respuesta arraigada, casi automática, a ambas preguntas es “sí”. Creo que la respuesta correcta es un cualificado pero firme “no”. Sin embargo, defender esta respuesta es algo que debo dejar para otra ocasión.²²

21 Véase Putnam (1967) y Fodor (1974).

22 Para más detalles, véase Kim (1993a). Para un mayor desarrollo, véase Kim, (1993b).

Bibliografía

- Alexander, S. *Space, Time, and Deity*. Vol. 2. London: Macmillan, 1927.
- Davidson, D. "Mental Events". *Essays on Actions and Events*. Oxford: Oxford University Press, 1980.
- Davidson, D. "Thinking Causes". *Mental Causation*. Eds. Heil, J. and Mele, A. Oxford: Clarendon Press, 1993. 3-17.
- Davidson, D. "Sucesos mentales". *Ensayos sobre acciones y sucesos*. Barcelona: Crítica, 1995. 263-287.
- Fodor, J. "Special Sciences, or the Disunity of Science as a Working Hypothesis". *Synthese* 28.2 (1974): 97-115.
- Fodor, J. "Making Mind Matter More". *Philosophical Topics* 17.1 (1989): 59-79.
- Hellman, G. and Thompson, F. "Physicalism: Ontology, Determination, and Reduction". *Journal of Philosophy* 72 (1975): 551-564.
- Kim, J. "Concepts of Supervenience". *Philosophy and Phenomenological Research* 65 (1984a): 153-176.
- Kim, J. "Epiphenomenal and Supervenient Causation". *Midwest Studies in Philosophy* 9.1 (1984b): 257-270.
- Kim, J. "The Myth of Nonreductive Materialism". *Proceedings of the American Philosophical Association* 63.3 (1989a): 31-47.
- Kim, J. "Mechanism, Purpose and Explanatory Exclusion". *Philosophical Perspectives* 3 (1989b): 77-108.
- Kim, J. "Supervenience as a Philosophical Concept". *Metaphilosophy* 21.1-2 (1990a): 1-27.
- Kim, J. "Explanatory Exclusion and the Problem of Mental Causation". *Information, Semantics, and Epistemology*. Ed. Villanueva, E. Oxford: Basil Blackwell, 1990b. 77-108.
- Kim, J. "'Downward Causation' in Emergentism and Nonreductive Physicalism". *Emergence or Reduction?* Eds. Beckermann, A., Flohr, H. and Kim, J. Berlin: De Gruyter, 1992. 119-138.
- Kim, J. "Postscripts on Mental Causation". *Supervenience and mind. Selected philosophical essays*. New York: Cambridge University Press. 1993b. 358-367.
- LePore, E. and Loewer, B. "Mind Matters". *Journal of Philosophy* 93 (1987): 630-642.
- LePore, E. and Loewer, B. "More on Making Mind Matter". *Philosophical Topics* 17.1 (1989): 175-191.
- McLaughlin, B. "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical". *Philosophical Perspectives* 3 (1989): 109-136.
- Menzies, P. "Against Causal Reductionism". *Mind* 97.388 (1988): 551-574.
- Morgan, C. *Emergent Evolution*. London: Williams and Norgate, 1923.

- Oppenheim, P. and Putnam, H. "Unity of Science as a Working Hypothesis". *Minnesota Studies in the Philosophy of Science*. Vol. 2. Eds. Feigl, H., Scriven, M. and Maxwell, G. Minneapolis: University of Minnesota Press, 1958. 3-36.
- Putnam, H. "Minds and Machines". *Dimensions of Mind*. Ed. Hook, S. New York: New York University Press, 1960. 57-80.
- Putnam, H. "Psychological Predicates". *Art, Mind and Religion*. Eds. Capitan W. and Merrill, D. Pittsburgh: University of Pittsburgh Press, 1967. 37-48.
- Segal, G. and Sober, E. "The Causal Efficacy of Content". *Philosophical Studies* 63.1 (1991): 1-30.
- Sperry, R. "A Modified Concept of Consciousness". *Psychological Review* 76.6 (1969): 532-536.