

V Jornadas de Investigación en Filosofía. Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación. Departamento de Filosofía, La Plata, 2004.

Fisicalismo y exclusión explicativa.

Herrera, Maximiliano.

Cita:

Herrera, Maximiliano (2004). *Fisicalismo y exclusión explicativa*. V *Jornadas de Investigación en Filosofía*. Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación. Departamento de Filosofía, La Plata.

Dirección estable: <https://www.aacademica.org/000-094/61>

ARK: <https://n2t.net/ark:/13683/ebYP/aXp>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<http://creativecommons.org/licenses/by-nc-nd/2.5/ar/>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. *Acta Académica* fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

FISICALISMO Y EXCLUSIÓN EXPLICATIVA

Maximiliano Herrera

UNLP

Una de las empresas teóricas más importantes en la filosofía de la mente es el intento de preservar alguna característica distintiva de lo mental dentro del mundo físico. El fisicalismo no reductivo, una posición que intenta retener a lo mental qua mental sin caer en el dualismo, es hoy poco menos que la posición ortodoxa.

Sin embargo ciertos aspectos de lo mental siguen siendo difíciles de salvar, uno de ellos es la eficacia causal de lo mental. Jaegwon Kim ha argumentado que el fisicalismo no reductivo no puede tener éxito en su empresa, y que cualquier fisicalista que espere salvar la eficacia causal de lo mental debe aceptar el reduccionismo.

Tal vez el argumento de Kim más influyente a estos efectos sea el de la exclusión causal. El objetivo de este trabajo es bosquejar y evaluar brevemente el argumento de Kim, así como mencionar las alternativas posibles al problema generado por la exclusión.

El problema de la exclusión puede presentarse exponiendo un conjunto de tesis, cada una de las cuales es altamente plausible, pero que tomadas en conjunto generan inconsistencia.¹

[1] *la clausura causal de lo físico*: Todas las propiedades físicas conforman un sistema causalmente cerrado. Por lo que todo evento físico ha de tener una causa física.

[2] *la no identidad de lo mental y lo físico*: junto a las propiedades físicas existen propiedades mentales que no son ni idénticas ni reducibles a las propiedades físicas.

[3] *la dependencia de lo mental de lo físico*: las propiedades mentales no flotan libremente. Están relacionadas con las propiedades físicas y dependen de ellas.

[4] *la eficacia causal de lo mental*: las propiedades mentales tienen poderes causales.

La tesis [1] forma parte del núcleo conceptual del fisicalismo. No aceptar la clausura causal de lo físico, permitiría la intromisión de entidades misteriosas como

¹ Esta es una manera habitual de presentar el problema, cfr. Horgan (2001), Loewer (2002), Menzies (2001).

causas de ciertos eventos físicos. De esto se deriva que las propiedades físicas por sí solas son suficientes para causar eventos físicos. Y, más aun, podríamos concluir que todas las propiedades causales son propiedades físicas. Si se acepta esto la contradicción con la tesis [4] es evidente. Por lo tanto quien quiera mantener juntas las tesis 1-4 deberá encontrar alguna forma de compatibilizar la posibilidad de que las propiedades mentales tengan genuinos poderes causales en un mundo que es fundamentalmente físico. Este es, resumido, uno de los problemas que se presentan a la hora de considerar a la acusación mental dentro del fisicalismo no reductivo. Pero no es todavía el problema de la exclusión tal como Kim lo presenta. Para introducir el problema consideremos el marco que ha quedado bosquejado con la enunciación de las cuatro tesis.

El fisicalismo es la tesis según la cual todo lo que existe en el mundo es, en última instancia, físico. Pero aun así, puede asumirse la existencia de propiedades con cualidades distintivas, como las propiedades mentales. La forma de conectar estas dos clases de propiedades es a través de una relación de dependencia entre ambas. Tal vez la noción preferida para entender esta relación sea la noción de superveniencia.

Superveniencia: Las propiedades mentales supervienen sobre las propiedades físicas si, necesariamente, para cualquier objeto x y para cada propiedad M que x posea, existe una propiedad física de base P , tal que x posee P , y siempre que se da P se da M

La definición de superveniencia puede usarse para formular mas concretamente la tesis [3], a la vez que permite dar cuenta de una versión del fisicalismo que mantenga las dos primeras tesis. En primer lugar, todo lo que existe lo hace en virtud de ser (o de depender de lo) físico. Una vez que se han determinado las propiedades de base, quedan fijadas todas las propiedades que habrán de sobrevenir sobre ellas. En segundo lugar, la superveniencia es una relación que no afirma la identidad entre las propiedades de base y las propiedades supervenientes. Una vez que hemos aceptado esta imagen de la relación entre lo mental y lo físico, parece que podemos aceptar que esas propiedades mentales poseen poder causal.

Lo primero que hay que notar es que la posibilidad de que un evento mental sea causa de otro evento mental, presupone que un evento mental puede ser causa de un evento físico.² Supongamos que la propiedad mental M causa otra propiedad mental M^* . Ahora bien, de acuerdo a la tesis [3], lo mental depende de lo físico, y la noción de

² Kim (1998) p. 43

supervenencia que hemos aceptado como la mejor manera de entender esta dependencia, dice que debe existir una propiedad física de base que sea tenida siempre que se tiene una propiedad mental. Por lo tanto M* debe tener como base una propiedad física F*, y para que M pueda ser causa de M* debe causar aquella propiedad física F*. Pero que M* pueda ser causalmente responsable de F* contradice la tesis [1], según la cual todo evento (o propiedad) físico ha de tener una causa física.

¿Qué propiedad física podría ser causa de F*? La respuesta parece obvia: M posee una propiedad física de base. Y es esta propiedad F la que puede considerarse como causa de F*. Dado que al responder así le negamos toda eficacia causal a lo mental, la pregunta es ¿puede haber alguna manera en la que M y F realicen el trabajo causal juntas? Esto es lo que niega precisamente el Principio de Exclusión.

Principio de exclusión: No puede existir más que una causa completa e independiente del mismo evento.

La idea detrás de este principio es que si un evento tiene una causa suficiente, ningún otro evento puede ser causa suya. Pensar que M y F pueden ser ambas causas de F* es un caso de sobredeterminación causal, y existen razones para rechazar la sobredeterminación. En principio debemos suponer que ambas son causas independientes. Esto implica, por un lado, que si la causa física no se hubiera dado, entonces la causa mental por sí sola habría sido causalmente responsable. Como hemos visto, se viola la clausura causal de lo físico. Por otro lado si la causa mental no hubiese ocurrido entonces la causa física habría hecho todo el trabajo. Estos dos casos contrafácticos muestran la implausibilidad de la sobredeterminación.

Una primera conclusión es que si se acepta el principio de exclusión y la clausura causal de lo físico debemos considerar que es la propiedad física F la única causa de F*.

Por lo tanto si se acepta el marco del fisicalismo no reductivo, debemos concluir que lo mental resulta causalmente ineficaz.

Existen estrategias alternativas que buscan resolver el problema planteado por el argumento de la exclusión. Dos de estas aceptan el principio de exclusión, y por lo tanto aceptan que el fisicalismo no reductivo y la causación mental no pueden coexistir. La tercera estrategia mantiene la compatibilidad del no reduccionismo y la eficacia causal de lo mental. Retomemos las tesis presentadas al principio para bosquejar un mapa de

estas alternativas.³ De acuerdo a que tesis se niegue tendremos una posición sustantiva distinta.

La primera opción es negar la clausura causal del mundo físico. Quien rechaza esta tesis se compromete con el emergentismo. Si bien puede considerarse que el emergentismo es compatible con el fisicalismo dado que las propiedades mentales dependen de las propiedades físicas, el problema surge cuando evaluamos sus poderes causales. Las propiedades emergentes tienen poderes causales que son distintos de los poderes causales de las propiedades físicas de las que dependen. Dado que como hemos visto la causación de lo mental sobre lo físico presupone la causación de lo físico sobre lo mental, las propiedades emergentes harían todo el trabajo causal solas. Es decir, los poderes causales de M no dependen de F, y por tanto, M puede por si sola causar tanto M* como F*:

Una segunda opción es rechazar la tesis [2], esto es negar la no identidad de lo mental y de lo físico. La posición típica es la Teoría de la identidad que mantiene la eficacia causal de las propiedades mentales, porque estas son idénticas a propiedades físicas⁴.

Estas dos opciones son formas de un incompatibilismo que mantiene la eficacia causal de lo mental, pero termina rechazando el fisicalismo no reductivo.

La tercera estrategia, que ejemplifica la otra clase de incompatibilismo, consiste en negar la eficacia causal de lo mental. Esto es lo que define al epifenomenalismo. Las propiedades mentales son causadas por propiedades físicas pero ellas no causan nada. También puede negarse la eficacia causal pero mantener cierta relevancia causal de lo mental dentro del marco del no reduccionismo.

La última alternativa consiste en negar que todas las propiedades causales sean propiedades físicas. No se niega que el mundo sea un sistema físico cerrado, sino que se acepta que junto a las causas físicas existen las causas mentales. Esta es la estrategia compatibilista propia de todo aquel que intenta mantener el marco original del no reduccionismo.⁵

Dejemos de lado las estrategias incompatibilistas, y veamos como defender la posición compatibilista.

³ Para una presentación de estas diferentes estrategias, cfr. Sabatés (2001)

⁴ Cfr. Por ejemplo Bickle (1998) y Crane (1995)

⁵ Cfr. Sabatés (1997)

La cuestión central para el compatibilista es encontrar una opción donde M y F puedan ser ambas causas de F*. Como hemos visto esto es un caso de sobredeterminación causal. Barry Loewer (2002) ha argumentado que ninguna de las razones dadas por Kim para rechazar la sobredeterminación es persuasiva. En primer lugar Kim trata a todos los casos de sobredeterminación según el modelo del ejemplo de los dos asesinos (donde ambos disparan a la vez y causan conjuntamente la muerte de un tercero). En este caso las dos causas son metafísica y nomológicamente independientes. Pero hay que distinguir estos casos de los casos típicos postulados por el fisicalismo no reductivo. Un ejemplo de estos casos es: un evento neurofisiológico N causa un evento mental M y ambos causan E. en este caso M depende de N dado que es físicamente implicado por N, por lo que ambos no pueden independientes. El principio de exclusión causal dice que ningún evento no puede tener dos causas independientes por lo que no se aplicaría a estos casos.

En segundo lugar según Loewer, Kim piensa la causación como una relación en la cual la causa genera o produce el efecto. Si N y M producen E, deben hacerlo independientemente, y el argumento contra la determinación depende de esta noción de causación.

Loewer propone usar una versión modificada del análisis contrafáctico de David Lewis. La idea central es que M causa E si M y E son eventos distintos, y si M no se hubiera dado no se habría dado E. Esta idea permite la sobredeterminación: F es causa de E, i.e., si F no se hubiera dado no se habría dado E; M depende de F porque es físicamente implicado por F; y M cuenta como causa de E porque existe una cadena de dependencia contrafáctica entre F y E. La relación $M > E$ se sigue por transitividad a partir de $F > E$ y $F > M$.

Terry Horgan (2001) también acepta una concepción contrafáctica de la causalidad. La idea central es que el argumento de Kim considera que las propiedades son causales o no de manera absoluta. Horgan propone considerar a las nociones de causación y de explicación causal como gobernadas por parámetros contextuales implícitos, que involucran patrones sistemáticos de dependencia contrafáctica. Un mismo fenómeno puede ser explicado por una variedad de explicaciones causales diferentes, que incluyen una diversidad de patrones de dependencia contrafáctica en diferentes niveles ontológicos y descriptivos. Y estas diferentes explicaciones no compiten entre sí, ni se excluyen.

Por otro lado, los parámetros contextuales determinan que clase de explicación causal es apropiada en cada contexto, y del mismo modo determinan la propia noción de propiedad causal. Las propiedades que cuentan como causales son aquellas que figuran en la clase relevante de explicación causal.

Estas ideas le permiten a Horgan tratar la sobredeterminación: las propiedades mentales y las propiedades físicas no sobredeterminan causalmente sus efectos porque ellas figuran en distintos patrones de dependencia contrafáctica en distintos niveles teóricos.

Lo que Kim objeta a las concepciones contrafácticas de la causalidad es que no pueden distinguir entre los casos contrafácticos genuinos y casos epifénomenicos. Según Kim ni la causación $M \rightarrow M^*$ ni $M \rightarrow F^*$ son genuinas relaciones causales, sino que son análogas a las sombras arrojadas por un auto en movimiento, es decir, no hay conexiones causales entre la sombra de un auto en un momento dado y la sombra del auto un instante después. M solo cumple un papel epifenómeno, dado que F se anticipa a M como causa de F^* .

Por otra parte en los análisis contrafácticos se supone que M depende de F y es causa de F^* porque existe una cadena de dependencia contrafáctica entre F y F^* . Pero es plausible pensar que esta dependencia de M sobre F no es una relación causal, por lo que la cadena de dependencia no sería una cadena causal. Y si asumimos, como en el análisis de Loewer, que M es múltiple realizable sobre diferentes bases físicas, podemos pensar que no todas esas propiedades de base podrían haber causado F^* . La idea es que toda propiedad mental es causalmente eficaz, independientemente de las peculiaridades físicas del sistema en la que aquella está instanciada. Pero comúnmente las generalizaciones que se refieren a propiedades múltiplemente realizables involucran cláusulas *caeteris paribus*, destinadas a prevenirnos contra ciertas excepciones a esas generalizaciones. Pero entonces estas generalizaciones se refieren a un grupo relevante de casos. Alguno de los realizadores físicos de M podría no pertenecer a este grupo relevante. Dicho con más precisión, sean (F1, F2, Fn) los realizadores físicos de M; y si tenemos una generalización como "Los Ms causan los Es, *caeteris paribus*", estamos suponiendo la verdad de ciertos enunciados como "F1 causa E", "F2 causa E", pero también que hay un cierto realizador, Fn, que no entra en relación causal con E.

En cuanto a la respuesta de Horgan, la caracterización de las nociones de causalidad y de explicación causal como relativas a intereses y sensibles al contexto,

parecen comprometernos con el irrealismo explicativo. Es una intuición común entre quienes defienden el realismo de lo mental, suponer que el hecho de que haya genuinas relaciones causales que involucren propiedades mentales es lo que permite la existencia de explicaciones que mencionen términos mentales. Pero si consideramos, como Horgan, que una propiedad es causal solo si figura en una clase relevante de explicación, perdemos buena parte de esa intuición. Y si bien podemos mantener explicaciones que mencionen términos mentales por cuestiones metodológicas, nada nos impide pensar que no se trata de genuinas explicaciones, sino sólo de racionalizaciones de ciertos procesos.

En mi opinión el compatibilismo no es una buena estrategia para resolver el problema de la exclusión, y esto por tres motivos: 1) los análisis contrafácticos de la causación son dudosos, 2) es implausible que exista una masiva sobredeterminación, y 3) algunas de las respuestas, como la de Horgan, dejan sin tratar los aspectos más difíciles del problema.

Parece, entonces, que debemos aceptar alguna de las estrategias incompatibilistas. Asumiendo que el emergentismo no es una posición deseable, nos quedan o mantener la identificación de lo mental con lo físico o negar la eficacia causal, manteniendo una cierta relevancia explicativa de lo mental. Mas allá de la opción que uno elija, el desafío de Kim muestra que el fisicalismo no reductivo no puede mantener el marco original, que había sido diseñado para resolver los problemas de la causación mental.

Bibliografía

Bickle, J. (1998) *Psychoneural Reduction. The New Wave*: Cambridge, Mass. The MIT Press.

Crane, T. (1995) "The Mental Causation Debate" *Proceedings of the Aristotelian Society* LXIX.

Horgan, T. (2001) "Causal Compatibilism and the Exclusion Problem" *Theoria* 16/40: 95-116.

Kim, J. (1989 a) "Mechanism, Purpose, and Explanatory Exclusion" en Tomberlin, J. (ed) *Philosophical Perspectives vol. 3*: Atascadero, Cal. Ridgeview Press.

- Kim, J. (1989b) "The Myth of Non reductive Materialism" *Proceedings and Address of the APA* 63: 31-47.
- Kim, J. (1998) *Mind in a Physical World*: Cambridge, Mass. The MIT Press.
- Kim, J. (2002) "Responses" *Philosophy and Phenomenological Research* LXV/3: 671-680.
- Loewer, B. (2002) "Comments on Jaegwon Kim's *Mind in a Physical World*" *Philosophy and Phenomenological Research* LXV/3:655-662.
- Menzies, P. (2001) "The Causal Efficacy of the Mental States" en Monnoyer, (ed) *The Structure of The World*: Vrin Publishers.
- Sabatés, M. (1996)"Kim on the Metaphysics of Explanation" *Manuscrito* XIX/2: 93/110.
- Sabatés, M. (1997) "Should a Cognitive Psychologist Worry about the causal Inefficacy of the Mental?" en Niggemeyer, B: (ed.) *The Cognitive Level*: Duisburg, LAUD Verlag
- Sabatés, M. (2001) "Varieties of Exclusion" *Theoria* 16/40:13-42.
- Sabatés, M. (2002) "*Mind in a physical World?*" *Philosophy and Phenomenological Research* LXV/3:663-670.