
LIBERTAD Y DETERMINISMO: A PROPÓSITO DE LA RELEVANCIA CAUSAL DE LO MENTAL

VICENTE RAGA ROSALENY¹

ABSTRACT. One extended belief is that of our freedom, one that lives side by side with an image of a deterministic world. In cognitive studies, this question has to do with causal physicalism. I intend to reveal here that this does not hold, by taking Fodor's twin earths argument and showing that its conclusions collide with the notion of multiple actuality, one postulated for mental properties by reason of their apparent similarity to the functional ones. However, I will conclude that the difficulties posed by my critique corroborate the strong image of a physicalist world.

KEY WORDS. Physicalism, cognitivism, determinism, freedom, mental properties, functional properties, multiple actuality, restricted correlate, causality, context.

I.
Constantemente, en el curso de mi vida, he tenido la experiencia de creer y desear cosas diversas, y siempre me ha parecido, tanto que mi deseo tenía relevancia causal, como que tenía y tengo posibilidades alternativas a aquello que de hecho he acabado haciendo. En suma, subyaciendo a ambas convicciones, tengo la idea de que soy libre.

Sin embargo, por desgracia, junto a estas halagüeñas consideraciones convive una poderosa imagen del mundo que chocaría con esta intuición primera. Me refiero a la concepción determinista del mundo, según la cual, dado el conjunto completo de condiciones antecedentes en el momento *t* y las leyes de la física, no podría hacer otra cosa que lo que de hecho hago.

Esta antinomia se resuelve habitualmente en el campo de la filosofía de la mente en virtud del influjo del punto de vista dominante, el del fisicismo causal, que sostiene dos principios que se pretenden metafísicamente incuestionables: la clausura causal del mundo físico (esto es, que todo efecto físico tiene una explicación completa en términos físicos) y la exclusión explicativa (que buscando evitar la sobredeterminación postula

Facultat de Filosofia, Departament de Metafísica i Teoria del Coneixement, Universitat de València, España. / vicente.raga@uv.es

la no existencia de dos o más explicaciones, completas e independientes, del mismo efecto).

Pues bien, mi deseo sería poner en cuestión la estrategia de la posición dominante, que busca armonizar sus consideraciones relativas al estatus de las propiedades físicas, como propiedades básicas del mundo, con la relevancia causal autónoma de lo mental. Para ello presentaré una convincente crítica a esta imagen de la clausura causal del mundo físico, con la consecuente defensa, aunque de modo distinto a como lo plantea el fisicismo causal, que concluye en dependencia, de la relevancia causal de lo mental y, con ello, de la posibilidad de la libertad. Por último, cuestionaré alguno de los argumentos y supuestos de esta crítica, que sigue el hilo conductor de algunas de las ideas centrales de la obra de Corbí y Prades, *Minds, Causes and Mechanisms*, para así mostrar la fuerza de la imagen del mundo fisicista, que tan cerca está de nuestras intuiciones básicas (así como las dificultades de una plausible noción de libertad).

Para llegar a estas conclusiones, he de dar un pequeño rodeo y a partir del argumento de las tierras gemelas de Putnam y Burger, y las respuestas de Fodor a éste, enunciaré la noción de contenido restringido y el principio de intrinsicidad, que articulan, en parte, la primacía de las propiedades físicas en la posición dominante, así como el papel, en última instancia dependiente, de lo mental y, con relación a todo ello, el rol del contexto o condiciones del entorno. En seguida mostraré primero el choque entre la múltiple realizabilidad postulada para las propiedades mentales, por su cercanía a las propiedades funcionales, donde la multirrealización parece evidente, y las nociones fodorianas. Después, en la misma línea (en ambos casos, con relación a la importancia del contexto), al hilo de una crítica a la noción de las metadisposiciones causales descondicionalizadas, expondré el cuestionamiento de la noción misma de correlato restringido, así como la del principio de intrinsicidad y, por ende, la supuesta ruptura de la clausura causal del mundo físico. Luego mostraré el argumento de los autores antedichos, defendiendo que las propiedades mentales no pueden tener mecanismo implementador, frente a su existencia en las propiedades funcionales, noción ésta estrechamente relacionada con las anteriores y con cuyo rechazo, junto con el de los otros supuestos fisicistas enumerados, indicaría la vía expedita para la defensa de la libertad. Para concluir, trataré de poner objeciones relativas al éxito de su argumentación exponiendo cómo suponen aquello que todavía deberían demostrar, con lo que la posición dominante mantendría, si bien que reducida, su plausibilidad y se evidenciaría la necesidad de argumentos adicionales.

He dicho antes, siguiendo a Corbí y Prades, que uno de los supuestos que opera en el trasfondo de la imaginación del fisicista causal es el de que las propiedades físicas del mundo son básicas, pero cabe preguntarse, ¿cómo se definen estas propiedades o qué significa que sean básicas?

Pues bien, aquí estoy yo, Vicente, con los pies en la Tierra (espero) y allí mi gemelo molecular, Vicente gemelo, con los pies sobre la Tierra gemela, cuya diferencia, es cosa sabida, radica tan solo en que mientras que el agua de Tenerife, y en general la de la Tierra, es H_2O , la de la Tierra gemela es XYZ. Por abreviar, ya que la historia es conocida, la idea a la que atiende Fodor es la de que yo y mi gemelo, pese a ser idénticos en nuestras propiedades físicas, tendríamos diferentes propiedades mentales: yo deseo agua y mi gemelo desea agua gemela.

De todo esto podríamos inferir, por un lado, que las propiedades mentales se instancian relacionalmente, esto es, remitiendo al contexto y, por otro, que las propiedades mentales, a primera vista, tienen relevancia causal autónoma. Pues bien, la apuesta de Fodor, la de la posición dominante, consistiría, primero, en respetar tanto el carácter relacional de las propiedades mentales, de las propiedades no-básicas, así como la relevancia causal autónoma de las propiedades funcionales y mentales. Y, segundo, defender que las propiedades básicas, las propiedades físicas, se instancian individualmente, esto es, sin atender a contexto, y que la taxonomía de los estados mentales, dictada por el sentido común o la psicología tradicional, no atrapa en realidad poderes causales. En suma, que las propiedades no-básicas serían dependientes de las básicas en cuanto a su relevancia causal.

Para llevar esto a cabo, plantea Fodor una relevante distinción en el interior de las propiedades mentales. Por un lado, aquello en que difieren los gemelos y que se instanciaría relacionalmente sería el contenido amplio de los estados mentales, que no atraparía poderes causales. Por otro, estaría el contenido restringido, compartido por los estados mentales de los gemelos, que se instanciaría individualmente y que sí atraparía poderes causales.

Este último contenido sería posible definirlo en términos de un, así lo llaman los autores que sigo, "test de intercambiabilidad" de gemelos físicos. Esto es, que si yo en un contexto nomológicamente posible adquiriera una propiedad Z, mi gemelo físico, en ese contexto, también la adquiriría.

Esa identidad de poderes causales a través de contextos, esas propiedades compartidas por los gemelos físicos, serían las propiedades básicas y hemos dicho que uno de los presupuestos de la posición dominante es que las propiedades físicas son básicas. Así pues, dos objetos con las mismas propiedades físicas intrínsecas tendrían los mismos poderes causales en todos los contextos (que es el principio de intrinsicidad).

Otra forma de enunciar este intento de disolución de la tensión entre la primacía de lo físico y la relevancia de lo mental sería la llamada supervinencia fuerte. Una idea básica, no sólo del fisicismo, es la de que las propiedades funcionales son múltiplemente realizables, esto es, que

tienen diversas ejemplificaciones en términos físicos, lo que conllevaría la idea de la existencia de una base de propiedades físicas para cada ejemplificación, a la que sobrevendría la propiedad funcional (algo así como un mecanismo implementador). La pretensión fisicista sería la de asimilar las propiedades mentales a propiedades funcionales, y el correlato restringido del argumento fodoriano a esa base de propiedades físicas, y así como la propiedad no-básica, o contenido amplio, no atraparía poderes causales, aquí diríamos que no podría ir más allá de los poderes causales de su mecanismo implementador o correlato restringido (de propiedades físicas).

II.

Sin embargo, estas propuestas chocan con ciertas tensiones manifestadas entre las exigencias fisicistas y sus concesiones al ámbito de lo mental. La realizabilidad múltiple es, como acabamos de decir, una de las características atribuibles a las propiedades mentales por su buscada semejanza con las propiedades funcionales y de ello se seguiría que el correlato restringido podría serlo de sujetos físicamente dispares. A nadie se le escapa, sin embargo, que en este caso habría contextos físicamente posibles diseñados de tal modo que captasen la disparidad física de los portadores del mismo correlato restringido, bloqueando el *output* relevante de la función elaborada por Fodor. Objeción ésta que contrastaría con la caracterización de las propiedades funcionales como múltiplemente realizables, pero es que frente a la opción fodoriana básica de la instanciación no relacional del correlato restringido, en el caso de las propiedades funcionales la remisión al contexto sería ineliminable, ya que éstas son relacionales (y pueden delimitar los contextos en que se realizan sus ejemplificaciones).

Así pues, aunque he indicado que la explicación psicológica requiere de la satisfacción de la realización múltiple y la crítica antedicha parecería relevante, se podrían esgrimir argumentos alternativos, como de hecho se ha intentado en el seno del fisicismo. De cualquier modo, tal como ahora lo articularé con la distinción entre disposiciones condicionales y metadisposiciones descondicionadas, el hecho es que el CE (las condiciones del entorno o contexto) reaparece y cuestiona el estatus del correlato restringido (y, con ello, las propiedades físicas como básicas, el principio de intrinsicidad, estrechamente relacionados con la restricción y, en última instancia, la clausura causal del mundo físico).

Puedo, pues, enunciar ahora, aceptando que sea D la disposición a producir el *output* apropiado en el CR (contexto de realización) relevante, dos versiones diversas: 1) el objeto X tiene la propiedad funcional F sí y sólo sí en el CE adecuado tiene D; 2) el objeto X tiene la metadisposición

MD, en virtud de la cual tiene D si se encuentra en el CE adecuado. Siendo la MD una disposición a realizar una disposición descondicionalizada, esto es, que se conservaría en cualquier circunstancia físicamente posible, la verdad del contrafáctico sería independiente de las condiciones y siempre que éstas fueran adecuadas se produciría D.

Hay, sin embargo, dos objeciones importantes a esta noción de la MD causal descondicionada, tal como lo plantean los autores que sigo. Por un lado, se da el hecho de que esta MD sólo se manifiesta en la disposición D, pero parece que una disposición a producir otra disposición debería de ser capaz, para que pudiera postularse su existencia, de producir otras disposiciones independientes (en eso consiste su carácter metadisposicional). Sin embargo, tal cosa supondría atender a circunstancias, condicionarse y, por su propia formulación, la MD causal descondicionalizado tendría este paso vedado.

Por otro lado, la MD instancia los poderes causales, fija para el contexto adecuado la disposición correspondiente, mediante condicionales causales básicos (condicionales con el subconjunto de las propiedades básicas del objeto, CR y CE, que fijan la disposición D productora del *output* relevante O), pero con éstos, podría argumentarse, sucedería como en el caso de la crítica a los correlatos restringidos fodorianos. De este modo, lo que la objeción indicaría sería que si no diseñamos un entorno tal que incluya una cláusula restringiendo las circunstancias posibles, lo que condicionaría la MD de manera inasequible para ésta, nos encontraríamos con que, pese a la afirmación de una persistencia a través de todos los contextos físicamente posibles, siempre sería posible diseñar una circunstancia que bloqueara el *output* relevante de uno de los conjuntos de propiedades con la misma MD y no del otro conjunto. No se podría, pues, pese a los reiterados intentos del fisicismo causal, eliminar la remisión al contexto.

Sin embargo, no debemos precipitarnos, pues aún queda un paso que dar en defensa de la relevancia causal autónoma de lo mental, y este es uno de los puntos más problemáticos de la crítica que he estado siguiendo.

III.

Pensemos en un objeto cotidiano, por ejemplo, las llaves de mi casa, en concreto la del buzón (pequeña, gris, oxidada y de muescas desgastadas), pues bien, es cierto que para un efecto determinado, por ejemplo, abrir el buzón, hay diversas propiedades de la llave irrelevantes (por ejemplo, el color), mientras que otras sí serían causalmente relevantes y, de hecho, estarían unidas por un vínculo de necesidad contrafáctica (VCR). Pero, ¿significa esto que las propiedades superfluas para ese efecto estarían desvinculadas de las relevantes? No, en virtud de la ya mentada estabili-

dad del objeto existiría otro vínculo (VE) tal que, propiedades relevantes e irrelevantes, para un efecto determinado, estarían vinculadas entre sí, siendo este vínculo también contrafácticamente relevante para el efecto (que la llave sea gris y no amarilla, no es relevante para el efecto, abrir mi buzón, pero es mi llave [que, entre otras cosas es gris] la que lo abre).

Sabemos ya, porque lo he adelantado, que las propiedades funcionales, para el fisicismo causal se realizan en términos de supervinencia fuerte, esto es, que hay una base física a la que sobreviene una ejemplificación de la propiedad funcional multirrealizable, en diversos contextos posibles, pero no he insistido todavía en el otro requisito que ha de cumplir, el de la minimalidad, si es que se desea evitar la sobredeterminación, como explícitamente pretende la posición dominante con el principio de exclusión explicativa ya mencionado (esto es, la exclusión de dos o más explicaciones completas, en términos causales, de un mismo efecto).

Así pues, y para evitar elementos superfluos que podrían producir sobredeterminación, cada una de las propiedades físicas de la única base mínima de la ejemplificación de una propiedad funcional debería ser condición necesaria para que ésta fuera condición suficiente del efecto. Asimismo, la base mínima debería ser contrafácticamente necesaria. Un elemento superfluo permitiría la existencia de, al menos, dos bases de supervinencia posibles, una con ese elemento y otra sin él; causas completas e independientes del mismo efecto.

Pienso ahora en mi llave, las propiedades físicas que constituyen la única base mínima de la ejemplificación de la propiedad funcional "ser una llave" tienen diversos parámetros, todos presentes en la ejemplificación, y podría argumentarse que el conjunto de parámetros P_1 a P_x de las propiedades físicas de esa ejemplificación constituiría una explicación completa del efecto "abrir mi buzón", tanto como el conjunto de parámetros P_2 a P_z . Pero es que entonces se produciría una multiplicidad de bases mínimas, que desembocarían, sin que el fisicismo pudiese evitarlo, en la sobredeterminación. Sin embargo, con la introducción de la noción de objeto, el mecanismo implementador se salvaría, en virtud del VE mencionado. Podemos, pues, distinguir entre los poderes causales de las propiedades funcionales y los de su ejemplificación, requisito éste para la instanciación de mecanismos implementadores y con lo que, al mismo tiempo, evitaríamos la amenaza de la sobredeterminación. La propiedad funcional tendría un mecanismo implementador físico de sus ejemplificaciones.

No sería éste, sin embargo, según la crítica de Corbí y Prades, el caso de los estados mentales, como por ejemplo el deseo de beber un vaso de agua o la creencia de que encima de mi mesa hay una botella. Pues, si bien es cierto que el estado mental requeriría un cierto detalle, para distinguirlo de otros (mi deseo de beber un vaso de agua del de estrangular a mis

compañeros en la mesa de departamento), también lo es que el estado mental debería contemplar una cierta flexibilidad, lo que ha de ver con la noción de contrafáctico de rastreo. Esto es, que si bien mi deseo de beber agua es nítido, cómo lo lleve a cabo variaría acorde con las variaciones del mundo externo, pudiendo realizarlo, asimismo, de múltiples maneras. Dicho esto, parecería difícil sostener la existencia de una VE, ya que éste requeriría de un sistema estable en el cerebro, de un mecanismo implementador de propiedades físicas lo suficientemente rico en detalles como para distinguirse de otros mecanismos implementadores de deseos y creencias, y ser condición suficiente de la conducta correspondiente a la suma de tal creencia con determinado deseo, y a la vez lo suficientemente pobre como para cumplirlo de muy diversas maneras. E incluso, refinando más el argumento, la flexibilidad del mecanismo implementador de una creencia cabría decir que debería ser tal que pudiese combinarse con los más variados mecanismos implementadores de posibles deseos y creencias que, combinados con la creencia susodicha, dieran lugar al mismo efecto o conducta.

A esta crítica añaden Corbí y Prades objeciones derivadas de la noción de objeto, noción *ad hoc* de la que requerirían para mantener una plausible versión del funcionalismo. Frente al VE que salvaría de la sobredeterminación a cualquier ejemplificación de las propiedades funcionales, los estados mentales, arguyen, no tendrían criterios independientes de identificación, más allá del hecho de que se realicen o no, con lo que, en tanto que meros conjuntos de propiedades físicas, las ejemplificaciones de las propiedades mentales incurrirían en sobredeterminación.

Y además, y volviendo a la crítica al principio de intrinsicidad que se derivaba del argumento de los gemelos físicos de Fodor, fueran objetos o conjuntos de propiedades, parecería que los estados mentales, entendiéndose que no podrían acotar contextos de realización como las propiedades funcionales, se verían ante el reto de dar cuenta de la múltiple realizabilidad, pudiéndose objetar, como en el caso del correlato restringido, que siempre podría diseñarse un contexto en el que uno de los mecanismos de implementación sí produjese la conducta esperada, y otro del mismo estado mental, no.

Sin embargo, y aquí entran mis objeciones a las críticas reseñadas: ¿hasta dónde tales argumentos logran hacer implausible una concepción de los seres humanos en tanto que máquinas detectoras complejas? Todavía, para eliminar la posibilidad de que las creencias tuviesen mecanismos implementadores, sería necesario probar que tales contenidos no son más que conjuntos de propiedades (y no objetos) y además nada impide que un alto grado de refinamiento y flexibilidad diera cuenta de las combinaciones de un mecanismo implementador de creencias con un amplio abanico de otros posibles deseos y creencias.

En cuanto al último y más fuerte de los argumentos, relativo a la posibilidad de que dos diversos mecanismos implementadores de un mismo contenido intencional, en contextos físicos dispares produjesen en un caso un resultado y no en el otro, cabría, por un lado, rechazar que la creencia, sin más argumentaciones, no fuese susceptible de acotar contextos de realización y, por otro, podría objetarse la directa asimilación del contraejemplo de los gemelos físicos al caso que nos ocupa, ya que, si bien, como recordará mi lector, nuestras creencias se realizan o no en coordinación con un amplio abanico de otros posibles deseos y creencias, no se sigue que del hecho de que pudiésemos diseñar contextos que captasen la disparidad física de los mecanismos implementadores, que no cupiese la posibilidad de que tales creencias y deseos combinados diesen en ambos casos el mismo resultado conductual (con lo que se salvaría la multirrealizabilidad de las propiedades mentales y la asimilación postulada de éstas a las propiedades funcionales, así como las subsiguientes consecuencias de esta exitosa maniobra, a saber, la existencia de mecanismos implementadores de los contenidos mentales y la dependencia causal de éstos con respecto a esos mecanismos físicos).

Pero entonces la relevancia causal autónoma de lo mental, defendida por los autores citados, volvería a ponerse en cuestión, y pese a la batería de críticas a que habría sido sometido el fisicismo causal, la antinomia moderna del principio de nuestro texto, entre determinismo y libertad, continuaría vigente, e inclinándose, si bien ahora con una plausibilidad reducida, del lado del primer término. Se evidencia así, como ya anticipaba al principio, la necesidad de una argumentación adicional que continúe en la línea marcada por estos autores, y que logre así inclinar la balanza del lado de la libertad, o bien renunciar a ella definitivamente.

NOTA

1 Departament de Metafísica i Teoria del Coneixement. Universitat de València. *Phrónesis-Analytic Philosophy Group*. Becario de investigación FPU del Ministerio de Educación y Ciencia (AP-2004-0353). E-mail: vicente.raga@uv.es. Domicilio Profesional: Avda. Blasco Ibañez num. 30, 46010 Valencia, España. Telefono: +34 96 386 48 48, Fax +34 96 386 44 32. Este trabajo ha sido llevado a cabo dentro del proyecto de investigación "Creencia, motivación y verdad" (BFF2003-08335-C03-01).

BIBLIOGRAFÍA

- Burge, T. (1986), "Individualism and psychology," *Philosophical Review* 95: 3-45.
- Corbí, J. E. (2000), "Detectar y percibir: Una crítica del fisicismo causal". La Laguna, (ponencia).
- Corbí, J. E.; Prades, J. L. (2000), *Minds, Causes and Mechanisms*. Oxford: Blackwell, pp. 24-108, 193-232.
- Fodor, J. A. (1994), *Psicosemántica*. Trad. de Óscar L. González-Castán. Madrid: Tecnos, pp. 52-89.
- (1995), "un argumento modal en favor del contenido estrecho", en Rabossi, E. (ed.), *Filosofía de la mente y ciencia cognitiva*. Barcelona: Paidós, pp. 331-355.
- Loewer, B.; Rey, G. (eds.) (1991), *Meaning in Mind. Fodor and his Critics*. Oxford: Blackwell, pp. 33-64, 263-271.
- Mellor, H. (1991), *Matters of Metaphysics*. Cambridge: Cambridge University Press, pp. 104-122.
- Putnam, H. (1975), "The meaning of 'meaning'" in *Mind, Language and Reality*. Cambridge: Cambridge University Press, pp. 215-271.