

# LA ESTRATEGIA DEL *EXPLANANDUM* DUAL FRENTE AL PROBLEMA DE LA INEFICACIA CAUSAL DE LO MENTAL\*

GUSTAVO FERNÁNDEZ ACEVEDO

facevedo@mdp.edu.ar

UNIVERSIDAD NACIONAL DE MAR DEL PLATA

## RESUMEN

La denominada 'estrategia del *explanandum* dual' ha sido utilizada en el pasado para enfrentar problemas de distinta clase, entre los que se cuentan el debate 'explicación vs. comprensión' y el problema de cómo las razones pueden explicar la conducta. Esta estrategia puede ser descripta someramente como un intento de resolver la rivalidad explicativa entre dos explicaciones por medio de la división del *explanandum*. En este artículo se analiza el uso que Ausonio Marras ha hecho de ella. Este autor ha sostenido que su utilización permite enfrentar el problema de la ineficacia causal de las propiedades mentales, problema que ha representado una seria amenaza en las últimas dos décadas para el programa del materialismo no reduccionista. Se intentará mostrar que la estrategia sugerida no puede ser exitosa frente a este problema a menos que pueda proveerse de una explicación exitosa de cómo las propiedades mentales pueden mantener su eficacia causal completa, y que la respuesta que se ofrece no satisface esta exigencia.

**PALABRAS CLAVE:** estrategia del *explanandum* dual, materialismo no reduccionista, causación mental, ineficacia causal.

## ABSTRACT

The 'dual *explanandum* strategy' has been used in the past to contend some quite different problems, including the 'explanation vs. understanding' debate and the problem of how reasons can explain behavior. This strategy can be roughly described as an intent to solve the explanatory rivalry between two explanations by means of the division of the *explanandum*. In this paper I want to analyze the use of this strategy by Ausonio Marras, who has claimed that it allows us to face the problem of causal inefficacy of mental properties. This problem has arisen in the last two decades in the context of non reductive materialism and has become a serious menace to this program. I try to show that this strategy can not work out this problem unless it can be provided with a successful account of the way that mental properties can keep their complete causal efficacy, and the answer he offers does not satisfy this demand.

\* Deseo agradecer las valiosas observaciones de los árbitros anónimos de *Análisis Filosófico* y de Manuel Comesaña, que me permitieron mejorar sustancialmente varios aspectos de este trabajo.

KEYWORDS: dual *explanandum* strategy, non reductive materialism, mental causation, causal inefficacy.

## 1. La eficacia causal de los sucesos mentales y su relación con la explicación psicológica

El problema de la eficacia causal de los sucesos mentales se ha revelado en las últimas décadas como uno de los desafíos más difíciles que debe enfrentar toda forma de materialismo no reduccionista que se pretenda viable<sup>1</sup>. Si bien esta clase de doctrinas pareció constituir una respuesta satisfactoria para tradicionales interrogantes relativos a la naturaleza de la relación entre lo mental y lo físico, una serie de argumentos surgidos a partir de la década del ochenta (Block, 1990; Jackson y Pettit, 1990, con un enfoque más general; muy especialmente Kim, 1989b, 1990, 1993a, 1998)<sup>2</sup> han tendido a mostrar que el marco conceptual proporcionado por aquellas resultaba incapaz de asegurar la eficacia causal de los sucesos mentales. Kim (1998) ha presentado la que quizás es la versión más directa e influyente de los argumentos tendientes a mostrar tal ineficacia causal, esto es, el denominado ‘argumento de la superveniencia’. La conclusión de este argumento es formulada por Kim a la manera de un dilema: si la superveniencia mente-cuerpo no tiene lugar, la causación mental es ininteligible; si tiene lugar, la causación mental es también ininteligible. Por lo tanto, la causación mental es ininteligible. La superveniencia, lejos de constituir una solución al problema mente-cuerpo, constituye una de las bases de este.

A su vez, la presunta impotencia causal de los sucesos mentales derivada de argumentos como el de la superveniencia ha sido considerada como premisa de otra serie de argumentos que conducen ya sea a la irrealidad de tales sucesos o bien a su irrelevancia explicativa. El razonamiento para sostener esta última tesis es muy directo: si las explicaciones mentalistas requieren de la causación mental, y esta es prohibida por los argumentos de la exclusión, tales explicaciones deberían ser desechadas<sup>3</sup>. Esta irrelevancia explicati-

<sup>1</sup> Por supuesto, también a cualquier forma de dualismo.

<sup>2</sup> Un artículo de Norman Malcolm (1968) constituye un antecedente lejano de estos argumentos.

<sup>3</sup> Este argumento tiene menos fuerza de la que parece tener en una primera aproximación. Cfr. Sabatés (2001).

va, obviamente, acarrearía consecuencias funestas para una ciencia de lo mental supuestamente autónoma; las explicaciones psicológicas cederían irremediabilmente su lugar a explicaciones puramente físicas o fisiológicas, las cuales harían referencia a propiedades causalmente eficientes. Este problema ha sido denominado en ocasiones ‘problema de la exclusión causal explicativa’<sup>4</sup>.

Una de las formas en la cual la explicación mentalista se ha visto conectada con el problema de la causación mental consiste en la denominada ‘estrategia del *explanandum* dual’ (o ‘de los dos *explananda*’). En líneas muy generales, este enfoque puede ser caracterizado como el intento de ‘resolver la rivalidad explicativa al sostener que las dos explicaciones no comparten en realidad los mismos *explananda*’ (Kim, 1991, p. 293)<sup>5</sup>. Esto es, no existiría una competencia entre las explicaciones mentalistas *vis-à-vis* las explicaciones biológicas (o físicas) ya que los *explananda* de ambos tipos de explicaciones no son idénticos.

En este trabajo examinaremos un reciente y ambicioso intento de lidiar con el problema de la exclusión causal explicativa por medio de la utilización de la estrategia de los dos *explananda*: el llevado a cabo por Ausonio Marras (1998, 2000). La línea argumental que intentaremos desarrollar será la siguiente: en primer lugar, examinaremos su versión de la estrategia del *explanandum* dual, desarrollada para enfrentar el problema de la competencia explicativa entre propiedades mentales y propiedades físicas dentro del marco del materialismo no reduccionista; sostendremos que esta versión, si se desea preservar los supuestos e intuiciones de tal marco, requiere enfrentar el problema de la eficacia causal de las propiedades mentales. En segundo lugar, examinaremos la aplicación de la estrategia del *explanandum* dual al argumento de la

<sup>4</sup> Al sostener una posición en algún sentido contraria, algunos autores (Baker, 1993; Burge, 1993) han afirmado que el problema de la presunta ineficacia causal de los estados mentales es un problema espurio, originado por una serie de supuestos metafísicos injustificables o, en el mejor de los casos, sumamente discutibles. Para estos autores, la comprensión de la causación mental debe basarse en el examen de la práctica explicativa y de las explicaciones científicas de lo mental que se consideran exitosas. La explicación psicológica no solo no se ve, de esta forma, afectada por los argumentos de la exclusión, sino que constituye el fundamento para la comprensión de la causación mental.

<sup>5</sup> Esta caracterización de Kim es más general que alternativas más específicas y, quizás, menos preferibles. Sabatés (2001) sugiere que la posibilidad de salvar la causación de sucesos mentales por parte de otros sucesos mentales permitiría desarrollar una estrategia promisoriosa del *explanandum* dual.

supervenencia propuesto por Kim, aplicación que requerirá, en nuestra opinión, de un esquema satisfactorio de la relación entre propiedades mentales y propiedades físicas. Por último, señalaremos lo que consideramos son limitaciones del esquema de esta relación propuesto por Marras, limitaciones que tornan su versión de la estrategia de los dos *explananda* insuficiente para enfrentar el problema de la eficacia causal de lo mental. Previamente reseñaremos en forma sintética algunas propuestas en la línea de la estrategia del *explanandum* dual, y señalaremos que es una alternativa menos homogénea que lo que usualmente se cree.

## 2. La estrategia del *explanandum* dual: algunos antecedentes

La utilización de la estrategia del *explanandum* dual tiene antecedentes relativamente lejanos. En Von Wright (1971) se encuentra una exploración de la posibilidad de que las explicaciones ‘teleológicas’ y las explicaciones ‘causales’ no compartan los mismos *explananda*, en cuyo caso no se plantearía un problema de competencia o incompatibilidad entre ambas. El planteo de Von Wright tiene lugar en el marco de la discusión entre concepciones antagónicas de las ciencias del hombre, en particular en el contexto del antiguo debate ‘explicación vs. comprensión’.

La estrategia ha sido empleada posteriormente por Dretske en un contexto muy distinto, esto es, en el marco del problema de cómo las razones, en tanto cierto tipo de sucesos mentales, contribuyen a explicar los efectos consistentes en *outputs* motores (esto es, el problema de la causación intencional o racional)<sup>6</sup>. En líneas generales, el enfoque de Dretske puede ser descrito de la siguiente manera. Los *explananda* de las explicaciones racionalizadoras son aquellos que se explican cuando se ofrece una ‘razón’, una conducta o una acción, algo que el agente ‘hace’. Dretske sostiene que un agente S hace A solo si algún estado interno de S,  $S_i$ , *causa* un cierto *resultado* R (típicamente, un *output* motor) asociado con A. Hay una diferencia crucial, en el enfoque de Dretske, entre hacer A (la causación de R por parte de  $S_i$ ) y R. La explicación de R, o la determinación de la causa de R, es algo muy diferente a la explicación de

<sup>6</sup> Su propuesta se encuentra en forma resumida en su (1990).

por qué  $S_i$  causó R. Si bien la neurobiología puede proporcionar una descripción exhaustiva de la etiología causal de R, incluyendo una descripción de *cómo*  $S_i$  causó R, no puede proporcionar una explicación de *por qué*  $S_i$  causó R. El trabajo de explicar esta relación causal es tarea indispensable y distintiva de la psicología; explicar por qué la estructura causal relacional  $S_i$  causó R es explicar por qué o cómo  $S_i$  ha llegado a unirse con R en el agente S, y el contenido de  $S_i$  explica, y es la causa de cómo este nexo causal ha llegado a ser como es. En opinión de Kim, la solución de Dretske al problema de la exclusión es un caso de lo que puede ser llamado estrategia ‘de los dos *explananda*’: las explicaciones biológicas y racionalizadoras no comparten los mismos *explananda*, por lo cual no existe competencia entre ellas ni exclusión mutua. Dado que las razones hacen racionales a las acciones en un sentido intencional, las explicaciones fisiológicas solo pueden explicar movimientos corporales descritos en términos puramente físicos; ya que las dos explicaciones parecen tener diferentes *explananda*, el conflicto potencial se desvanece<sup>7</sup>.

Más recientemente, otra variante de esta estrategia ha sido propuesta por Richard Montgomery (1995). Este autor pretende mostrar cómo, independientemente de los efectos que los argumentos de la exclusión puedan tener sobre ciertas clases de explicación psicológica, son inocuos contra otros esquemas explicativos frecuentemente utilizados en la psicología cognitiva. Las explicaciones psicológicas vulnerables a este tipo de argumentos son denominadas ‘explicaciones cartesianas’, las cuales se caracterizan por no admitir relaciones entre la mente y el mundo, como la verdad y falsedad o la exactitud e inexactitud en el *explanandum* o en el *explanans*, razón por la cual las explicaciones del error cognitivo están fuera del alcance del modelo cartesiano. Por el contrario, las explicaciones psicológicas que apelan a lo que este autor denomina ‘contenido fundado’ no pretenden explicar simplemente sucesos psicológicos estrechos, como ocurre en el modelo cartesiano, sino que en ellas se trata de explicar ciertos sucesos psicológicos híbridos consistentes en la relación semántica entre tres *relata*: un suceso psicológico al estilo cartesiano; la relación semántica que tal estado acarrea; y el mundo real, que determina el *status* semántico de ese contenido. Según Montgomery, su propuesta ‘elimina la competencia explicativa entre explicaciones psicológicas y neurofisiológicas al construirlas como intentos de explicar diferentes cosas’ (p. 235).

<sup>7</sup> Cfr. Kim (1990) para una crítica a esta estrategia.

Vicente (2002) presenta una caracterización alternativa de la estrategia. En su presentación, el fisicalismo es la doctrina que sostiene que los sucesos mentales *son* sucesos físicos. La admisión del fisicalismo parece la conclusión obligada del argumento de la exclusión, conjunto de premisas que sostiene: a) el principio de clausura causal del mundo físico<sup>8</sup>; b) la eficacia causal de lo mental (los sucesos mentales producen cambios en el mundo físico); y c) el principio de exclusión causal explicativa (no puede haber dos causas/explicaciones completas e independientes para un mismo suceso, excepto en casos de sobredeterminación). La estrategia del *explanandum* dual se basaría en un rechazo de la segunda premisa, al sostener que los sucesos mentales pueden ser causalmente potentes, si bien no causarían sucesos físicos. Obsérvese que esta caracterización de la estrategia no es explicativa (ni siquiera epistemológica), sino ontológica: no hace referencia alguna a la explicación<sup>9</sup>. Por otra parte, Vicente advierte que esta estrategia del *explanandum* dual es una manera de enfrentar al fisicalismo, no un intento de solución del problema de la exclusión *simpliciter*<sup>10</sup>.

Estos antecedentes deberían bastar para mostrar cómo la estrategia del *explanandum* dual, pese a la manera en que a veces es presentada (y a la caracterización preliminar que hemos ofrecido), dista de constituir una posición homogénea ante el problema de la causación mental; más aún, como lo muestra la utilización de Von Wright, ni siquiera puede ser considerada únicamente como *una respuesta* a este problema. No supondremos, por lo tanto, que la posición que examinaremos en lo sucesivo implique el compromiso con alguna de las tesis que hemos descripto. Por esta razón, sus posibilidades de éxito deberán ser evaluadas a la luz del problema específico en cuya solución la estrategia del *explanandum* dual pretende emplearse.

<sup>8</sup> '[C]ualquier suceso físico que tiene una causa en el momento *t* tiene una causa física en *t*. Este es el supuesto de que si rastreamos los antecedentes causales de un suceso físico, nunca necesitamos ir más allá del dominio físico. Negar este supuesto es aceptar la idea cartesiana de que algunos sucesos físicos necesitan causas no físicas, y si esto es verdad no puede haber en principio una teoría física completa y autosuficiente del dominio físico' (Kim, 1989, p. 280. *Cursivas del autor*).

<sup>9</sup> Quizás convendría considerarla más bien una estrategia de la 'causación dual'.

<sup>10</sup> Presenta como exponentes de esta estrategia a Scott Sturgeon ("Physicalism and Overdetermination", (*Mind*, p. 107, 1998), y "Conceptual Gaps and Odd Possibilities", (*Mind*, p. 108, 1999), y Jennifer Hornsby (*Simple Mindedness: In Defense of a Naive Naturalism in the Philosophy of Mind*, Harvard University Press, 1997).

En las siguientes secciones nos centraremos en la presentación que Marras hace de esta estrategia, dirigida a enfrentar el problema de la ineficacia causal y explicativa de lo mental dentro del marco del materialismo no reduccionista.

### 3. Marras y una estrategia alternativa de los dos *explananda*

#### 3.1. Realismo explicativo y exclusión explicativa

Se ha advertido en alguna oportunidad que, si bien inaceptable en la manera en que es formulado por Kim, el principio de exclusión explicativa encierra una intuición plausible<sup>11</sup>, razón por la cual parece posible una reformulación que evite los problemas de aquella y que a la vez preserve la intuición considerada válida. Sin embargo, el rechazo del principio y su reemplazo por una formulación no equivalente, pero que conserve su espíritu, resulta una cuestión sencilla.

En un artículo reciente Ausonio Marras (1998) desarrolla una propuesta en esta última dirección y presenta una crítica tanto al principio de exclusión explicativa cuanto a dos de sus supuestos, esto es, el criterio de individualización de los sucesos y las explicaciones y el principio del realismo explicativo. En su opinión, el principio de exclusión explicativa (PEE) enunciado por Kim (1989a), el cual sostiene que ningún suceso puede tener más que una explicación completa e independiente, carece de apoyo apropiado y conduce a consecuencias inaceptables; sin embargo, considera que puede ser defendido un principio estrechamente relacionado, pero significativamente diferente.

Marras comienza su crítica analizando el principio del realismo explicativo propuesto por Kim. Este principio sostiene lo siguiente:

RE:  $C$  es un *explanans* para  $E$  en virtud del hecho de que  $c$  mantiene con  $e$  alguna relación objetiva determinada  $R$  (Kim, 1987, p. 226).

La relación objetiva  $R$  es una relación que vincula a los sucesos en el mundo, y 'fundamenta' la relación explicativa entre las

<sup>11</sup> Cfr. Sabatés (1996).

proposiciones que la componen. Cuando la explicación es una explicación causal la relación objetiva  $R$  es la relación causal misma. Una explicación causal será verdadera o correcta en tanto  $e$  tenga lugar como efecto de una causa objetiva  $c$ .

Marras está en completo desacuerdo con la afirmación de Kim acerca de que el principio de exclusión explicativa (PEE) constituye una restricción general plausible para cualquier modelo realista de la explicación. Bajo el principio del realismo explicativo propuesto por Kim, observa, las explicaciones son individualizadas bajo un criterio estrictamente *extensional*: no importa cuáles sean las descripciones que correspondan a ' $c_1$ ' y ' $c_2$ ' en la afirmación de identidad ' $c_1 = c_2$ '; en tanto la afirmación de identidad sea verdadera, las dos afirmaciones, si bien lógicamente inequivalentes, proporcionan la misma explicación. Esta consecuencia es, para Marras, muy contraintuitiva. Puede estarse de acuerdo en que el terremoto causó el colapso del edificio si y solo si el suceso sobre el que se informa en la página 5 de la edición de hoy del *Globe and Mail* lo hizo, dado que el último suceso  *fue* el terremoto en cuestión. Pero, señala, si las explicaciones han de proveer comprensión, es al menos discutible que esas dos afirmaciones causales singulares provean la misma explicación de por qué el edificio colapsó. Dos afirmaciones causales singulares bien pueden ser extensionalmente equivalentes y carecer del mismo contenido explicativo. De hecho, observa, una afirmación causal puede ser verdadera y no ser en absoluto explicativa. Estas observaciones no constituyen una petición de principio, señala, ya que el propio Kim acepta el carácter *epistémico* de la noción de explicación. Debido a esto, prosigue, RE individualiza las explicaciones demasiado toscamente. Todo lo que requiere una forma razonable de realismo explicativo es la siguiente restricción:

RE\*:  $C$  es un *explanans* para  $E$  solo si el suceso  $c$  realmente causó el suceso  $e$  (Marras, 1998, p. 443. Cursivas del autor).

Este principio es más débil que el planteado por Kim. La causalidad provee el 'fundamento objetivo' para la explicación causal en el sentido de que una explicación de  $e$  en términos de  $c$  no será correcta a menos que  $c$  haya causado efectivamente  $e$ . Pero se requieren otras condiciones, epistémicas, para la corrección de la explicación, esto es, que la causa y el efecto sean descriptos en ciertos modos particulares y no en otros. Si bien un principio de exclusión causal:



PCX: No puede haber más que una única *causa* completa e independiente de cualquier suceso (Marras, 1998, p. 444. *Cursivas del autor*).

resulta admisible (exceptuando la sobredeterminación), bien puede ser aceptado mientras se rechaza PEE. Y, a menos que se asuma la forma de realismo explicativo propuesta por Kim, la cual hace a la mismidad de la causa necesaria y suficiente para la mismidad de la explicación, no puede inferirse PEE a partir de PCX.

### 3.2. La individualización de los *explananda*

Sugerir que la explicación no es extensional, sostiene Marras, es llamar la atención sobre el hecho de que la relación explicativa se produce entre sucesos *como un tipo*, o en la medida que ejemplifiquen una u otra propiedad. La forma canónica de una afirmación explicativa singular no es ‘*c* explica *e*’, sino ‘el que *c* sea *F* (o *qua F*) explica el hecho de que *e* sea *G*’, y en donde los tipos *F* y *G* son identificados como la causa y el efecto respectivamente. Los debates recientes sobre la causación mental, prosigue, han enfatizado el hecho de que mientras la relación causal puede ser extensional, relacionando sucesos sin importar la manera en que son descriptos, una afirmación singular causal no llega a ser explicativa a menos que identifique la causa en términos de sus propiedades ‘causalmente relevantes’. Marras considera que lo que hace a una propiedad causalmente relevante no es importante en esta discusión, sino que lo que importa es que una afirmación causal singular que pretende ser explicativa tendrá que individualizar la causa en términos de aquellas (entre sus indefinidamente muchas) propiedades que fueron relevantes para el efecto.

Este análisis se vincula con el problema de la exclusión de la manera que se expone a continuación. Sean dos explicaciones –una de ellas una explicación racionalizadora, y la otra una explicación neurológica– de un suceso singular *B*, por ejemplo, el que Jorge se levante del sillón:

E1: el que *c* sea una clase de suceso *R* causó *B*

E2: el que *c\** sea una clase de suceso *N* causó *B*<sup>12</sup>

<sup>12</sup> *R* y *N* deben ser interpretadas como propiedades mentales y físicas que permiten describir en modos alternativos al suceso *B*.

$c$  y  $c^*$  son sucesos que ocurren simultáneamente en Jorge. ¿Pueden E1 y E2 ser aceptados simultáneamente como explicaciones completas e independientes de  $B$ ? La respuesta de Kim, basándose en el principio de exclusión, es negativa. No hay necesidad de analizar, sostiene Marras, las posibles maneras en que  $c$  y  $c^*$  están relacionados. Para refutar la afirmación de Kim es suficiente con identificar al menos una posible interpretación de E1 y E2 bajo la cual ambas serían aceptables como explicaciones completas e independientes de  $B$ , sin violar la clausura del mundo físico.

Supóngase entonces, sostiene Marras, de acuerdo con los ampliamente difundidos enfoques no reduccionistas de la relación mente-cuerpo, que  $c = c^*$ , pero que  $R$ , al ser múltiplemente realizable, es distinta de  $N$ .  $B$  debe ser considerado un suceso físico; sin embargo, debe recordarse que  $B$  tiene la estructura de un suceso de una clase determinada. Entonces, las dos explicaciones E1 y E2 deberían ser reformuladas de manera acorde, para reflejar el hecho de que  $B$  no es un suceso individual *simpliciter*, sino un cierto *hecho* individual —el hecho de que  $e$  es un caso de un cierto tipo de sucesos. Tan pronto como se conjuga esta consideración con la observación previa de que la relación ‘causal’ de la cual se habla en E1 y E2 es realmente la relación ‘quausal’, resulta claro que la forma canónica de las explicaciones del tipo ejemplificado por E1 y E2 es realmente:

$c$  causó  $e$ , y esto fue en virtud del hecho de que  $c$  sea  $F$  y de que  $e$  sea  $G$

Esta es la forma en la cual la expresión *c qua F causó e qua G* es entendida de manera estándar en la literatura sobre la relevancia causal<sup>13</sup>.

<sup>13</sup> Lo que es entonces la cuestión en el problema de la causación mental no es tanto la *eficacia* causal de las propiedades mentales, sino más bien la *relevancia* de las propiedades mentales en la formulación de explicaciones causales. Llamemos a este tipo de relevancia explicativa *relevancia causal*, para distinguirla de otros tipos de relevancia explicativa que las propiedades mentales puedan tener (*i. e.*, para la formulación de explicaciones ‘racionalizadoras’), y no dejemos que la *relevancia* causal sea confundida con la *eficacia* causal, la cual pertenece propiamente a sucesos. La objeción epifenomenista puede ser reformulada como sigue: en el no reduccionismo, el fisicalismo de casos, las propiedades mentales son causalmente irrelevantes. Llamo a esto ‘objeción *epifenomenista de propiedades*’, para distinguirla de lo que podemos haber llamado ‘objeción *epifenomenista de sucesos*’, previamente considerada. (*Epifenomenismo de sucesos*, entonces, es la concepción según la cual los sucesos mentales son causalmente ineficaces; *epifenomenismo de propiedades* es la concepción según la cual las propiedades mentales son causalmente irrelevantes) (Marras, 1994, p. 473. *Cursivas del autor*).

En el caso objeto de discusión, es claro que no hay un solo tipo de suceso que el que ‘Jorge se levanta del sillón’ deba ejemplificar. Dada la naturaleza del ejemplo que se está examinando (un ejemplo de conducta intencional), podemos suponer razonablemente que el suceso en cuestión ejemplifica tanto la propiedad de ser una cierta clase de movimiento corporal cuanto la propiedad *distinta* de ser una cierta clase de acción intencional.

Sea entonces *A* la propiedad de levantarse intencionalmente del sillón, y *B* la propiedad fisiológica correspondiente. Entonces cada una de las dos explicaciones E1 y E2 puede ser reformulada canónicamente en dos formas:

- E1(a): *c* causó *e*, y esto fue en virtud del hecho de que *c* sea *R* y de que *e* sea *A*  
 E1(b): *c* causó *e*, y esto fue en virtud del hecho de que *c* sea *R* y de que *e* sea *B*  
 E2(a): *c* causó *e*, y esto fue en virtud del hecho de que *c* sea *N* y de que *e* sea *A*  
 E2(b): *c* causó *e*, y esto fue en virtud del hecho de que *c* sea *N* y de que *e* sea *B*

Cada una de estas explicaciones, a su modo, pretende explicar por qué *e* ha ocurrido. El hecho de que *c* lo causó, claramente, subdetermina cuál de esas explicaciones es correcta. Lo que necesitamos saber es: ¿en virtud de qué *c* causó *e*? Y para responder a esta cuestión se necesita saber, en primer lugar, cómo *e*, el suceso *explanandum*, es identificado como ejemplificación de un tipo determinado. Dado que, por hipótesis, *e* tiene tanto la propiedad *A* como la propiedad *B*, como se identifiquen tipos a los fines explicativos dependerá de razones contextuales –por ejemplo, si estamos interesados en explicar por qué una acción de tipo *A* ocurrió en esa ocasión, o por qué tuvo lugar un movimiento corporal de tipo *B*–. Si bien el suceso *explanandum* es uno y el mismo, los hechos *explanandum* son seguramente diferentes<sup>14</sup>.

Supongamos entonces que identificamos a *e* como un ejemplo de *A* –una acción de un determinado tipo–. ¿Causó *c*, en esa oportunidad, una acción de este tipo en virtud de poseer la propiedad

<sup>14</sup> La idea de que no se proporcionan explicaciones de sucesos considerados como un todo, sino de determinados aspectos del suceso, está ya presente en Hempel (1965). Sin embargo, Hempel no se compromete allí con una ontología determinada, en especial con una concepción particular de los sucesos y las propiedades.

racionalizadora  $R$  o en virtud de poseer la propiedad neurobiológica  $N$ ? Que la respuesta sea una u otra dependerá de si la posesión de la propiedad  $R$  por parte de  $c$  fue causalmente relevante para la posesión de  $A$  por parte de  $e$ , o si la posesión de  $N$  por parte de  $c$  lo fue.

Marras no ofrece aquí una concepción de la relevancia causal (si bien observa que posee una)<sup>15</sup>, pero considera que todo lo que necesita en ese contexto es que se acepte la siguiente condición —a la que considera no controvertible— para la relevancia causal:

CR: donde  $c$  causa  $e$ , y donde  $c$  es un  $F$  y  $e$  es  $G$ , que  $c$  sea  $F$  es causalmente relevante para que  $e$  sea  $G$  solo si el contrafáctico ' $\neg Fc \rightarrow \neg Ge$ ' es verdadero (1998, p. 448)<sup>16</sup>.

Sobre la base de CR puede considerarse, sostiene Marras, a E1(a)-E2(b) como verdaderos solo si los contrafácticos que implican son verdaderos. Correspondiendo cada uno a E1(a)-E2(b) tenemos:

(1a)  $\neg Rc \rightarrow \neg Ae$

(1b)  $\neg Rc \rightarrow \neg Be$

(2a)  $\neg Nc \rightarrow \neg Ae$

(2b)  $\neg Nc \rightarrow \neg Be$

En el enfoque no reduccionista que se ha supuesto, prosigue Marras, estos contrafácticos tienen interpretaciones naturales de acuerdo con las cuales (1a) y (2b) resultan ser verdaderos, mientras que (1b) y (2a) resultan ser falsos.

Marras considera que tenemos, entonces, dos explicaciones distintas y posiblemente verdaderas de la ocurrencia del suceso  $e$ : E1(a) y E2(b). Nuestra fe en los esquemas explicativos de la neurofisiología nos persuade de la verdad de E2(b); y, al suponer que nuestra creencia en la relevancia causal de las propiedades mentales está bien fundada, tenemos todas las razones para creer en la

<sup>15</sup> Describiremos esta propuesta en el apartado 3.4.

<sup>16</sup> Uno de los árbitros de *Análisis Filosófico* ha objetado el hecho de que se recurra a los contrafácticos para analizar la noción de relevancia causal y luego se evalúen contrafácticos por medio de la relación causal, estimando este procedimiento como circular e inaceptable. Esta observación es razonable; sin embargo, no necesitamos suscribir ni discutir las afirmaciones de Marras respecto de la pertinencia de los contrafácticos para la relevancia causal. Es suficiente, para nuestros propósitos, admitir que el suceso en cuestión ejemplifica propiedades físicas y mentales, y que las relaciones causales tienen lugar solamente entre propiedades del mismo nivel.

verdad de E1(a). Nótese, sin embargo, que E1(a) y E2(b) individualizan el suceso *explanandum* en diferentes modos –el primero como una cierta clase de acción intencional, y el segundo como una cierta clase de movimiento corporal–. *Relativamente a cómo sea individualizado el explanandum*, E1(a) y E2(b) no competirán el uno con el otro: E1(a) pretende explicar por qué un cierto tipo de acción ocurrió en una ocasión determinada, y E2(b) pretende explicar por qué cierto tipo de *movimiento corporal* lo hizo –y estos son hechos *explananda*–. Esto nos autoriza a decir, sostiene Marras, que *hay* un principio de exclusión explicativa PEE\*, distinto de PEE, el cual puede muy bien ser verdadero:

PEE\*: No puede haber más que una explicación completa e independiente de un *explanandum* dado (1998, p. 449).

donde ‘*explanandum*’ no quiere decir un suceso *simpliciter*, sino *el que un suceso sea de cierta clase*. Si tomamos a nuestros *explananda* como sucesos bajo una individualización por tipo –‘bajo una descripción’, en términos de Davidson– entonces la relación explicativa es no extensional; y bajo este esquema de cosas PEE es falso mientras PEE\* bien puede ser verdadero. Esta forma de tratar con el problema de la exclusión, sostiene Marras, es un caso de la estrategia denominada por Kim ‘de los dos *explananda*’<sup>17</sup>.

<sup>17</sup> Pese a que Kim, como bien ha observado uno de los árbitros de *Análisis Filosófico*, se ha caracterizado por proponer una ontología ‘fina’ de acontecimientos, según la cual deben ser entendidos como ejemplificaciones de propiedades, al parecer no ha considerado necesario suponer su propia ontología como premisa al momento de argumentar en favor de la exclusión explicativa. En efecto, en su (1990), al discutir las diferentes posibilidades en las cuales un suceso es determinado causalmente y, por lo tanto, puede darse una explicación causal de él, afirma lo siguiente: “Hemos estado hablando hasta ahora indistintamente de sucesos *simpliciter* y de sucesos que sean de cierta clase o que posean cierta propiedad, como causas y efectos. *Esto no hace diferencia: sin importar cómo individualicemos causas y efectos, cuando dos cosas son propuestas como la causa de un efecto singular, enfrentamos el mismo problema*. El ‘principio de exclusión explicativa’ completo que he elaborado en otra parte dice esto: puede haber como máximo una explicación *completa e independiente* de un *explanandum* único” (p. 126. Las cursivas son nuestras). Esta observación de Kim muestra que, al parecer, el principio de exclusión explicativa no debe ser entendido como un principio aplicable exclusivamente a los casos de sucesos entendidos como ejemplificaciones de propiedades y, por lo tanto, no válido para otras caracterización de los acontecimientos (por ejemplo, la de Davidson). No necesitamos comprometernos, por otra parte, con las conclusiones de Marras respecto del principio de exclusión explicativa modificado. Todo lo que se requiere, a los fines de la presente argumentación, es aceptar provisoriamente tales conclusiones a los efectos de

### 3.3. La dualidad de los *explananda* y el argumento de la superveniencia

El análisis de Marras expuesto hasta el momento combina consideraciones ontológicas, relativas a la naturaleza de la relación causal y a la caracterización de los sucesos, con reflexiones epistemológicas, concernientes a la naturaleza de lo que debe considerarse una explicación satisfactoria. Su objetivo principal ha sido mostrar cómo pueden ofrecerse dos explicaciones de un mismo suceso (conductual) sin que se plantee un problema de exclusión o competencia explicativa en los términos que Kim propone. Su posición requiere de una serie de supuestos complejos relativos a la naturaleza de la explicación, la individuación de los sucesos y la relación causal. Muchos de estos supuestos pueden ser discutidos y constituirse en bases para la réplica a los argumentos de Marras. Sin embargo, no objetaremos alguno de ellos en particular, sino que nos concentraremos en un aspecto que solo es mencionado en el análisis de este autor.

Al describir el suceso davidsoniano ‘Jorge se levantó del sillón’, que ejemplifica dos propiedades (una propiedad racionalizadora y una propiedad neurobiológica)<sup>18</sup>, Marras afirma que no es necesario examinar las maneras en que tales propiedades pueden estar relacionadas. En su opinión, como hemos visto, es suficiente encontrar al menos una posible manera en que ambas explicaciones sean aceptables como completas e independientes sin violar la clausura del mundo físico. Esta suposición parece sumamente cuestionable. Marras acepta que la propiedad racionalizadora es distinta de la propiedad neurofisiológica y múltiplemente realizable por propiedades de esta última clase; sin embargo, estos rasgos alcanzan para excluir algunas posiciones con respecto a la relación entre tales tipos de propiedades (por ejemplo, algunas clases de dualismo, como el interaccionismo o el paralelismo, o la teoría de la identidad), pero no son suficientes para descartar otras, cuyas consecuencias para el plano explicativo no son menores. Un emergentista, por ejemplo, podría aceptar que la relación que vincula a las propiedades menta-

---

evaluar las posibilidades de éxito de la estrategia del *explanandum* dual desarrollada por Marras.

<sup>18</sup> Quizás N podría ser descripta más apropiadamente como una propiedad física (un movimiento corporal). Sin embargo, dado que esta cuestión no jugará ningún rol fundamental en los argumentos que siguen, ignoraremos las eventuales consecuencias que de ella puedan derivarse.

les con las propiedades físicas es una relación de realización (incluyendo la realizabilidad múltiple); sin embargo, aceptaría también la causación descendente, de lo mental a lo físico, como uno de los rasgos centrales de su posición. En este último caso, es evidente que la aceptación de esta doctrina condiciona fuertemente lo que seremos capaces de afirmar acerca del rol que juegan las propiedades mentales en las explicaciones de la conducta. La posición adoptada implica, por otra parte, reconocer que las propiedades mentales son capaces de causar otras propiedades mentales. Esta tesis puede ser objetada; podría afirmarse, como Kim lo ha hecho a partir del argumento de la superveniencia, que las propiedades mentales no solo carecen de eficacia causal para producir propiedades físicas, sino también para producir otras propiedades mentales.

Si estas consideraciones son correctas, solo podríamos aceptar la variante de la estrategia del *explanandum* dual propuesta por Marras en caso de que pueda darse una respuesta satisfactoria a las cuestiones ontológicas planteadas por la causación mental, esto es, encontrar una forma de enfrentar los argumentos que conducen a la exclusión causal de lo mental y sugerir una relación entre lo mental y lo físico que preserve los poderes que se pretende adscribirles en el marco del materialismo no reduccionista. Esta exigencia es consistente con nuestra sugerencia inicial según la cual, de acuerdo con las últimas (y más sólidas) formulaciones, el problema de la exclusión explicativa en el caso de los sucesos o propiedades mentales se plantea más bien como una consecuencia del problema ontológico<sup>19</sup>. En lo que sigue examinaremos, por lo tanto, su respuesta al argumento de la superveniencia y su propuesta relativa a la relación entre propiedades mentales y propiedades físicas.

En su (1998) Marras, aunque de manera más bien implícita, parece sostener que puede resolverse el problema de la relevancia explicativa sin enfrentar la cuestión de fondo, esto es, la de relación entre propiedades mentales y físicas. Sin embargo, en 2000, en su nota crítica de *Mind in a Physical World*, Marras aplica algunas de las conclusiones expuestas en su (1998) al argumento de la superveniencia desarrollado por Kim. Las describiremos a continuación.

Marras considera al argumento de la superveniencia propuesto por Kim ‘elegante y poderoso’, sin embargo, cuenta con que hay maneras de enfrentarlo. Kim, observa, niega que el ‘fiscalismo de casos’ sostenido por Davidson sea suficiente para responder a la

<sup>19</sup> Cfr. al respecto Sabatés (2001).

pregunta que formula el problema de la exclusión causal/explicativa, esto es: (P) 'Dado que todo suceso físico que tiene una causa tiene una causa física, ¿cómo es posible también una causa mental?' Al negar esta posibilidad, Kim debe suponer que, o bien la relación causal no es extensional, o bien los términos de la relación causal no son sucesos en el sentido davidsoniano (particulares espacio temporales concretos e irrepetibles). Si se adopta una ontología de sucesos davidsonianos, prosigue, algunos de los cuales son tanto mentales cuanto físicos, parece ser suficiente para responder P; la causa mental de un suceso es idéntica como caso (*token identical*) con su causa física, salvo que se suponga que la relación causal no es extensional. Bajo esta interpretación, la afirmación singular causal '*c* causó *e*' (en la cual *c* y *e* son sucesos davidsonianos) es interpretada como una formulación elíptica para la afirmación singular 'quausal' '*c* causó *e* en virtud de que *c* es *F* y *e* es *G*', y en la cual *c* y *e* están relacionados por la relación compleja 'causa-en-virtud-de- $\langle F, G \rangle$ ', donde  $\langle F, G \rangle$  es un par ordenado de propiedades causalmente prominentes de la causa y el efecto respectivamente. De esta manera, continúa Marras, (P) puede ser reformulada en los siguientes términos:

(P\*) Dado que todo suceso físico que tiene una causa es causado por un suceso en virtud de las propiedades físicas de este, ¿cómo puede ser causado por el mismo suceso en virtud de sus propiedades mentales?

Marras observa que Kim, al negar que el fisicalismo de casos de Davidson sea suficiente para responder (P), quiere decir en realidad que no es suficiente para responder (P\*). Esta lectura del problema, considera, se ajusta bien a la afirmación de Kim de que el problema de la causación mental involucra, en última instancia, la cuestión de la eficacia causal de las propiedades mentales. Y, desde esta perspectiva, el problema de la causación mental se transforma en cómo evitar el epifenomenismo de tipos, ya que el enfoque de Davidson es suficiente para evitar el epifenomenismo de casos.

Como alternativa, Kim puede suponer que si bien la relación causal es extensional, los términos de la relación son sucesos en el sentido de 'tropos' o 'ejemplificaciones de propiedades', y no en el sentido que les asigna Davidson. Si bien en el argumento de la superveniencia, observa Marras, Kim hace referencia a ejemplificaciones de *M* y de *P*, él mismo señala en otro trabajo que las ejem-



plificaciones de propiedades pueden ser interpretadas como tropos (una ejemplificación de  $F$  es un suceso), o, de manera alternativa, como algo (un suceso u objeto) que ejemplifica o posee  $F$ . Ambas formulaciones son, según Marras, equivalentes, ya que puede suponerse que los tropos causan lo que causan en virtud de sus propiedades constitutivas. De plantearse el problema en términos de tropos y sus propiedades constitutivas, este planteo sería el siguiente: dado que los tropos son individualizados por sus propiedades constitutivas, si las propiedades mentales son distintas de las propiedades físicas, ¿cómo un tropo físico que tiene una causa puede ser causado tanto por otro tropo físico cuanto por un tropo mental?

Provistos con las conclusiones precedentes, continúa Marras, hay que retornar al argumento de la superveniencia y considerar las relaciones causales  $M\text{-}a\text{-}M^*$  y  $M\text{-}a\text{-}P^*$ , las cuales, de acuerdo con la conclusión del argumento, son solo aparentes en la medida en que surgen de procesos causales genuinos que ocurren de  $P\text{-}a\text{-}P^*$ . Sea entonces  $c$  un suceso davidsoniano que ejemplifica tanto  $M$  como  $P$ , y sea  $e$  un suceso davidsoniano que ejemplifica tanto  $M^*$  como  $P^*$  (y en donde  $M$  y  $M^*$  sobrevienen a partir de  $P$  y  $P^*$  respectivamente); puede preguntarse entonces:

1. ¿ $c$  causó  $e$  en virtud de que  $c$  sea  $M$  y  $e$  sea  $M^*$ , o  $c$  causó  $e$  en virtud de que  $c$  sea  $P$  y  $e$  sea  $M^*$ ?
2. ¿ $c$  causó  $e$  en virtud de que  $c$  sea  $M$  y  $e$  sea  $P^*$ , o  $c$  causó  $e$  en virtud de que  $c$  sea  $P$  y  $e$  sea  $P^*$ ?

Marras considera que las respuestas plausibles a estas preguntas son:

1. En virtud de que  $e$  sea  $M^*$  (en tanto sea identificada como un suceso mental de tipo  $M^*$ ),  $c$  causó  $e$  en virtud de que  $c$  sea  $M$ , y no en virtud de que  $c$  sea  $P$ . Aún si  $c$  no hubiera sido  $P$  (sino, por ejemplo,  $P'$ , otra propiedad física en la base de superveniencia de  $M$ ), habría causado igualmente un ejemplo de  $M^*$  (por ejemplo, un suceso  $e'$  que ejemplificara una propiedad física diferente,  $P'^*$ , en la base de superveniencia de  $M^*$ ). Pero si  $c$  no hubiera sido  $M$ , tampoco habría sido  $P$  (ya que  $M$  sobreviene a partir de  $P$ ); cualquier cosa que  $c$  pudiera haber causado, entonces, no habría sido un ejemplo de  $P^*$  en la base de superveniencia de  $M^*$  y, por lo tanto, no hubiera causado ejemplo alguno de  $M^*$ .

2. En la medida en que  $e$  sea identificada como de tipo  $P^*$ ,  $c$  causó  $e$  en virtud de que  $c$  sea  $P$  y no en virtud de que sea  $M$ . Si  $c$  hubiera sido  $M$  sin ser  $P$  (sino, por ejemplo,  $P'$ , una base distinta de superveniencia de  $M$ ),  $c$  no habría causado un ejemplo de  $P^*$  (sino, por ejemplo, un ejemplo de  $P^{*'}).$

Marras observa que la causación en-virtud-de depende generalmente de cómo la causa y el efecto son individualizados. La 'quausación' es una relación explicativa no extensional, y qué descripciones puedan ser utilizadas para identificar a los términos de la relación depende de manera crucial de los intereses explicativos. Si lo que se quiere explicar es un ejemplo de un movimiento corporal, una explicación neurofisiológica que apele a causas neuronales será apropiada; pero si lo que se desea es explicar el mismo suceso como un caso de una acción intencional, se pretenderá una explicación en términos de una causa mental.

### 3.4. Un esquema de la dependencia psicofísica

En la sección anterior hemos descrito la manera en que Marras considera que puede hacerse frente al argumento de la superveniencia. Su propuesta se basa esencialmente en la aplicación de la estrategia del *explanandum* dual tal como ha sido presentada para enfrentar el problema de la exclusión explicativa. Puede parecer extraño, prima facie, que el mismo esquema que ha sido de utilidad para salvar las explicaciones de la conducta que apelan a propiedades mentales en su presunta competencia con explicaciones que apelan a propiedades físicas sirva *también* para responder al argumento de la superveniencia. Esta objeción es anticipada por Marras, quien observa al respecto lo siguiente. Kim podría reclamar, señala, que este esquema confunde cuestiones metafísicas acerca de la causación con cuestiones epistemológicas acerca de la explicación causal; en otros términos, la pregunta acerca de qué propiedades de un suceso son *causalmente eficaces* con la pregunta acerca de qué propiedades son *causalmente relevantes* (relevantes para explicaciones causales). Pero esto no es así, sostiene. En primer lugar, porque una propiedad no sería explicativamente relevante a menos que fuera causalmente eficaz. En segundo lugar, porque qué propiedades de un suceso son causalmente eficaces es una cuestión en sí misma relativa a qué *tipo* de efecto está en discusión. Propone que se considere una variante del ejemplo de

Ernest Sosa: un sonoro disparo de arma de fuego bien dirigido causa simultáneamente una herida de bala y sordera en la víctima. En la concepción de Davidson de los sucesos que ha supuesto, hay una causa y un efecto, cada uno individualizado en términos de diferentes propiedades. Qué propiedades de la causa fueron causalmente eficaces para el efecto solo puede ser determinado con referencia al *tipo* de efecto que se esté considerando: que la víctima se vuelva sorda, o que la víctima sufra una herida de bala. Marras considera que se debería sostener que si el efecto es un caso de lo último fue determinado porque la causa fue un disparo bien dirigido, mientras que si el efecto es un caso de lo primero fue determinado porque la causa fue un suceso sonoro.

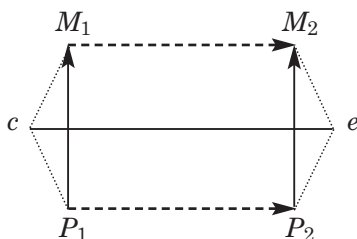
Marras concluye el análisis de este problema al sugerir que Kim está en lo correcto al afirmar que la eficacia causal de las propiedades mentales sobrevinientes no es básica, sino que ‘surge’, o ‘depende de’ la eficacia causal de las propiedades físicas subvenientes. Sin embargo, sostiene, dependiente o ‘heredada’, la eficacia causal es, con todo, eficacia causal; no habría razones para suponerla meramente ‘aparente’ o ‘epifenoménica’, ni tampoco para suponer que se está ante un caso de sobredeterminación, ya que una relación causal singular entre dos sucesos davidsonianos es descripta en diferentes niveles (los niveles sobreviniente y subveniente). Lo que implica la dependencia de los poderes causales de lo mental sobre los poderes causales de lo físico no es que los primeros no sean reales, sino que no son primitivos y deben ser *explicados* en términos de los segundos.

Aun en el caso de que pueda ser considerada satisfactoria, la respuesta de Marras al argumento de la superveniencia descansa, en última instancia, en una explicación de la relación entre propiedades mentales y físicas. Dicho en otros términos, sus afirmaciones referentes a los poderes causales de lo mental requieren de una justificación, y esta justificación debe consistir en una respuesta definida acerca de la relación entre tales propiedades. Esta tarea es afrontada por Marras en un artículo anterior (1997b)<sup>20</sup>. Examinaremos ahora este aspecto de su propuesta.

En este artículo Marras analiza la necesidad de proveer una relación de dependencia psicofísica que rescate el estatus empírico y causal de las leyes psicológicas. Lo que se necesita en última ins-

<sup>20</sup> Algunas de las ideas desarrolladas en este artículo se encuentran esbozadas en sus (1994 y 1997a).

tancia es, en su opinión, una explicación de la idea básica implícita en los puntos de vista physicalistas de que todos los hechos y propiedades (incluidos los hechos y propiedades mentales) dependen, o son determinados por, hechos y propiedades físicas. Su propuesta (esquemática, según su propia admisión) puede ser caracterizada a través del siguiente diagrama:



En la figura,  $c$  y  $e$  representan sucesos individuales (causa y efecto respectivamente), mientras que la flecha que los une representa la relación causal.  $M_1$  y  $P_1$  representan respectivamente la propiedad mental y la propiedad física ejemplificadas por  $c$ , y  $M_2$  y  $P_2$  representan la propiedad mental y la propiedad física (corporal) ejemplificadas por  $e$ . Las líneas oblicuas representan la relación de ejemplificación, las flechas verticales continuas representan la relación de dependencia/supervenencia, y las flechas segmentadas horizontales la relación legal (ya sea estricta o no estricta).

La idea de la dependencia metafísica que debe ser desarrollada puede ser esbozada a partir de las siguientes observaciones:

1. Las propiedades mentales intencionales son construidas como propiedades funcionales de nivel superior, definidas en términos de su rol funcional característico en el sistema en el cual son ejemplificadas. Los estados que ejemplifican esas propiedades son típicamente construidos como disposiciones, donde la atribución de una disposición es la atribución de un estado causal, esto es, un estado que, bajo condiciones apropiadas, es causalmente responsable de determinados patrones de conducta.
2. Tal propiedad funcional  $F$ , en tanto es ejemplificable, es también realizable. Esto es, decir que  $F$  es realizable es decir que existe algún sistema de entidades de nivel inferior que es capaz de jugar el rol distintivo de  $F$ .
3. Debe suponerse también que las propiedades funcionales no son solamente realizables, sino que son *físicamente* realiza-

bles. Para cada propiedad funcional  $F$  hay un sistema de entidades físicas y, en última instancia, microfísicas, que son capaces de jugar el rol causal distintivo de  $F$ , como una consecuencia nomológica del tipo de entidades físicas que son y la manera en que se encuentran organizadas.

4. La estructura física subyacente, o sistema de entidades físicas que realizan la propiedad funcional  $F$  puede ser llamada *mecanismo físico* (uno entre muchos posibles). Este mecanismo implementa las generalizaciones implicadas por la atribución de la propiedad funcional y la disposición asociada con esa propiedad.

En opinión de Marras, si estas observaciones apuntan en la dirección correcta se trata del comienzo de una comprensión de cómo los poderes causales de las propiedades mentales sobrevinientes son determinados por, y dependientes de, los poderes causales de las propiedades físicas subvenientes.

No obstante, esta estrategia debe afrontar, según Marras, algunas objeciones.

En primer lugar, podría objetarse que el esquema propuesto despoja a las propiedades causales de poderes causales propios, ya que estos serían apropiados [*pre-empted*] por los poderes causales de las propiedades físicas que implementan los mecanismos de los cuales aquellas dependen. Marras encuentra esta objeción algo extraña; adscribir poderes causales a algo es típicamente postular la existencia de un mecanismo físico a través del cual pueda ejercer sus poderes causales. Por ejemplo, cuando los científicos conjeturan que fumar causa cáncer, y buscan el mecanismo físico subyacente para sostener esta conjetura, resultaría extravagante sostener, una vez que se lo ha hallado, que la existencia de tal mecanismo muestra que realmente fumar no causa cáncer, sino que lo hacen las propiedades físicas del mecanismo subyacente. De manera similar, concluye, cuando las creencias y deseos causan la conducta lo hacen a través de las propiedades físicas subvenientes de los mecanismos que implementan las conexiones creencia-deseo-conducta.

Una segunda objeción podría ser la afirmación de que el estatus dependiente, de nivel superior, de las propiedades mentales, privaría a las leyes psicológicas de cualquier estatus causal genuino, ya que ellas serían verdaderas solo debido a que ciertas leyes físicas que se aplican a los ejemplos relevantes de ejemplificaciones de propiedades mentales serían verdaderas. Marras repone que no hay razón para suponer que el estatus dependiente de las leyes psicoló-

gicas las hace menos causales; las leyes causales relacionan tipos de sucesos cuyas ejemplificaciones son pares de causa-efecto, y las leyes psicológicas no hacen esto menos que las leyes físicas, aun cuando lo hacen *porque* las propiedades mentales son ontológicamente dependientes de las propiedades físicas.

En tercer lugar, podría objetarse que la existencia de leyes físicas que se aplican a los mecanismos físicos subyacentes torna a las leyes y a las explicaciones psicológicas estrictamente superfluas. Cualquier interacción causal que involucrara a un suceso mental podría ser explicada directamente subsumiendo el suceso bajo leyes físicas, en virtud de las propiedades físicas del suceso. De esta manera, parecería redundante buscar una explicación del mismo suceso subsumiéndolo bajo leyes causales cuyo estatus causal es, en el mejor de los casos, derivado. Más aún, observa, la dualidad de explicaciones de un mismo suceso parece ser incompatible con el principio de exclusión explicativa.

Marras considera que hay varias respuestas a esta tercera objeción. En primer lugar, mientras la dependencia de las propiedades mentales sobre las propiedades físicas confiere estatus causal a las leyes psicológicas, no implica reducibilidad de lo mental a lo físico<sup>21</sup>. En su opinión, la relación de superveniencia preserva la distinción entre niveles de descripción y apoya la realizabilidad múltiple de las propiedades mentales. En consecuencia, las leyes psicológicas y físicas diferirán en alcance y generalidad; las leyes físicas y las leyes psicológicas, en una oportunidad dada, subsumirán clases no equivalentes de transacciones causales.

Por otra parte, las explicaciones son en general reconocidas como no extensionales y dependientes del contexto. La relación explicativa tiene lugar entre hechos, los que implican propiedades o tipos de sucesos, por lo que explicar por qué ha tenido lugar un determinado suceso en el momento *t* es explicar por qué un cierto *tipo* de suceso ha ocurrido en el momento *t*. Y, al solicitar una explicación de un suceso, debe preguntarse por qué *un tipo de suceso dado* ha ocurrido en esta oportunidad, y esta pregunta dependerá

<sup>21</sup> En su (1993), Marras desarrolla detallados argumentos en contra de la tesis de que la superveniencia de las propiedades mentales a partir de las propiedades físicas pueda implicar la reducibilidad de las primeras a las segundas. Cabe observar, no obstante, que el análisis de las posibilidades del reduccionismo se basa en el modelo clásico de reducción propuesto por Nagel (1961), basado en la existencia de leyes puente bicondicionales. Quedan por determinar, por lo tanto, las posibilidades de modelos alternativos de reducción.

del aspecto del suceso que produzca perplejidad o interés, esto es, una cuestión dependiente del contexto.

Esta propuesta, observa Marras, es capaz de explicar cómo las explicaciones racionalizadoras pueden ser explicaciones causales. Difiere radicalmente del enfoque de Davidson en cuanto a que permite mantener la psicología dentro del dominio de las ciencias naturales, mientras que preserva su integridad como una ciencia especial. Permite reivindicar, en síntesis, las principales motivaciones que inspiran a las formas de materialismo no reduccionista: la creencia en la dependencia metafísica de lo mental sobre lo físico, junto con la creencia en la autonomía metodológica de la ciencia de la mente<sup>22</sup>.

#### 4. Los límites de la propuesta

Esta extensa descripción de la propuesta de Marras ha tenido como objetivo, tal como hemos sostenido en el apartado 3.3, mostrar que, en última instancia, su versión de la estrategia del *explanandum* dual desemboca en la necesidad de explicar la causación psicofísica.

La propuesta de Marras plantea dos cuestiones de interés para analizar: 1) si el esquema propuesto salva completamente los poderes causales de lo mental, o los preserva únicamente en su propio nivel, con la consecuencia de que la conducta no estaría determinada por lo mental, sino solo por sucesos físicos; y 2) si el problema de la exclusión no se plantea también con respecto a la causación de un suceso mental por otro suceso mental. Estas dos cuestiones son centrales para determinar si el intento de Marras puede ser considerado exitoso. Las examinaremos sucesivamente.

La primera cuestión que analizaremos consiste en la posibilidad de que las propiedades mentales realizadas por propiedades físicas solo posean eficacia causal en su propio nivel; en otros términos, que no sean capaces de afectar causalmente a otras propiedades que no sean mentales.

Es evidente que el esquema propuesto por Marras no puede ser calificado de epifenomenista (o cuasi-epifenomenista), por dos razones obvias: en primer lugar, porque lo mental no depende cau-

<sup>22</sup> Marras parece adjudicar un cierto carácter programático al esquema propuesto. Hasta donde hemos podido ver, no hay elaboraciones posteriores de la propuesta, por lo que parece razonable juzgarla aún reconociendo tal carácter provisorio y mejorable.

salmente de lo físico, sino que sobreviene a partir de él; en segundo lugar, porque si bien lo mental no tiene efectos sobre el mundo físico, es capaz de causar la ejemplificación de otras propiedades mentales. La propuesta de Marras, si es considerada exitosa, permitiría salvar las explicaciones de la ejemplificación de propiedades mentales (el hecho *explanandum*) que serían causadas por la ejemplificación de otras propiedades mentales<sup>23</sup>.

Ahora bien, hemos visto que, para Marras, el contrafáctico  $\neg Rc \rightarrow \neg Be$  (donde  $R$  es una propiedad racionalizadora y  $B$  es un movimiento corporal) es falso, dado que solo una propiedad neurofisiológica es capaz de causar  $B$ ; las propiedades mentales no causan la ejemplificación de propiedades físicas. Asimismo es notorio, en el esquema de la dependencia psicofísica propuesto por Marras, que no hay ninguna otra conexión entre propiedades mentales y propiedades físicas que no sea la relación de dependencia - realización. Las propiedades mentales solo mantienen relaciones nomológico-causales con otras propiedades mentales. No hay causación descendente; las propiedades mentales no afectan causalmente a las propiedades físicas del suceso que ejemplifica ambas clases de propiedades. Si esto fuese así, lo que deseamos y creemos no tendría influencia causal sobre el mundo físico, en tanto este último está determinado (causado) únicamente por propiedades físicas. Las propiedades puramente físicas de una acción (esto es, los cambios fisiológicos y mecánicos acaecidos en el cuerpo) solo pueden ser explicadas con referencia a propiedades puramente físicas.

<sup>23</sup> Uno de los árbitros de *Análisis Filosófico* ha considerado que la idea de que las acciones tienen una parte física requiere argumentación. Esta observación parece correcta, partiendo del supuesto de que no existe una caracterización única del concepto de acción: '[F]rente a la hipótesis de que las acciones son eventos particulares que ejemplifican diversas propiedades, está la hipótesis opuesta de que no hay tales acciones individuales, sino que las acciones son en sí mismas propiedades o universales cuyas instancias son los *agentes* que las ejecutan. Mientras en el primer caso el predicado 'cerrar la ventana' designaría una propiedad que se instanciaría en acciones particulares que también instancian muchas otras propiedades, en el segundo caso, el predicado designaría a la acción misma cuyas instancias particulares son agentes' (Nino, 1987, p. 11). Bajo la primera hipótesis puede admitirse que las acciones tienen una parte física (ejemplifican una propiedad física) y que esta parte física tiene una causa y, potencialmente, una explicación causal. La elección por una u otra de las posibles caracterizaciones de la acción (o de otras alternativas) requiere de un análisis detallado que no podemos desarrollar aquí. Supondremos, por lo tanto, que la caracterización de las acciones como una clase de sucesos que ejemplifican conjuntos de propiedades, entre ellas propiedades físicas, es admisible a los fines del presente argumento.



¿En qué afecta esto a la propuesta de Marras? En escasa medida o de manera decisiva, dependiendo de los propósitos de la propuesta. Si la estrategia del *explanandum* dual limitara su alcance a salvar el rol explicativo y el poder causal de las propiedades (sin fijar una preferencia acerca del nivel en el cual estos poderes causales son ejercidos), podríamos considerar su intento como exitoso. Sin embargo, el objetivo general de la propuesta parece ser más ambicioso. El intento de Marras parece ser claramente construir una forma de salvar la causación mental y las explicaciones psicológicas de una manera enteramente compatible con el materialismo no reduccionista. Transcribimos al respecto una observación ilustrativa:

Mi objetivo aquí es mostrar que hay una versión de materialismo no reduccionista que es capaz de sobrevivir a un importante tipo de objeción de la cual se ha dicho que el monismo anómalo de Davidson es vulnerable: la objeción *epifenomenista*. Esta objeción consiste en que el monismo anómalo de Davidson es incapaz de proveer una solución satisfactoria al problema de la causación mental –el problema de explicar cómo los sucesos mentales entran en relaciones causales con otros sucesos– (1994, pp. 466-467).

Estas afirmaciones contribuyen a esclarecer el sentido de la propuesta de Marras. No se trata simplemente de mitigar las consecuencias indeseables de los argumentos de la exclusión (como el irrealismo o la irrelevancia explicativa), ni de salvar solo parcialmente la causación mental. Se trata de mostrar cómo los argumentos de la exclusión son erróneos, y cómo lo mental puede conservar los poderes causales que ordinariamente le adscribimos. Si este es el objetivo fundamental de la propuesta, no puede menos que concluirse que, en el mejor de los casos, solo se ha logrado parcialmente.

Examinemos ahora la segunda objeción a la propuesta de Marras. Hay autores que conciben a lo mental como causalmente aislado (Sabatés, 1997). Esto es, los sucesos o propiedades mentales no participan en ninguna clase de interacción causal, ni con otros sucesos o propiedades mentales o con sucesos o propiedades físicos; dependen (no causalmente) de sucesos físicos, vía superveniencia mereológica. En este modelo obviamente no puede existir sobre-determinación para los sucesos o propiedades mentales, ya que estos solo están determinados por una relación de superveniencia, y no causalmente. Ahora bien, en el modelo de Marras las propiedades mentales están determinadas (causalmente) por otras propiedades

mentales y por propiedades físicas (por superveniencia/realización). Esta situación conduce a pensar si se trata de un posible caso de sobredeterminación.

Sabatés (2001) observa que una situación problemática de posible sobredeterminación entre dos determinantes (uno causal y otro por superveniencia) solo puede plantearse si se admite un principio de exclusión ontológica que sostenga que no pueden existir dos conjuntos de propiedades completos e independientes cada uno de los cuales sea suficiente para la ejemplificación de una propiedad singular. Tal principio es inevitable, observa, si se desea defender un principio de exclusión para los casos de causación de una propiedad mental a otra propiedad mental. Si este principio es aceptado, lo mental pierde completamente sus poderes causales; la base de superveniencia física de la propiedad mental excluye a la propiedad mental que es su supuesta causa.

Sin embargo, señala Sabatés, este principio no puede aceptarse sin alguna cualificación. No es posible admitir que cualquier par de relaciones de necesidad [*necessitation*] compitan o se excluyan mutuamente. Por ejemplo, que la estrella de la tarde sea visible hoy necesita [*necessitates*] que Venus sea visible hoy, pero que la estrella de la tarde sea visible hoy no parece impedir que el suceso consistente en que la noche sea clara en Kansas cause que Venus sea visible hoy. Por lo tanto, para que el principio de exclusión de determinantes tenga posibilidades de ser plausible se debe admitir que las relaciones que pueden excluirse mutuamente son solo aquellas que son tanto metafísica cuanto conceptualmente contingentes. Si una de las relaciones que presuntamente entran en competencia no se encuentra vinculada de manera contingente al efecto, la exclusión no tiene lugar. En este caso, señala Sabatés, la carga de la prueba recae sobre el defensor de la exclusión: es él quien debe mostrar que la relación de superveniencia entre lo mental y lo físico es una relación metafísica/conceptualmente contingente. Si esto no puede lograrse, podría mantenerse la causación de lo mental a lo mental. En este caso, se salvaría la mitad de la causación mental, dando lugar a una interesante posibilidad para una estrategia del *explanandum* dual.

Ahora bien, el esquema de dependencia psicofísica propuesto por Marras no apela solamente a una relación de superveniencia indeterminada entre lo mental y lo físico; como hemos visto, apela a la relación de realización física entre ambas familias de propiedades.

¿Debe ser entendida la relación de realización física como una relación contingente? Una caracterización estándar de esta noción es la siguiente: un suceso  $e$  que sea F realiza el que  $e$  sea G solo si i)  $e$  es F, ii)  $e$  es G, iii) para todo  $e$  es físicamente necesario que si  $e$  es F entonces  $e$  es G y, iv) el que  $e$  sea F explica el que  $e$  sea G. Sin embargo, esta caracterización no capta el sentido que esta noción posee para algunos funcionalistas, que sostienen que un estado físico puede realizar un estado funcional de manera contingente. El que exista una situación de competencia, y posible sobredeterminación, entre ambos determinantes dependerá entonces de cómo se conciba la relación de realización física. Si bien no hay argumentos concluyentes para mostrar que se produce de manera efectiva una competencia entre determinantes, no hay tampoco razones para pensar que el peligro de la ineficacia causal de lo mental ha desaparecido completamente en virtud del esquema propuesto por Marras<sup>24</sup>.

La aceptabilidad de la estrategia del *explanandum* dual propuesta por Marras, en síntesis, descansa en la plausibilidad de sus puntos de vista referentes a las relaciones entre propiedades mentales y propiedades físicas. Hemos intentado mostrar que el esquema de la dependencia psicofísica propuesto por Marras presenta aspectos dudosos y cuestiones sin resolver; en particular, si nuestra interpretación es correcta, no consigue asegurar un rol para las propiedades mentales en la producción de propiedades físicas, por lo que lo mental no jugaría un rol activo en la producción de la conducta. En tal caso, su versión de la estrategia de los dos *explananda* no podría ser considerada exitosa.

Sin embargo, el fracaso al menos parcial de esta versión de la estrategia de los dos *explananda* frente al problema de la ineficacia causal de las propiedades mentales no tiene necesariamente que ser extensivo a todas las posibles aplicaciones de aquella. Su fertilidad

<sup>24</sup> Kim, en las versiones más recientes de sus argumentos en favor de la exclusión causal (en especial, en el argumento de la superveniencia) parece inclinarse no solo por afirmar la ineficacia causal de los sucesos mentales en relación con los sucesos físicos, sino por sostener la ineficacia causal completa (esto es, los sucesos mentales serían incapaces de causar nada). Es ilustrativa su afirmación sobre la causación de un suceso mental por otro: 'en el caso de la supuesta causación  $M-M^*$ , la situación es casi como una serie de sombras proyectada por un automóvil en movimiento: no hay conexión causal entre la sombra del auto en un instante y su sombra en un instante posterior, siendo cada una de ellas un efecto del movimiento del auto' (1998, p. 45). El peligro de sobredeterminación no causal desaparece bajo esta descripción, pero al costo de concebir lo mental como completamente inerte.

futura y la amplitud de su campo de aplicación merecen ser objeto de indagaciones posteriores.

## Bibliografía

- Baker, Lynn (1993), "Metaphysics and Mental Causation", en J. Heil y A. Mele (eds.), 1993.
- Block, Ned (1990), "Can the Mind Change the World?", en C. y G. Macdonald (eds.), 1995.
- Burge, Tyler (1993), "Mind-Body Causation and Explanatory Practice", en J. Heil y A. Mele (eds.), 1993.
- Dretske, Fred (1990), "Does Meaning Matter?", en C. y G. Macdonald (eds.), 1995.
- Heil, John, y Alfred Mele (eds.) (1993), *Mental Causation*, New York, Oxford University Press.
- Hempel, Carl (1965), "Aspectos de la explicación científica", en C. Hempel (1965), *La explicación científica. Estudios sobre la filosofía de la ciencia*, Barcelona, Paidós, 1988.
- Jackson, Frank y Philip Pettit (1990), "Program Explanation: a General Perspective", *Analysis*, 50, pp. 107-117.
- Kim, Jaegwon (1987) "Explanatory Realism, Causal Realism, and Explanatory Exclusion", en P. French, T. Uehling and H. Wettstein (eds.) *Midwest Studies in Philosophy*, vol. 12, Notre Dame, University of Notre Dame Press.
- (1989a) "Mechanism, Purpose, and Explanatory Exclusion", *Philosophical Perspectives*, 3, *Philosophy of Mind and Action Theory*, James E. Tomberlin (ed.), Atascadero, Cal., Ridgeview Publishing Company. Reimpreso en J. Kim (1993b).
- (1989b), "The Myth of Nonreductive Materialism", *Proceedings and Addresses of the American Philosophical Association*, 63, pp. 31-47. Reimpreso en J. Kim (1993b)
- (1990), "Explanatory Exclusion and the Problem of Mental Causation", en Cynthia y Graham Macdonald (eds.), 1995.
- (1991) "Dretske on How Reasons Explains Behavior", en Brian McLaughlin (ed.), *Dretske and his Critics*, Oxford, Basil Blackwell, 1991. Reimpreso en J. Kim (1993b).
- (1993a), "The Nonreductivist's Troubles with Mental Causation", en John Heil y Alfred Mele (eds.), *Mental Causation*, Oxford, Oxford University Press, 1989. Reimpreso en J. Kim (1993b).

- (1993b), *Supervenience and Mind*, Cambridge, Cambridge University Press.
- (1998), *Mind in a Physical World. An Essay on the Mind-Body Problem and Mental Causation*, Cambridge, Mass., The MIT Press.
- Macdonald, Cynthia y Graham Macdonald (eds.) (1995), *Philosophy of Psychology. Debates on Psychological Explanation*, Volume One, London, Blackwell.
- Malcolm, Norman (1968), "The Conceivability of Mechanism", *Philosophical Review*, 77, pp. 45-72. Reimpreso en *Free Will*, Gary Watson (ed.), Oxford University Press, 1982.
- Marras, Ausonio (1993), "Psychophysical Supervenience and Nonreductive Materialism", *Synthese*, 95, pp. 275-304.
- (1994), "Nonreductive Materialism and Mental Causation", *Canadian Journal of Philosophy*, 24, 3, pp. 465-493.
- (1997a), "The Causal Relevance of Mental Properties", *Philosophia*, 25, 1-4, pp. 389-400.
- (1997b), "Metaphysical Foundations of Action Explanation", en G. Holmström Hintikka y R. Tuomela (eds.), *Contemporary Action Theory*, vol. 1, Dordrecht, Kluwer.
- (1998), "Kim's principle of Explanatory Exclusion", *Australasian Journal of Philosophy*, 76, pp. 439-451.
- (2000), "Critical Notice of *Mind in a Physical World* by Jaegwon Kim", *Canadian Journal of Philosophy*, 30, pp. 137-160.
- Montgomery, Richard (1995), "Non-Cartesian Explanations Meet the Problem of Mental Causation", *Southern Journal of Philosophy*, 33, 2, pp. 221-241.
- Nino, Carlos S. (1987), *Introducción a la filosofía de la acción humana*, Buenos Aires, EUdeBA.
- Sabatés, Marcelo (1996), "Kim on the Metaphysics of Explanation", *Manuscrito*, XIX, 2, p. 93-110.
- Sabatés, Marcelo (1997), "Should a Cognitive Psychologist Worry About the Causal Inefficacy of the Mental?", en B. Niggemeyer (ed.), *The Cognitive Level*, Duisburg, LAUD Verlag.
- Sabatés, Marcelo (2001), "Varieties of Exclusion", *Theoria*, Vol. 16, N° 40, pp. 13-42.
- Vicente, Agustín (2002), "The Dual *Explanandum* Strategy", *Crítica*, 34, 101, pp. 73-96.
- Von Wrigth, Georg H. (1971), *Explicación y comprensión*, Madrid, Alianza, 1987.