

¿HAY PROPIEDADES PSICOLÓGICAS QUE NO DEPENDEN DE LO FÍSICO?

Francisco Javier Vidal. Universidad de Navarra

«Esto ocurre mucho en filosofía; se argumenta que 'todo lo que encontramos' es tal y tal, pero resulta que el argumentador ha excluido de su idea de 'encontrar' el tipo de cosa que dice que no 'encontramos'».

G. E. M. ANSCOMBE, *Causality and Determination*.

1. La superveniencia psicofísica.

Casi todo materialismo, también llamado 'fiscalismo', en la filosofía contemporánea de la mente comienza por sostener la tesis, introducida por J. Kim (cfr, p.e, 1990), de la *superveniencia psicofísica*, o sobrevenir de las propiedades psicológicas en las físicas:

(SP) Para toda propiedad psicológica P , (i) si algún x tiene P entonces necesariamente existe alguna propiedad física F tal que x tiene F , y (ii) necesariamente todo lo que tiene F tiene P .

Por ejemplo, para todo caso de dolor orgánico, (i) si algún organismo siente dolor entonces necesariamente hay alguna propiedad neurofisiológica N tal que ese organismo tiene N , y (ii) necesariamente todo organismo que tiene N siente dolor.

A veces se discute si (SP) expresa una mera *dependencia* entre lo psicológico y lo físico o una estricta *determinación* de lo psicológico por lo físico. J. Heil (cfr, 1992, 58-76) considera que (i) expresa mera dependencia, mientras que (ii), dado (i), expresa estricta determinación. La razón que podemos dar es la siguiente. (i) formula un principio de extrema liberalidad sobre la conexión psicofísica. Supongamos que es verdadera la tesis de la *realización física de lo psicológico*: las propiedades físicas y sólo las propiedades físicas realizan, en algún sentido de 'realizar', las propiedades psicológicas. Resulta que (i) expresa algo compatible con la tesis de la *realización múltiple* de lo psicológico en lo físico:

(RM) Para toda propiedad psicológica P , hay un conjunto de propiedades físicas heterogéneas, $F1, F2, F3 \dots$, tal que cada propiedad F es suficiente, pero no necesaria, para realizar P ; así P implica la disyunción $\langle F1 \text{ o } F2 \text{ o } F3 \dots \rangle$ y esta disyunción implica P .

Por ejemplo, (RM) se hace cargo de una hipótesis muy plausible sobre el dolor orgánico: la misma propiedad psicológica, el mismo tipo de dolor, es realizado por una propiedad neurofisiológica $N1$ en los hombres, una propiedad $N2$ en los reptiles, una propiedad $N3$ en las aves, etc...

Pues bien, según (i), si algún x tiene P , entonces necesariamente existe *alguna* propiedad física F tal que x tiene F , pero no es verdad que necesariamente existe *una y sólo una* propiedad F tal que x tiene F . Luego (i) es formulable como: si algún x tiene

P entonces necesariamente existe una u otra propiedad física F , de un conjunto $\langle F1, F2, F3 \dots \rangle$, tal que x tiene F . Por tanto, (i) es compatible con (RM): si algún x tiene P entonces necesariamente existe una u otra propiedad física F , del conjunto relevante para la realización de P , tal que x tiene F y F realiza P . De este modo de (i) se sigue:

(i') Para todo x e y , si x es indiscernible de y en todas sus propiedades psicológicas P , no es necesario que x sea indiscernible de y en todas sus propiedades físicas F .

Pero (ii), dado (i), expresa un principio de fuerte constricción de lo psicológico por lo físico. La prueba de ello es que (ii) expresa algo incompatible con la tesis, innominada en la literatura filosófica, de la *superviniencia múltiple* de lo psicológico en lo físico:

(SM) Para toda propiedad psicológica P , (i) si algún x tiene P entonces necesariamente existe alguna propiedad física F tal que x tiene F , y (ii') necesariamente todo lo que tiene F tiene una u otra propiedad P de un conjunto $\langle P1, P2, P3, \dots \rangle$ de propiedades psicológicas del mismo rango.

Formulo (SM) para expresar la siguiente hipótesis: si una propiedad física es lo bastante compleja como para que en ella sobrevenga una propiedad psicológica, es razonable pensar que en ella no necesariamente tiene que supervenir esa propiedad psicológica sino sólo que necesariamente tiene que hacerlo cualquier propiedad psicológica del mismo rango. Por ejemplo, si en un organismo una propiedad neurofisiológica N es tal que en ella superviene una propiedad psicológica $P1$ como el dolor orgánico, es razonable suponer que en N no necesariamente tiene que supervenir $P1$ sino sólo que necesariamente tiene que hacerlo una propiedad $P2$ como la sensación de frío o una propiedad $P3$ como un escozor o cualquier otra propiedad del mismo rango fenomenológico.

Pues bien, (ii), a diferencia de (ii'), expresa un principio de determinación de toda diferencia psicológica entre dos individuos en términos de alguna diferencia física entre ellos: no hay diferencia psicológica sin diferencia física. Según (ii), si, al tener P , algún x necesariamente tiene alguna propiedad física F , entonces necesariamente todo lo que tiene la propiedad física F , la F de x , tiene *una y sólo una* propiedad psicológica P , la P de x . Luego (ii) es obviamente incompatible con (SM): si, al tener P , algún x necesariamente tiene alguna propiedad física F , entonces necesariamente todo lo que tiene F , la F de x , *no* tiene una u otra propiedad psicológica P , de un conjunto $\langle P1, P2, P3, \dots \rangle$, sino que necesariamente tiene la P de x . Así de (ii) se sigue:

(ii') Para todo x e y , si x es indiscernible de y en todas sus propiedades físicas F , entonces necesariamente x es indiscernible de y en todas sus propiedades psicológicas P .

2. Un argumento contra la superviniencia psicofísica.

No es el único objetivo de esta exposición formular un argumento contra la superviniencia psicofísica. Aunque el título parecía mencionar sólo ese objetivo, la idea es ir también un paso atrás: se trata tanto de mostrar que algunas propiedades psicológicas no son dependientes de lo físico como de mostrar que *hay* cierto tipo de propiedades psicológicas. Como se verá, la idea es refutar también algunas objeciones de J. Fodor («A Modal Argument for Narrow Content», 1991) contra la *existencia* de ese tipo de propiedades en términos de la crítica de L. R. Baker (*Explaining Attitudes*, 1995). Ahora bien, resulta que si las propiedades de ese tipo existen, entonces no supervienen en propiedades físicas. Precisamente porque, de existir, no hay superviniencia psicofísica de tales propiedades,

me interesa vivamente justificar la tesis de que existen. Por eso, a continuación ofrezco un argumento contra la superveniencia psicofísica de ese tipo de propiedades, suponiendo, que es lo que después estará en cuestión, que en efecto existen.

Vamos a introducir, de entrada, un género de propiedades psicológicas indisputadas, las *propiedades de contenido*, así nominadas porque se individualizan por su contenido intencional: por ejemplo, *el pensamiento de que está lloviendo*. En la literatura filosófica es frecuente referirse a la ejemplificación de estas propiedades en términos de *actitudes proposicionales* de un sujeto. La idea es que la atribución de una propiedad de contenido a un sujeto menciona la actitud psicológica del sujeto -p.e, *tener un pensamiento*- hacia una proposición u objeto evaluable semánticamente según verdad y falsedad -p.e, *está lloviendo*-.

Permítaseme recurrir a un experimento imaginativo del tipo que se ha hecho popular en la filosofía de la mente desde el célebre experimento de H. Putnam (1975) sobre el agua. Las claves del experimento se hallan en T. Burge (cfr, 1979, 1986) y en Fodor (cfr, 1988, cap. 2) pero los detalles son de L. R. Baker (cfr, 1995, cap. 2) y míos. Lo que se trata de mostrar es que existen intuitivamente dos tipos de propiedades de contenido, tesis que el propio Fodor no discute. Lo que hace Fodor es negar, y se lo discutiremos después, que uno de esos tipos de propiedades sea un tipo de propiedad *psicológica*.

Tomemos la palabra 'jade'. En español la palabra 'jade' denota dos tipos de piedras o minerales, la jadeita y la nefrita, con estructuras distintas pero que son extremadamente semejantes en su aspecto, en especial, por su color verde oscuro en los casos paradigmáticos. De este modo, a pesar de su similitud, la jadeita es una piedra que se encuentra principalmente en Centroamérica y, en parte por su escasez, es mucho más valiosa que la nefrita, que se encuentra por todo el mundo. De hecho, aunque los expertos en minerales y en joyas conocen perfectamente sus diferencias estructurales, lo cierto es que la palabra 'jade' se refiere a las dos. Esta circunstancia refrenda la tesis de que el lenguaje es una institución social: a la hora de mantener o cambiar la referencia de la palabra, ha pesado más el uso social del hablante común, que no es capaz de distinguir la jadeita de la nefrita, que el criterio de los expertos, a quienes se reserva un juicio decisorio sólo cuando, por ejemplo, hay que establecer el valor de un trozo de jade.

Imaginemos ahora que existe una comunidad lingüística, posiblemente en Centroamérica, en la que, como en la nuestra, también los expertos conocen perfectamente la distinción entre la jadeita y la nefrita y, además, la palabra 'jade' sólo denota la jadeita, debido a un peso social excesivo de los expertos por el valor de la jadeita. Por lo tanto, la piedra menos valiosa, la nefrita, recibe otro nombre y un trozo de nefrita en ningún caso constituye un trozo de jade. Imaginemos también que un hablante común de esta comunidad es tan incapaz como nosotros de distinguir la jadeita de la nefrita. Aunque, dado el carácter social del lenguaje, un hablante de esa comunidad que usa la palabra 'jade' sólo denota la jadeita, él tiene las mismas creencias sobre la jadeita y la nefrita que las que tenemos nosotros: apenas, que se trata de piedras de color verde oscuro.

Tomemos un hablante común de cada comunidad, la nuestra y la imaginada. Los dos están diciendo, esto es, emitiendo los sonidos: «Este trozo de jade es verde oscuro». Ahora evaluemos semánticamente la emisión de cada uno. Resulta que, dado el carácter social del lenguaje, las dos emisiones tienen condiciones de verdad diferentes. La emisión de «Este trozo de jade es verde oscuro» por nuestro hablante es verdadera si, y sólo si, *este trozo de jadeita o de nefrita es verde oscuro*. La emisión de «Este trozo de jade es verde

oscuro» por el hablante imaginado es verdadera si, y sólo si, *este trozo de jadeita es verde oscuro*. Así, supongamos que el trozo de jade al que ambos se refieren es un trozo de nefrita. Entonces, la emisión de nuestro hablante es verdadera mientras que la emisión del hablante imaginado es falsa.

Pero, el habla con sentido expresa, entre otras cosas, los pensamientos de los hablantes. Supongamos que individuamos los pensamientos en términos de las *creencias* que tienen los hablantes. Entonces, como nuestro hablante y el imaginado tienen las mismas creencias sobre la jadeita y la nefrita, cuando dicen algo sobre el jade expresan los mismos pensamientos: tanto uno como otro, al decir «Este trozo de jade es verde oscuro», expresan *el pensamiento de que este trozo de piedra-jade es verde oscuro*. Por el contrario, supongamos que individuamos los pensamientos en términos del *lenguaje* que usan los hablantes en su comunidad. Entonces, dado el carácter social del lenguaje, cuando ambos dicen algo sobre el jade expresan pensamientos diferentes: diciendo «Este trozo de jade es verde oscuro», nuestro hablante expresa *el pensamiento de que este trozo de jadeita o de nefrita es verde oscuro* mientras que el hablante imaginado expresa *el pensamiento de que este trozo de jadeita es verde oscuro*.

No voy a discutir la cuestión de si ambos tipo de pensamiento existen. Si aceptamos una distinción intuitiva entre las creencias de un sujeto y su uso del lenguaje, podemos atribuir ambos tipos de pensamiento. Tenemos, por tanto, dos tipos de propiedades de contenido. En un caso, las propiedades de contenido son *internas* a los sujetos, en el sentido de que se individualizan en términos de las *creencias* de los sujetos sobre el mundo independientemente de cómo se relacionen de hecho con él: tanto nuestro hablante como el imaginado *creen* (y entonces piensan) que un trozo de piedra-jade es verde oscuro, aunque uno vive en una comunidad en que las cuestiones de valor no han impedido seguir refiriéndose con 'jade' a la jadeita y a la nefrita mientras que el otro vive en una comunidad en que eso es lo que ha ocurrido. Este tipo de propiedades de contenido, como el pensamiento -compartido por los dos hablantes- de que un trozo de piedra-jade es verde oscuro, son los llamados *contenidos estrechos*. En el otro caso, las propiedades de contenido son *externas* a los sujetos en el sentido de que se individualizan en términos de cómo los sujetos se *relacionan* de hecho con el mundo independientemente de sus creencias sobre él: sólo por usar el lenguaje de sus comunidades, nuestro hablante piensa que un trozo de jadeita o de nefrita es verde oscuro y el imaginado piensa que un trozo de jadeita es verde oscuro, de manera que sus pensamientos difieren en términos del peso relativo del valor de la jadeita en sus respectivas comunidades, es decir, en términos de sus distintas *formas de vida*, aunque ambos tengan las mismas creencias sobre el jade. Este tipo de propiedades de contenido, como el pensamiento -exclusivo de nuestro hablante- de que un trozo de jadeita o de nefrita es verde oscuro, son los llamados *contenidos amplios*.

Admitamos, sin más, que los contenidos estrechos son propiedades psicológicas. La tesis de Fodor, que discutiremos en las próximas secciones, es que los contenidos amplios no son propiedades *psicológicas*, no que no los haya. Pero, para formular el argumento contra la superveniencia psicofísica, vamos a conceder que sí lo son.

Pues bien, supongamos que nuestro hablante, digamos, Oscar1, y el hablante imaginado, Oscar2, son duplicados físicos: ambos son físicamente indiscernibles, esto es, indiscernibles en todas sus propiedades físicas. Por lo tanto, cuando ambos dicen «Este trozo de jade es verde oscuro» ambos se encuentran en el mismo estado físico. Más aún, como

cada emisión de sonidos es una ejemplificación de un patrón físico de sonidos, la indiscernibilidad física de Oscar1 y Oscar2 es una explicación del hecho de que cada vez que Oscar1 emite «Este trozo de jade es verde oscuro», también lo hace Oscar2, ejemplificando el mismo patrón físico. Pero, si hablan con sentido, entonces cada vez que Oscar1 expresa un pensamiento al realizar su emisión, también lo hace Oscar2. Supongamos inicialmente que los pensamientos expresados en las emisiones son contenidos estrechos.

Ahora, se trata de preguntarse si los contenidos estrechos supervienen en algunas propiedades físicas de Oscar1 y Oscar2, esto es, si (SP) es válido de los contenidos estrechos. No podemos dar por sentado, como hicimos para distinguir contenidos, que, en un sentido, la emisión de «Este trozo de jade es verde oscuro» tanto por Oscar1 como por Oscar2 expresa *el mismo pensamiento*, que este trozo de piedra-jade es verde oscuro. Ello implicaría una petición de principio: si hay que evaluar (SP) para los contenidos estrechos, hay que mostrar precisamente si, *por el hecho de ser físicamente indiscernibles*, Oscar1 y Oscar2 necesariamente expresan o no *el mismo pensamiento* al realizar la misma emisión. Lo que sí debemos dar por sentado es que tanto Oscar1 como Oscar2 tienen las mismas creencias sobre el jade antes de tener el pensamiento expresado por sus respectivas emisiones de «Este trozo de jade es verde oscuro». De otro modo, es también evidente, por no decir una pura perogrullada, que, teniendo creencias diferentes, tendrían pensamientos diferentes. Así, pues, la tesis es que es plausiblemente verdadero

(SPa) Para todo contenido estrecho, (i) si algún x tiene ese contenido estrecho entonces necesariamente existe alguna propiedad física F tal que x tiene F , y (ii) necesariamente todo lo que tiene F tiene ese contenido estrecho.

La razón para sostener (SPa) va a ser, en efecto, de mera plausibilidad respecto de su verdad. Las cuestiones son: (i) si Oscar1 tiene el pensamiento de que un trozo de piedra-jade es verde oscuro, ¿entonces necesariamente hay alguna propiedad física F tal que Oscar1 tiene F ?, y (ii) ¿necesariamente todo lo que tiene F , *ex hypothesi*, Oscar2 dada la indiscernibilidad física con Oscar1, tiene ese pensamiento? Según (a), el pensamiento de Oscar1 de que un trozo de piedra-jade es verde oscuro ha de ser el contenido estrecho expresado por su emisión. En ese caso, el pensamiento es interno a Oscar1 en el sentido de que depende de, más precisamente, *superviene en*, las creencias particulares, idiosincrásicas, de Oscar1 sobre el jade. Pero, parece plausible suponer que, siendo interno e idiosincrásico el pensamiento de Oscar1 por superviniencia en sus creencias o, como lo formula figurativamente Putnam, estando el pensamiento de Oscar1 «en su cabeza», (i) entonces necesariamente hay alguna propiedad física F tal que es una propiedad física interna «no más allá de la piel»- y, al tener ese pensamiento, Oscar1 ejemplifica F . El mismo carácter interno e idiosincrásico del pensamiento de Oscar1 parece justificar la afirmación de que (ii) necesariamente todo lo que tiene la propiedad física interna F que Oscar1 tiene cuando piensa que un trozo de piedra-jade es verde oscuro, también tiene ese pensamiento: como, *ex hypothesi*, Oscar2 tiene las mismas propiedades físicas internas que Oscar1 y, por tanto, tiene F entonces necesariamente tiene el pensamiento de que un trozo de piedra-jade es verde oscuro.

He hablado de plausibilidad porque, sin demostrar nada, se ve que (SPa) pueda ser verdadero. Pero, lo relevante es que se vea, bajo el supuesto de que los contenidos amplios son propiedades psicológicas, que (SP) no es válido de tales contenidos. De nuevo, cada vez que Oscar1 dice «Este trozo de jade es verde oscuro», Oscar2 también lo dice,

y ambos expresan algún pensamiento. Supongamos ahora que los pensamientos expresados en las emisiones son contenidos amplios.

¿Por qué es imposible que los contenidos amplios supervengan en algunas propiedades físicas de Oscar1 y Oscar2? Antes que nada, no podemos asumir, como al distinguir contenidos, que, en un sentido, el pensamiento expresado por la emisión de Oscar1 es el pensamiento de que un trozo de jadeita o de nefrita es verde oscuro mientras que el pensamiento expresado por la emisión de Oscar2 es el pensamiento de que un trozo de jadeita es verde oscuro, esto es, no podemos asumir de entrada que, en un sentido, la misma emisión por Oscar1 y por Oscar2 expresa *pensamientos diferentes*. Sería un caso de petición de principio: si hay que evaluar (SP) para los contenidos amplios, hay que mostrar precisamente si, *por el hecho de ser físicamente indiscernibles*, Oscar1 y Oscar2 necesariamente expresan o, por el contrario, *no expresan el mismo pensamiento* al realizar la misma emisión. Pero, sí podemos asumir que Oscar1 y Oscar2 son habitantes de diferentes comunidades lingüísticas, comunidades en que el significado de las palabras es distinto: no obstante, para que esta diferencia no afecte la indiscernibilidad física de Oscar1 y Oscar2, cabe estipular que la conducta lingüística, el conjunto de emisiones físicas verbales, de los habitantes de ambas comunidades no se distingue en nada y, por tanto, ambas comunidades de habla son físicamente indiscernibles. Del mismo modo que si Oscar1 realiza la emisión «Este trozo de jade es verde oscuro» Oscar2 también lo hace, independientemente de cuál sea el pensamiento que expresa cada uno, para cada habitante de una comunidad que realiza la emisión «Este vaso es de cerveza» hay un habitante de la otra comunidad, físicamente indiscernible del primero, que también lo hace, independientemente de cuáles sean los pensamientos expresados. Pues bien la tesis es que

(SPb) Para todo contenido amplio, (i) si algún x (tiene ese contenido amplio entonces necesariamente existe alguna propiedad física F tal que x tiene F , y (ii) necesariamente todo lo que tiene F tiene ese contenido amplio; no puede ser verdadero. Las cuestiones son obviamente: (i) si Oscar1 tiene el pensamiento de que un trozo de jadeita o de nefrita es verde oscuro, ¿entonces necesariamente hay alguna propiedad física F tal que Oscar1 tiene F ?, y (ii) ¿necesariamente todo lo que tiene F , *ex hypothesi*, Oscar2 dada la indiscernibilidad física con Oscar1, tiene ese pensamiento? Según (b), el pensamiento de Oscar1 de que un trozo de jadeita o de nefrita es verde oscuro ha de ser el contenido amplio expresado por su emisión. Por lo tanto, el pensamiento es externo a Oscar1 en el sentido de que depende de, realmente *superviene en*, la institución social del lenguaje en su comunidad, sea cual sea la constitución interna e idiosincrásica de Oscar1. Con otras palabras: el pensamiento es externo a Oscar1 en el sentido de que superviene en la *propiedad relacional* de Oscar1 consistente en usar la palabra 'jade' según su significado en una comunidad lingüística determinada, sean cuales sean las *propiedades internas* de Oscar1 consistentes en tener creencias idiosincrásicas sobre el jade.

Pero, si el pensamiento de Oscar1 es externo por superviniencia en su uso social del lenguaje o, como dice Putnam, si el pensamiento de Oscar1 «no está en su cabeza», es falso, de entrada, que (ii) necesariamente todo lo que tiene la propiedad física interna F que Oscar1 -si fuera así- tiene cuando piensa que un trozo de jadeita o de nefrita es verde oscuro, también tiene ese pensamiento. En efecto, (1) hemos introducido el supuesto de que Oscar1 y Oscar2 son habitantes de diferentes comunidades lingüísticas, donde

las palabras tienen significados distintos, de tal manera que esa diferencia de pertenencia no afectara a su indiscernibilidad física. Ahora bien, (2) el concepto de contenido amplio implica que el pensamiento de Oscar1 superviene en el uso con sentido del lenguaje en su comunidad. La conclusión es que, dados (1) y (2), es falso que (ii) como, *ex hypothesi*, Oscar2 tiene las mismas propiedades físicas internas que Oscar1 y, por tanto, tiene *F* entonces necesariamente tiene el pensamiento de que un trozo de jadeita o de nefrita es verde oscuro. Si, según (1) Oscar2 es un habitante de una comunidad lingüística diferente de la de Oscar1 en la que, en concreto, la palabra 'jade' tiene un significado distinto, y según (2) cualesquiera que sea el pensamiento de Oscar2 sobre un trozo de jade, ese pensamiento superviene en el uso social de la palabra 'jade' en su comunidad, entonces: falsificando (ii), resulta que, a pesar de que, *ex hypothesi*, Oscar1 y Oscar2 son físicamente indiscernibles y, por tanto, Oscar2 tiene la propiedad física interna *F* de Oscar1, necesariamente Oscar2 tiene un *pensamiento diferente* sobre un trozo de jade del que tiene Oscar1. En concreto, si en la comunidad de Oscar2 la palabra 'jade' sólo se refiere a la jadeita, resulta que, al emitir «Este trozo de jade es verde oscuro», es verdad tanto que Oscar2 es físicamente indiscernible de Oscar1 como que Oscar2 expresa el pensamiento de que un trozo de jadeita es verde oscuro y *no* el pensamiento de Oscar1 de que un trozo de jadeita o de nefrita es verde oscuro.

Ahora bien, el mismo tipo de razonamiento se aplica a (i) en (SPb). Si el pensamiento de Oscar1 es externo por superviniencia en su uso social del lenguaje, entonces es falso que (i) necesariamente existe alguna propiedad física interna *F* tal que, al tener ese pensamiento, Oscar1 ejemplifica *F*. Primero, mantengamos (1) pero añadamos la estipulación (1') de que dentro de una *misma* comunidad lingüística no puede haber dos hablantes físicamente indiscernibles no sólo en todas sus propiedades físicas sino tampoco en alguna de ellas: aunque cada hablante de una comunidad tiene un duplicado físico completo en la otra, dos hablantes de una misma comunidad no son duplicados físicos en ninguna propiedad. Segundo, estipulemos que, mínimamente, el rango que cubre *x* en la formulación de (i) en (SPb) es el de los hablantes o sujetos pensantes de una comunidad humana: de otro modo, si el rango de *x* es cualquier individuo u objeto del mundo, (i) sería una pura trivialidad y, lo que es más importante, sería inútil en la formulación de leyes psicofísicas, especialmente de leyes psicofísicas extensibles a toda la especie humana. Es el caso que Luis1 pertenece a la misma comunidad lingüística que Oscar1 y, al realizar ahora la emisión «Este trozo de jade es verde oscuro», el pensamiento que expresa es el contenido amplio, que un trozo de jadeita o de nefrita es verde oscuro. La conclusión es que, dados (1') y (2), es falso que (i) si algún *x*-habitante de la comunidad lingüística de Oscar1, p.e, Oscar1 o Luis1- piensa que un trozo de jadeita o de nefrita es verde oscuro, entonces necesariamente existe alguna propiedad física interna *F* tal que *x*-Oscar1 o Luis1- tiene *F*. Si, según (1') no hay ninguna propiedad física interna *F* tal que, siendo habitantes de la misma comunidad lingüística, Oscar1 y Luis1 la ejemplifiquen por duplicado al emitir «Este trozo de jade es verde oscuro», y según (2), al realizar la emisión, Oscar1 y Luis1 expresan el mismo pensamiento en tanto que el pensamiento expresado superviene en el uso social de ese tipo de emisión en su comunidad, entonces: falsificando (i), resulta que, a pesar de que Oscar1 y Luis1 tienen el mismo pensamiento al realizar la misma emisión, que un trozo de jadeita o de nefrita es verde oscuro, necesariamente, al tener ambos ese pensamiento, Oscar1 ejemplifica una *propiedad física diferente* que la que Luis1 ejemplifica.

La consecuencia lógica es que (SPb) es falso y, por ello, que si los contenidos amplios son propiedades psicológicas no supervienen en propiedades físicas.

3. ¿Son los contenidos amplios propiedades psicológicas?

La cuestión de si las propiedades de contenido amplio son psicológicas es, para J. Fodor, la cuestión de si los contenidos amplios son propiedades pertenecientes a la taxonomía de la psicología. Pero, la prueba de su pertenencia a la taxonomía de la psicología es pasar el test del *principio de individuación causal de las propiedades*:

(IC) Las propiedades se individualúan en ciencia en términos de sus *poderes causales*.

Por lo tanto, la cuestión de si las propiedades de contenido amplio son psicológicas es la cuestión de si los contenidos amplios se individualúan en psicología en términos de sus poderes causales, específicamente, como propiedades psicológicas, en términos de sus efectos causales en la *conducta* de los agentes. Por su parte, Fodor no tiene duda alguna de que la psicología científica es una extensión teórica de la psicología 'popular', es decir, de la psicología de sentido común que usamos cada día al interpretar a un agente, y que, en consecuencia, la cuestión de si algunas propiedades pertenecen a la taxonomía de la psicología es, siguiendo (IC), la cuestión de si tales propiedades tienen los poderes causales que tienen las propiedades de *nuestra* psicología 'popular'. Así, no es discutible que las propiedades atribuidas a un agente por nuestra psicología son causalmente eficaces en producir su conducta. De hecho, comenta:

«Si no fuese literalmente verdadero que mi deseo es causalmente responsable de mi logro, y mi picazón es causalmente responsable de mi rascadura, y mi pensamiento es causalmente responsable de mi emisión, si nada de esto es literalmente verdadero, entonces prácticamente todo lo que creo acerca de cualquier cosa es falso y es el fin del mundo» (1989, 156).

Ahora bien, Fodor sostiene que las propiedades de contenido consistentes en tener un deseo, sentir picazón o tener un pensamiento, que son relevantes por su eficacia causal, son los contenidos estrechos. Luego, según (IC), las propiedades de la psicología son los contenidos estrechos. Esto no impide que nos expresemos *como si* los contenidos amplios fuesen propiedades psicológicas y nos refiramos a ellos en términos de tener un deseo, sentir picazón o tener un pensamiento. Fodor acepta que digamos, como lo hacemos en la cotidianidad, que Oscar2 tiene el pensamiento de que un trozo de jadeita es verde oscuro, en un sentido en que este pensamiento es un contenido amplio. Con todo, la idea es que, al interpretar a Oscar2 como un agente, no le atribuimos ese pensamiento o no se lo atribuimos como una propiedad psicológica porque no consideramos que sea causalmente responsable de la conducta de Oscar2.

Imaginemos la siguiente situación. Oscar1 y Oscar2 están participando en un concurso, se encuentran delante de una vitrina con distintos minerales, hay un trozo paradigmático de *jadeita* y no hay ningún otro mineral de color verde oscuro. El presentador les pide que señalen con la mano cuál de los trozos expuestos es un trozo de jade, y después lo digan. El resultado es que ambos señalan simultáneamente el trozo de jadeita y dicen simultáneamente «Este es el trozo de jade». Entre el público hay un experto en minerales de cada comunidad lingüística y el presentador pide a cada experto que verifique si el trozo señalado es en efecto un trozo de jade: ambos confirman que la respuesta de

los concursantes es verdadera. Finalmente, el presentador pide a los concursantes que justifiquen su conducta en términos de sus conocimientos y sus creencias, por qué señalaron ese trozo y dijeron lo que dijeron: ambos responden que no saben nada sobre el jade y que la única justificación es que creen que el jade es de color verde oscuro y que sólo el trozo señalado es de ese color.

Nosotros, psicólogos aficionados, que también estábamos entre el público, hemos sacado algunas conclusiones. Supongamos que se nos ha contado la historia sobre Oscar1 y Oscar2, y sobre sus comunidades lingüísticas, que he relatado en la sección anterior. Entonces, si aceptamos todos los hechos, podríamos incluso afirmar que la emisión

(1) «Este es el trozo de jade» por Oscar1 *expresa* la propiedad ser el pensamiento de que *este* es el trozo de jadeíta o de nefrita; mientras que (1) por Oscar2 expresa la propiedad ser el pensamiento de que *este* es el trozo de jadeíta.

Al atribuir tanto una como otra propiedad, a Oscar1 y a Oscar2 respectivamente, se mantiene el resultado verificado por los expertos: según las condiciones de verdad de cada emisión en la comunidad respectiva, ambas emisiones son verdaderas. Pero, precisamente el hecho de que ambas emisiones tengan *el mismo* valor de verdad excluye la atribución de pensamientos *diferentes* al interpretar a Oscar1 y a Oscar2 como *agentes*, esto es, causalmente responsables de su conducta. El razonamiento de Fodor es como sigue. De entrada, es evidente que si la emisión (1) tanto por Oscar1 como por Oscar2 expresa la propiedad (c) ser el pensamiento de que *este* es el trozo de piedra-jade, entendida como el contenido estrecho que superviene en sus creencias, entonces al atribuir *la misma* propiedad a ambos se preserva la interpretación semántica de las emisiones, esto es, se mantiene el resultado verificado por los expertos: ambas emisiones son verdaderas. Digamos, pues, que según (a), (b) y (c) la emisión (1) tiene el mismo valor de verdad o, en el contexto del concurso, tiene la misma propiedad intencional (d) ser una respuesta ganadora tanto al ser realizada por Oscar1 como por Oscar2. ¿Qué sucede si no nos preguntamos por la emisión (1) en términos de los pensamientos que *expresa* o puede expresar sino que, como psicólogos de la conducta de Oscar1 y Oscar2, nos preguntamos por la emisión (1) en términos de los pensamientos que la *causan* o han causado en cuanto es descrita como (d), esto es, en cuanto es descrita *intencionalmente*? Para responder esta cuestión, como las propiedades psicológicas o intencionales deben ser taxonómicas en psicología, tenemos que aplicar el *principio de individuación de los poderes causales*:

(IPC) Los poderes causales en ciencia se individuán en términos de sus *efectos*.

Así, la respuesta es que, según (IPC), la *causa* psicológica de la emisión (1) en Oscar1 y en Oscar2 como ejemplificación de *la misma* propiedad intencional (d) tiene que ser *la misma* propiedad intencional o psicológica de Oscar1 y Oscar2, en este caso, la propiedad psicológica (c). De ahí que, según (IC)-(IPC), cualesquiera que sean los pensamientos expresados por (1), incluso pudiendo ser (a) y (b), con todo, ni (a) ni (b) son propiedades psicológicas.

El argumento de Fodor es que los contenidos amplios no forman parte del repertorio causal de un agente. Ahora bien, como hemos visto, los contenidos amplios son propiedades externas o *relacionales*. Pero, la tesis de Fodor no consiste en discutir que cualesquiera propiedades relacionales no pasan el test de (IC)-(IPC). Al contrario, sostiene que «la taxonomía en términos de propiedades relacionales es frecuente en las ciencias» (1991, 12). Consideremos el caso de una ciencia especial como la astronomía. Resulta

que la propiedad de *ser un meteoro* es una propiedad relacional de un fragmento rocoso y, no obstante, es taxonómica en astronomía. ¿Es una propiedad relacional? Basta suponer que existen dos fragmentos rocosos indiscernibles en todas sus propiedades físicas internas pero tales que uno es un meteoro y el otro no: la diferencia está en que el primero, pero no el segundo, tiene una historia causal que lo relaciona con un cuerpo celeste. ¿Es taxonómica en astronomía? La evidencia de que sí cumple (IC)-(IPC) es que, por su carácter relacional, el fragmento que es un meteoro, pero no el otro, puede ser la causa de un efecto específico como el origen de un cráter en un planeta (cfr, *ibid*, 20). Así, como señala L. R. Baker (1995, 43), Fodor trata de defender tanto que los contenidos amplios, propiedades relacionales, no son taxonómicos en psicología como que algunas propiedades relacionales son taxonómicas en las ciencias especiales.

Uno de los objetivos de Baker es mostrar que (A) es falso. La argumentación de Baker será desarrollada, más extensivamente por mi parte, en la sección 4. Así, queda claro que el razonamiento de Fodor en defensa de (A) tiene la siguiente forma:

(1-IC) En todas las taxonomías científicas las propiedades se individualúan en términos de sus poderes causales.

(2) En psicología las propiedades se individualúan en términos de su contenido intencional.

(3) Una diferencia de contenido amplio no es suficiente para que haya una diferencia relevante en los poderes causales.

En consecuencia

(4) En psicología las propiedades se individualúan en términos de su contenido estrecho (cfr, Baker, 1995, 44) ((2) es una obviedad de sentido común que no voy a precisar) El procedimiento de Fodor, en el curso de probar (3) para concluir (4), más que evaluar los contenidos amplios según su adecuación global al test (IC)-(IPC), que es lo que antes he presentado en esbozo, es evaluar los contenidos amplios según su adecuación parcial a un test –esencialmente, pues hay otro test menos importante, el test del *cambio-de-contexto*- que él considera constitutivo de (IC)-(IPC), esto es, constitutivo de lo que es ser una propiedad con poderes causales y, por tanto, tener valor taxonómico en ciencia: es el test de la *conexión-no-conceptual* (NC). En resumen, para probar (3) Fodor tiene que probar que, al no pasar (NC), la diferencia entre tener un contenido amplio y no tenerlo (o tener uno diferente) no es una diferencia en poder causal ni, por ello, en valor taxonómico en términos de la eficacia causal del contenido amplio en producir una diferencia en la conducta de un agente. Básicamente, la idea es mostrar en la sección 4 -a partir de Baker- que, aunque (NC) sea constitutivo de (IC)-(IPC), (A) es falso, en cuyo caso, la consecuencia lógica es que los contenidos amplios son realmente propiedades psicológicas.

4. El test de la *conexión-no-conceptual*.

La tesis de Fodor es que (NC) expresa una condición *necesaria* que los contenidos amplios tendrían que satisfacer, y no lo hacen, si una diferencia de contenido amplio va a ser una diferencia de poder causal en términos de una diferencia de conducta en los agentes. Probar que los contenidos amplios pasan (NC) es, por tanto, hacer colapsar la premisa (3). En la presentación de (NC) Fodor comienza por introducir un esquema

trivialmente verdadero de toda diferencia causal en términos de los efectos. Consideremos una situación en la que intervienen dos causas, dos sucesos, *C1* y *C2*, y se producen, respectivamente, dos efectos, también dos sucesos, *E1* y *E2*; entonces es trivialmente verdadero que

C1 difiere de *C2* en que *C1* tiene una propiedad causal *CP1* mientras que *C2* tiene una propiedad causal *CP2*.

E1 difiere de *E2* en que *E1* tiene una propiedad resultante *EP1* mientras que *E2* tiene una propiedad resultante *EP2*.

La diferencia entre *C1* y *C2* es responsable de la diferencia entre *E1* y *E2* en el sentido de que si *C1* hubiera tenido *CP2* en lugar de *CP1*, entonces *E1* habría tenido *EP2* en lugar de *EP1*, y si *C2* hubiera tenido *CP1* en lugar de *CP2*, entonces *E2* habría tenido *EP1* en lugar de *EP2* (cfr, 1991, 9-10).

La cuestión es: ¿en qué casos de ejemplificación del esquema la diferencia entre tener *CP1* y tener *CP2* es una diferencia del poder causal de *C1* y de *C2* en términos de poder producir la diferencia de propiedades entre *E1* y *E2*? Esto es, ¿cuándo la diferencia entre *CP1* y *CP2* es una diferencia causalmente eficaz? La respuesta de Fodor es que hay una diferencia de poder causal sólo si, ejemplificando el esquema, la diferencia entre las propiedades causales no está conectada conceptualmente con la diferencia entre las propiedades resultantes (cfr, ibid, 19). Así,

(NC1) La diferencia entre *CP1* y *CP2* es una diferencia de poder causal sólo si, ejemplificando el esquema, no es una verdad conceptual que como *C1* tiene *CP1* mientras que *C2* tiene *CP2*, causan efectos, *E1* y *E2*, que difieren en que uno tiene *EP1* y el otro *EP2*.

La argumentación de Fodor es que los contenidos amplios pueden ejemplificar el esquema pero no satisfacen (NC1) por el hecho de que es una verdad conceptual que causan la conducta que causan. ¿Qué significa que la conexión causal entre los contenidos amplios y sus efectos es una verdad conceptual y que, por ello, no hay poder causal en juego?

Volvamos de nuevo al concurso en que participan Oscar1 y Oscar2, de manera que la emisión

(1) «Este es el trozo de jade» por Oscar1 exprese la propiedad ser el pensamiento de que *este* es el trozo de jadeíta o de nefrita, y por Oscar2 exprese la propiedad ser el pensamiento de que *este* es el trozo de jadeíta.

Ahora, es natural individuar la conducta de un agente en términos de sus pensamientos o contenidos intencionales: decimos de alguien que, al alzar la mano, está respondiendo que sí si piensa que alzar la mano es responder que sí y está respondiendo que no si piensa que alzar la mano es responder que no. Entonces, un defensor del poder causal de los contenidos amplios podría proponer individuar la conducta

(2) señalar con la mano el trozo de jade en términos de (a) y (b), de modo que (2) tenga la propiedad (e) ser un señalamiento con la mano del trozo de jadeíta o de nefrita si se trata de la conducta de Oscar1, que tiene el pensamiento (a), y la propiedad (f) ser un señalamiento con la mano del trozo de jadeíta si se trata de la conducta de Oscar2, que tiene el pensamiento (b). Pues bien, en tanto que Oscar1 señala con la mano un trozo de jadeíta o de nefrita mientras que Oscar2 señala con la mano un trozo de jadeíta, el causalista de los contenidos amplios podría sostener que la diferencia entre las propiedades intencionales (a) y (b) es una diferencia en el poder causal de los pensamientos de Oscar1 y Oscar2 en términos de la diferencia resultante entre las propiedades

intencionales (e) y (f) de su conducta (2). En efecto, la situación de Oscar1 y Oscar2 ejemplifica trivialmente el esquema, en concreto, es trivialmente verdadero que la diferencia entre los pensamientos ($C1$ y $C2$) de Oscar1 y Oscar2 es responsable de la diferencia entre su conducta ($E1$ y $E2$) en el sentido de que si el pensamiento de Oscar1 hubiera sido el pensamiento de que *este* es el trozo de jadeita ($CP2$) en lugar del pensamiento de que *este* es el trozo de jadeita o de nefrita (CPI), entonces la conducta de Oscar1 habría sido el señalamiento del trozo de jadeita ($EP2$) en lugar del señalamiento del trozo de jadeita o de nefrita (EPI), y lo mismo en el caso del pensamiento de Oscar2.

Pero, el argumento de Fodor es precisamente que, a pesar de ejemplificar el esquema, los contenidos amplios de Oscar1 y Oscar2 no tienen poder causal en su conducta porque no satisfacen (NC1). La razón es que hay una conexión conceptual entre el pensamiento de Oscar1 como ha sido descrito (CPI) y la conducta (2) de Oscar1 como ha sido descrita (EPI), y lo mismo en el otro caso, una conexión análoga a la que hay entre una causa descrita como *la causa del efecto E* y el efecto descrito como *E*. (1) es una verdad conceptual, esto es, independiente de cualesquiera que haya sido de hecho la causa, que la causa del efecto *E* causó el efecto *E* –diríamos, además, que es una verdad necesaria-. (2), por el hecho de que esa atribución de causalidad es una verdad conceptual, no es la atribución a la causa de un *poder* causal diferente del resto de los poderes causales que tiene la causa: ser *la causa del efecto E* no es la propiedad causal de la causa que fue eficaz en la producción del efecto. De la misma manera, (1) es una verdad conceptual, esto es, independiente de cualesquiera que haya sido de hecho el contenido causante, que el pensamiento de Oscar1 descrito como el pensamiento de que *este* es el trozo de jadeita o de nefrita causó la conducta de Oscar1 descrita como el señalamiento del trozo de jadeita o de nefrita. El contenido causante de la conducta puede haber sido otro, pero el contenido así descrito es necesariamente y *a priori* la causa de la conducta así descrita: *bajo las descripciones CPI y EPI , $C1$ es a priori la causa necesaria de $E1$, aunque en absoluto la propiedad causal del pensamiento de Oscar1 responsable de su conducta pueda ser otra propiedad de contenido que CPI .* (2), como la descripción EPI sólo juega el rol conceptual de describir la conducta de Oscar1 en términos de la descripción CPI de su pensamiento, no se trata, en la atribución de causalidad al pensamiento de Oscar1 en términos del contenido amplio, de una atribución de un poder causal diferente del resto de los poderes causales de ese pensamiento, sino que se trata de una atribución trivial reformulable en términos de la genuina eficacia causal de alguna otra propiedad de contenido del pensamiento: el contenido amplio de que *este* es el trozo de jadeita o de nefrita no es la propiedad causal del pensamiento de Oscar1 que fue eficaz en la producción del señalamiento del trozo de jade porque no lo es bajo cualquier descripción de la conducta de Oscar1 que no sea implicada trivialmente por esa descripción de su pensamiento.

Por lo tanto, es una verdad conceptual que los pensamientos que, por ejemplo, *sólo* difieren en referirse o bien a la jadeita y/o la nefrita o bien a la jadeita tienen efectos conductuales que *sólo* difieren en hacer algo relativo o bien a la jadeita y/o la nefrita o bien a la jadeita. En consecuencia, los contenidos amplios no satisfacen (NC1): es una verdad conceptual que se diferencian en tener los efectos que tienen por ser lo diferentes que son, lo cual es una trivialidad causalmente irrelevante (cfr, Baker, 1995, 45-6).

Sin embargo, como había puesto de manifiesto S. Stich, la satisfacción de (NC1) no es suficiente para eliminar todos los casos en que se da una trivialidad causalmente

irrelevante. Supongamos, por ejemplo, que el jade –jadeita y/o nefrita- es la piedra más deseada por Oscar1. Entonces, (1') no es una verdad conceptual que los pensamientos que sólo difieren en que uno es relativo a la piedra más deseada por Oscar1 y el otro es relativo a la jadeita tienen efectos conductuales que sólo difieren en que uno es hacer algo relativo al jade –jadeita o nefrita- y el otro es hacer algo relativo a la jadeita: es evidente que la descripción del pensamiento de Oscar1 como el pensamiento de que *este* es el trozo de su piedra más deseada no está conectada conceptualmente, esto es, *a priori*, con la descripción de la conducta de Oscar1 como un señalamiento del trozo de jadeita o de nefrita. Pero, (2'), como sucedía en el caso de la conexión conceptual, la atribución de causalidad al pensamiento de Oscar1 así descrito no es la atribución de un poder causal diferente del resto de los poderes causales de ese pensamiento. Basta considerar que, al preguntársele en el concurso cuál es el trozo de jade, el pensamiento descrito como el pensamiento de que *este* es el trozo de su piedra *más* deseada no puede ser la propiedad causal que fue causalmente responsable del señalamiento del trozo de jade: la prueba es que si el jade fuera la piedra *menos* deseada por Oscar1, y bajo el supuesto de que sus demás creencias permanecen estables, el pensamiento de Oscar1 todavía habría sido causalmente responsable del señalamiento del trozo de jade, por lo que el pensamiento de Oscar1 tanto bajo la descripción *este es el trozo de su piedra más deseada* como bajo la descripción *este es el trozo de su piedra menos deseada* no registra la propiedad de contenido de su pensamiento que es causalmente responsable de su conducta.

La consecuencia es que la diferencia entre un pensamiento relativo a la piedra más deseada y un pensamiento relativo a la jadeita no es una diferencia en los poderes causales de ambos pensamientos. Aunque Fodor no corrige expresamente la formulación de la condición necesaria de una diferencia de poder causal como está en (NC1), sí que introduce el requisito de que, al contrario que

(*) si ser relativa a la piedra más deseada por Oscar1 es una propiedad que tiene toda conducta relativa a la jadeita y/o la nefrita, entonces si alguien tiene pensamientos relativos a la jadeita y/o la nefrita, entonces su conducta es relativa a la piedra más deseada por Oscar1, las propiedades *psicológicas* estén involucradas en condicionales análogos a (*) que, sin embargo, no sean *conceptualmente necesarios*. En efecto, (*) es conceptualmente necesario: (i) hay una conexión conceptual entre los pensamientos relativos a la jadeita y/o nefrita y las conductas relativas a la jadeita y/o nefrita y (ii) *no* hay un mundo posible en el que ser relativa a la piedra más deseada por Oscar1 es una propiedad de toda conducta relativa a la jadeita y/o nefrita pero, aunque M. Cerezo está pensando sobre la jadeita y/o nefrita, su conducta *no* es relativa a la piedra más deseada por Oscar1. Podría parecer que todo condicional análogo a (*) que involucre propiedades psicológicas es trivialmente conceptualmente necesario, pero no es así. Como propone Fodor, supongamos que está comprobado que pensar sobre topología causa dolor de cabeza. Pues bien, resulta que el condicional

(**) si ser una sensación de dolor es una propiedad de todo dolor de cabeza, entonces si alguien tiene pensamientos relativos a la topología, entonces tiene una sensación de dolor no es conceptualmente necesario (cfr, 1991, 20-21). Tenemos que (i') es evidente que no hay una conexión conceptual sino empírica –si la hay- entre los pensamientos relativos a la topología y los dolores de cabeza, y que (ii') *hay* una multitud de mundos posibles en que si ser una sensación de dolor es una propiedad de los dolores de cabeza,

entonces si M. Cerezo está pensando sobre topología, *no* tiene una sensación de dolor, a saber: (ii') es verdadero en todos los mundos posibles en que precisamente los pensamientos relativos a la topología no causan dolores de cabeza y, por tanto, no causan sensaciones de dolor, porque empíricamente, como una cuestión de hecho, no es así.

Pues bien, a partir del requisito introducido por Fodor, podemos formular la condición necesaria de una diferencia de poder causal como debe figurar en el test (NC) del siguiente modo:

(NC2) Dos propiedades causales, *CP1* y *CP2*, tienen diferentes poderes causales sólo si ni (I) ni (II), siendo verdaderas, son verdades conceptuales:

(I) las causas *C1* y *C2*, que difieren en que *C1* tiene *CP1* en tanto que *C2* tiene *CP2*, tienen efectos, *E1* y *E2*, que difieren en que *E1* tiene *EPI* en tanto que *E2* tiene *EP2*; y

(II) si *Q* es una propiedad que tiene todo lo que tiene *EPI*, entonces si *C1* tiene *CP1*, entonces *E1* tiene *Q* (cfr, Baker, 1995, 46-7).

Obviamente el requisito ulterior es (II), no satisfecho por (*) pero sí por (**) bajo el supuesto de Fodor de que hay una conexión empírica entre pensar sobre topología y tener dolor de cabeza. Primero, es una verdad conceptual que si ser relativa a la piedra más deseada por Oscar1 es una propiedad (*Q*) que tiene toda conducta que tiene la propiedad de ser relativa a la jadeita y/o nefrita (*EPI*), entonces si un pensamiento de M. Cerezo (*CI*) es relativo a la jadeita y/o nefrita (*CP1*), entonces la conducta de M. Cerezo (*E1*) tiene la propiedad de ser relativa a la piedra más deseada por Oscar1 (*Q*). Segundo, bajo el supuesto de Fodor, es una verdad empírica, no conceptual, que si ser una sensación de dolor es una propiedad (*Q*) que tiene todo estado mental *x* que tiene la propiedad de ser un dolor de cabeza (*EPI*), entonces si un pensamiento de M. Cerezo (*CI*) tiene la propiedad de ser relativo a la topología (*CP1*), entonces un estado mental *x* de M. Cerezo (*E1*) tiene la propiedad de ser una sensación de dolor (*Q*).

Por supuesto, la tesis de Fodor es que, como los contenidos amplios no satisfacen (NC2), no son propiedades psicológicas individuadas en términos de sus poderes causales. Mi propósito es mostrar –a partir de Baker– que los contenidos amplios sí satisfacen (NC2), es decir, (I) y (II).

Estamos otra vez en el concurso, pero ahora, entre los minerales expuestos, lo que hay es un trozo de *nefrita*, no de jadeita, y no hay ningún otro mineral de color verde oscuro. El presentador pide a Oscar1 y Oscar2 que señalen con la mano cuál de los minerales es un trozo de jade, y después lo digan. El resultado es que ambos señalan simultáneamente el trozo de nefrita y dicen simultáneamente «Este es el trozo de jade». Entonces, intervienen los expertos de cada comunidad presentes entre el público, y la situación cambia totalmente respecto al caso anterior. Recordemos que, como en la comunidad de Oscar1, en la comunidad de Oscar2 los expertos conocen la diferencia entre la jadeita y la nefrita, pero en esta última la palabra 'jade' sólo se refiere a la jadeita, no a la jadeita y a la nefrita. El experto de la comunidad de Oscar1 verificará que la respuesta de Oscar1 es verdadera en tanto que, por la interpretación semántica que rige en su comunidad, un trozo de nefrita es un trozo de jade. Por el contrario, el experto de la comunidad de Oscar2 verificará que la respuesta de Oscar2 es *falsa* en tanto que, por la interpretación semántica que rige en su comunidad, un trozo de nefrita *no* es un trozo de jade -no cae en la extensión de 'jade' - y la respuesta de Oscar2 ha sido decir que sí lo hace, que un trozo, que es de nefrita aunque él no lo sabe, es el trozo de jade. Finalmente, el presentador pide a Oscar1

y Oscar2 que justifiquen su conducta en términos de sus conocimientos y sus creencias, por qué señalaron ese trozo y dijeron lo que dijeron: ambos responden que no saben nada sobre el jade y que la única justificación es que creen que el jade es de color verde oscuro y que sólo el trozo señalado es de ese color. Así, Oscar2 no se ha despistado o precipitado: en términos de sus conocimientos y de sus creencias *tenía que decir* «Este es el trozo de jade» señalando con la mano el trozo de nefrita.

Ahora, como por sus condiciones de verdad en cada comunidad, las emisiones de Oscar1 y Oscar2 tienen valores de verdad diferentes, una es una respuesta verdadera y otra falsa, y, como por la participación en el concurso, tienen propiedades intencionales *diferentes*, una es una respuesta ganadora y otra perdedora, el resultado es: según (IPC) –y, por ello, (IC)– el hecho de que ambas emisiones tengan valores de verdad y propiedades intencionales *diferentes* excluye la atribución del *mismo* pensamiento al interpretar a Oscar1 y Oscar2 como *agentes*, esto es, causalmente responsables de su conducta. Según (IC)-(IPC), las *causas* psicológicas de las emisiones de Oscar1 y Oscar2 como ejemplificación de propiedades intencionales *diferentes* tienen que ser *diferentes* propiedades intencionales o psicológicas de sus pensamientos, en este caso, los contenidos amplios de que *este* es el trozo de jadeíta o de nefrita y de que *este* es el trozo de jadeíta, respectivamente. Entonces, hay que sostener contra Fodor que los contenidos amplios, el tipo de contenidos por los que se *diferencia* todo pensamiento sobre el jade de Oscar1 y Oscar2, son propiedades causales de los pensamientos de Oscar1 y Oscar2 sobre el trozo de jade porque son los contenidos que pueden ser causalmente responsables de la *diferencia* entre sus respuestas, de que la respuesta de Oscar1 sea ganadora y la respuesta de Oscar2 sea perdedora, respectivamente. Por tanto, según (IC)-(IPC), los contenidos amplios son realmente propiedades *psicológicas* (cfr, Baker, 1995, 47-8).

Pero, como hemos visto, para Fodor la prueba de que los contenidos amplios individuán los pensamientos en términos de sus poderes causales y, así, pasan el test (IC)-(IPC), debe concretarse en la satisfacción del test (NC), en mi versión, (NC2). Por eso, obviando la satisfacción directa de (IC)-(IPC), empecemos por mostrar que los contenidos amplios de los pensamientos de Oscar1 y Oscar2 ejemplifican el esquema de Fodor. Así, a partir de la clasificación causal

(C1) el pensamiento de Oscar1,

(C2) el pensamiento de Oscar2,

(CP1) ser un pensamiento relativo a *este* trozo de jadeíta o de nefrita,

(CP2) ser un pensamiento relativo a *este* trozo de jadeíta,

(E1) «Este es un trozo de jade»,

(E2) «Este es un trozo de jade»,

(EP1) ser una respuesta ganadora,

(EP2) ser una respuesta perdedora; obtenemos que, en efecto, la diferencia entre los

pensamientos (C1 y C2) de Oscar1 y Oscar2 es responsable de la diferencia entre sus emisiones de «Este es el trozo de jade» (E1 y E2) en el sentido de que si el pensamiento de Oscar1 hubiera sido el pensamiento de que *este* es el trozo de jadeíta (CP2) en lugar del pensamiento de que *este* es el trozo de jadeíta o de nefrita (CP1), entonces la emisión de Oscar1 habría sido una respuesta perdedora (EP2) en lugar de una respuesta ganadora (EP1), y lo mismo en el caso del pensamiento de Oscar2.

De este modo, como ejemplificación del esquema, es, al menos trivialmente, verdadero que (DC) la diferencia entre ser relativo a *este* trozo de jadeita o de nefrita (*CP1*) y ser relativo a *este* trozo de jadeita (*CP2*) es una diferencia causal en el sentido de que como el pensamiento de Oscar1 (*C1*) tiene *CP1* mientras que el pensamiento de Oscar2 (*C2*) tiene *CP2* causan efectos, sus emisiones de «Este es el trozo de jade» (*E1* y *E2*), que difieren en que *E1* de Oscar1 es una respuesta ganadora (*EP1*) y *E2* de Oscar2 es una respuesta perdedora (*EP2*).

Sin embargo, queda claro que para Fodor la diferencia causal entre *CP1* y *CP2* es una diferencia de *poder* causal o una diferencia en su *eficacia* causal en la conducta de los agentes sólo si (DC) es más que trivialmente verdadero, y es más que trivialmente verdadero sólo si *CP1* y *CP2* satisfacen (NC2). Equivalentemente, todo condicional análogo a (*) si ser una respuesta premiada es una propiedad de toda emisión «Este es el trozo de jade» que es una respuesta ganadora, entonces si alguien piensa que *este* es el trozo de jadeita o de nefrita, entonces su emisión «Este es el trozo de jade» es una respuesta premiada, *no* debe ser conceptualmente necesario. Como es evidente, si no es verdad que ser una respuesta premiada es una propiedad de toda emisión «Este es un trozo de jade» que es una respuesta ganadora, es irrelevante: se trata de un supuesto. Pues bien, mostremos que *CP1* (y *CP2* en un condicional análogo) en (*) satisface (NC2).

Primero, la cláusula sencilla: es manifiesto que (I) las causas, *C1* y *C2*, que difieren en que *C1* tiene *CP1* en tanto que *C2* tiene *CP2* tienen efectos, *E1* y *E2*, que difieren en que *E1* tiene *EP1* en tanto que *E2* tiene *EP2*, no es una verdad conceptual en el caso de (DC):

(1), obviamente, no hay una conexión conceptual entre el pensamiento de Oscar1 como el pensamiento de que *este* es el trozo de jadeita o de nefrita (*CP1*) y su emisión como una respuesta ganadora (*EP1*). La descripción *CP1* con la mera referencia a un trozo de jadeita o de nefrita no está conceptualmente conectada, es decir, *a priori*, con la descripción *EP1* con la mera referencia al hecho de ganar. No es una verdad conceptual, esto es, independiente de cualesquiera que haya sido de hecho el contenido causante de la respuesta ganadora de Oscar1 (*EP1*), que el pensamiento de Oscar1 (*C1*) descrito como el pensamiento de que *este* es el trozo de jadeita o de nefrita (*CP1*) causó la emisión (*E1*) descrita como una respuesta ganadora (*EP1*). Ese pensamiento así descrito no es necesariamente y *a priori* la causa de la emisión así descrita: *bajo las descripciones CP1 y EP1, C1 no es a priori una causa necesaria de E1;*

(2) aunque la descripción *EP1* no tiene el uso de describir la emisión de Oscar1 en términos de la descripción *CP1* de su pensamiento, ¿se trata, en la atribución de causalidad al pensamiento de Oscar1 en términos del contenido amplio *CP1*, de la atribución de un *poder* causal diferente del resto de los poderes causales de ese pensamiento o de una atribución trivial que demanda la *eficacia* causal de alguna otra propiedad de contenido del pensamiento? Simplemente consideremos que si la jadeita, pero no la nefrita, fuese la piedra a la que se refiere la palabra 'jade' de Oscar1, y bajo el supuesto de que sus demás creencias permanecen estables, el pensamiento de Oscar1 como el pensamiento de que *este* es el trozo de jadeita, y sólo como tal, habría sido causalmente responsable de que su emisión fuera una respuesta perdedora y no, por el contrario, una respuesta ganadora: por ello, no hay otra descripción del pensamiento de Oscar1, es decir, no hay otra propiedad de contenido de *C1* diferente de *CP1*, que sea causalmente responsable de

que su emisión sea una respuesta ganadora y no una respuesta perdedora. Así, los pensamientos de Oscar1 y Oscar2 que *sólo* difieren en que *C1* es relativo a *este* trozo de jadeita o de nefrita y *C2* es relativo a *este* trozo de jadeita tienen efectos conductuales, como las emisiones de ambos, que *no sólo* difieren en que *E1* es una conducta relativa a *este* trozo de jadeita o de nefrita y *E2* es una conducta relativa a *este* trozo de jadeita: como revela de suyo (DC), *CP1* y *CP2* tienen efectos que pueden diferir en múltiples formas, por ejemplo, en que *E1* es una respuesta ganadora y *E2* es una respuesta perdedora;

(i) El resultado es que *CP1* satisface (I) de (NC2).

Segundo, mostremos que (II) si *Q* es una propiedad que tiene todo lo que tiene *EP1*, entonces si *C1* tiene *CP1*, entonces *E1* tiene *Q*, *no* es una verdad conceptual en el caso del condicional (*). Pues bien, *hay* multitud de mundos posibles en que precisamente un pensamiento relativo a *este* trozo de jadeita o de nefrita *no* causa una emisión que es una respuesta ganadora y, por tanto, que es una respuesta premiada, porque empíricamente, como una cuestión de hecho, no es así. Por ejemplo, imaginemos un mundo posible, el del presentador del concurso, en el que la palabra 'jade' se refiere a la jadeita, a la nefrita y a un tercer mineral *Z* que también es de color verde oscuro. Ahora la situación es tal que el único mineral de color verde oscuro es un trozo paradigmático de *Z*. Como siempre, el presentador le pide a Oscar1 que señale cuál de los trozos expuestos es el trozo de jade, y después lo diga. El resultado es que Oscar1 simultáneamente señala el trozo de *Z* y dice «Este es el trozo de jade». Cuando es consultado, el experto de la comunidad de Oscar1 verifica que la respuesta de Oscar1 es *falsa* en tanto que, por la interpretación semántica que rige en su comunidad, un trozo de *Z* *no* es un trozo de jade –no cae en la extensión de 'jade'– y la respuesta de Oscar1 ha sido decir que sí lo hace, que un trozo, que es de *Z* aunque él no lo sabe, es el trozo de jade: por las condiciones de verdad de la emisión de Oscar1 en su comunidad, sólo un trozo de jadeita o de nefrita es un trozo de jade. Cuando el presentador pide a Oscar1 que justifique su emisión en términos de sus conocimientos y sus creencias, por qué dijo lo que dijo, Oscar1 responde que no sabe nada sobre el jade y que la única justificación es que cree que el jade es de color verde oscuro y que sólo el trozo señalado es de ese color. Así, Oscar1 no se ha extralimitado: en términos de sus conocimientos y sus creencias *tenía que* decir «Este es el trozo de jade» señalando con la mano el trozo de *Z*. Entonces, hay una multitud de mundos posibles en los que, como en el recién imaginado, (*) es empíricamente falso: el pensamiento de que *este* es el trozo de jadeita o de nefrita *no* causa de hecho una emisión que es una respuesta ganadora y, por tanto, una respuesta premiada sino que, al contrario, causa de hecho una emisión que es una respuesta perdedora y, por tanto, no premiada. En nuestro experimento imaginativo original, es una verdad empírica, *no conceptual*, que si ser una respuesta premiada es una propiedad (*Q*) que tiene toda emisión «Este es el trozo de jade» que tiene la propiedad de ser una respuesta ganadora (*EP1*), entonces si un pensamiento de *M. Cerezo* (*CI*) tiene la propiedad de ser relativo a *este* trozo de jadeita o de nefrita (*CP1*), entonces la emisión «Este es el trozo de jade» de *M. Cerezo* (*E1*) tiene la propiedad de ser una respuesta premiada (*Q*).

(ii) El resultado es que *CP1* satisface (II) de (NC2),

(iii) El resultado es que *CP1* satisface (NC2),

(iv) El resultado es que *CP1* satisface (IC)-(IPC),

(v) El resultado es que *CP1* revoca la premisa (3) de (A),

- (vi) El resultado es que (A) es falso,
(vii) El resultado es que *CPI*, un contenido amplio, es una propiedad *psicológica*
(cfr, Baker, 1995, 48-51).

Bibliografía

- Baker, L. R. (1995): *Explaining Attitudes*, Cambridge, Cambridge University Press.
Burge, T. (1979): «Individualism and the Mental», *Midwest Studies in Philosophy* 4: 73-121.
Burge, T. (1986): «Individualism and Psychology», *Philosophical Review* 45: 3-45.
Fodor, J. (1988): *Psychosemantics*, Cambridge Mass, MIT Press.
Fodor, J. (1991): «A Modal Argument for Narrow Content», *Journal of Philosophy* 88: 5-26.
Heil, J. (1992): *The Nature of True Minds*, Cambridge, Cambridge University Press.
Kim, J. (1990): «Supervenience as a Philosophical Concept», *Metaphilosophy* 12: 1-27.
Kim, J. (1993): *Supervenience and Mind*, Cambridge, Cambridge University Press.
Putnam, H. (1975): «The Meaning of 'Meaning'», reeditado en Putnam, H. (1975): *Mind, Language, and Reality*, Cambridge, Cambridge University Press.

* * *

Francisco Javier Vidal
Departamento de Filosofía
Universidad de Navarra
31080 Pamplona