

Taula (UIB) núm. 15, 1992

## Conexionismo y el epifenomenalismo de lo mental

Josep E. Corbí

Universitat de València

### Resumen

*En este artículo discuto una serie de argumentos presentados por J. Fodor, Z. Pylyshyn y B. McLaughlin donde se defiende que el surgimiento de los modelos conexionistas no representa realmente una amenaza para la hipótesis del lenguaje del pensamiento, dado que esos modelos no pueden dar cuenta de la sistematicidad del pensamiento ni de la eficacia causal de las propiedades mentales. A lo largo del artículo intento mostrar, sin embargo, que los argumentos mencionados no aciertan a excluir la posibilidad de dar cuenta, dentro de un modelo conexionista, de la sistematicidad y la eficacia causal de lo mental. Este hecho afectaría ciertamente a los mejores argumentos en defensa de la hipótesis del lenguaje del pensamiento, dado que tales argumentos tiene la estructura de una inferencia en favor de la mejor explicación.*

### Abstract

*I discuss some arguments presented by J. Fodor, Z. Pylyshyn and B. McLaughlin where they defend that the emergence of connectionist models does not constitute a threat to the language of thought hypothesis, since they cannot account for the systematicity of thoughts and the causal efficacy of mental properties. I argue, however, that the arguments supplied do not rule out the possibility of accounting, within connectionist models, for the systematicity and the causal efficacy of the mental; and this fact represents a major challenge to the best arguments for the language of thought hypothesis, which have the structure of inferences to the best explanation.*

## 1. Introducción

En este artículo, se discuten algunos argumentos de J. Fodor y B. McLaughlin en los que se defiende que el surgimiento de los modelos conexionistas no constituye una amenaza para la hipótesis del lenguaje del pensamiento (HLP). Intentaré mostrar, por el contrario, que, si bien los modelos clásicos y conexionistas no son necesariamente incompatibles, el surgimiento de estos últimos socava los mejores argumentos en favor de HLP. En este sentido, argüiré que los argumentos de Fodor y McLaughlin fracasan en su intento de excluir la posibilidad de construir una teoría representacional de la cognición que, además de incorporar los modelos conexionistas, prescindiera de HLP.

Así, en la primera sección, ofrezco una caracterización de HLP y presento los argumentos que Fodor propone en su defensa y que, como veremos, descansan básicamente en la carencia de explicaciones alternativas. En la segunda sección, describo la crítica de Fodor y McLaughlin al modelo de Smolensky en tanto que posible alternativa a HLP. Su crítica se centra en la imposibilidad de que este modelo dé cuenta de la eficacia causal de las propiedades mentales, por lo que la discusión de este tópico nos obliga a analizar las condiciones que, según Fodor, debe reunir una propiedad para que sea reconocida como causalmente eficaz. En la tercera sección, se desarrolla una crítica a las condiciones que Fodor establece para la eficacia causal y se presenta una propuesta alternativa, que, entre otras cosas, permite hacer compatible el modelo de Smolensky con la eficacia causal de las propiedades mentales. Este hecho me lleva, finalmente, a concluir que el surgimiento de los modelos conexionistas parece socavar los mejores argumentos en favor de HLP.

### 1. La Hipótesis del Lenguaje del Pensamiento

La Hipótesis del Lenguaje del Pensamiento va más allá del Realismo Intencional. Así, un realista intencional asume: (i) «...Que hay creencias y deseos y que hay una cuestión de hecho acerca de sus contenidos intencionales.» (ii) «Podría ocurrir que creer y desear fuesen estados del cerebro.» (iii) «...Que las creencias y los deseos tienen roles causales y que la conducta pública es, típicamente, el efecto de complejas interacciones entre estos estados mentales.»<sup>1</sup>

HLP añade, sin embargo, «que *los estados mentales* - y no sólo sus objetos proposicionales- *tienen típicamente estructura constituyente*». <sup>2</sup>Un estado tiene estructura constituyente si y sólo si tiene partes constituyentes. La noción de 'parte constituyente' puede, a su vez, analizarse en términos de transportabilidad en el seno de una semántica combinatoria. Diremos, en este sentido, que una parte de un estado mental es una parte constituyente sólo si es transportable a otro estado mental, es decir, si realiza una aportación estable al contenido de los estados mentales de los que forma parte. Por tanto,

<sup>1</sup> Todas las notas en este párrafo pertenecen a Fodor (1987), p. 135.

<sup>2</sup> Fodor (1987), p. 136.

lo que HLP defiende es que los estados mentales (ya sean tipos o casos particulares), y no sólo sus objetos proposicionales, tienen partes constituyentes y, en definitiva, estructura constituyente.<sup>3</sup> Por otro lado, si los estados mentales particulares son, como sugiere la tesis (ii), estados cerebrales particulares, entonces podemos añadir que los estados cerebrales particulares tendrán también partes constituyentes.

Los argumentos en favor de HLP descansan, según Fodor, en el siguiente prejuicio metafísico:

«No creo que haya mecanismos intencionales. Es decir, no creo que los contenidos por sí mismos determinen roles causales. En consecuencia, tiene que ser posible contar la historia completa de la causación mental (la historia completa de la implementación de las generalizaciones que las psicologías del deseo/creencia articulan) *sin referirse a las propiedades intencionales que se subsumen bajo tales generalizaciones.*» (Fodor (1987), p. 139)

Podemos, ciertamente, distinguir entre una interpretación débil y una interpretación fuerte de la tesis que en esta cita se propone. Según la interpretación débil, podríamos decir que todo lo que pretende afirmar Fodor cuando dice que no hay mecanismos intencionales es que la historia causal completa acerca de los procesos mentales podría contarse sin apelar a sus propiedades intencionales.

Hay razones para sospechar, sin embargo, que esta es una interpretación demasiado débil; pues, en primer lugar, el prejuicio metafísico se reduciría a un principio general, y más bien inocuo, de supervinencia de lo mental respecto a lo físico; y, en segundo lugar, la referencia a la inexistencia de mecanismos intencionales parecería como inadecuada e irrelevante en un contexto en el que se pretende fundamentar HLP mediante una inferencia a la mejor explicación, ya que el hecho de que la historia causal completa pueda contarse sin apelar a propiedades intencionales no excluye por sí mismo la posibilidad de que surjan historias causales alternativas en las que intervengan propiedades intencionales.

De acuerdo con esto, podríamos proponer una interpretación más interesante, y también más fuerte, del texto de Fodor y, para ello, podríamos dejarnos guiar por el significado de los términos que forman la expresión 'mecanismo intencional'. La idea de mecanismo es, en general, la idea de un proceso causal regular desarrollado en el seno de un sistema relativamente cerrado. Mas, en este contexto, ¿qué añade el adjetivo 'intencional' al sustantivo 'mecanismo'? La segunda parte de la cita sugiere la siguiente respuesta: Un mecanismo intencional es un proceso causal regular que se desarrolla en el seno de un sistema relativamente cerrado en el que intervienen propiedades intencionales que condicionan causalmente su desarrollo. He de reconocer que, a mi entender, Fodor se compromete con el prejuicio metafísico en su versión fuerte, pues en él descansan los

<sup>3</sup> Así, dice Fodor:

«De manera similar, 'mutatis mutandis', el 'P' que expresa la proposición P en la fórmula 'P' es un caso del mismo tipo que el 'P' que expresa la proposición P en la fórmula 'P & Q'. Si quisiéramos ser un poco más precisos, podríamos decir que la historia de HPL equivale a decir que (i) (algunas) fórmulas mentales tienen fórmulas mentales como partes; y (ii) las partes son 'transportables': las mismas partes que aparecen en muchas fórmulas.» (Fodor (1987), p. 137)

<sup>4</sup> Cf. Fodor (1987), appendix.

tres argumentos que desarrolla Fodor en apoyo de HLP.<sup>4</sup> Los tres argumentos tienen la estructura de una afirmación del consecuente o, dicho de otro modo, la estructura de una inferencia a la mejor explicación. Esta estructura argumentativa es, sin embargo, bastante frágil, pues sólo resulta válida en la medida en que no se descubra otro antecedente que, respaldado por evidencia independiente, dé lugar al mismo consecuente. En concreto, lo que vamos a discutir es si los modelos conexionistas nos proporcionan elementos adecuados para generar esa historia alternativa. Sin embargo, con el fin de saber lo que esta historia alternativa se supone que debe contarnos, necesitamos describir, previamente, los rasgos de nuestras capacidades cognitivas para cuya explicación se postula HLP.

En este sentido, Fodor considera que una teoría satisfactoria de la cognición debe ser capaz de dar cuenta de la responsabilidad causal de lo mental, así como de la productividad y sistematicidad de nuestros pensamientos, sin violar nuestro prejuicio metafísico fuerte. En este contexto, Fodor entiende, como cabría esperar, que HLP constituye en la actualidad la mejor hipótesis disponible para dar cuenta de estos rasgos de lo mental.<sup>5</sup>

En mi opinión, existen al menos dos estrategias diferentes para poner en cuestión la inferencia a la mejor explicación en la, que descansa HLP. En primer lugar, podríamos poner en duda que HLP pueda proporcionarnos una buena explicación de los rasgos básicos de nuestras capacidades cognitivas. Esta línea de argumentación no intentaría atacar la posibilidad de dar cuenta de la sistematicidad y productividad del pensamiento gracias a HLP; sino que subrayaría la existencia de rasgos de nuestro pensamiento para los que HLP difícilmente podría constituir un modelo realista. Pensemos, por ejemplo, en el problema del esquema y en la explosión computacional que a la que conducen todos los intentos de solución de este problema en el seno de un modelo clásico. Esta explosión computacional es especialmente importante porque entra en conflicto con el número de neuronas y la velocidad a la que se producen las transmisiones en nuestro cerebro; de donde parece seguirse que lo que sabemos acerca de nuestro cerebro parece minar la plausibilidad HLP en tanto que modelo realista de nuestra mente, dado que no puede dar cuenta de algunos de los rasgos de nuestro pensamiento dentro de condiciones neurológicamente plausibles.

Una segunda estrategia, frente al argumento de Fodor, intentaría proponer una historia alternativa acerca de la eficacia causal de los estados mentales, así como de su productividad y sistematicidad. La posibilidad de narrar esta historia alternativa se ha hecho más plausible desde el surgimiento de los modelos conexionistas. En lo que sigue me limitaré a discutir este último punto, es decir, si hay alguna razón de principio en contra de la posibilidad de que los modelos conexionistas puedan constituir la base para esa historia alternativa, es decir, para un tratamiento de la productividad y sistematicidad del pensamiento, así como de la responsabilidad causal de las propiedades intencionales, que prescinda de HLP.

---

<sup>5</sup> Fodor afirma en *Psychosemantics* que, apoyándonos en la productividad y sistematicidad de los pensamientos, puede construirse el siguiente argumento en apoyo de HPL: (a) Las capacidades lingüísticas son productivas y sistemáticas en virtud de tener partes constituyentes; (b) los pensamientos son también productivos y sistemáticos; (c) por tanto, una inferencia a la mejor explicación nos permite afirmar que los pensamientos también tienen partes constituyentes.

## 2. Fodor y McLaughlin: el epifenomenalismo de lo mental en el modelo de Smolensk

En esta sección, consideraremos, en primer lugar, la crítica de Fodor y McLaughlin al modelo de Smolensky como un modelo que pueda dar cuenta de la eficacia causal de las propiedades intencionales, lo que nos llevará, en segundo término, a la necesidad de ofrecer una caracterización más cuidadosa de las condiciones que propone Fodor para determinar la eficacia causal de una propiedad. Será precisamente la revisión de estas condiciones (que se llevará a cabo en la sección tercera) la que nos permitirá reivindicar el modelo de Smolensky frente a las objeciones de Fodor y McLaughlin.

### 2.1. Fodor y McLaughlin: Crítica del modelo de Smolensky

Conviene advertir, para empezar, que, en este artículo, no entraré a discutir en detalle el intento de Fodor y Pylyshyn (1988) de presentar los modelos clásicos y conexionistas como complementarios, si bien situando los modelos conexionistas a un nivel implementacional, y no cognitivo, de manera que estos modelos constituirían, como mucho, una de las implementaciones posibles que pueden recibir los modelos clásicos. Brevemente, diría que, incluso si asumimos la tesis de Fodor y Pylyshyn acerca de que los modelos conexionistas no se ubican a un nivel cognitivo, este hecho por sí mismo no excluye la posibilidad de elaborar una teoría representacional de la cognición cuyos estados básicos estén definidos a un nivel conexionista y donde las propiedades cognitivas se entiendan, sin postular HLP, como surgiendo de estados de un nivel inferior. En otras palabras, lo que parece crucial es determinar no si los estados conexionistas están a un nivel cognitivo, sino si hay alguna razón de principio en contra de la posibilidad de elaborar una teoría conexionista de la cognición que, a pesar de prescindir de HLP, dé cuenta de los rasgos básicos de nuestras capacidades cognitivas.<sup>6</sup>

De hecho, Fodor y McLaughlin intentan presentar esa razón de principio en su escrito «Connectionism and the Problem of Systematicity; Why Smolensky's Solution Doesn't Work». En este texto, se indica que HLP se postula con el fin de dar cuenta de la eficacia causal de los estados mentales en tanto que estados mentales, sin violar el principio metafísico fuerte que excluye la existencia de mecanismos intencionales; y se argumenta, por el contrario, que los modelos conexionistas no pueden preservar la eficacia causal de los contenidos mentales, una vez que se acepta el prejuicio metafísico fuerte.

En concreto, Fodor y McLaughlin centran sus críticas en la propuesta de Smolensky para dar cuenta de la sistematicidad de nuestros pensamientos dentro de un modelo conexionista. En general, podemos decir que Smolensky elabora su aparato conceptual en términos de vectores de productos de tensores y de superposición de vectores, si bien no es necesario entrar en los detalles. En lo que ahora nos interesa, basta con asumir, siguiendo a Fodor y McLaughlin, «...Que cada árbol clásico puede codificarse mediante un vector complejo de tal manera que los constituyentes del árbol se corresponden de algún modo regular con los componentes del vector.»<sup>7</sup>

<sup>6</sup> Para una discusión detallada de este punto, cf. Corbí (en prensa).

<sup>7</sup> Fodor y McLaughlin (1990), p. 197.

Otro punto importante que aceptan las dos partes implicadas es que, aunque los vectores complejos deben corresponder a pautas efectivas de actividad a lo largo de las unidades de la máquina, no hay ninguna necesidad de que los componentes del vector queden reflejados efectivamente en tales pautas de actividad, pues, entre otras razones, un mismo vector complejo puede descomponerse de modos diferentes; si bien algunos modos de descomposición pueden resultar más fructíferos que otros, desde el punto de vista de la simulación de árboles clásicos:

«... Un punto sobre el que el propio Smolensky es explícito y resulta vital en este respecto: Los componentes de un vector complejo no necesitan responder a pautas de actividad en la unidades efectivas de la máquina. En términos del propio Smolensky, los estados de actividad de los rellenadores y las unidades de rol pueden ser ‘imaginarios’, si bien los vectores de actividad últimos -los que no sirven ellos mismos como rellenadores o como componentes de rol de estructuras más complejas- deben de ser pautas de actividad que se den efectivamente a lo largo de las unidades de la máquina.» (Fodor and McLaughlin (1990), p. 197)

De todo esto se sigue, según Fodor y McLaughlin, que los componentes de los vectores complejos no son causalmente responsables del procesamiento que lleva a cabo la máquina conexionista, pues sería realmente milagroso que algo que no se da físicamente tuviese poderes causales, al menos si aceptamos nuestro prejuicio metafísico en cualquiera de sus versiones. Este hecho lleva a Fodor y McLaughlin a concluir que si, como parece seguirse del análisis de Smolensky, los componentes vectoriales son causalmente inertes y, por tanto, no pueden explicar la evolución de un sistema, entonces todavía no se ha mostrado cómo un modelo conexionista puede dar cuenta, en términos no-clásicos, de la sistematicidad y de la eficacia causal del pensamiento:

«[Smolensky] reconoce que los constituyentes de las representaciones mentales complejas no juegan ningún papel en la determinación de lo que ocurre cuando se instancian las representaciones... Parece, de este modo, que Smolensky debe de esconder en su manga una solución *no*-clásica para el problema de la sistematicidad; alguna solución que *no* dependa de asumir que los procesos mentales sean sensibles a su estructura constituyente.... No queda claro como los constituyentes (‘componentes’) de los vectores complejos pueden explicar su evolución [de la máquina] cuando, ‘*ex hypothesi*’, son como mucho causalmente inertes si no meramente imaginarios.» (Fodor and McLaughlin (1990), p. 200-1)

Me gustaría descartar, para empezar, una interpretación de la posición de Smolensky que parece resultar algo atractiva para Fodor y McLaughlin, y según la cual Smolensky estaría afirmando que los componentes vectoriales son entidades imaginarias.

Es cierto, como hemos visto, que Smolensky afirma explícitamente que los componentes vectoriales son entidades ‘imaginarias’ (si bien haciendo uso de las comillas), y esta indicación lleva a Fodor y McLaughlin a decir que los componentes vectoriales son «... como mucho causalmente inertes si no meramente imaginarios.» (aquí las comillas ya han desaparecido).

Estoy, por supuesto, de acuerdo con Fodor y McLaughlin en que entidades o propiedades meramente imaginarias, como 'ser una sirena', son causalmente inertes, dado que nunca se han instanciado. Sin embargo, no estoy seguro de que los componentes vectoriales sean imaginarios en este mismo sentido.

A primera vista, parece que el hecho de que puedan elaborarse diferentes análisis composicionales para el mismo vector complejo no impide que tengamos criterios para determinar, en el seno de cada análisis composicional, si un cierto componente vectorial está o no instanciado. Obviamente, los componentes vectoriales sólo serían causalmente responsables cuando estuviesen instanciados, mas, en contra de lo que ocurre con 'ser una sirena', podemos afirmar sin reparo que hay instanciaciones de componentes vectoriales.

Fodor y McLaughlin podrían replicar, sin duda, que se necesita un poco más para garantizar que una propiedad sea causalmente eficaz, y que estas condiciones adicionales no quedan satisfechas por la propuesta de Smolensky. Esta es precisamente la cuestión que pretendo examinar en las páginas que siguen.

## 2.2 Fodor: Condiciones suficientes para la eficacia causal

Una discusión detallada del punto anterior requiere, sin duda, una caracterización previa de cuándo se puede decir que una propiedad es causalmente responsable, y cuándo podemos excluir una propiedad como causalmente irrelevante. En este sentido, podemos empezar considerando la propuesta de Fodor en «Making Mind Matter More», donde se establece explícitamente una condición suficiente para que una propiedad sea causalmente responsable:

*Condición I:* «P es causalmente responsable si es una propiedad en virtud de la cual los individuos se subsumen bajo leyes causales.» (Fodor (1990b), p. 143)

Fodor entiende que esta condición suficiente no es circular a pesar del uso de la palabra 'causal' en el 'explanans', pues, en su opinión, puede que, al fin y al cabo, todas las leyes sean causales o, en cualquier caso, siempre podría ofrecerse una caracterización no-circular de ley causal en términos, por ejemplo, de sucesión temporal, asimetría, etc.

De todos modos, la condición I necesita complementarse con una segunda condición que tenga en cuenta la distinción entre las leyes de la ciencia básica y las leyes de las ciencias especiales. En este sentido, habría que decir que, mientras que en la ciencia básica no tiene sentido plantearse ninguna cuestión adicional acerca de cómo se implementan sus leyes, las ciencias especiales están siempre abiertas a la cuestión acerca de los mecanismos de la ciencia básica que explican cómo el antecedente de una ley no-básica causa su consecuente. Esta situación se deriva directamente de nuestro compromiso con el prejuicio metafísico fuerte, y daría lugar a la siguiente condición para la responsabilidad causal de una propiedad:

*Condición II:* «Pero una diferencia metafísicamente interesante entre leyes básicas y no-básicas es que, en el caso de las últimas pero no de las primeras, es necesario que haya un mecanismo en virtud del cual la satisfacción de su antecedente cause la satisfacción de su consecuente. Si 'Fs causan Gs' es básica, entonces no hay ninguna respuesta a la cuestión de cómo Fs causan Gs; simplemente lo hacen, y el hecho de que

lo hagan forma parte de los hechos que-no-requieren-explicación-adicional acerca del modo en el que el mundo está organizado.» (Fodor (1990b), p. 144)

En mi opinión, sin embargo, el análisis de Smolensky concerniente a la conducta de un sistema conexionista en términos de vectores de productos tensoriales y de superposición de vectores, así como su descomposición reflejando la estructura de los árboles clásicos, satisface la condición suficiente para la eficacia causal indicada por Fodor. En primer lugar, cabe indicar que las propiedades de los componentes vectoriales se subsumen bajo leyes causales si las propiedades de las partes constituyentes clásicas lo hacen, pues hemos asumido que puede establecerse una correspondencia estable entre ambas. Podemos, por tanto, concluir que las propiedades de los componentes vectoriales satisfacen la condición I.

En segundo término, el procesamiento al nivel de las pautas de conectividad entre las unidades puede verse como el mecanismo que garantiza la validez de las leyes que vinculan las propiedades de los componentes vectoriales, dado que, según el modelo de Smolensky, los vectores complejos encuentran su instanciación en las pautas de conectividad entre unidades. En definitiva, podemos concluir que las propiedades de los componentes vectoriales satisfacen también la condición II para la existencia de eficacia causal.

Se hace necesaria, en consecuencia, una lectura más restringida de la condición II si Fodor y McLaughlin quieren negar la eficacia causal de las propiedades de los componentes vectoriales. Nos referiremos a esta versión fuerte de la condición II como 'Condición II\*'. Así, podríamos decir que la condición II\* requiere la instanciación física de las propiedades causalmente responsables, de manera que pueda establecerse una relación sistemática entre los componentes de los tipos de estado no-físico y físico que se instancian. Parece claro que el análisis de Smolensky no satisface esta condición instanciadora fuerte pues, como vimos, no se puede establecer ninguna relación sistemática entre los componentes vectoriales y los componentes de las pautas de actividad.

Asumo, en definitiva, que Fodor y McLaughlin se comprometen con la condición II\*, dado que este compromiso les permite concluir: (i) que Smolensky sólo consigue dar cuenta de la sistematicidad en un modelo conexionista al precio de la inercia de las propiedades intencionales; y (ii) que HLP es la mejor explicación disponible tanto para la sistematicidad como para la eficacia causal de los contenidos mentales, pues HLP no sólo satisface la condición I, sino que se compromete directamente con la condición II\*, en la medida en que reivindica que los estados mentales mismos, y no sólo sus objetos intencionales, tienen partes constituyentes.

### **3. Un tratamiento alternativo de la responsabilidad causal**

En esta sección, intentaré mostrar que la condición II\* es demasiado fuerte como condición necesaria para la responsabilidad causal. En segundo término, defenderé un tratamiento alternativo de la responsabilidad causal que, aparte de concordar con nuestras intuiciones, parece disolver el conflicto entre los modelos conexionistas y la responsabilidad causal de las propiedades intencionales.

### 3.1. Una objeción a la condición II\*

Una propiedad no-física es causalmente responsable, según la condición II\*, sólo si puede establecerse, al menos en principio, una relación sistemática entre la instanciación de esta propiedad y la instanciación de ciertas propiedades físicas<sup>8</sup>. Pienso, sin embargo, que una condición necesaria tan fuerte entra en conflicto con nuestras intuiciones acerca de la responsabilidad causal.

Para empezar, podemos indicar que nosotros adscribimos responsabilidad causal a todo tipo de propiedades en las ciencias especiales sin tener en cuenta si realmente puede establecerse esa relación sistemática. Se pueden señalar, además, razones de principio para dudar acerca de la posibilidad de establecer tales relaciones tipo-tipo, pues, en múltiples ocasiones, tales relaciones exigirán la introducción de disyunciones indefinidamente abiertas de propiedades físicas, y este hecho nos llevará a algo muy cercano a una mera correlación caso-caso.

Recordemos, por ejemplo, cuando la selección nacional de Alemania Occidental marcó el gol de la victoria en la final de la Copa del Mundo de 1990. Ríos de alegría recorrieron las calles de Alemania, los jugadores se abrazaban entusiasmados. Pero las consecuencias de este gol irían mucho más lejos: incrementaría el orgullo nacional de los alemanes, triplicaría la prima que recibirían jugadores y entrenadores, aumentaría la cotización de los jugadores alemanes en general, se organizarían recepciones multitudinarias. Nadie en su sano juicio dudaría de la eficacia causal de ese gol y, en general, de los goles que dan la victoria en la final de una Copa del Mundo.

Y, sin embargo, este hecho tan trivial parece poner en apuros al fiscalista fuerte, a la condición II\*. Pensemos, ¿podemos, acaso, especificar un conjunto de propiedades físicas que constituyan condición suficiente de la realización de un gol? A primera vista, la respuesta es sencilla: basta con que el balón traspase la raya de la portería. Aunque habría que ser un poco más cautelosos: no basta con traspasar la raya, tendrían que concurrir otras circunstancias, por ejemplo, que el hecho en cuestión ocurriese durante un partido de fútbol y no durante un entrenamiento. Hay que reconocer, con todo, que también hay entrenamientos tan serios que se llaman partidos de entrenamiento y donde los goles sí que se computan. Lo que esta ambigüedad parece sugerir es que también podemos encontrarnos con dificultades a la hora indicar un conjunto de propiedades físicas que constituyan condición suficiente de la realización de un partido de fútbol. De hecho, habría que distinguir entre partidos de entrenamiento, partidos oficiales, partidos amistosos, partidos de liga, partidos de copa etc. Todas estas distinciones tienen su relevancia causal: la prima no es la misma si se gana la Liga que si se gana la Copa, si se vence en un partido amistoso que si el partido es oficial; las emociones que agitan a las masas en los estadios y el aspecto que presentan sus gradas son bastante diferentes. Distinguimos con facilidad entre todos estos tipos de partido atendiendo a los decretos de los diferentes órganos rectores del fútbol nacional e internacional, pero ¿cómo indicar condiciones físicas suficientes para hacer tales distingos?

<sup>8</sup> Las propiedades físicas que estamos considerando están constituidas muy a menudo por conjuntos disyuntivos de propiedades físicas.

Habría que recordar, por otra parte, que no estábamos hablando de un gol cualquiera ni de un partido cualquiera, estábamos hablando del gol que daba la victoria en la gran final. Los goles que han dado la victoria en la final de las diferentes ediciones de la Copa del Mundo de Fútbol han sido de muy distinta factura, se han visto rodeados de condiciones muy diferentes, los organismos organizadores de la Copa y del partido han ido variando, también las normas del fútbol que obligan a conceder o invalidar ciertos goles etc. Parece imposible encontrar un conjunto de propiedades físicas que sea condición suficiente de la instanciación de la propiedad 'ser el gol de la victoria de una final de la Copa del Mundo de Fútbol' y, sin embargo, esta propiedad tiene una exagerada eficacia causal. Esto no quiere decir que para que algo sea un gol de la gran final no sea *necesario* que ocurran ciertos eventos físicos. Ese no es el problema que aquí se discute. Lo que más bien se sugiere es que no hay un conjunto de eventos físicos que sea condición *suficiente* para la instanciación de tales goles.

En términos generales, podríamos decir que lo que en principio se presentaba como una relación tipo-tipo entre un conjunto de condiciones físicas y la propiedad de 'ser un gol' o 'ser el gol que da la victoria en la final de la Copa del Mundo de Fútbol', se aproxima cada vez más a un relación caso-caso. Así, para determinar que algo sea un gol no basta con indicar que una forma esférica con ciertas características físicas ha traspasado una raya entre dos postes y un largero, hace falta observar la completa distribución de organismos en un rectángulo, hay que extender nuestras consideraciones más allá del estadio y describir los elementos físicos implicados en la multitud de decisiones institucionales que preceden a un partido, a la organización de una competición, etc; de manera que pequeñas variaciones en lugares muy alejados del estadio pueden hacer que algo no sea un gol o que no sea un gol válido. Además, lo que en un momento era un gol válido podría dejar de serlo por la alteración del reglamento del fútbol y, en consecuencia, lo que un momento del tiempo eran condiciones físicas suficientes pueden dejar de serlo, a no ser, por supuesto, que en la enumeración de esas condiciones se incluyan más y más elementos del mundo, con lo que nos iremos acercando peligrosamente al establecimiento de una relación caso-caso en la medida en que las condiciones físicas de instanciación de propiedades como 'ser un gol' parecen incluir aspectos muy extensos del mundo si no el universo entero, pues pequeñas alteraciones físicas en algún lugar del mundo pueden provocar que lo que era un gol válido deje de serlo. Al final, lo único que parece que podemos decir es que en un mundo físicamente idéntico a este en todos sus aspectos coincidiría con este en todas sus propiedades futbolísticas. Pero esto no es más que el principio general de supervivencia de las propiedades no-físicas del mundo respecto de sus propiedades físicas.

Se podría replicar, con todo, que no está claro que esta circunstancia se dé en ciencias naturales especiales como la geología o la biología. Pienso, sin embargo, que estas ciencias especiales se verían afectadas también por mi argumento, aunque en menor grado. Así, en la medida en que mutaciones en las propiedades físicas de objetos individuados no-físicamente son compatibles con la estabilidad de sus propiedades no-físicas, puede haber casos en los que cambios físicos en otras áreas del universo hagan que un conjunto de condiciones físicas deje de ser suficiente para la posesión de ciertas propiedades no-físicas. Este hecho introduce, en definitiva, los dos factores que, en el

caso de las categorías económicas, parecían socavar el carácter tipo-tipo de la relación que la condición II\* postula, a saber: la relatividad a un cierto tiempo y la necesidad de incorporar más y más rasgos del mundo.

De todos modos, el hecho de que teorías distintas parcelen el mundo de modo diferente es compatible, a mi entender: (i) con el carácter básico de las taxonomías físicas y, por tanto, con la posibilidad de establecer diversos tipos de conexiones entre las diferentes taxonomías, incluso si ya no podemos esperar que estas conexiones sean tan fuertes como se sugiere en la condición II\*; y (ii) con la responsabilidad causal de las propiedades no-físicas. Así, en la próxima sección, presentaré un tratamiento de la responsabilidad causal que aspira a integrar todos estos elementos.

### 3.2 *Un tratamiento alternativo*

Me gustaría proponer el siguiente conjunto de requisitos como una condición suficiente para la responsabilidad causal de una propiedad:

- (a) que la propiedad forme parte de una ley causal
- (b) que la propiedad esté físicamente instanciada o, al menos, sea superviniente respecto a las propiedades físicas del mundo, y
- (c) que no exista ningún conjunto alternativo de leyes causales más generales y rigurosas que explique aproximadamente el mismo tipo de fenómenos, y en los que la propiedad en cuestión no aparezca.

El requisito (a) reproduce literalmente la condición I establecida por Fodor. A primera vista, este requisito parece excluir la posibilidad de que las propiedades intencionales sean causalmente eficaces, dado que se tiende a pensar que no hay leyes estrictas en psicología ni tampoco leyes psicofísicas, pues todas las generalizaciones que se puedan proponer en estos campos han de incluir cláusulas 'ceteris paribus'. Donald Davidson es, como sabemos, el filósofo que ha planteado con más rigor esta objeción. No obstante, Fodor ha propuesto una respuesta a esta objeción que encuentro satisfactoria: el argumento de Davidson es tan potente que afectaría no sólo a la psicología sino a todas las ciencias especiales, dado que todas ellas incluyen inevitablemente ese tipo de cláusulas, las cuales, por otra parte, no pueden eliminarse, al igual que ocurre en psicología, manteniéndose dentro del vocabulario de cada una de esas ciencias. Por tanto, si la estrategia de Davidson fuese correcta no sólo las propiedades intencionales dejarían de satisfacer este requisito, sino también las propiedades definidas por la geología, la biología, etc. En estas condiciones, me parece más apropiado ampliar nuestro concepto de ley causal y entender que la presencia de las cláusulas 'ceteris paribus' no descarta la posibilidad de que las propiedades intencionales, o las propiedades de cualquier ciencia especial, satisfagan este primer requisito.<sup>9</sup>

El requisito (c) nos ayuda a excluir del ámbito de las propiedades causalmente eficaces algunas propiedades que, a pesar de aparecer en algunas generalizaciones

<sup>9</sup> Con esto no pretendo que se hayan resuelto todos los problemas en torno al carácter legaliforme de las generalizaciones psicológicas, sino únicamente indicar cómo se podría evitar una de las objeciones más inmediatas que se suele plantear (Cf. Fodor (1988), pp. 15 ff).

verdaderas, no nos sentimos inclinados a reconocerlas como causalmente eficaces. A mi entender, eso ocurre cuando disponemos de una teoría alternativa que da cuenta de esas generalizaciones en virtud de un sistema de principios más generales y rigurosos. Consideremos, por ejemplo, la siguiente generalización: «Si un niño tiene fiebre alta y le duele la garganta, probablemente alguno de sus hermanos se encontrará en la misma situación unos días después». Concedamos que esta generalización, o alguna semejante, es verdadera, con lo que las propiedades 'tener fiebre alta' o 'doler la garganta' cumplen el requisito (a) y, sin embargo, no me parece nada correcto decir que la causa de que un niño tenga fiebre alta y le duela la cabeza el viernes sea que su hermano estaba en esas circunstancias a principios de semana. Entiendo, por el contrario, que la propiedad causalmente eficaz, en este caso, es la transmisión de ciertos gérmenes bacteriológicos de un hermano a otro. Considero que lo que me induce a pensar que esta última es la verdadera causa es que nos encontramos con una teoría de la transmisión de enfermedades infecciosas que no sólo es más general y rigurosa que la recogida en la generalización con la que iniciábamos el ejemplo, sino que además da cuenta de aproximadamente el mismo tipo de fenómenos: el contagio de enfermedades.

Es importante resaltar, sin embargo, que el requisito (c) no disminuye las expectativas de que las propiedades intencionales tengan eficacia causal porque estamos muy lejos de disponer de una teoría alternativa que explique aproximadamente el mismo tipo de fenómenos que la psicología intencional.

Respecto al requisito (b), cabría empezar advirtiendo que, tras la discusión anterior acerca de los modos en que algunas propiedades de las ciencias especiales se relacionan con las propiedades físicas, debemos estar preparados para reconocer la existencia de diferentes grados de instanciación física compatible con nuestro prejuicio metafísico acerca de la inexistencia de mecanismos intencionales. En algunos casos encontraremos una correlación tipo-tipo entre conjuntos disyuntivos cerrados de propiedades físicas y propiedades no-físicas, pero, en otros casos, ese conjunto disyuntivo se hará más extenso y menos cerrado.<sup>10</sup>

¿Cuál es, no obstante, el límite en la flexibilidad de las posibles conexiones entre las propiedades físicas y no-físicas que estoy dispuesto a tolerar? La respuesta que, en mi opinión, se sigue del debate anterior es que la supervivencia de las propiedades no-físicas respecto a las propiedades físicas define los límites del requisito (b). De hecho, este requisito define la versión débil del prejuicio metafísico, que, a su vez, caracteriza a la posición fisicalista más endeble y la única versión del prejuicio metafísico que realmente considero digna de tal nombre. En realidad, este último principio forma parte del suelo rocoso de mi visión del mundo; en cierto modo, no puedo concebir ningún tipo de evidencia que pudiese convencerme de su falsedad. En cambio, no sólo me resulta plenamente comprensible que la tesis fisicalista fuerte sea falsa, sino que además he intentado mostrar por qué pienso que lo es. Otra cuestión que no discuto

---

<sup>10</sup> En este contexto, el carácter cerrado de un sistema es claramente una cuestión de grado: por ejemplo, podríamos decir que las taxonomías biológicas están más cerradas desde un punto de vista físico que las categorías económicas.

aquí es si la condición fiscalista fuerte puede interpretarse no tanto como un requisito, sino como una estrategia de investigación tendente a incrementar el nivel de cohesión entre las diferentes teorías que componen nuestra concepción materialista del mundo.

La satisfacción de estos tres requisitos es, por supuesto, una cuestión de grado. Se puede esperar, por tanto, que la cuestión acerca de la responsabilidad causal de una cierta clase de propiedades no sea en todos los casos un asunto de todo o nada. En este sentido, es posible encontrar combinaciones diferentes en la satisfacción de los requisitos (a)-(c), de manera que cuanto menos quede satisfecho cada uno de ellos, menos nos sentiremos tentados a adscribir responsabilidad causal a la clase de propiedades que en cada caso se considere.

Finalmente, parece obvio que, si atendemos a mi tratamiento de la responsabilidad causal, los modelos conexionistas no son incompatibles con la responsabilidad causal de las propiedades intencionales, pues la propuesta de Smolensky no sólo satisface los requisitos (a) y (c), sino también el requisito (b) en una de sus versiones más restrictivas, pues, si bien no se concede la versión más fuerte de este requisito, al menos se defiende la existencia de relaciones tipo-tipo entre los vectores complejos y las pautas de actividad.

#### 4. Conclusiones

Hemos indicado, al inicio, que el surgimiento de los modelos conexionistas puede representar una amenaza seria para HPL, pues todos los argumentos en defensa de HPL tienen la estructura de una inferencia a la mejor explicación y, en consecuencia, el nacimiento de un tratamiento alternativo, basado en evidencia independiente, tiende ciertamente a invalidar tales argumentos.

Fodor y McLaughlin replican, sin embargo, que una teoría conexionista de la cognición que abandone HPL no puede constituir tal planteamiento alternativo; pues tales modelos no son compatibles con la responsabilidad causal de las propiedades intencionales, dado que sólo HPL puede garantizar la eficacia causal de los estados mentales en tanto que estados mentales.

He intentado mostrar, sin embargo, que el argumento de Fodor y McLaughlin descansa en una concepción demasiado restrictiva de la responsabilidad causal, que conduce a negar la responsabilidad causal de muchas propiedades no-físicas definidas en el seno de ciencias especiales tales como la economía y la biología, a pesar de nuestra tendencia intuitiva a reconocerlas como causalmente relevantes.

Este hecho me llevó a proponer un conjunto alternativo de requisitos que definen una condición suficiente para la responsabilidad causal. Según esta propuesta, (i) se respeta el prejuicio metafísico débil concerniente a la inexistencia de mecanismos intencionales; (ii) quedan excluidas, en cambio, las propiedades que no resultan interesantes desde un punto de vista explicativo o que forman parte de teorías no-empíricas o metafísicas; (iii) las propiedades postuladas por la economía y la biología serán igualmente causalmente responsables; (iv) los modelos conexionistas son, en principio, compatibles con la responsabilidad causal de las propiedades intencionales; por lo que podemos concluir (v) que los argumentos de Fodor y McLaughlin no

consiguen hacer frente a la amenaza que la emergencia de los modelos conexionistas supone para la hipótesis del lenguaje del pensamiento.

## Bibliografía

- CORBI, J.E. (en prensa), «Classical and Connectionist Models: Levels of Description», *Synthese*.
- FODOR, J. (1987), *Psychosemantics*, Cambridge, Mass., The MIT Press.
- FODOR, J. (1990), «Making Mind Matter More» en: Fodor, J., *Theory of content and other essays*, Cambridge, Mass., The MIT Press, pp. 137-159.
- FODOR, J. & PYLYSHYN, Z.W. (1988), «Connectionism and Cognitive Architecture: A Critical Analysis» in *Cognition*, v. 28, pp. 3-71.
- FODOR, J. & MCLAUGHLIN, B. (1990), «Connectionism and The Problem of Systematicity; Why Smolensky's Solution Doesn't Work» in *Cognition*, v. 35, pp. 183-204.
- SMOLENSKY, P. (1987), «The Constituent Structure of Mental States: A Reply to Fodor and Pylyshyn» en: *Southern Journal of Philosophy*, 26, 137-160.
- SMOLENSKY, P. (1988 b), «Connectionism, Constituency and the Language of Thought.» en: Loewer, B. & Rey, G. (eds.), *Meaning in Mind: Fodor and his Critics*, Oxford, Basil Blackwell.