

## **Problemas de la causalidad mental: el nuevo dualismo**

Carlos J. Moya. Universitat de València (carlos.moya@uv.es)

Congreso “Problemas filosóficos de la relación mente-cuerpo”

València, 27-29 de noviembre de 2013

Desde tiempos muy antiguos, el desarrollo de las ciencias de la naturaleza ha tendido a entrar en conflicto con convicciones muy básicas del sentido común. Pensemos por ejemplo en la afirmación de Demócrito según la cual lo único que realmente existe son átomos y vacío, frente a nuestra creencia en la existencia real de muchas otras propiedades, seres y objetos. Este conflicto, sin embargo, se agudizó extraordinariamente con el nacimiento de la Nueva Ciencia matemática de la naturaleza, con la astronomía de Copérnico y Kepler y la física de Galileo y Newton. La concepción del universo como un gigantesco mecanismo regido por leyes necesarias parecía amenazar, entre otras, la creencia en la libertad de la voluntad y en la responsabilidad moral de los seres humanos. En este contexto, el dualismo antropológico cartesiano tenía la gran ventaja de conciliar la concepción determinista-causal del mundo de la Nueva Ciencia con las convicciones del sentido común. Así, en relación con los seres humanos, no había realmente conflicto o contradicción entre ambas perspectivas; ambas podían ser verdaderas, y según Descartes de hecho lo eran; el conflicto se evitaba al sostener que no eran verdaderas de la misma entidad, sino de entidades o sustancias distintas, ya que un ser humano no era en realidad una única entidad, sino la unión pasajera de dos entidades o sustancias, el cuerpo y el alma o la mente. Las leyes necesarias que regían la naturaleza física regían también el cuerpo humano, como una parte más de dicha naturaleza física, y las tesis de la física serían también verdaderas de él, como de cualquier otra entidad física. Sin embargo, las convicciones del sentido común acerca de la libertad, la racionalidad, la conciencia, la responsabilidad moral, etc., eran también verdaderas, sólo que no del cuerpo sino del alma o la mente, una entidad no física y no sujeta por lo tanto a las leyes de la naturaleza.

La solución cartesiana al mencionado conflicto de perspectivas sobre el ser humano era atractiva y elegante. Sin embargo, era internamente inestable. Como es bien sabido, el

talón de Aquiles del dualismo cartesiano fue el problema de la causalidad mental, en especial la influencia de la mente y sus estados sobre el cuerpo y el comportamiento. ¿Cómo dar sentido, desde la concepción dualista, a explicaciones tan razonables de mi comportamiento como “fui a casa de mi amigo porque quería felicitarle por su cumpleaños” y otras muchas por el estilo? La explicación da cuenta de mi acción, que se desarrolla en el mundo físico, espaciotemporal, atribuyéndola a un deseo mío, que es un estado mental. Pero desde el punto de vista del dualismo cartesiano, esta influencia de mi deseo sobre mi comportamiento es un misterio.

Una posición, como el dualismo cartesiano, que no pueda dar cuenta de esa clase de explicaciones es seriamente defectuosa, porque se trata de explicaciones que cumplen funciones muy importantes en nuestra vida: nos permiten comprender el comportamiento propio y ajeno al presentarlo como racionalmente justificado; nos permiten prever las acciones de otras personas y coordinar nuestras acciones con las suyas al atribuirles determinadas razones para actuar. La acción llevada a cabo por razones es, además, parte de lo que significa ser un agente racional, libre y responsable. Por lo tanto, el escollo de la causalidad mental tendía a amenazar, retrospectivamente, la propuesta de Descartes para conciliar la ciencia y el sentido común. Si el alma y el cuerpo son tan distintos que no puede haber influencia causal de la una sobre el otro, si nuestras deliberaciones, creencias, deseos y decisiones no son capaces de mover nuestros cuerpos, ¿cómo podemos sostener que nuestras acciones, que involucran el movimiento de nuestros cuerpos, pueden ser libres? Las acciones humanas parecen estar causadas, pues, por fenómenos físicos y no por nuestras creencias, valores y decisiones; y esas relaciones causales parecen estar regidas por leyes físicas necesarias, no por nuestra (supuestamente) libre voluntad.

Nuestra aceptación cotidiana de las explicaciones de la acción mediante razones, junto con las propiedades que van asociadas a ellas, como la libertad, la racionalidad y la responsabilidad, presupone la convicción de que aquello que creemos, deseamos, esperamos, valoramos, etc., puede ejercer, y de hecho ejerce, una influencia sobre nuestro comportamiento, llevándonos a actuar de un modo determinado y no de otro. Llamaremos a esta influencia de nuestros estados y procesos mentales sobre nuestro comportamiento “causalidad mental”. En el texto siguiente, Fodor expresa claramente

la importancia que tiene en nuestra vida la creencia en la causalidad mental, y el sentimiento de vértigo que se apodera de nosotros cuando imaginamos su inexistencia:

Si no es literalmente verdad que mi deseo es causalmente responsable de que trate de alcanzar algo, o que el prurito que siento es causalmente responsable de que me rasque, o que mi creencia es causalmente responsable de que diga algo..., si nada de eso es literalmente verdadero, prácticamente todo lo que creo sobre cualquier cosa es falso, y es el fin del mundo.<sup>1</sup>

En el texto anterior, Fodor incluye bajo el concepto de causalidad mental la influencia de estados mentales de cualquier tipo sobre el comportamiento. Pero según una clasificación hoy habitual (que desde luego Fodor conoce perfectamente), podemos distinguir dos grandes grupos de estados mentales.

Tenemos, por una parte, las sensaciones de diversos tipos, caracterizadas por una peculiar forma de aparecer al sujeto o de ser experimentadas por él. El dolor y el placer y sus diversos tipos, las sensaciones olfativas, gustativas, etc., serían ejemplos de este grupo. Podríamos denominar esta clase de estados “estados fenomenológicos”, debido a que al menos una parte central de su naturaleza consiste en el modo peculiar en que aparecen o se presentan al sujeto y son sentidos o experimentados por él.

El otro gran grupo de estados mentales está constituido por todos aquellos estados que tienen un contenido semántico o intencional, que representan o refieren a algo distinto del estado en cuestión. En este grupo encontramos las creencias, deseos, esperanzas, propósitos, temores, etc. Así, el contenido de una creencia es aquello que es creído, el de un deseo lo que es deseado, etc. Brentano consideró la intencionalidad como la característica distintiva de lo mental. Al referirse a ella, Brentano habló de “inexistencia intencional”, de “referencia a un contenido”, de “dirección hacia un objeto (por el que no hemos de entender una realidad en este caso)”, de “objetividad inmanente”. Al

---

<sup>1</sup> Fodor, J., “Making mind matter more”, *Philosophical Topics*, 17 (1989), pp. 59-80, p. 77.

matizar de esta forma el concepto de lo intencional, Brentano apunta al hecho de que, por ejemplo, lo que alguien cree puede ser falso, o lo que alguien desea puede no existir o incluso ser imposible de lograr, sin que ello prive a estos estados de contenido o dirección hacia un objeto. Advertamos que, según este criterio, el significado, el contenido semántico de lo que decimos o escribimos es también una propiedad mental, en tanto que es intencional en el sentido indicado. Así, el uso significativo del lenguaje hablado o escrito es también un fenómeno mental.

Esta clasificación no pretende ser exhaustiva. Es plausible pensar que hay estados mentales, como por ejemplo los sentimientos y ciertas emociones, que tienen a la vez fenomenología e intencionalidad. Y tampoco cubre todo aquello que llamamos "mental". Por ejemplo, ciertos rasgos de carácter (la generosidad, la envidia, el egoísmo, etc.) o capacidades (la inteligencia, la sociabilidad, el talento musical, etc.) son propiedades mentales, pero no parecen tener fenomenología ni intencionalidad, aunque sí pueden tenerlas sus manifestaciones típicas (como el deseo de que otra persona fracase en todo aquello que emprende, en el caso de la envidia).

Sin embargo, podemos considerar la fenomenología y la intencionalidad, en el sentido que hemos indicado, como propiedades mentales fundamentales, que definen al menos dos grupos básicos de estados mentales, los estados fenomenológicos y los estados intencionales.

Correspondiendo a esta distinción entre estados fenomenológicos y estados intencionales, podemos también distinguir dos tipos de causalidad mental, que podemos denominar, siguiendo esa terminología, causalidad fenomenológica y causalidad intencional; estas formas de la causalidad mental designan, respectivamente, la influencia causal sobre el comportamiento de los estados fenomenológicos y de los estados intencionales. Un ejemplo del primer tipo sería la queja o el grito causados por un dolor repentino. Un ejemplo del segundo tipo sería volver a mi despacho al recordar que he dejado el ordenador encendido.

En esta ponencia dejaremos de lado la causalidad fenomenológica y nos ocuparemos en exclusiva de la causalidad intencional.

Es importante, antes de proseguir, precisar lo que vamos a entender por “causalidad intencional”, que, recordémoslo, es una variedad de la causalidad mental. La causalidad intencional (y la causalidad mental en general) no consiste en que un estado o suceso mental cause un estado o suceso físico, sino en que lo cause *en virtud de* sus propiedades mentales: en el caso de un estado mental intencional, en virtud de su significado o su contenido intencional. Para aclarar este punto, podemos recurrir a un ejemplo de Dretske.<sup>2</sup> Supongamos que el canto de las notas más agudas de un aria de ópera por parte de una soprano causa la rotura de una copa de fino cristal. Las palabras cantadas por la soprano tienen significado y, en este sentido, el canto es un suceso mental; este suceso mental causa un suceso físico, la rotura de la copa; sin embargo, no estamos ante un caso de causalidad mental, sino de causalidad puramente física, ya que la rotura se produjo en virtud de propiedades físicas del canto, como la longitud de onda del sonido y la vibración del aire, mientras que sus propiedades mentales fueron totalmente irrelevantes. Si las palabras cantadas hubieran tenido otro significado o ninguno en absoluto, la copa se habría roto igualmente.

Frente a la posición dualista cartesiana, que concebía las creencias, deseos, intenciones, etc., como modos del atributo pensamiento, que era la propiedad esencial de una sustancia no física, el alma o yo pensante, la concepción dominante desde hace ya mucho tiempo es de carácter monista-materialista. En las posiciones más radicales, la relación entre las propiedades mentales y las propiedades físicas se concibe en términos de identidad: las propiedades mentales *son* (en el sentido de “son idénticas a”) propiedades físicas; concretamente, propiedades del sistema nervioso central. La teoría de la identidad de propiedades goza hoy de escaso favor entre los teóricos, debido, entre otros problemas, a que una misma propiedad mental, por ejemplo, el deseo de comer una naranja, puede realizarse físicamente (neurológicamente) de múltiples formas en distintos cerebros, e incluso en un mismo cerebro a través del tiempo. Esta identidad de propiedades conlleva la identidad de sucesos o estados particulares mentales y físicos, pero no a la inversa.

---

<sup>2</sup> Dretske, F., *Explaining Behavior: reasons in a world of causes*, Cambridge MA, MIT Press, 1988, p. 79.

Posiciones más moderadas, y más plausibles, se inclinan, no por la identidad, sino por la superveniencia de las propiedades mentales sobre las propiedades físicas. Un grupo de propiedades (por ejemplo, mentales) sobreviene fuertemente a otro grupo de propiedades (por ejemplo, neurológicas) en el caso de que, necesariamente, si se dan estas últimas en alguna situación o entidad, se dan también las primeras. La relación de superveniencia es una relación de dependencia metafísica necesaria entre propiedades. Por ejemplo, si las propiedades mentales sobrevienen fuertemente a las propiedades neurológicas, dos personas con cerebros neurológicamente idénticos hasta sus últimos detalles en un momento dado tendrán también las mismas creencias, deseos, sensaciones, intenciones, etc. en ese momento. Así, para este materialismo sobre la mente, las propiedades mentales dependen metafísicamente de las propiedades físicas, más concretamente neurológicas, y no pueden darse sin ellas. La mente, podemos decir, sobreviene al cerebro, depende metafísicamente de él y no puede existir sin él. En adelante nos centraremos sobre todo en esta visión más moderada del materialismo, en términos de superveniencia de las propiedades mentales sobre las propiedades neurológicas, más que en términos de identidad.

Normalmente los partidarios de la superveniencia de propiedades aceptan la identidad de estados y sucesos particulares mentales y físicos, de modo que cada estado mental particular es numéricamente idéntico a un estado físico particular: son uno y el mismo estado. Así, un mismo estado, por ejemplo, una determinada configuración neuronal, puede tener tanto propiedades físicas como propiedades mentales, siendo a la vez mental y físico. No existen, pues, frente al dualismo, sustancias particulares que sean sólo mentales o tengan sólo propiedades mentales. Todo particular mental (que tenga propiedades mentales) es también físico (tiene propiedades físicas), aunque no a la inversa. Las propiedades físicas son el sustrato necesario para las propiedades mentales y las determinan.

Podría parecer que esta perspectiva materialista sobre la mente y las propiedades y estados mentales ha resuelto definitivamente el problema de la causalidad mental que socavó las bases del dualismo cartesiano. Según esta nueva perspectiva, no hay realmente dificultad en entender cómo los sucesos y estados mentales pueden causar sucesos y estados físicos (por ejemplo, cómo nuestras creencias, deseos, intenciones...

actúan sobre nuestros músculos y nos llevan a actuar): pueden hacerlo porque esos estados mentales son a la vez estados físicos (neurofisiológicos).

En este punto, sin embargo, algunos autores han hablado justificadamente de “la venganza de Descartes”,<sup>3</sup> porque el problema de la causalidad mental reaparece una vez más, bajo nuevas formas, en el seno del materialismo. Kim escribe:

Uno de los desarrollos sorprendentes en la filosofía de la mente durante las dos últimas décadas aproximadamente ha sido el retorno del problema mente-cuerpo – no como un problema para los dualistas substanciales, una raza que prácticamente ha desaparecido de la escena, sino como una amenaza para los fisicalistas [materialistas] que aspiran a tomar en serio lo mental.<sup>4</sup>

Recordemos que el problema de la causalidad mental consiste en entender cómo un suceso mental puede dar lugar a uno físico *en virtud de sus propiedades mentales* (su intencionalidad, su contenido intencional, en el caso de los estados intencionales), no de sus propiedades físicas. Y esto es lo que el nuevo materialismo tiene problemas en explicar. En concepciones como las de Donald Davidson o Jaegwon Kim, que representan formas distintas de materialismo sobre la mente, basadas en relaciones de superveniencia, las propiedades mentales tienden a convertirse en epifenómicas y causalmente irrelevantes, de modo que el trabajo causal lo llevan a cabo las propiedades físicas a las que aquellas sobrevienen y de las que dependen.

Kim, por ejemplo, acepta la realización múltiple de las propiedades mentales, pero considera que esta característica es compatible con su reducción a lo físico. Para Kim, las propiedades mentales han de concebirse como propiedades funcionales, caracterizadas por su papel causal en la economía de los organismos que las poseen; las propiedades mentales son realizadas, en distintos organismos, por diversas propiedades físicas, presumiblemente neurofisiológicas, que desempeñan de hecho los papeles

---

<sup>3</sup> Kim, J., *Mind in a Physical World*, Cambridge MA, MIT Press, 2000, p. 46.

<sup>4</sup> *Ibid.*, p. 56.

causales asociados a las propiedades mentales. Así, los poderes causales de las propiedades mentales son idénticos a los poderes causales de las propiedades físicas que las realizan y de las que dependen. Como Kim lo expresa:

En una posición reduccionista de esta clase... los poderes causales de las propiedades mentales resultan ser precisamente los de sus realizadores físicos, y las propiedades mentales no aportan al mundo poderes causales nuevos.<sup>5</sup>

Pero, según esto, en la propuesta de Kim, las propiedades mentales tienden a aparecer como puramente epifenómicas, causalmente prescindibles. De este modo, la propuesta de Kim parece eliminar la causalidad mental e intencional, más que dar cuenta de ella.

Mi hipótesis en esta ponencia es que, en la base del materialismo contemporáneo, y de sus problemas en relación con la causalidad mental, hay una nueva forma de dualismo, que cabría llamar, un tanto paradójicamente, un dualismo materialista. Se trata de una posición muy extendida en la actualidad, no sólo entre filósofos, psicólogos y neurocientíficos, sino también entre no especialistas. Podemos encontrarla con cierta facilidad en revistas de divulgación científica, e incluso en revistas de información general. En esta ocasión, no se trata del dualismo entre la mente o el yo pensante, como una entidad no física, y el cuerpo, sino entre una parte del cuerpo, el cerebro, y el resto de él. Mientras que en el dualismo cartesiano el yo pensante se identificaba con el alma o la mente (“este yo, es decir, el alma, por la cual soy lo que soy”),<sup>6</sup> ahora el yo se identifica con el cerebro. Somos nuestro cerebro. El resto del cuerpo no es esencial a nuestra identidad psicológica. Así, la importancia del cerebro, que no pretendemos negar, ha sido sin embargo magnificada en detrimento de la de otras partes del cuerpo, que tienden a aparecer como meros apéndices periféricos al servicio del cerebro. En el marco de este nuevo dualismo materialista, somos cerebros que, en lugar de en una

---

<sup>5</sup> *Ibid.*, p. 118.

<sup>6</sup> Descartes, R., *Discurso del Método*, parte IV.

cubeta, como en el famoso experimento mental de Putnam, habitan en un cuerpo físico, que llamamos nuestro y que los mantiene vivos.

Hay ciertamente razones y causas importantes que inclinan hacia esa posición. La investigación sobre las relaciones entre determinadas enfermedades psicológicas y psiquiátricas y determinadas lesiones o carencias cerebrales, el desarrollo progresivo de las neurociencias y la neuropsicología y el creciente prestigio que las acompaña, así como, en el plano opuesto, el aura de misterio que rodea el cerebro humano, entre otros factores, favorecen el auge de este nuevo dualismo.

El nuevo dualismo materialista consta de dos tesis fundamentales y conectadas entre sí. En primer lugar, cada uno de nosotros, cada ser humano, como sujeto particular de estados mentales, es esencialmente su cerebro, y sólo accidentalmente el resto de su cuerpo. En segundo lugar, los estados y procesos mentales particulares son estados neurológicos, estados del sistema nervioso central. Esta segunda tesis no se compromete con la tesis, más radical, de la identidad de tipos o propiedades mentales y neurológicas; la identidad se limita a estados o sucesos particulares, mentales y neurológicos. Sin embargo, la tesis asume que las propiedades mentales de esos estados o sucesos particulares sobrevienen a sus propiedades físicas en el sentido, antes indicado, de que dependen metafísicamente de ellas.

De la primera tesis (la identidad yo-cerebro) se desprende la inexistencia de un yo unitario, ya que el cerebro, con el que se identifica el yo, no es una entidad unitaria, sino un sistema complejo, formado por múltiples módulos y subsistemas, y funciona a través de una multitud de procesos paralelos y distribuidos, desarrollados por esos módulos y subsistemas. Así, del mismo modo que Hume no encontró el yo entre sus experiencias, ya que se supone que el yo era una entidad persistente en el tiempo mientras que las experiencias son pasajeras y cambiantes, Michael Gazzaniga,<sup>7</sup> un importante neurocientífico, no encuentra tampoco un yo unitario y persistente en el cerebro, sino sólo una diversidad de procesos, conexiones y sucesos. ¿Y qué sucede con la creencia, o mejor dicho la profunda convicción que cada uno tenemos, en la existencia de un yo unitario? Para Gazzaniga, la creencia en el mismo es una ilusión fabricada por un módulo

---

<sup>7</sup> Cf. Gazzaniga, M., *Who's in Charge? Free will and the science of the brain*, New York, HarperCollins, 2011, pp. 102-105.

cerebral que él denomina “el intérprete” y que fabula sin cesar explicaciones coherentes de nuestros actos, generando así la idea ilusoria de un sujeto único y persistente.<sup>8</sup>

Ahora bien, propiedades de nuestra imagen de sentido común como la libertad y la responsabilidad presuponen la existencia de un yo o sujeto unitario, al que se atribuyen esas propiedades. Por lo tanto, la inexistencia de un yo unitario conlleva asimismo el escepticismo sobre esas propiedades.

La segunda tesis, según la cual los estados y procesos mentales son estados neurológicos, arroja serias sospechas sobre la realidad de la causalidad mental. Desde el punto de vista del nuevo dualismo materialista, cualquier proceso o suceso mental (y el estado neurológico al que es idéntico) es precedido, aunque sea por milisegundos, y posibilitado por procesos o sucesos neurológicos de los que no somos conscientes. Así, estos procesos neurológicos son a fin de cuentas procesos físicos y están regidos por leyes que conectan propiedades físicas. Pero entonces las propiedades intencionales, como el significado o el contenido, de los estados y procesos mentales, propiedades que, en esta perspectiva, sobrevienen a sus propiedades neurológicas, comienzan a aparecer como causalmente irrelevantes. Las relaciones causales entre estados mentales tienen lugar en virtud de las propiedades físicas de esos estados, no de sus propiedades mentales. Nuestras creencias, por ejemplo, tienen propiedades mentales, como un contenido intencional, y son por tanto estados mentales, pero si causan nuestro comportamiento es en virtud de que son también estados físicos y tienen propiedades físicas (neurológicas). La eficacia causal de la mente, de la que el dualismo cartesiano no pudo dar cuenta, sigue inexplicada en el marco del materialismo y del nuevo dualismo cuerpo-cerebro. Una nueva forma de epifenomenismo amenaza de nuevo la causalidad mental. Es, como apuntamos más arriba, la venganza del dualismo cartesiano frente al monismo materialista.

Daniel Dennett expresa claramente la idea según la cual, en el marco del nuevo dualismo materialista, propiedades como el significado (o el contenido intencional) son (y han de ser) causalmente inertes:

---

<sup>8</sup> Cf. *Ibid.*.

El cerebro es, ante todo y sobre todo, un ingenio sintáctico, que se puede considerar con fruto como una imitación de un ingenio semántico, pero en el cual los significados mismos nunca anulan, predominan, y ni siquiera influyen, sobre el flujo bruto, mecanicista o sintáctico, de la causalidad local en el sistema nervioso. (Un ingenio semántico, sostengo, es una imposibilidad mecánica –como lo es una máquina de movimiento perpetuo...).<sup>9</sup>

Pero si la causalidad mental es una ilusión, una mera apariencia, ello afecta también a la realidad de propiedades como las ya indicadas: libertad, responsabilidad o agencia racional, ya que la realidad de estas propiedades presupone la realidad de la causalidad mental. Si la decisión que tomamos tras considerar razones a favor y en contra de ella no es el resultado del contenido semántico de esas razones, sino la apariencia superficial de un suceso neurológico causado a su vez por procesos neurológicos subyacentes, más allá de mi control consciente, no cabe considerar esa decisión como fruto de mi voluntad libre, ni está justificado atribuirme responsabilidad moral por ella. La creencia en la libertad sería, como ya sostuvo Spinoza, una ilusión derivada de la conciencia que tenemos de nuestros deseos, apetitos y otros estados mentales y de nuestra ignorancia de los procesos que subyacen a ellos y los causan.

En el marco de las neurociencias, la tesis según la cual las causas reales de nuestras acciones no son nuestras decisiones conscientes, sino sucesos neurológicos inconscientes, ha sido defendida por distintos autores sobre la base, supuestamente, de experimentos, entre los que destacan los realizados por Benjamin Libet,<sup>10</sup> reproducidos posteriormente por otros psicólogos y neurólogos. Estos experimentos prueban, supuestamente, que la decisión o la intención consciente de mover una parte del cuerpo está precedida, en unos 300 milisegundos, por un suceso cerebral llamado “potencial preparatorio” (que, a su vez, precede al movimiento corporal en unos 500

---

<sup>9</sup> Dennett, D, “Ways of establishing harmony”, en B. P McLaughlin (ed.), *Dretske and his Critics*, Oxford, Blackwell, 1991, pp. 118-130, p. 119

<sup>10</sup> Cf. Libet, B., “Unconscious cerebral initiative and the role of conscious will in voluntary action”, *Behavioral and Brain Sciences* 8 (1985), pp. 529-566.

milisegundos); este suceso neurológico (el potencial preparatorio), totalmente oculto a nuestra conciencia, sería la causa real *tanto* de la decisión consciente *como* del movimiento corporal. Dicho en términos más crudos: no es mi voluntad consciente, sino mi cerebro, el que toma las decisiones que luego se tornan conscientes, generando así la ilusión de que soy yo quien las toma. Pero la influencia causal de nuestros estados y actos mentales conscientes sobre nuestro comportamiento, es decir, la causalidad mental, es un supuesto y requisito de la libertad, de modo que, si los experimentos mencionados prueban realmente la inexistencia de la causalidad mental, prueban también la inexistencia de la libertad, una consecuencia que muchos teóricos e investigadores en neurociencias extraen explícitamente.

Alguien podría pensar que la libertad (y la responsabilidad moral) no se ven amenazadas por este tipo de argumento si se acepta una posición compatibilista, según la cual libertad y responsabilidad moral son compatibles con el determinismo, con la determinación causal de todo el acontecer, incluidas nuestras decisiones y acciones. Sin embargo, esto es un error, ya que para el compatibilista una decisión y una acción libres requieren un tipo determinado de causas, que incluyen sobre todo las razones que el sujeto considera, con su contenido intencional, pero no son compatibles con una causalidad puramente física o neurológica.

El nuevo dualismo materialista cuestiona también abiertamente la idea de responsabilidad entendida como merecimiento de premio o castigo. El siguiente es un texto de Richard Dawkins, biólogo y célebre autor de una famosa obra de divulgación de la teoría evolucionista y de sus consecuencias en la vida psicológica y social, *El gen egoísta*:

La retribución como principio moral es incompatible con una concepción científica de la conducta humana. Como científicos, creemos que los cerebros humanos, aunque puedan no funcionar del mismo modo que los ordenadores, están con seguridad igualmente gobernados por las leyes de la física. Cuando un ordenador funciona mal, no lo castigamos. Averiguamos cuál es el problema y lo resolvemos, normalmente sustituyendo un componente dañado... Una concepción

verdaderamente científica, mecanicista, del sistema nervioso, ¿no convierte en un sinsentido la idea misma de responsabilidad...? ... Asignar culpa y responsabilidad es un aspecto de la útil ficción de los agentes intencionales que construimos en nuestros cerebros...<sup>11</sup>

Pero entonces la profunda convicción de que actuamos en función de lo que creemos, valoramos, preferimos y decidimos, en función del contenido intencional y semántico de esos estados y actos mentales, la causalidad mental, en suma, es una ficción. Y si la intencionalidad misma es causalmente inerte, entonces, por el llamado principio de Samuel Alexander, según el cual existir implica ser causa de algo, la intencionalidad misma es también una ilusión y no existe realmente. La intencionalidad, el contenido intencional, el significado, son ficciones cerebrales. Para Dennett, por ejemplo, la creencia en la intencionalidad y el significado, y en su eficacia causal, deriva de que, aunque en realidad somos mecanismos físicos complejos, somos buenas imitaciones o remedos de máquinas semánticas, es decir, de mecanismos que, *per impossibile*, se movieran en virtud de significados, lo que nos genera la ilusión de que realmente lo somos. Lo que digo o pienso no tiene realmente significado ni refiere a nada, aunque me parezca que lo hace.

Pero en algún punto un filósofo ha de detenerse, reconocer que determinadas consecuencias son obviamente falsas o absurdas y preguntarse en qué punto del camino que conduce a ellas ha cometido un error.

Creo que las dos tesis principales del nuevo dualismo materialista son falsas. El yo unitario no es una ilusión; su carácter unitario no es una construcción ilusoria del cerebro, sino que es función de la unidad de acción de un organismo corpóreo completo que interactúa como un todo con su entorno. La unidad de nuestro yo es nuestra unidad como agentes psicofísicos, no como cerebros. Por lo que respecta a la segunda tesis, los estados mentales no son (numéricamente) idénticos a estados neurofisiológicos. Podemos sustentar esta afirmación en el principio leibniziano según el cual los idénticos son indiscernibles, en el sentido de que comparten todas sus propiedades. Los estados

---

<sup>11</sup> Dawkins, R., "Let's all stop beating Basil's car", en [www.edge.org/q2006](http://www.edge.org/q2006).

del cerebro como tales no tienen intencionalidad; no refieren a nada ni son acerca de nada. Los estados mentales intencionales, en cambio, lo son: refieren esencialmente a circunstancias, personas, objetos y situaciones externas a ellos. Y podemos tener estados con una referencia externa constitutiva (estados intencionales) en la medida en que somos agentes corpóreos, que interactúan con su entorno natural y social. Los estados mentales no son estados del cerebro, sino de un agente corpóreo entero. La intencionalidad no es una ilusión, ni una ficción del cerebro. Considero también falso el supuesto de la superveniencia de las propiedades mentales respecto de las propiedades neurológicas. Las teorías externistas sobre el significado y sobre el contenido<sup>12</sup>, en la estela de Putnam y Burge, según las cuales el significado de nuestro lenguaje y el contenido intencional de nuestros estados mentales dependen parcialmente del entorno natural y social con el que interactuamos, hablan claramente a favor de esa conexión entre intencionalidad, corporeidad y acción, y en contra de la reducción, e incluso de la superveniencia, de lo mental a lo cerebral. Podemos expresar esta idea mediante la famosa afirmación de Putnam según la cual “los significados no están en la cabeza”. Por eso no somos nuestros cerebros; el cerebro, con sus procesos neuronales paralelos y distribuidos, está al servicio de la continuidad y el florecimiento del organismo entero del que forma parte. Y esto es también lo que cabría esperar de la teoría de la evolución, de modo que esta teoría no favorece tampoco el nuevo dualismo materialista.

Naturalmente, estas consideraciones no constituyen una respuesta acabada al problema de la causalidad mental ni al nuevo dualismo materialista; pero apuntan en la dirección en que dicha respuesta podría plausiblemente ser elaborada. Del mismo modo que Kant no preguntó si la ciencia era posible, sino cómo era posible, creo que nuestra tarea no es tampoco preguntar si la causalidad mental es posible, sino cómo es posible.

---

<sup>12</sup> Véase H. Putnam, “The meaning of ‘meaning’”, en *Philosophical Papers, vol. II: Mind, Language and Reality*, Cambridge, Cambridge University Press, 1975, pp. 215-71 y T. Burge, “Individualism and the Mental”, en P. A. French y otros (eds.), *Midwest Studies in Philosophy, vol. 4, Studies in Metaphysics*, Minneapolis, University of Minnesota Press, 1979, pp. 73-121.