

DONALD DAVIDSON: EL MONISMO ANÓMALO

Por Carlos J. Moya

1. Teorías filosóficas de la mente



Composition with Red, Yellow, and Blue (Piet Mondrian, 1935-1942)

El problema de la naturaleza de la mente y su relación con el mundo físico ha despertado el interés de los filósofos al menos desde los tiempos de Platón. El referente principal para la reflexión moderna y contemporánea sobre este venerable problema es, sin embargo, Descartes, quien funda el dualismo, ya propuesto por Platón, sobre nuevas bases. El dualismo cartesiano descansa sobre la indubitabilidad de los enunciados en que un sujeto se atribuye a sí mismo, con sinceridad, un estado mental, como sucede, paradigmáticamente, en el enunciado "yo pienso". Mientras que un enunciado como "yo pienso (creo) que estoy corriendo" es indudablemente verdadero en cada momento en que lo concibo o lo emito sinceramente, un enunciado como "yo estoy corriendo", en el que me atribuyo un estado físico, puede ser falso (si, por ejemplo, sueño que estoy corriendo o sufro una alucinación). Para Descartes, existen entidades y eventos particulares no físicos: el yo que piensa y sus pensamientos (sus estados mentales conscientes). El dualismo cartesiano es una concepción fuertemente antirreductiva de la mente: lo mental es esencial y radicalmente distinto de lo físico. Como es sabido, el principal escollo del dualismo de sustancias o entidades cartesiano es el problema de la interacción psicofísica. En el marco del dualismo cartesiano, la verdad de un enunciado tan simple como "entré en la librería porque deseaba comprar *El hereje* de Delibes" se convierte en un misterio. El ocasionalismo de Malebranche o la teoría

de la armonía preestablecida de Leibniz son intentos fallidos de resolver este problema sobre el supuesto de que no existe interacción causal real entre lo mental y lo físico.

La oposición al dualismo ha caracterizado la reflexión contemporánea sobre la mente. El siglo XX ha sido testigo del nacimiento de diversas teorías materialistas de carácter reductivo. Para el conductismo lógico, defendido entre otros autores por R. Carnap y Carl G. Hempel, los términos que expresan propiedades mentales, como "creer", "desear" o "sentir dolor", pueden definirse, sin pérdida de significado, en términos de comportamiento físico efectivo o de disposiciones para dicho comportamiento, de modo que un análisis correcto del vocabulario psicológico no contendría, idealmente, ningún término o expresión psicológica. Sobre esta base, habría de ser posible sustituir, en una oración, todas las expresiones de carácter mental que aparezcan en ella por expresiones relativas a la conducta física observable y obtener así una oración semántica y lógicamente equivalente a la primera. Así, el lenguaje psicológico sería reductible a un lenguaje puramente físico y la psicología sería en principio reductible a las ciencias físicas. Sin embargo, hay razones para pensar que dicha empresa está condenada al fracaso: siempre cabe concebir situaciones en que una oración psicológica es verdadera y la oración física supuestamente equivalente a ella es falsa o viceversa, de modo que no hay realmente equivalencia lógica (y mucho menos semántica) entre oraciones de ambos tipos. Dicho de otro modo: el conductismo lógico parece sencillamente falso y, al menos en sus crudas formulaciones iniciales, es hoy considerado como una propuesta teórica

inviabile.

Hacia los años sesenta del pasado siglo surgió con fuerza, de la mano de autores como U. T. Place y J. J. Smart, una nueva teoría materialista, conocida en esos momentos como "materialismo del estado central" y mejor conocida hoy como "teoría de la identidad de propiedades". Sus proponentes aceptan desde el principio que las expresiones y oraciones psicológicas no tienen el mismo significado que las expresiones y oraciones en lenguaje físico, de modo que renuncian gustosamente al programa del conductismo lógico. La reducción de lo mental a lo físico propuesta por esta nueva forma de monismo materialista no descansaba en definiciones o análisis semánticos. La hipótesis central de esta teoría consistía en la identidad fáctica de las propiedades psicológicas con propiedades neurofisiológicas. Esta hipótesis no podía ser confirmada mediante análisis semánticos, sino sólo mediante la investigación científica en psicología y neurofisiología. Se trataba de hallar leyes verdaderas que establecieran la identidad fáctica de cada propiedad psicológica con una determinada propiedad neurofisiológica, de modo que cualquier organismo que poseyera esta última poseería también la primera y viceversa. Así, del mismo modo que la ciencia ha descubierto que los genes son en realidad fragmentos de ADN, que el agua es en realidad un agregado de moléculas de H₂O, o que la temperatura de un fluido es la energía cinética media de las moléculas que lo componen, la ciencia podría llegar a descubrir que la experiencia de dolor es en realidad la excitación de un determinado conjunto de fibras nerviosas o que la creencia de que la Tierra es redonda es en realidad una determinada configuración neuronal en el cerebro. Desde esta perspectiva, la interacción causal psicofísica deja de ser un misterio: la mente puede influir causalmente en el mundo físico porque ella misma es parte del mundo físico.

Sin embargo, si la hipótesis de la identidad de propiedades mentales y físicas es realmente una hipótesis empírica, hay importantes razones empíricas para pensar que es falsa. Una misma propiedad mental podría tener realizaciones físicas diversas en distintas especies animales, e incluso, en virtud de la plasticidad del cerebro y el sistema nervioso, en distintos individuos de una misma especie. Por otra parte, si hay en el universo seres vivos con una organización física distinta de la nuestra, y con un sistema de control interno muy distinto de nuestro sistema nervioso, la teoría tendría que negarles propiedades mentales, aun cuando mostraran claros signos de inteligencia y sensibilidad. Las propiedades mentales pueden tener bases físicas muy diversas y ser así multirrealizables desde el punto de vista físico, lo cual es un escollo probablemente insalvable para la reducción de la psicología a la neurofisiología.

La reflexión sobre las dificultades de la teoría de la identidad de propiedades, así como el creciente interés de los filósofos en la inteligencia artificial, se hallan en la base de una concepción de la mente que comienza a perfilarse en algunos trabajos de H. Putnam a finales de los años sesenta: el funcionalismo. De acuerdo con esta concepción, hallarse en un estado mental de cierto tipo (digamos, tener un determinado deseo) es hallarse en un estado capaz de desempeñar un determinado papel causal: un estado que posee causas y efectos típicos, que incluyen otros estados mentales. Así, por ejemplo, la creencia de que está lloviendo es un estado causado típicamente por la percepción de la lluvia y que causa a su vez, en unión con el deseo de no mojarse, la decisión de coger el paraguas y la conducta correspondiente. Desde este punto de vista, sistemas físicamente muy distintos pueden compartir esa creencia, a condición de que se hallen en un estado capaz de desempeñar ese papel causal. Por ejemplo, un robot, adecuadamente programado, podría tener la creencia de que está lloviendo, puesto que tener esa creencia no es más que hallarse en un estado con las relaciones y disposiciones causales apropiadas. En la versión del funcionalismo inspirada por la inteligencia artificial (el funcionalismo computacional), un ordenador es un modelo de la mente humana y de su relación con el cuerpo. Ordenadores físicamente distintos pueden ser funcionalmente equivalentes si son capaces de desempeñar las mismas tareas. Así, de dos ordenadores distintos puede decirse que están, digamos, sumando 55 más 67 si, tras la introducción y procesamiento de

estos signos, generan como resultado el signo "122", no importa cuán diferentes puedan ser desde el punto de vista físico. La respuesta del funcionalismo computacional al venerable problema de la relación entre la mente y el cuerpo es provocadoramente simple: dicha relación puede entenderse adecuadamente como la relación entre el programa lógico (el *software*) y el soporte físico (el *hardware*) en un ordenador digital. Si esta relación no nos parece misteriosa, tampoco debería parecernos la relación entre la mente y el cuerpo.

Frente a la opinión de algunos de sus partidarios, el funcionalismo es una teoría reductiva de la mente. Para esta concepción, y a diferencia del conductismo lógico y la teoría de la identidad, las propiedades mentales no son propiedades del comportamiento ni propiedades neurofisiológicas, sino propiedades funcionales. Pero ésta es una tesis reductiva: de acuerdo con ella, tener una mente no es más que tener una determinada organización funcional y hallarse en un estado mental de cierto tipo no es más que hallarse en un estado funcional determinado. No hay ninguna característica propia de lo mental de la que no se pueda dar cuenta en términos de características funcionales. La tarea de la psicología, desde este punto de vista, consiste en averiguar la organización funcional (el "programa") de determinados sistemas físicos complejos, entre los que se encuentran los seres humanos.

Si, como hemos indicado, la teoría de la identidad de propiedades es indebidamente restrictiva en la atribución de propiedades mentales, el funcionalismo parece, por el contrario, demasiado liberal. De acuerdo con esta teoría, hay demasiadas entidades dotadas de propiedades mentales que, intuitivamente, no las poseen. Por mi parte, no siento ninguna inclinación a atribuir una mente a los ordenadores. Me parecen simples máquinas, en el sentido en que Descartes (falsamente, en mi opinión) consideraba simples máquinas a los animales no humanos. Y la razón es la misma: no hay (más cautamente, no parece haber) en ellos experiencias conscientes. En realidad, el hecho de la experiencia consciente, de la conciencia, es el gran escollo con que tropieza el funcionalismo. Pero este hecho es precisamente lo que, para Descartes, constituía el rasgo distintivo de lo mental. Tampoco el conductismo lógico es capaz de dar cuenta de este rasgo. En cuanto a la teoría de la identidad, se halla en principio más acorde con nuestra tendencia natural a atribuir conciencia únicamente a seres con una naturaleza orgánica determinada, aun cuando también esta teoría ha de ver el hecho de que tales seres poseen experiencias conscientes como un hecho bruto e inexplicado. En realidad, en el ámbito de la filosofía de la mente, el problema de la conciencia sigue siendo el más recalcitrante frente a todos los intentos de resolverlo. Y el monismo anómalo no es una excepción.

2. El monismo anómalo

Como teoría filosófica de la mente, el monismo anómalo davidsoniano se concibe a sí mismo como una alternativa a las teorías que hemos descrito en el apartado anterior. A diferencia del dualismo, rechaza la existencia de sustancias (entidades o eventos particulares) no físicas y comparte así la inspiración materialista del conductismo lógico, la teoría de la identidad de propiedades y el funcionalismo. Sin embargo, se aparta drásticamente de estas últimas teorías al rechazar de plano la orientación reductiva que las caracteriza. Aun cuando, de acuerdo con el monismo anómalo, los eventos mentales particulares sean eventos físicos bajo otras descripciones, los conceptos y propiedades mentales no son reductibles o idénticas a propiedades de otro tipo, sean conductuales, neurofisiológicas o funcionales, de modo que no es posible hallar equivalencias semánticas o nomológicas entre conceptos o propiedades mentales y conceptos o propiedades físicas. Según el monismo anómalo, no existen leyes estrictas, de carácter psicofísico o psicológico, sobre cuya base pudiera predecirse la vida mental de un sujeto (así, pues, creo que la expresión "monismo anómico" habría sido más exacta y menos confundente para caracterizar la concepción defendida por Davidson). El monismo anómalo es, así, una forma de materialismo no reductivo acerca de lo mental. Dos

tesis fundamentales constituyen el núcleo de esta teoría. En primer lugar, la tesis monista según la cual cada evento mental particular es un evento físico particular, presumiblemente un evento neurofisiológico. En segundo lugar, la tesis antirreductiva del anomalismo de lo mental, según la cual las propiedades que hacen de un evento particular un evento mental no están conectadas mediante leyes estrictas con las propiedades físicas de dicho evento, ni pueden conectarse entre sí mediante leyes estrictas de carácter puramente psicológico. Aun cuando Davidson reconoce la existencia de generalizaciones verdaderas de carácter psicofísico y psicológico, sostiene que estas generalizaciones no son susceptibles del refinamiento progresivo que pudiera dotarlas del poder predictivo propio de las leyes de las ciencias físico-naturales.



Composition
(Piet Mondrian, 1938-39)

La tesis monista de la identidad entre eventos particulares (no entre propiedades o tipos de eventos) mentales y físicos deriva de tres premisas fundamentales. La primera es la existencia de interacciones causales entre eventos y sucesos mentales y físicos (Principio de Interacción Causal psicofísica). Esta premisa es ciertamente plausible. De hecho, es la plausibilidad de esta premisa lo que hace de la interacción entre el alma y el cuerpo un escollo de primer orden para el dualismo cartesiano. Sin embargo, Davidson tuvo que defender esta premisa, en sus ensayos sobre filosofía de la acción, frente a aquellos filósofos, influidos por la obra de Wittgenstein, según los cuales las explicaciones de la acción humana en términos de razones (creencias, deseos, propósitos, etc.) no eran explicaciones causales. La segunda premisa es el principio, de inspiración humeana, según el cual toda relación causal es un caso particular de una ley general estricta (Principio del Carácter Nomológico de la Causalidad). La verdad de esta segunda premisa no es evidente. Sin embargo, no encontramos en Davidson una defensa detallada de la misma, lo cual ha sido objeto justificado de crítica. Tal vez Davidson pensó que Hume y sus seguidores habían ya proporcionado argumentos suficientes en favor de su verdad. De estas dos premisas parece deducirse que existen leyes estrictas psicofísicas que subyacen a las relaciones causales entre eventos mentales y físicos. Sin embargo, la tercera premisa niega esta conclusión y sostiene que no existen tales leyes estrictas de carácter psicofísico (Principio del Anomalismo de lo Mental). Si añadimos a esta premisa la inexistencia de leyes psicológicas estrictas, tenemos de hecho la segunda tesis central del monismo anómalo. Esta es una premisa controvertida. Sería sin duda cuestionada cuando menos por los materialistas reductivos, los psicofisiólogos y los neuropsicólogos, y sería asimismo contemplada con sospecha por quienes piensan que la psicología es una ciencia nomológica rigurosa, comparable a las ciencias físico-naturales. Sin embargo, dejaremos de lado, por el momento, las razones que Davidson esgrime en defensa de la misma y aceptaremos provisionalmente su verdad por mor del argumento. Así, esta premisa, en unión con las dos primeras, lleva a Davidson a sostener que las leyes que subyacen a las relaciones causales entre eventos particulares mentales y físicos habrán de ser leyes físicas. Y si estas leyes han de poder fundamentar las relaciones causales psicofísicas, los eventos mentales involucrados en estas relaciones (creencias, deseos, intenciones particulares de un sujeto en un momento

La segunda tesis central del monismo anómalo, el carácter anómalo de lo mental, que sustenta el carácter fuertemente antirreductivo de esta teoría, coincide, como hemos indicado, con la tercera premisa, la inexistencia de leyes psicofísicas, si le añadimos la inexistencia de leyes psicológicas. La principal razón por la que Davidson sostiene que no existen leyes psicológicas estrictas deriva del hecho de que el ámbito de lo psicológico, a diferencia del ámbito de lo físico, no es causalmente cerrado. Muchos sucesos de carácter no psicológico (como la presencia de objetos físicos que estimulan nuestros órganos sensoriales) tienen efectos psicológicos (como estados perceptivos y creencias sobre la presencia de tales objetos). Pero si el ámbito de lo psicológico no es causalmente cerrado, no podemos esperar que haya leyes psicológicas con un poder predictivo comparable al de las leyes físicas. Esta deficiencia predictiva podría subsanarse, al menos en principio, si hubiera leyes psicofísicas que conectasen los conceptos psicológicos con los físicos, como esperaban los materialistas reductivos. Pero, como ya hemos indicado, Davidson rechaza, no ya la mera existencia, sino la posibilidad misma de tales leyes. Atendamos ahora a las razones de este rechazo.

Estas razones descansan en una determinada concepción de lo mental, muy distinta de la que movió a Descartes a sostener la radical diferencia entre la *res cogitans* y la *res extensa*. Así, mientras que este último considera la conciencia subjetiva de nuestros propios pensamientos, con la certeza, inmediatez e infalibilidad que le son propias, como el rasgo definitorio de lo mental y hace reposar sobre él su esencial diferencia frente a lo físico, Davidson, siguiendo a Brentano, considera como característica esencial y distintiva de los estados mentales la intencionalidad, la referencia a un contenido u objeto: "... El rasgo distintivo de lo mental no es que sea privado, subjetivo o inmaterial, sino que exhibe lo que Brentano llamó intencionalidad" (1982: 211).

No es que el fenómeno de la intencionalidad fuese desconocido para Descartes. El filósofo francés sabía perfectamente que todo *cogitare* conlleva un *cogitatum*. Lo pensado, el objeto del pensamiento, sin embargo, era para Descartes interno a la mente, y como tal inmediata e infaliblemente presente a la conciencia, de modo que la inexistencia externa de la situación u objeto pensado, e incluso la inexistencia del mundo físico mismo, no tenían consecuencia alguna para la naturaleza y existencia del objeto intencional del pensamiento. La intencionalidad en Descartes, podríamos decir, es secundaria con respecto a la conciencia y queda absorbida en ésta. Para Davidson, por el contrario, la intencionalidad es una característica primaria de la mente, y la conciencia ha de entenderse, en todo caso, sobre la base de aquélla. Asimismo, Davidson rechaza de plano la idea según la cual los objetos intencionales de nuestros pensamientos son internos a la mente misma. Para Davidson, que la mente es esencialmente intencional significa que está esencialmente abierta al mundo, a lo que no es ella misma. Aquello en lo que pensamos no es una entidad distinta de los objetos y situaciones comunes, sino estos objetos y situaciones mismas. Existe una estrecha conexión entre la intencionalidad de la mente y el significado del lenguaje. En ambos casos hay una referencia a fenómenos públicos, abiertos a la comprensión e interpretación intersubjetivas. La comprensión del significado de las expresiones de otras personas y la atribución de creencias y otros estados mentales son dos aspectos inseparables del proceso por el que interpretamos y explicamos el comportamiento de nuestros semejantes. Esta perspectiva objetiva, de tercera persona, desde la que Davidson se aproxima a la cuestión de la naturaleza de la mente, y que comparte con muchos otros pensadores contemporáneos, contrasta con la perspectiva subjetiva, de primera persona, que caracteriza la aproximación cartesiana. Como veremos, las ventajas de esa perspectiva no dejan de tener un precio.

La cuestión que nos ocupa ahora es la relación entre la intencionalidad y la segunda tesis central del monismo anómalo: la inexistencia de leyes estrictas, en especial de carácter psicofísico, y la correspondiente irreductibilidad de los conceptos y propiedades mentales a conceptos y propiedades físicas. La irreductibilidad de lo mental a lo físico se debe, para

Davidson, a la irreductibilidad de la intencionalidad como rasgo esencial de lo mental. No todos los filósofos de la mente aceptarían esta irreductibilidad. El programa de naturalización de la intencionalidad, promovido en la actualidad por pensadores como J. A. Fodor, F. Dretske o R. Millikan, trata de establecer condiciones al menos suficientes, de carácter no intencional, para la ejemplificación de propiedades intencionales y semánticas. En el marco de este programa, se trata de explicar la intencionalidad y el significado a partir del proceso de evolución biológica o de vínculos de carácter nómico entre eventos físicos y fisiológicos. Aun cuando buenas dosis de ingenio han sido empleadas en el desarrollo de este programa de inspiración reductiva, hay escollos hasta ahora insuperables para su plena realización. Estos escollos guardan estrecha conexión con la normatividad propia de la posesión de estados intencionales: nociones normativas como las de justificación, racionalidad, error o corrección parecen a la vez necesarias para la comprensión de la intencionalidad e incapaces de ser explicadas en términos no intencionales. De hecho, la normatividad de lo intencional es, para Davidson, una razón conceptual para la imposibilidad de establecer leyes psicofísicas reductivas. La atribución de actitudes intencionales, de creencias, deseos, propósitos, etc., en el proceso de interpretación del lenguaje y el comportamiento no verbal está regida necesariamente por lo que Davidson denomina el "ideal constitutivo de la racionalidad" (1982: 223). Este principio no desempeña papel alguno en la atribución de propiedades físicas, y ésta es una razón fundamental por la que no podemos esperar hallar leyes psicofísicas estrictas. El siguiente texto de Davidson expone con claridad esta idea central:

Mi estrategia general para tratar de mostrar que no hay leyes psicofísicas estrictas depende, en primer lugar, del énfasis en el carácter holista del ámbito cognitivo. Cualquier esfuerzo dirigido a incrementar la exactitud y el poder de una teoría de la conducta nos fuerza a tomar en cuenta directamente zonas cada vez más amplias del sistema entero de creencias y motivos del agente. Pero al inferir este sistema a partir de los datos, imponemos necesariamente condiciones de coherencia, racionalidad y consistencia. Estas condiciones no tienen eco alguno en la teoría física, y ésta es la razón por la que no podemos esperar hallar algo más que vagas correlaciones entre fenómenos psicológicos y físicos (1982: 231).

En los fundamentos de la tesis de la inexistencia de leyes psicofísicas estrictas convergen las investigaciones de Davidson en los campos de la filosofía de la acción, la filosofía de la mente y la filosofía del lenguaje. La explicación racional de la acción requiere la atribución al agente de estados intencionales, creencias y deseos, apropiadamente descritos y conectados conceptual y causalmente con la acción que tratamos de explicar. En esta tarea, la acción, si ha de ser realmente explicada, ha de aparecer como racionalmente justificada a la luz de los estados intencionales del agente, además de ser causada por tales estados. De este modo, el supuesto de la coherencia y racionalidad básica del agente preside ya la tarea de explicar y entender sus acciones particulares. Davidson advierte progresivamente que la explicación de la acción en términos de razones es en realidad un aspecto de un proceso más general del que forma parte la comprensión de las emisiones lingüísticas de las personas. Esta comprensión nos proporciona claves importantes para nuestra atribución a las personas de estados intencionales que den cuenta de sus acciones, y esta atribución a su vez nos proporciona claves esenciales para la comprensión del significado de sus emisiones. Estados mentales y significados son atribuidos conjuntamente y en mutua dependencia bajo el supuesto general de la racionalidad, teniendo en cuenta también las claves que nos proporciona el comportamiento, de modo que los estados mentales se atribuyen siempre de modo global y no atomista, en grandes bloques, y no uno por uno. La atribución de una creencia o un deseo sólo es inteligible sobre el trasfondo de otras creencias y deseos que implícitamente atribuimos y suponemos en el agente. Este carácter holista del proceso de interpretación, regido por el supuesto general de coherencia racional (el "ideal constitutivo de la racionalidad"), se pone de manifiesto cada vez con más claridad a medida que segmentos más amplios y pautas más generales del comportamiento de una persona van siendo objeto de nuestro interés, así como en los casos

en que las acciones ajenas, o las propias, se tornan opacas, extrañas o incomprensibles. En estos casos, nos vemos llevados a revisar nuestras atribuciones previas de creencias, deseos, intenciones y valores a esas personas y a replantear nuestra interpretación. Así, "cuando usamos los conceptos de creencia, deseo, y demás, hemos de estar preparados, a medida que los datos se acumulan, a ajustar nuestra teoría a la luz de consideraciones de consistencia global: el ideal constitutivo de la racionalidad controla parcialmente cada fase del desarrollo de lo que ha de ser una teoría en evolución" (Davidson 1982: 223). La adscripción a los seres humanos de un alto grado de consistencia y racionalidad en su vida mental y en sus acciones no es una máxima metodológica opcional, sino un requisito indispensable para la interpretación de sus acciones a la luz de creencias, deseos y otros estados intencionales. Así, la concepción davidsoniana de la intencionalidad, y por tanto de la naturaleza de lo mental, está presidida por la noción general de racionalidad que rige el proceso de interpretación. Y una consecuencia de esta aproximación a la mente es una concepción holista de la misma: la naturaleza y contenido de cada estado mental depende de su lugar en el contexto de una teoría global acerca de sus creencias, deseos y conceptos, así como del significado de sus palabras. Como escribe Davidson:

No cabe asignar creencias a una persona una a una sobre la base de su conducta verbal, sus elecciones, u otros signos locales, por claros y evidentes que sean, puesto que damos sentido a creencias particulares sólo en la medida en que son coherentes con otras creencias, con preferencias, intenciones, esperanzas, miedos, expectativas y demás... El contenido de una actitud proposicional deriva de su lugar en esa pauta global (1982: 221).

Estas restricciones no se aplican a la atribución de propiedades físicas a sucesos con vistas a explicarlos. No se trata, obviamente, de que la teorías físicas no estén regidas por normas de racionalidad y coherencia, sino de que no necesitan (en realidad, no deben, si desean tener éxito en su tarea explicativa y predictiva) adscribir el seguimiento de estas normas a los objetos y sucesos físicos que estudian para explicar su comportamiento y relaciones causales con otros objetos y sucesos. Esta diferencia entre la atribución de propiedades físicas y mentales, y entre la explicación de los sucesos físicos y la interpretación del comportamiento intencional, es, pues, el principal fundamento de la segunda tesis central del monismo anómalo, la inexistencia de leyes psicofísicas y la correspondiente irreductibilidad de las propiedades mentales a propiedades físicas. Como Davidson lo expresa sucintamente:

Es un rasgo de la realidad física que el cambio físico puede ser explicado por leyes que lo conectan con otros cambios y condiciones descritos en términos físicos. Es un rasgo de lo mental que la atribución de de fenómenos mentales ha de responder al trasfondo de razones, creencias e intenciones del individuo. No puede haber conexiones estrechas entre ambos campos si cada uno ha de ser fiel a su propia fuente de evidencias (1982: 222).

Así concluimos la exposición del monismo anómalo y de los argumentos que sustentan sus tesis centrales. En el siguiente apartado analizaremos algunas de sus dificultades. Sin embargo, antes de pasar a ello, quisiera llamar la atención sobre un rasgo de la filosofía de Davidson que no habrá pasado desapercibido a algunos lectores: su cercanía a una concepción hermenéutica de las ciencias humanas y sociales. Es notable hallar, en este representante de la filosofía analítica anglosajona, ecos claramente perceptibles de la antigua polémica decimonónica entre el *Verstehen* y el *Erklären*, entre los defensores del carácter fundamentalmente nomológico de la explicación en las ciencias humanas y sociales y los defensores de la naturaleza esencialmente interpretativa de estas ciencias. Y no menos notable es hallar a Davidson, discípulo de Quine (un filósofo de orientación claramente empirista y cientificista), tan cerca del segundo grupo. Notable pero no del todo sorprendente, porque la concepción davidsoniana de la interpretación debe mucho a algunos aspectos de la obra de Quine, en especial a sus observaciones sobre el significado, los conceptos intencionales

y el proceso de traducción.

3. Tensiones internas

El monismo anómalo tiene un indudable atractivo. Permite conciliar una concepción naturalista y materialista de los seres humanos como organismos complejos, resultantes del proceso de evolución biológica y sometidos como éste a leyes físico-químicas, y una concepción humanista según la cual los seres humanos poseen características que los distinguen radicalmente de otros seres vivos, como la racionalidad, la libertad y la responsabilidad moral, de modo que su conducta intencional ha de ser entendida y explicada en términos muy distintos de los que usamos para entender y explicar los procesos físicos, e incluso la conducta animal. Conceptos como razones, decisiones, valores, propósitos, más que las leyes causales físico-naturales, parecen los apropiados cuando se trata de dar cuenta de las acciones conscientes y voluntarias de los seres humanos. Ambas concepciones, naturalista y humanista, nos resultan convincentes, pero, ¿cómo pueden ser ambas verdaderas? ¿Cómo puede uno y el mismo ser estar sometido a la necesidad natural y no estar sometido a ella? El monismo anómalo pretende disipar la apariencia de contradicción entre ambas concepciones. Como Kant, a quien se refiere explícitamente, Davidson trata de mostrar que no existe realmente contradicción entre la necesidad natural y la libertad. Nuestras creencias, deseos, intenciones y decisiones son eventos físicos, son parte del mundo natural, y no sucesos inmateriales carentes de propiedades físicas. Así, pueden interactuar causalmente con la realidad física sin que esa interacción involucre los misterios asociados al dualismo substancial cartesiano. Sin embargo, en tanto que poseen propiedades mentales, en tanto que son descritos con verdad como actitudes psicológicas dotadas de contenido, no pueden ser subsumidos bajo leyes estrictas y escapan así al nexa nomológico que rige las relaciones causales en el mundo natural. Este mismo espíritu de conciliación preside la defensa davidsoniana de la tesis según la cual las razones que justifican y explican racionalmente nuestras acciones son al mismo tiempo causas de dichas acciones. Así, frente a la afirmación de diversos filósofos según la cual sólo las explicaciones en términos de razones, y no las explicaciones causales, son apropiadas para dar cuenta de la acción intencional humana, Davidson sostiene que no hemos de optar realmente entre unas y otras, sino que las explicaciones racionales de la acción son al mismo tiempo explicaciones causales.

Sin embargo, la articulación y la estructura argumentativa de esta atractiva propuesta metafísica de conciliación entre materialismo y mentalismo, entre naturalismo y humanismo, no está exenta de fuertes tensiones internas. Veamos alguna de ellas.

En primer lugar, una dificultad frecuentemente señalada por los críticos concierne a la respuesta del monismo anómalo al problema de la interacción psicofísica. Aun cuando el monismo anómalo permite entender cómo un evento mental puede causar un evento físico (y viceversa), es dudoso que permita entender cómo aquello que hace de un determinado evento un evento mental, a saber, el hecho de que se trata de una actitud psicológica (una creencia, un deseo, etc.) hacia un determinado contenido intencional (aquello que se cree, se desea, etc.), puede tener alguna importancia en su eficacia causal sobre el mundo físico, y en particular sobre el cuerpo del sujeto. Esta dificultad puede plantearse como sigue. En primer lugar, parece correcto suponer que no todas las propiedades de un evento u objeto particular son pertinentes para dar cuenta de sus relaciones causales. Supongamos que depositamos una manzana roja, procedente de La Rioja, en el plato de una balanza y, como consecuencia de ello, el indicador de la balanza se desplaza hasta la señal de los 200 gramos. Este es un caso claro de relación causal. Pero parece asimismo claro que el color de la manzana o su procedencia no tienen importancia para el efecto que tiene lugar al depositarla en el plato, a saber, el movimiento del indicador. Depositar sobre el plato una manzana verde, procedente de El Bierzo, con la misma masa (o peso) habría tenido exactamente el mismo

efecto, y depositar una manzana roja y procedente de La Rioja, pero de menor masa, habría tenido un efecto diferente. Podemos decir que el color y la procedencia de la manzana son, con respecto a este efecto, propiedades epifenoménicas o causalmente inertes. Pensemos ahora en la concepción nomológica de la causalidad que constituye una premisa del monismo anómalo, según la cual toda relación causal entre eventos es un caso particular de una ley general estricta. Una ley, a su vez, puede entenderse plausiblemente como un enunciado universal que conecta propiedades, o bien clases de eventos en tanto que poseen dichas propiedades. En este contexto, el hecho de que el color y la procedencia de la manzana sean propiedades epifenoménicas puede explicarse diciendo que no existen leyes que conecten tales propiedades con el movimiento del indicador en el marco de una explicación física verdadera. La masa, en cambio, es causalmente importante porque, presumiblemente, hay leyes físicas que conectan la masa con el movimiento y permiten explicar la relación causal en cuestión.

Pensemos ahora en la relación entre mi deseo de comprar manzanas y mi acción de entrar en una frutería para conseguirlo. Esta relación es un caso de interacción psicofísica, puesto que mi acción de entrar en la frutería involucra el movimiento físico de mi cuerpo. Estoy convencido de que mi deseo de comprar manzanas es lo que explica que entre en la frutería. El monismo anómalo puede dar cuenta de la relación causal entre estos dos eventos. Sin embargo, según esta teoría, no existen leyes que conecten propiedades mentales con propiedades físicas, o propiedades mentales entre sí. Así, en el marco del monismo anómalo, las propiedades mentales de mi deseo, el hecho de que se trate de un deseo con cierto contenido, aparecen como tan prescindibles, desde el punto de vista causal, como el color o la procedencia de la manzana en el ejemplo anterior. Sólo las propiedades neurofisiológicas y físicas del deseo, sólo el hecho de que constituya un estado determinado del cerebro, parecen causalmente importantes para dar cuenta de la acción, puesto que sólo éstas pueden figurar en leyes estrictas. A su vez, sólo las características físicas de ésta como mero movimiento físico, y no su carácter de acción voluntaria e intencional, parecen importantes en el marco de su relación causal con el deseo. Que el monismo anómalo dé lugar a la misma conclusión acerca de las propiedades mentales de nuestros deseos y creencias que acerca de propiedades como el color y la procedencia de la manzana en el ejemplo anterior contradice fuertemente nuestras convicciones, puesto que, mientras que vemos con claridad que las últimas son causalmente inertes en ese contexto, estamos en cambio firmemente convencidos de que el contenido de mi deseo fue decisivo para llevarme a entrar en la frutería. En suma, la respuesta del monismo anómalo al problema de la interacción psicofísica no resulta satisfactoria y no corresponde a nuestras convicciones sobre la eficacia causal de nuestros deseos, creencias y decisiones.

Un modo de evitar estas consecuencias sería rechazar la concepción nomológica de la causalidad. Alternativamente, cabría rechazar el anomalismo de lo mental y defender la existencia de leyes psicofísicas. Sin embargo, cualquiera de estas dos alternativas privaría a la teoría de un apoyo esencial para mantenerse en pie, puesto que, como hemos visto, la concepción nomológica de la causalidad y el anomalismo de lo mental son premisas centrales del monismo anómalo.

Además del escollo del epifenomenismo, el monismo anómalo se enfrenta, en mi opinión, con una segunda dificultad, relacionada con la tesis monista de la identidad de cada evento mental particular con un determinado evento físico (presumiblemente neurofisiológico). La dificultad deriva del modo en que lo mental es concebido en el seno de esta teoría. Recordemos que, para Davidson, la atribución de propiedades mentales, en el marco de la interpretación del lenguaje y del comportamiento, está regida por el ideal constitutivo de la racionalidad. La tesis según la cual la racionalidad es constitutiva de lo mental deriva de la reflexión davidsoniana sobre las condiciones de posibilidad de la interpretación y esta reflexión informa

asimismo su concepción de la mente. Los estados intencionales, creencias, deseos, etc., son concebidos en el marco de este proceso interpretativo: son estados que atribuimos a las personas para interpretar su lenguaje y entender y explicar su conducta. Ahora bien, sólo si adscribimos un alto grado de racionalidad al sujeto que tratamos de interpretar podremos llevar a cabo inteligiblemente esta tarea. Una importante consecuencia de todo ello es una concepción fuertemente holista de las actitudes intencionales. Según esta concepción, como vimos, el contenido de, digamos, una creencia de un sujeto depende de su lugar en el marco global de su psicología y de sus relaciones con el resto de sus creencias, deseos, esperanzas y otras actitudes intencionales. Así, un cambio en alguna de las creencias, deseos o valores de una persona supone reajustes y modificaciones, sin límites claros, en el resto de sus estados mentales, bajo la presión del ideal constitutivo de racionalidad. Pero de esta concepción fuertemente holista de la mente se desprende que los estados intencionales poseen condiciones de individuación e identidad muy difusas. No existe una respuesta determinada a la pregunta acerca de si un sujeto posee o no una determinada creencia, puesto que la respuesta depende a su vez de la respuesta a muchas otras preguntas relativas al resto de sus creencias, deseos, propósitos, miedos, valores, etc. Pero esto constituye un serio obstáculo para una teoría que sostiene que cada evento mental particular es idéntico a un determinado evento físico. A diferencia de los estados físicos, los eventos mentales no poseen, en el marco de dicha teoría, una identidad suficientemente robusta para que la tesis monista resulte plausible. El carácter holista del ámbito psicológico, regido por el ideal constitutivo de la racionalidad, es un apoyo central a la premisa del anomalismo de lo mental y la inexistencia de leyes psicofísicas. Podemos ver ahora, sin embargo, cómo este apoyo al anomalismo tiende a minar las bases del monismo.

Advirtamos que, desde nuestra perspectiva subjetiva, nuestros estados mentales no presentan condiciones de individuación e identidad difusas, como las que se desprenden de la perspectiva interpretativa de la tercera persona desde la que Davidson desarrolla su concepción de lo mental. Esto sugiere que el monismo anómalo pierde de vista aspectos importantes de la vida mental humana. Tal vez, frente al monismo anómalo, la racionalidad no constituya la clave para una concepción adecuada de la intencionalidad. Tal vez, asimismo, la intencionalidad no agote la esencia de lo mental. En cualquier caso, el logro de una comprensión teórica satisfactoria de la mente parece un objetivo aún lejano, aun cuando el monismo anómalo nos haya acercado a él.

4. Observaciones finales

El monismo anómalo ocupa un lugar destacado entre las propuestas teóricas dirigidas a la comprensión de la naturaleza de la mente y su relación con el mundo físico. En el presente trabajo hemos presentado, en primer lugar, un panorama sucinto de tales propuestas, con vistas a entender el lugar que el monismo anómalo ocupa en relación con ellas y los rasgos distintivos que lo caracterizan. En segundo lugar, hemos expuesto, de manera sucinta, las tesis centrales de esta posición y la estructura argumentativa sobre la que descansan, con el propósito de facilitar su comprensión al lector no especialista en filosofía. Sobre esta base, hemos puesto de manifiesto, finalmente, algunas de las dificultades a las que se enfrenta el monismo anómalo, tanto internas como externas a su propia estructura conceptual. Estas dificultades son sólo una muestra de las muchas que se presentan en el intento de resolver el venerable problema de la relación psicofísica. La perspectiva objetiva, de tercera persona, que Davidson adopta frente a este problema, y que comparte con otras de las propuestas que hemos presentado, entra en conflicto con rasgos de la mente que sólo se nos presentan desde nuestra perspectiva subjetiva, de primera persona. De esta última recibe el cartesianismo su fuerza y su poder de atracción. Una integración coherente de ambas perspectivas es un paso preliminar indispensable en el desarrollo de una teoría filosófica satisfactoria de la mente. Sin embargo, aun ese paso preliminar está lejos de haber sido dado.

NOTA BIBLIOGRÁFICA

Todas las citas en el texto del artículo refieren a la recopilación de artículos de Donald Davidson, *Essays on Actions and Events*, Clarendon Press, Oxford, 1982. Hay traducción al español de esta obra, a cargo de O. Hansberg, J. A. Robles y M. Valdés, bajo el título *Ensayos sobre acciones y sucesos*, UNAM/Critica, México y Barcelona, 1995. Hay también traducción de algunos de los artículos, incluido "Mental Events", a cargo de M. Candel, bajo el título *Filosofía de la Psicología*, Anthropos, Barcelona, 1994.

Otra importante recopilación de artículos de Davidson es *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford, 1984. Hay traducción al español, a cargo de G. Filippi, bajo el título *De la verdad y la interpretación*, Gedisa, Barcelona, 1990. Es preferible, sin embargo, recurrir a la versión original inglesa. En esta recopilación se contienen algunas de las ideas centrales de Davidson sobre el proceso de interpretación, cuya importancia para el monismo anómalo he tratado de destacar en este trabajo. Ha sido también traducida al español, por O. Fernández Prat (Cátedra, Madrid, 2003), una tercera recopilación de artículos de Davidson, *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford, 2001.

Existe una amplia bibliografía en inglés acerca de diversos aspectos de la obra de Donald Davidson. Acerca de su filosofía de la acción y de la mente sigue siendo importante *Actions and Events. Perspectives on the Philosophy of Donald Davidson* (E. LePore y B. McLaughlin, eds.), Blackwell, Oxford, 1985, y la más reciente *The Philosophy of Donald Davidson* (L. E. Hahn, ed.), Open Court, La Salle, 1999. El libro de S. Evnine, *Donald Davidson* (Polity Press, Cambridge, 1991) es una útil introducción general a la filosofía de Davidson. De especial interés para la cuestión de la interacción psicofísica en el monismo anómalo, a la que nos hemos referido en el texto, es la primera parte, titulada "Davidson, Anomalous Monism, and Mental Causation", del libro *Mental Causation* (J. Heil y A. Mele, eds.), Clarendon Press, Oxford, 1993. Puede verse también, sobre esta cuestión, mi artículo "Reason and Causation in Davidson's Theory of Action Explanation", *Critica*, vol. 30, 1998, pp. 29-43.

En castellano ha aparecido también un considerable número de artículos y capítulos de libros sobre diversos aspectos de la obra de Davidson. Una breve introducción, en español, al conjunto de su filosofía es mi "Introducción a la filosofía de Davidson: mente, mundo y acción", en D. Davidson, *Mente, mundo y acción*, Paidós/UAB, Barcelona, 1992, pp. 9-45. Los artículos incluidos en esta obra fueron seleccionados expresamente por el propio Davidson. De Carlos E. Caorsi es *De una teoría del lenguaje a una teoría de la acción intencional. Una introducción a la filosofía de Donald Davidson* (Factótum, Salamanca, 2001). A M. Hernández Iglesias debemos el estudio *La semántica de Davidson. Una introducción crítica* (Visor, Madrid, 1990). Carlos E. Caorsi ha editado asimismo una colección de artículos sobre diversos aspectos de la filosofía de Davidson: *Ensayos sobre Davidson*, Fundación de Cultura Universitaria, Montevideo, 1999, que contiene algunos trabajos sobre el monismo anómalo, incluido mi "Davidson y la eficacia causal de la mente" y el ensayo de J. Fló, "Observaciones sobre el monismo anómalo".