

(Versión para imprenta)

## DAVIDSON Y LA EFICACIA CAUSAL DE LA MENTE\*

Carlos J. Moya

Departamento de Metafísica y Teoría del Conocimiento  
Universidad de Valencia

### I

El 'monismo anómalo', la solución propuesta por Davidson al viejo —y recalcitrante— problema de la relación entre mente y cuerpo,<sup>1</sup> posee un indudable atractivo. Permite, *prima facie*, satisfacer varias intuiciones de indudable poder de convicción. Así, tendemos a pensar que la conducta intencional humana no está sometida al determinismo causal. El monismo anómalo sustenta esta intuición al defender que no existen leyes estrictas que conecten lo mental con lo físico. Tendemos también a pensar que esa conducta intencional depende realmente de las creencias y los deseos de los agentes. El monismo anómalo recoge esta intuición en la premisa de la interacción causal entre lo mental y lo físico y, de modo más específico, en la tesis según la cual las razones (típicamente, pares de creencias y deseos) son causas de las acciones. El anomalismo de lo mental y la interacción causal psicofísica parecen incompatibles si se sostiene además que toda interacción causal está respaldada por una ley estricta. El monismo anómalo disipa la aparente incompatibilidad al sostener que los eventos mentales y la conducta intencional tienen descripciones físicas bajo las cuales ejemplifican leyes asimismo físicas. El monismo anómalo responde también a nuestra intuición materialista según la cual todo fenómeno físico (y la conducta intencional lo es) ha de tener una causa física al sostener que los eventos mentales son (también) eventos físicos, en la medida en que poseen descripciones físicas verdaderas (al menos aquellas bajo las cuales son subsumidos por leyes físicas estrictas). Finalmente, el monismo anómalo da cuenta de nuestras intuiciones mentalistas según las cuales los fenómenos mentales no son meramente fenómenos físicos, al sostener que, si bien todo particular mental es un particular físico, no existen correspondencias lógicas ni nomológicas, y mucho menos identidad, entre los tipos que subsumen los particulares mentales y los tipos que subsumen los particulares físicos. El discurso mental, incluyendo en él la descripción intencional de la conducta, es autónomo e irreducible al discurso físico, aunque todo particular mental es un particular físico.

Como teoría de la identidad psicofísica, el monismo anómalo hereda las ventajas del materialismo para dar cuenta de la interacción entre la mente y el cuerpo. Al mismo tiempo, como teoría no reductiva, el monismo anómalo hereda las ventajas del mentalismo implícito en la

---

\* La investigación conducente a este trabajo ha sido financiada por la DGICYT del Ministerio español de Educación y Ciencia, en el marco del proyecto PB93-1049-C03-02. Agradezco a esta institución su estímulo y ayuda.

<sup>1</sup> Davidson ha desarrollado esta propuesta en distintos trabajos, en especial en su importante artículo 'Mental Events', así como en 'Psychology and Philosophy' y 'The Material Mind'. Todos ellos se hallan recopilados en *Essays on Actions and Events*, Clarendon Press, Oxford, 1982, pp. 207-259. Existe traducción castellana de esta obra en editorial Crítica (grupo Grijalbo).

concepción cotidiana de los seres humanos para justificar nuestra atribución a los mismos de libertad, responsabilidad y racionalidad.

El atractivo del monismo anómalo, sin embargo, se vio pronto empañado por la sospecha de que, en realidad, la convicción de que nuestra conducta intencional depende realmente de nuestras creencias y deseos no quedaba salvaguardada por él. El monismo anómalo no parecía respetar lo suficiente la llamada 'causalidad mental', la eficacia causal del contenido de nuestros estados mentales en la guía de nuestra conducta. Aunque el monismo anómalo tiene perfecto derecho a sostener que los *eventos* mentales son causalmente eficaces, no parece asignar un papel causal relevante al carácter propiamente *mental* de esos eventos. El monismo anómalo parecía conceder a las propiedades mentales un papel de meros epifenómenos, de acompañantes ineficaces de los procesos causales físicos implicados en la producción de la conducta. El monismo anómalo, en fin, parecía constituir una nueva forma de epifenomenismo.<sup>2</sup> Y si esto era así, otras virtudes del monismo anómalo parecían desvanecerse. Por ejemplo, si las propiedades mentales, tales como la deliberación y la decisión, son epifenoménicas, carecemos de una justificación metafísica suficiente de la libertad y la responsabilidad. Las intuiciones materialistas quedaban salvaguardadas a costa del sacrificio de las intuiciones mentalistas. Por otra parte, la inanidad causal de las actitudes intencionales suscita dudas sobre su realidad y, así, afecta a la cuestión del realismo intencional. Si lo mental no introduce ninguna diferencia causal en el mundo, hay razones para dudar de su realidad.

Esta tendencia epifenomenista que diversos autores han advertido en el monismo anómalo deriva, en parte, del principio del carácter nomológico de la causalidad,<sup>3</sup> que es una de sus premisas, así como de la premisa del anomalismo de lo mental, según la cual no existen leyes estrictas psicológicas o psicofísicas. Según Davidson —y según casi todo el mundo— hay relaciones causales entre lo mental y lo mental y entre lo mental y lo físico (ésta es la primera premisa del monismo anómalo). Pero las leyes que subyacen a estas relaciones causales contienen sólo descripciones físicas y relacionan únicamente propiedades físicas. De modo que, aun cuando un evento mental cause otro evento mental (aunque, digamos, mi deseo de tomar un café cause mi intención de entrar en una cafetería) y un evento mental cause un evento físico (digamos, mi intención de entrar en una cafetería, junto con mi creencia de que hay una cafetería en la calle X, cause el movimiento de mi cuerpo en dirección a la calle X), las propiedades mentales de los eventos mentales en cuestión, en especial su carácter de actitudes hacia determinados contenidos semánticos e intencionales, no parecen ser causalmente relevantes; sólo las propiedades que permiten subsumir dichos eventos bajo leyes físicas estrictas, sólo, pues, las propiedades físicas de los eventos mentales (sus propiedades neurofisiológicas o microfísicas), parecen tener relevancia causal. De ahí la tendencia al epifenomenismo de lo mental.

---

<sup>2</sup> La tendencia del monismo anómalo hacia el epifenomenismo ha sido señalada por diversos autores. Cf., entre otros, T. Honderich, 'The Argument for Anomalous Monism', *Analysis* 42 (1982), pp. 59-64; M. Johnston, 'Why Having a Mind Matters', en E. LePore y B. McLaughlin (eds.), *Actions and Events. Perspectives on the Philosophy of Donald Davidson*, Blackwell, Oxford, 1985, pp. 408-26; J. Kim, 'Epiphenomenal and Supervenient Causation', *Midwest Studies in Philosophy* 9 (1984), pp. 257-70; J. Fodor, 'Making Mind Matter More', *Philosophical Topics* 17 (1989), pp. 59-80; F. Dretske, 'Reasons and Causes', *Philosophical Perspectives* 3 (1989), pp. 1-15.

<sup>3</sup> En 'Mental Events' (p. 208), Davidson expone este principio del modo siguiente: '... Donde hay causalidad, ha de haber una ley: los eventos relacionados como causa y efecto caen bajo leyes deterministas estrictas.'

Tras varios años de silencio acerca de la acusación de epifenomenismo insistentemente dirigida contra el monismo anómalo, Davidson asumió finalmente la defensa de su posición en un artículo denominado 'Thinking Causes'.<sup>4</sup> Davidson insiste, en primer lugar, en que, dada una concepción extensional de la relación causal,<sup>5</sup> como una relación que vincula eventos particulares, no importa cómo puedan ser descritos, 'no tiene sentido literal... hablar de un evento que cause algo en tanto que mental, o en virtud de sus propiedades mentales, o en tanto que es descrito de un modo u otro'.<sup>6</sup>

En su respuesta al artículo de Davidson,<sup>7</sup> Brian McLaughlin subraya (con razón, en opinión de quien esto escribe) que la concepción extensional de la relación causal, como tal, no implica que carezca de sentido hablar de que un evento cause algo en virtud de alguna de sus propiedades. Según McLaughlin, la concepción extensional de las relaciones causales (E) es consistente con la siguiente tesis (llamémosla P):

Si el evento *c* causó el evento *e*, *c* causó *e* en virtud de alguna de las propiedades de *c*.<sup>8</sup>

Tiene perfecto sentido, frente a Davidson, sostener a la vez E y P. Sin embargo, puesto que E no implica lógicamente P, es también posible sostener E y negar P. Según McLaughlin, 'esto parece ser lo que Davidson hace en "Thinking Causes"'. Parece sostener que, cuando un evento causa otro, no hay nada en los eventos en virtud de lo cual ello es así'. Sin embargo, McLaughlin señala correctamente que esta posición compromete a Davidson con la tesis, ciertamente implausible, según la cual las relaciones causales entre eventos particulares son hechos brutos, primitivos, que no admiten ulterior explicación.<sup>9</sup>

Por otra parte, Davidson sugiere que sostener P puede deberse a una confusión entre eventos particulares y tipos de eventos. Para Davidson, sólo los eventos particulares, y no los tipos de eventos, o las propiedades en cuyos términos clasificamos los eventos particulares, pueden intervenir en relaciones causales. Sin embargo, McLaughlin señala que el hecho de que *c* causó *e* en virtud de que *c* tenía la propiedad *F* no implica que el estado de cosas consistente en que *c* tuviera *F* (o que la propiedad *F*, podríamos añadir) causó *e*. P es perfectamente compatible con la tesis davidsoniana según la cual sólo los eventos, y no las propiedades o los estados de cosas, son causas y efectos. Para mostrar esto, McLaughlin subraya el paralelismo entre la relación causal y la relación 'pesar menos que'. Supongamos que *a* pesa 10 Kg. y *b* pesa 11 Kg. Podemos decir que *a* pesa menos que *b* en virtud de que pesa 10 Kg. Pero esto no implica que el estado de cosas consistente en que *a* pese 10 Kg. (o que la propiedad de pesar 10 Kg.) pese menos que *b*. Los estados de cosas (y las propiedades) no tienen peso. Así, 'un defensor de P ['C1' en el texto de

---

<sup>4</sup> En J. Heil y A. Mele (eds.), *Mental Causation*, Clarendon Press, Oxford, 1993, pp. 3-17.

<sup>5</sup> Sobre la concepción davidsoniana de las relaciones causales cf., p. ej., sus trabajos 'Causal Relations' y 'Hempel on Explaining Actions', ambos en *Essays on Actions and Events*.

<sup>6</sup> 'Thinking Causes', p. 13.

<sup>7</sup> B. P. McLaughlin, 'On Davidson's Response to the Charge of Epiphenomenalism', en Heil y Mele (eds.), *Mental Causation*, pp. 27-40.

<sup>8</sup> *Ibid.*, p. 31.

<sup>9</sup> *Ibid.*, pp. 32-3.

McLaughlin, C.M.] puede sostener, consistentemente, que los estados de cosas [o las propiedades, C.M.] no son nunca causas (o efectos), que sólo los eventos son causas y efectos'.<sup>10</sup>

De hecho, el propio Davidson hace una importante concesión a sus críticos cuando, en su artículo, admite que es posible que las propiedades mentales de un evento no supongan ninguna diferencia para sus relaciones causales. Al conceder esto, Davidson está reconociendo implícitamente que no es un sinsentido hablar de la relevancia de las propiedades para las relaciones causales. Siguiendo en el terreno de las concesiones, Davidson reconoce asimismo que el monismo anómalo es compatible con el epifenomenismo, aunque insiste en que no lo implica. Vale la pena citar por extenso al propio Davidson:

Pero, ¿no podría suceder que las propiedades mentales de un evento no introduzcan diferencia alguna en sus relaciones causales? En algo semejante a esto están pensando los críticos cuando dicen que, según *AM+P* [el monismo anómalo más las premisas que conducen a él, C.M.] lo mental es inerte. Desde luego, la idea según la cual las propiedades mentales no suponen diferencia causal alguna es consistente con la tesis de que no hay leyes psicofísicas (estrictas o no) y con el supuesto de que toda relación causal singular entre dos eventos está sustentada por una ley (física) estricta; es también consistente con la tesis de que los eventos mentales (i.e., eventos identificados mediante propiedades mentales) están causalmente relacionados con eventos físicos. Así, *AM+P* es *consistente* con la concepción (epifenomenista) según la cual las propiedades causales de los eventos no introducen diferencia alguna en las relaciones causales. Pero esto no es suficiente para desacreditar *AM+P*, pues no se sigue de ello que *AM* implica el carácter causalmente inerte de lo mental. Lo que los críticos han de mostrar es que *AM* (o *AM+P*) implica la impotencia de las propiedades mentales, y esto es algo que no veo forma alguna de establecer.<sup>11</sup>

Aun cuando Davidson tenga razón al afirmar que el monismo anómalo no *implica* el carácter causalmente inerte de las propiedades mentales, su reconocimiento de que el monismo anómalo es *consistente* con dicho carácter muestra claramente que el monismo anómalo tiene alguna cuenta que saldar con la acusación de epifenomenismo. Una doctrina compatible con el epifenomenismo de lo mental, una doctrina que no incluye o implica una explicación positiva de la relevancia causal de las propiedades mentales es, cuando menos, incompleta o insatisfactoria.<sup>12</sup>

En 'Thinking Causes', Davidson trata de complementar el monismo anómalo ofreciendo una explicación positiva de la relevancia causal de las propiedades mentales. Esta explicación descansa sobre todo en la superveniencia de las propiedades mentales sobre las propiedades físicas y en la existencia de leyes no estrictas que conectan lo mental con lo físico. Estos no son aspectos nuevos de la filosofía de Davidson. La tesis de la superveniencia es ya expuesta y defendida en 'Mental Events'<sup>13</sup> y, en diversos lugares, Davidson ha aceptado la existencia de generalizaciones verdaderas, no susceptibles de ser convertidas en leyes estrictas, que conectan lo mental con lo físico. El recurso de Davidson a estos elementos no es, pues, un expediente *ad hoc*. Sin embargo, hay razones para pensar que estos aspectos de la filosofía davidsoniana no ofrecen una respuesta suficiente a la objeción de que el monismo anómalo no asigna un papel causal relevante a las propiedades mentales. De acuerdo con la concepción davidsoniana de la superveniencia, si dos

---

<sup>10</sup> *Ibid.*, p. 34.

<sup>11</sup> 'Thinking Causes', p. 13.

<sup>12</sup> Cf. J. Kim, 'Can Supervenience and "Non-Strict Laws" Save Anomalous Monism?', en Heil y Mele (eds.), *Mental Causation*, pp. 19-26, esp. pp. 20-21.

<sup>13</sup> Cf. nota 1.

objetos o eventos difieren en alguna propiedad mental, han de diferir en alguna propiedad física, pero esto no supone que cualesquiera dos objetos o eventos que difieran en una misma propiedad mental hayan de diferir en una misma propiedad física. La superveniencia davidsoniana no conlleva correspondencias nomológicas entre propiedades mentales y físicas. No conlleva la posibilidad de reducir los conceptos mentales a conceptos físicos. Según Kim, esta concepción anómica, no reductiva, de la superveniencia no es suficiente para librar al monismo anómalo de la acusación de epifenomenismo.<sup>14</sup> De acuerdo con Davidson, 'si se da la superveniencia, las propiedades psicológicas suponen una diferencia en las relaciones causales de un evento, porque esas propiedades son relevantes para las propiedades físicas y las propiedades físicas son relevantes para las relaciones causales'.<sup>15</sup> Sin embargo, pretender que de este modo se asegura el papel causal de las propiedades mentales es un error. En el texto citado, Davidson está usando el término 'relevantes' en sentidos distintos. Que las propiedades mentales son relevantes para las propiedades físicas significa que éstas han de cambiar si las primeras cambian, dada la superveniencia de lo mental sobre lo físico. Que las propiedades físicas son relevantes para las relaciones causales significa que si dos eventos están causalmente relacionados, tienen propiedades físicas nomológicamente relacionadas. Por lo tanto, que las propiedades mentales son relevantes para las propiedades físicas y que éstas son relevantes para las relaciones causales no supone que las propiedades mentales son relevantes para las relaciones causales en el mismo sentido en que lo son las propiedades físicas. Que un evento no pueda cambiar en sus propiedades mentales sin sufrir un cambio en sus propiedades físicas y con él en sus relaciones causales no supone que sus propiedades mentales afecten a las relaciones causales en que pueda hallarse. Sólo sus propiedades físicas lo hacen. Así, sería concebible que un evento que ahora posee propiedades tanto físicas como mentales preservase todas sus propiedades físicas y perdiese todas sus propiedades mentales sin que sus relaciones causales se vieran afectadas. Es cierto que ambos eventos no podrían ser el mismo, ya que uno tiene propiedades de las que el otro carece. Pero esta observación no afecta al núcleo central de la cuestión, a saber, que son las propiedades físicas de un evento las que soportan la carga de sus relaciones causales. No parece, pues, que las propiedades mentales de un evento supongan una diferencia para sus relaciones causales. La superveniencia no ha conjurado el peligro del epifenomenismo.

Tampoco parece capaz de hacerlo el recurso a leyes no estrictas que vinculen lo mental con lo mental y lo mental con lo físico. Davidson acepta, como indicamos, este tipo de leyes (por ejemplo, los principios de la teoría de la decisión desempeñan un papel importante en su teoría de la acción y de la interpretación de las acciones). Supongamos, pues, que *a* (un evento mental) causa *b* (otro evento mental) y que ciertas propiedades mentales de *a* y *b* se hallan conectadas en una ley no estricta (por ejemplo, un principio de la teoría de la decisión). Ahora bien, de acuerdo con el principio del carácter nomológico de la causalidad, *a* y *b* han de tener también propiedades físicas conectadas en una ley (física) estricta. Es claro que estas propiedades físicas *son* relevantes para la relación causal entre ambos eventos. ¿Realmente estamos autorizados a suponer que las propiedades mentales de *a* y *b* son *también* relevantes para su relación causal si aceptamos que esas propiedades mentales sobrevienen a propiedades físicas de *a* y *b*, algunas de las cuales se hallan conectadas, a diferencia de las propiedades mentales, por leyes estrictas? Kim ha señalado que la

---

<sup>14</sup> Cf. Kim, 'Can Supervenience...'

<sup>15</sup> 'Thinking Causes', p. 14.

presencia de dos tipos de leyes que subyacen a la relación causal entre dos eventos plantea el problema de la 'exclusión explicativa': las propiedades (físicas) que figuran en la ley estricta privarían de relevancia causal a las propiedades mentales que figuran en la ley no estricta.<sup>16</sup> Davidson subraya, a mi entender correctamente, que 'si los conceptos mentales no son reductibles a conceptos físicos, no hay razón para suponer que perderíamos interés en las explicaciones formuladas en términos mentales sólo porque tuviésemos una explicación física completa'.<sup>17</sup> Sin embargo, considero que Davidson nos debe un análisis detallado de las relaciones entre ambos tipos de explicaciones. Insistir en su independencia e irreductibilidad puede inducir con demasiada facilidad la sospecha de que las propiedades mentales son meramente ficciones pragmáticamente útiles pero metafísicamente injustificadas. En un artículo reciente, Manuel García-Carpintero señala que 'la posición de Davidson consiste en que nuestra creencia sobre la eficacia causal de la mente es similar a la del supersticioso [que cree que viajar en martes causa desgracias], porque no es razonable creer en la existencia de leyes causales *strictas* que conecten lo mental y lo físico'.<sup>18</sup> Particularmente, no creo que ésta sea la posición de Davidson, pero resulta sintomático que su obra pueda suscitar interpretaciones de este tipo. Esto significa que Davidson no ha hecho lo suficiente para justificar el papel causal de las propiedades mentales y que la objeción de epifenomenismo dirigida contra su posición no carece de fundamento.

## II

Quisiera ahora ofrecer algunas observaciones sobre el temprano artículo de Davidson 'Acciones, razones y causas'. Treinta años separan este célebre trabajo de 'Thinking Causes'. Sin embargo, el germen de los problemas a los que Davidson se enfrenta en este último se halla ya presente en el primero. Puede, pues, ser útil retroceder hasta él. Se oponía Davidson en este artículo a una tradición, entonces floreciente, relacionada con autores como Collingwood, Dray y, sobre todo, Wittgenstein, según la cual las explicaciones de la acción intencional mediante razones no eran explicaciones causales y las razones no constituían causas de la acción. Esta tesis era una de las manifestaciones de una tendencia general, propia de la tradición en cuestión, a reivindicar la autonomía de las características específicamente humanas frente al mundo físico y de las ciencias sociales frente a las ciencias de la naturaleza. Esta tendencia se remontaba hasta el siglo XIX, aunque las motivaciones que dan lugar a ella pueden rastrearse en períodos anteriores de la cultura y de la filosofía occidental. Aunque los partidarios de la unidad de la ciencia, contrarios a esta tradición, vieron en el artículo de Davidson un importante apoyo para su posición, el desarrollo posterior de la filosofía de Davidson puso de manifiesto que la autonomía de lo humano seguía desempeñando en ella un importante papel. El monismo defendido por Davidson alcanzaba a los particulares, objetos y eventos, pero no se extendía a las características generales y a las propiedades de dichos particulares. A este respecto, Davidson se ha mostrado tan contrario a la reducción de lo intencional como pudieron serlo los autores a los que se opuso en 'Acciones, razones y causas'.

---

<sup>16</sup> 'Can Supervenience...', p. 25.

<sup>17</sup> 'Thinking Causes', p. 16.

<sup>18</sup> M. García-Carpintero, 'La diferencia entre explicar e interpretar, según Davidson', en M. Torreveano (comp.), *Filosofía analítica hoy*, Universidad de Santiago de Compostela, 1991, pp. 83-119, esp. p. 115.

La tesis central de Davidson en este trabajo es que las explicaciones de la acción intencional mediante razones (o racionalizaciones, como también las denomina Davidson) son 'una especie de explicación causal'.<sup>19</sup> Las racionalizaciones se distinguen de otras explicaciones del género causal, al menos, por el elemento de justificación del explanandum (la acción) contenido en ellas: 'Si la racionalización es, como pretendo argüir, una especie de explicación causal, la justificación... es al menos una propiedad diferencial'.<sup>20</sup> Este elemento de justificación de la acción se añade, pues, como una característica específica, al rasgo común que las racionalizaciones comparten con otras especies de explicación causal. Justificación y causalidad son dos condiciones independientes, ambas necesarias (aunque no conjuntamente suficientes, como veremos) para que una razón explique una acción en el marco de una racionalización. Aunque la formulación de la razón puede tomar muchas formas, si la razón ha de explicar la acción ha de constituir, en último término, una 'razón primaria'. Y para que una razón sea primaria ha de cumplir las dos condiciones necesarias mencionadas: ha de justificar la acción y, además, ha de causarla. La condición de justificación queda expuesta en C1:

C1. *R* es una razón primaria por la que un agente llevó a cabo la acción *A* bajo la descripción *d* sólo si *R* consiste en una actitud favorable del agente hacia acciones que tengan cierta propiedad y en una creencia del agente según la cual *A*, bajo la descripción *d*, tiene esa propiedad.<sup>21</sup>

Para entender la necesidad de esta condición, supongamos que queremos saber por qué Juan abandonó la fiesta de cumpleaños de Pedro y que alguien nos responde lo siguiente: 'Juan se marchó de la fiesta porque se lo estaba pasando muy bien'. Sin información adicional, la razón que se nos ofrece de la conducta de Juan no puede explicarla, porque no podemos hallar en ella el elemento de justificación que C1 pretende analizar. Si suponemos en Juan, como es natural, una actitud favorable hacia acciones que le permitan pasarlo bien y la creencia, sugerida en la "explicación", de que permanecer en la fiesta le permitiría pasarlo bien, no podemos entender que Juan se marchara de la fiesta. La razón aducida justificaría que Juan se quedara en la fiesta, no que se marchara. Naturalmente, en el complicado y sutil contexto pragmático en que se formulan las racionalizaciones, el *tono* de voz con que se nos dio esa "explicación" pudo muy bien sugerir precisamente que Juan se estaba *aburriendo* mortalmente, lo que restauraría la condición de justificación requerida.

Según Davidson, sin embargo, considerar que C1 es lo único que está implicado en la racionalización es un error. Porque una razón para una acción puede cumplir C1 y no explicar la acción.<sup>22</sup> Supongamos, en el ejemplo anterior, que Juan efectivamente se estaba aburriendo y que se marchó de la fiesta. De ello no se sigue sin más que se marchara de la fiesta *porque* se estaba aburriendo. De ahí que haya que introducir una segunda condición que una razón primaria ha de cumplir, a saber, la razón en cuestión ha de ser la razón *por la cual* el agente llevó a cabo la acción. Según Davidson, el modo más natural y obvio de especificar la demanda contenida en esta condición es considerarla como la exigencia de un vínculo causal entre la razón y la acción. De ahí la condición C2:

---

<sup>19</sup> 'Actions, Reasons, and Causes', p. 3.

<sup>20</sup> *Ibid.*, p. 9.

<sup>21</sup> *Ibid.*, p. 5.

<sup>22</sup> Cf. *ibid.*, p. 9.

C2. Una razón primaria de una acción es su causa.<sup>23</sup>

Según Davidson, ninguno de los argumentos esgrimidos contra C2 resulta decisivo.<sup>24</sup> Y, a falta de un argumento decisivo en contra, C2 debería aceptarse como la forma más natural de expresar la condición requerida de eficacia de la razón en la generación de la acción.

Davidson, pues, considera la justificación y la causalidad como dos condiciones necesarias, y recíprocamente independientes, de las racionalizaciones. Que la razón justifique la acción es sólo una condición de que la razón explique la acción. Es necesario también que la razón cause la acción.<sup>25</sup>

De acuerdo con las dos condiciones mencionadas, una racionalización tendría la siguiente forma:

(R) El agente *S* llevó a cabo la acción *A* (bajo la descripción *d*) *porque* *S* tenía una actitud favorable hacia acciones que tuvieran cierta propiedad y creía que la acción *A* (bajo la descripción *d*) tenía esa propiedad.

Que la razón cumpla únicamente la condición de justificación (C1) no nos permitiría, aunque el agente de hecho hiciera *A*, afirmar que hizo *A porque* tenía esa razón. Lo único que el cumplimiento de la condición C1 nos permitiría sería un esquema del tipo siguiente:

(J) El agente *S* llevó a cabo la acción *A* (bajo la descripción *d*) y *S* tenía una actitud favorable hacia acciones que tuvieran cierta propiedad y creía que la acción *A* (bajo la descripción *d*) tenía esa propiedad.

En efecto, aunque el esquema (J) sea aplicable a un agente, éste pudo haber llevado a cabo la acción *A por otra razón*. Lo que nos permite sustituir el primer 'y' del esquema (J) por un 'porque', obteniendo así el esquema (R), es que la razón causara la acción. Todo lo que se desprende de que el agente tuviera de hecho una razón del tipo descrito en los esquemas anteriores es que el agente tenía una actitud favorable hacia la acción *A* o tal vez que era razonable para él llevar a cabo esa acción, de modo que, aun cuando de hecho realizara la acción en cuestión, no podemos concluir que lo hizo por esa razón. Las razones, *como argumentos* en favor de una acción, parecen impotentes para dar lugar a la acción.<sup>26</sup> Necesitamos, pues, el vínculo causal, la eficacia causal de esas razones, si hemos de explicar la acción.

La argumentación davidsoniana, hasta aquí, parece correcta, si la hemos interpretado bien. Ahora bien, si tenemos el vínculo causal, si las razones causaron la acción, el contenido de las

---

<sup>23</sup> *Ibid.*, p. 12.

<sup>24</sup> Sobre estos argumentos, en especial el llamado 'argumento de la conexión lógica' cf. p. ej. mi *The Philosophy of Action: An Introduction*, Polity Press, Cambridge, 1990, cap. 10.

<sup>25</sup> Ambas condiciones, sin embargo, no son conjuntamente suficientes debido a la posibilidad de las cadenas causales desviadas.

<sup>26</sup> Este es, básicamente, el reproche que Hempel hizo a la concepción de la explicación de la acción de William Dray. Cf. C. G. Hempel, 'Explanation in Science and in History', en W. H. Dray (ed.), *Philosophical Analysis and History*, Harper & Row, New York, 1966, pp. 95-126.

razones, del que depende la justificación de la acción, queda privado de papel causal. En el fondo es indiferente, *para que la acción se lleve a cabo*, que las razones tengan un contenido tal que la acción aparezca como razonable a la luz de las mismas, *aun cuando de hecho tengan ese contenido*. Dada la relación causal como una condición adicional e independiente de la condición de justificación, el contenido de las razones comienza a aparecer como causalmente inerte. El epifenomenismo de lo mental comienza a dibujarse ante nosotros.

### III

De hecho, el resultado anterior es también congenial con la concepción davidsoniana de la causalidad y con su distinción entre relaciones causales entre eventos (descritas en enunciados causales singulares) y explicaciones causales. Esta concepción de la causalidad está ya algo más que implícita en 'Acciones, razones y causas', pero Davidson la desarrolla con mayor detalle y profundidad en otros artículos.<sup>27</sup>

Para Davidson, como ya sugerimos en el primer apartado, las relaciones causales son relaciones *extensionales* entre eventos particulares y los juicios que las expresan son verdaderos, si lo son, sea cual fuere la manera en que describamos ambos eventos. Los juicios causales singulares están formulados en un lenguaje extensional. En cambio, las explicaciones causales son relaciones entre enunciados, algunos de los cuales expresan leyes. La verdad de algunos de estos enunciados, que constituyen el explanans, implica lógicamente la verdad de un enunciado que describe el explanandum, o, si las leyes involucradas en la explicación son de carácter probabilista, presta cierto grado de probabilidad a la verdad de ese enunciado (en lo sucesivo, omitiremos esta referencia a la explicación probabilista).<sup>28</sup> En el contexto de la explicación causal, es esencial la manera en que se describan los eventos. Si de las leyes y la descripción de la causa y de ciertas circunstancias del caso se puede deducir un enunciado que describe el efecto, esto no supone que una descripción distinta del mismo efecto pueda también ser deducida de (explicada por) dichas leyes y descripciones. El lenguaje de la explicación causal es, pues, *intensional*.

La conexión entre la relación causal y la explicación causal es la siguiente: si se da una relación causal entre dos eventos, esto es, si un juicio causal singular es verdadero, entonces hay una explicación causal de esa relación causal, de modo que, de ciertas leyes y de una descripción adecuada de la causa, es posible deducir una descripción del efecto. Las explicaciones causales contienen aquellas relaciones nómicas entre propiedades que dan razón de la verdad de los enunciados causales singulares, es decir, indican qué propiedades de la causa y del efecto son relevantes para la existencia de una relación causal entre ambos.

Atendamos al siguiente enunciado causal singular:

El incendio del edificio *E* se debió al cortocircuito producido en el conmutador eléctrico de la tercera planta de dicho edificio.

---

<sup>27</sup> Cf. nota 5.

<sup>28</sup> La deuda de Davidson con la concepción hempeliana de la explicación científica es aquí bastante clara.

Supongamos que este enunciado es verdadero. Y supongamos que el incendio del edificio *E* fue el causante de la ruina de la compañía de seguros *S*. Entonces es verdad que:

El evento que causó la ruina de la compañía de seguros *S* se debió al cortocircuito en el conmutador eléctrico de la tercera planta del edificio *E*.

La explicación causal del incendio en cuestión contendrá leyes físicas sobre la electricidad, la combustión, etc., que, en unión con una descripción del cortocircuito en términos de la teoría física a la que pertenecen esas leyes y de las circunstancias particulares del caso (la presencia de material combustible y de comburente adecuado, etc.) permitirán la deducción de una descripción del incendio en términos de la teoría física en cuestión. Ciertamente no podríamos deducir en esa explicación un enunciado que describe el incendio en términos de sus efectos sobre la ruina de una compañía de seguros. Esa descripción sencillamente no es deducible de los enunciados que constituyen el explanans. Las relaciones del incendio con la compañía de seguros no son causalmente relevantes en su producción.

Esta concepción davidsoniana de la causalidad se ajusta bien a las relaciones causales entre eventos físicos. Sabemos qué significa describir la causa y el efecto de formas distintas y cómo ciertas descripciones serían apropiadas para la explicación causal, a diferencia de otras. Según Davidson, esta concepción de la causalidad se aplica también a las relaciones entre razones y acción. Esta aplicación, sin embargo, no está exenta de problemas. Veamos por qué.

Es claro que, en el marco de esta concepción, los casos particulares del esquema (R), es decir, las explicaciones de acciones particulares mediante razones del agente, corresponden a enunciados causales singulares. Un caso particular del esquema (R) sería, por ejemplo, el siguiente:

(A) Juan fue a la iglesia porque tenía una actitud favorable hacia acciones que tuvieran la propiedad de agradar a su madre y creía que ir a la iglesia tenía esa propiedad.

En el contexto cotidiano, la explicación anterior tendría una forma más simple, tal como:

(A') Juan fue a la iglesia porque quería agradar a su madre,

o bien:

(A'') Juan fue a la iglesia porque creía que hacerlo agradaría a su madre.

Aceptemos que (A') y (A'') tienen implícitamente la forma de (A) y que son verdaderos. (A), (A') o (A'') son, en la concepción davidsoniana, enunciados causales singulares. Dicen simplemente que hay dos eventos, uno de los cuales es la razón, el otro es la acción, y el primero causó el segundo. Pero esto supone que la descripción que damos de la causa y el efecto no afecta a la verdad del enunciado y que hay otras descripciones de la causa (el deseo de Juan, en el momento *t*, de agradar a su madre y su creencia, en *t*, de que ir a la iglesia agradaría a su madre) tales que, en unión con leyes apropiadas, harían posible deducir el efecto (la acción de Juan, en el momento *t+1*, de ir a la iglesia) descrito también en el vocabulario apropiado.

Ahora bien, aunque tenemos una idea clara de qué supone dar una descripción alternativa de un evento físico (podemos describir el incendio del edificio *E* como el evento que causó la ruina

de la compañía *S*) preservando la verdad del enunciado causal singular, no tenemos, en el caso de (A), (A') o (A''), una idea clara de cómo ofrecer descripciones alternativas de las razones que no consistan en meras paráfrasis en términos intencionales, más o menos sinónimas (por ejemplo, sustituir 'desear' por 'querer' o por 'tener una actitud favorable'). Davidson sugiere, y ésta es una sugerencia que debemos aceptar, que hay una historia causal física que parte del estado del cuerpo de Juan en el momento *t*, que describe el estado de su sistema nervioso, de sus músculos, etc., y que llega hasta el conjunto de movimientos físicos en que consiste la acción de Juan de ir a la iglesia. Una parte esencial de esta historia causal, que hace de ella una explicación causal, son las leyes neurofisiológicas o microfísicas que dan razón del vínculo causal entre los diversos eventos de esta historia. Sin embargo, que una descripción neurofisiológica de ciertos eventos cerebrales en el cuerpo de Juan sea una descripción alternativa de la creencia o el deseo de Juan es una hipótesis controvertida, que depende de la verdad del monismo. Y mucho más controvertido es que sea esa descripción alternativa la que, en unión con leyes neurofisiológicas y físicas apropiadas, dé razón del vínculo causal entre las razones de Juan y su acción de ir a la iglesia, y explique causalmente esa acción. No es, en cambio, una tesis controvertida que 'lo que causó la ruina de la compañía *S*' sea una descripción alternativa del incendio del edificio *E*. Ni es tampoco controvertido que lo que da razón del vínculo causal entre el cortocircuito y el incendio sea una explicación causal en términos de la teoría física, que incluye al menos leyes físicas y una descripción adecuada del cortocircuito y el incendio.

Esta diferencia entre los enunciados causales singulares sobre eventos físicos y las explicaciones de la acción mediante razones deberían hacernos dudar de que estas últimas sean realmente enunciados causales singulares, que esperan una explicación causal en términos de ciertas teorías científicas. Un propósito de la explicación causal en el marco científico-natural es, entre otros, hacer inteligibles los vínculos causales entre eventos particulares, al subsumirlos bajo categorías generales que puedan conectarse con leyes y teorías científicas. Las teorías científicas en cuyos términos se formulan las explicaciones causales son una fuente de inteligibilidad de los nexos causales singulares, que de otro modo serían hechos brutos y primitivos. Sin embargo, en el caso de (A), (A') o (A''), la descripción intencional es precisamente la que nos permite hacernos inteligible la conexión causal entre la razón y la acción. Si de lo que se trata es de entender la acción de Juan bajo la descripción 'ir a la iglesia' y no bajo la descripción 'moverse en dirección noroeste', es una descripción intencional de la causa lo que necesitamos, y esto es lo que nos proporcionan (A), (A') o (A''). Estas no son meros enunciados causales singulares, sino que contienen ya los términos adecuados para dar razón de la conexión causal. La libertad que tenemos para describir en otros términos la relación entre las razones y la acción está severamente limitada por el hecho de que la descripción bajo la cual el agente consideró su acción era la descripción bajo la cual el propio agente la consideraba deseable. Si 'ir a la iglesia' y 'moverse en dirección noroeste' son descripciones alternativas de la acción de Juan, es sólo la primera descripción la que Juan tuvo en cuenta y le llevó a realizar esa acción: es la descripción 'ir a la iglesia' aquella bajo la cual Juan creyó que su acción agradaría a su madre, lo que, en unión con su deseo de agradar a su madre, hizo (causó) que Juan se decidiera a ir a la iglesia, no a moverse en dirección noroeste, aun cuando ir a la iglesia fuese, en esas circunstancias, idéntico a moverse en dirección al noroeste. Los eventos no se preocupan de las descripciones bajo las cuales causan o son causados por otros eventos. Pero los agentes sí se preocupan de las descripciones bajo las cuales sus acciones se relacionan con sus razones. Y la explicación de estas acciones no puede pasar por alto este hecho. Las descripciones de las razones y de la acción son importantes, no sólo para que la acción esté *justificada* por esas

razones, sino también para que esas razones lleven al agente a realizar la acción, para la propia relación *causal*. La intensionalidad que esto introduce en las explicaciones de acciones particulares mediante razones muestra que éstas no son equivalentes a enunciados causales singulares. Tampoco son equivalentes a explicaciones causales en el sentido en que Davidson las entiende, ya que no disponemos en este caso de leyes científicas que nos permitan deducir la acción descrita en los términos en que el agente la consideraba deseable. Las explicaciones de acciones particulares mediante razones atraviesan oblicuamente la distinción davidsoniana entre enunciados causales singulares y explicaciones causales. Nos las tenemos aquí, por así decirlo, con una relación *causal intensional*.

Davidson reconoce, desde luego, la existencia de un elemento intensional en las racionalizaciones. Pero este elemento intensional se reduce para él a la relación de *justificación* entre las razones y la acción, sin afectar a su relación *causal*. Creo que esto es un error. El deseo de Juan de agradar a su madre, sea o no idéntico a un evento físico en el cerebro de Juan, causa que Juan vaya a la iglesia precisamente bajo la descripción 'el deseo de agradar a mi madre', no bajo la descripción física, si la tiene. Es el *contenido* del deseo de Juan, *lo que* Juan desea, lo que causa su acción. La relación causal entre razones y acción es, pues, intensional. Frente a Davidson, no sólo tiene perfecto sentido, sino que parece además verdadero, decir, en el caso de la relación entre razón y acción, que la razón causa la acción 'en tanto que es descrita' de cierta manera. Davidson tiene razón al considerar que la descripción de las razones y de la acción es esencial para que la acción sea justificada por las razones en cuestión. Sin embargo, no la tiene al sostener que la descripción de las razones y de la acción no afecta a la relación causal entre ambas. El contenido de la razón que justifica la acción es también, cuando el agente actúa por esa razón, lo que causa la acción. Esta presencia esencial del contenido, tanto en la relación de justificación como en la relación causal, sugiere que, en el caso de las racionalizaciones, las condiciones de justificación y de causalidad no son independientes.

Pero, ¿cómo podría afectar a la relación causal entre dos eventos la descripción que demos de uno y otro? Si el cortocircuito causó el incendio, la descripción que demos de ambos eventos no puede añadir ni quitar nada a la conexión causal entre ambos. Esto es sin duda correcto, pero no puede aplicarse sin más a la relación entre razón y acción. Porque en este caso no se trata de que haya dos eventos externos a quienes los describen de un modo u otro e independientes de ellos, sino que el propio agente que actúa por una razón concibe él mismo la razón y la acción bajo cierta descripción. Es la perspectiva del propio agente lo que introduce la diferencia frente a las relaciones causales ordinarias entre eventos físicos.

#### IV

De acuerdo con lo dicho, creemos que el hecho de que Davidson considere las condiciones de justificación y de causalidad como condiciones separadas y recíprocamente independientes de las racionalizaciones, así como su concepción de las racionalizaciones como enunciados causales singulares y de la relación entre la razón y la acción como una relación causal extensional, se hallan en la raíz misma de los problemas de epifenomenismo con que se enfrenta el monismo anómalo. La cuestión de las relaciones entre justificación y causalidad, sin embargo, no tiene solamente un alcance interno a la obra davidsoniana. De la comprensión adecuada de estas relaciones depende la posible solución al problema de la causalidad mental, un problema que afecta, no sólo a la

filosofía de Davidson, sino a la mayor parte de la filosofía de la mente de nuestros días. El problema de la causalidad mental puede entenderse como el problema de reconciliar justificación y causalidad en la relación entre razones y acciones (o entre razones y razones). El problema recibe diferentes formulaciones. La eficacia causal del contenido es una de ellas. ¿Cómo puede el contenido intencional o semántico de un deseo y una creencia causar una acción? El problema guarda una estrecha relación con las cuestiones tratadas en 'Acciones, razones y causas'. La condición de justificación, tal como Davidson la analiza en C1, consiste en que los contenidos de la actitud favorable y de la creencia del agente guarden con una determinada descripción de la acción (anticipada en la creencia del agente) una relación tal que la acción, bajo esa descripción, aparezca como una consecuencia razonable, o al menos inteligible, de las razones del agente. *Grosso modo*, el agente ha de creer que la acción *A*, bajo la descripción *d*, tiene una propiedad que la hace deseable. Davidson insiste en que, aun cuando esto sea cierto del agente, y el agente lleve a cabo *A*, la realización de *A* no queda aún racionalizada (i.e., explicada mediante esas razones del agente). Es también necesaria la condición de causalidad.

Veamos entonces cómo surge el problema del epifenomenismo. Supongamos que aceptamos, con Davidson, que las condiciones de justificación y de causalidad son *independientes* y que ambas son necesarias para que la razón racionalice la acción. La relación de justificación es una relación entre los contenidos de las razones y la descripción bajo la cual la acción, si se lleva a cabo, es intencional. La relación de justificación, pues, es intensional. La relación causal, por su parte, es una relación extensional entre eventos. Dado el principio del carácter nomológico de la causalidad, algunas de las descripciones de estos eventos ejemplifican una ley física estricta, ya que no hay leyes estrictas de carácter no físico. Esto significa que los contenidos de las razones y la descripción intencional (mental) de la acción no son relevantes para la relación causal, que encuentra respaldo suficiente en leyes que conectan únicamente propiedades físicas. De este modo, el contenido de las razones y la descripción intencional de la acción (la descripción bajo la cual el agente la consideraba deseable) se desligan de la estructura causal involucrada en las racionalizaciones, tendiendo con ello a convertirse en epifenómenos causalmente inertes.

Nuestra perspectiva apunta al rechazo de un aspecto básico de la concepción davidsoniana de las racionalizaciones: la independencia entre las condiciones de justificación y de causalidad. Nuestra insistencia en que la relación causal entre las razones y la acción es intensional, en que es el contenido mismo de sus creencias y deseos el que causa la acción del sujeto es, dado que el contenido mismo es indiscutiblemente relevante para la justificación de la acción, una invitación a unificar causalidad y justificación. En 'Acciones, razones y causas' Davidson se plantea, de pasada, la posibilidad de aceptar esta unificación, pero la descarta rápidamente:

Resulta central para la relación entre una razón y una acción explicada por ella la idea de que el agente llevó a cabo la acción *porque* tenía esa razón. Desde luego, podemos incluir también esta idea en la justificación; pero entonces la noción de justificación se torna tan oscura como la noción de razón hasta que podamos explicar la fuerza de ese 'porque'.<sup>29</sup>

Aceptemos, con Davidson, que la fuerza de ese 'porque' es causal. Sin embargo, si esa relación causal es intensional, si es una relación entre el contenido de las razones y la acción (descrita de cierta forma), es concebible entonces que las razones causen la acción precisamente

---

<sup>29</sup> 'Actions, Reasons, and Causes', p. 9.

como argumentos. Es concebible que las razones causen la acción precisamente porque, desde el punto de vista del agente, proporcionan argumentos en favor de la misma, e decir, la justifican. Esto evita que el problema del epifenomenismo de lo mental llegue siquiera a plantearse.

La consideración más importante ofrecida por Davidson para mantener la condición causal independiente de la condición de justificación es que 'una persona puede tener una razón para una acción, y llevar a cabo esa acción, sin que esa razón sea la razón por la que lo hizo.'<sup>30</sup> La diferencia entre este caso y el caso en el que dicha razón es la razón por la que el agente llevó a cabo la acción parece ser, pues, únicamente una diferencia causal. En el segundo caso, a diferencia del primero, la razón causó la acción, aunque en *ambos* casos la razón justifica la acción. Creemos, sin embargo, que la diferencia entre ambos casos es *también* una diferencia en la justificación de la acción. El hecho de que un agente tenga razones que podrían aducirse para justificar una acción si ésta se llevase a cabo (razones que cumplan la condición C1 u otra similar) no significa, sin más, que la acción, si se lleva a cabo, esté justificada por esas razones a menos que esté también causada por ellas, a menos que el agente lleve a cabo la acción *por* esas razones. La justificación de una acción particular por ciertas razones involucra la idea de que la acción se debió a esas razones. La existencia de una relación formalmente correcta entre la descripción de las razones y la descripción de la acción es sólo un aspecto necesario de la justificación de la acción. La justificación de la acción requiere también que el agente lleve a cabo la acción, bajo cierta descripción, debido al contenido de sus razones, debido a lo que cree y desea. El contexto de la justificación es un contexto explicativo: la acción sólo está justificada por determinadas razones si esas razones explican la realización de la acción. Que un agente pueda utilizar su conocimiento de las relaciones formales entre razones y acción para presentar su acción como aceptable o para disculparse ante otros no significa sin más que su acción esté realmente justificada por ese solo hecho.

Las ventajas de la propuesta que planteamos son claras. Si la justificación de la acción mediante las razones del agente es ya una relación causal-explicativa, la cuestión de que el contenido o las propiedades intencionales de las razones no sean causalmente relevantes no llega a surgir, porque si no es el contenido de las razones lo que causa la acción, ésta no está justificada y, por tanto, tampoco explicada por ellas.

Ofreceremos a continuación tres consideraciones en favor de nuestra propuesta.

En primer lugar, los casos de cadenas causales desviadas muestran precisamente la interdependencia entre justificación y causalidad en el marco de la acción intencional. Ciertas formas en que las razones de un agente pueden causar su acción afectan a la *justificación* de su acción. Pensemos en el conocido ejemplo de Roderick Chisholm.<sup>31</sup> Un joven desea heredar una fortuna y cree que matar a su tío le permitirá heredar una fortuna. Esta creencia y este deseo (estas razones) le excitan de tal modo que conduce imprudentemente y mata a un peatón que, casualmente, resulta ser su tío. No podemos decir aquí que su acción de atropellar a su tío esté justificada, aunque pueda estar causada, por las razones en cuestión, precisamente porque la cadena causal no opera a través de las relaciones entre los contenidos de sus razones y la descripción apropiada de su acción, no opera en el contexto de la justificación. La acción no fue realizada como conclusión del contenido de las razones, no fue la consecuencia racional que el agente extrajo de

---

<sup>30</sup> *Ibid.*

<sup>31</sup> R. Chisholm, 'Freedom and Action', en K. Lehrer (ed.) *Freedom and Determinism*, Random House, New York, 1966, pp. 11-44, esp. pp. 29-30.

ese contenido, sino que fue causada por el estado de nerviosismo generado en el agente por la perspectiva de matar a su tío.

En segundo lugar, pensemos en otro ejemplo. Juan es un probo funcionario que cree que la función pública no debe ponerse al servicio de intereses privados. Juan descubre que su compañero Pedro, que compite con él por un puesto superior, ha agilizado el expediente de un pariente en detrimento de otros expedientes. En consecuencia, Juan decide poner el hecho en conocimiento de sus superiores. Como justificación moral de su acción, ante sus otros compañeros, ofrece su deseo de mantener la función pública incontaminada de intereses privados. Supongamos, no obstante, que, en el fondo, lo que ha decidido a Juan a denunciar a Pedro ante sus superiores ha sido la perspectiva de quitarse de enmedio a un competidor. Aunque la razón que Juan ofrece a sus compañeros es una razón que Juan tiene para denunciar a Pedro, no fue la razón por la que denunció a Pedro. Así, aun cuando sus compañeros acepten la justificación de Juan, la razón aducida por éste no hace de su acción una acción moralmente correcta, no la justifica desde el punto de vista moral precisamente porque no fue la razón por la que Juan actuó.

Podría objetarse que nuestra argumentación descansa en la construcción de un concepto de justificación adecuado a nuestros propósitos. En cierto modo, en *algún* sentido, podría decirse, la razón que Juan ofrece a sus compañeros justifica su acción. Esto es cierto. Hay *algún* sentido en que lo que Juan dice a sus compañeros es una justificación, a saber, que la razón ofrecida por él es formalmente coherente con la acción (cumple C1 o una condición similar). Pero reducir la justificación a este aspecto borra una diferencia muy importante, la diferencia entre una acción moralmente justificada y una que no lo está, como la de Juan. Esta es una diferencia que no hemos construido, sino que utilizamos de hecho en ciertos contextos. La noción de justificación explicativa no es el resultado de una estipulación por nuestra parte, sino un concepto que funciona de hecho en nuestra vida.

Hay, pues, al menos un contexto en que la justificación de una acción sólo se produce si la justificación es explicativa: se trata del contexto de la justificación moral. Para que una acción esté moralmente justificada por una razón o consideración moral que el agente ofrece en su favor exigimos, no sólo que la acción, bajo la descripción que el agente da de ella, sea formalmente coherente con la razón en cuestión, sino que el agente llevase a cabo la acción *por* esa razón, que la razón fuese causalmente efectiva en el proceso que llevó al agente a actuar de ese modo. Aunque el agente considerase esa razón como una razón para actuar así, su acción no está moralmente justificada por ella si la razón no fue efectiva en la realización de la acción.

Creemos, sin embargo, y ésta constituye nuestra tercera consideración, que el concepto de justificación explicativa es también operativo en otros contextos. La justificación moral trata de establecer una determinada relación entre la acción que se pretende justificar y ciertos principios morales presupuestos en la razón que el agente ofrece para su acción. Un aspecto de esta relación es la coherencia de la acción con los principios en cuestión. Pero esto no basta. Pues una acción puede ser coherente con tales principios sin que éstos desempeñasen papel alguno en el proceso que llevó al agente a actuar así. En este último caso, la acción no es realmente moral bajo la perspectiva de tales principios, no constituye una aplicación de los mismos. La justificación de la acción mediante razones, en contextos no morales, ofrece asimismo fundamentos para considerar la acción como racional, adecuada, prudente, etc., a la luz de las razones ofrecidas en su favor y de ciertos principios presupuestos en ellas. Pues bien, si la acción se limita a ser coherente con las razones ofrecidas, sin que éstas llevasen realmente al agente a realizar la acción, ésta no es realmente racional bajo la perspectiva de tales razones, no está realmente justificada por ellas.

Naturalmente, podemos equivocarnos, tanto en contextos morales como no morales, y aceptar la moralidad o la racionalidad de una acción a la luz de ciertas razones ofrecidas por el agente cuando la acción no es en realidad moral o racional a la luz de tales razones.

Si lo dicho hasta aquí es correcto, hay una noción sustantiva de justificación de la acción que incluye la efectividad causal y que es equivalente a la explicación. No es necesario introducir la causalidad como una condición independiente de la justificación, fundada en principios ajenos a la propia justificación, y no estamos obligados a iniciar así el camino que, eventualmente, conduce a Davidson a tropezar con los escollos del epifenomenismo. Nuestra sugerencia es que Davidson, que toma en serio la autoridad del sujeto sobre el contenido de sus razones, que acepta que dichas razones causan la acción y que la acción está justificada por su relación con ellas, debería considerar la posibilidad de abandonar la independencia entre justificación y causalidad, aceptando así nuestra propuesta. Si no me equivoco, conseguiría con ello eludir el epifenomenismo, integrar en su teoría la relevancia causal del contenido y reforzar el realismo intencional ya presente en su obra. Hemos de reconocer, sin embargo, que la propuesta en cuestión, según la cual la justificación incluye la causalidad y no es independiente de ella, no está exenta de problemas y tiene repercusiones cuyo tratamiento no podemos desarrollar aquí. Su defensa incluiría al menos una respuesta a la cuestión de las relaciones entre la explicación intencional y la explicación física de la conducta, así como al problema de la 'exclusión explicativa' de Jaegwon Kim. Sería también importante analizar sus consecuencias para el concepto de causalidad, pues seguramente es incompatible con la validez universal del principio del carácter nomológico de la causalidad. Pero estos asuntos deben quedar para otra ocasión.