

VOLUMEN 18 • NÚMERO 34

# *euph*Yía

REVISTA DE FILOSOFÍA

VARIA

¿Es adecuado el rawlseanismo para plantear los fundamentos constitucionales anti-discriminación?, Enrique Camacho-Beltrán

Hayek and the concept of freedom: a summary, Facundo Guadagno  
Apuntes sobre el concepto de resistencia a la luz de los casos CRAC-PC y EZLN,  
Hugo Martínez García

Sobre el estatus moral de la mujer en Kant: una lectura desde el carácter moral,  
Luis Moises López-Flores

DOSSIER: ÉTICA AMBIENTAL Y ANIMAL

Introducción de los editores, Victor Hugo Salazar Ortiz y Daniel Oviedo Sotelo  
Carencia social de una ética animal: la educación oculta, Elba Castro Rosales y  
Javier Reyes Ruiz

Biopolítica, necropolítica, zootecnia y domexecración: el poder de la muerte,  
Esther Adriana Arvizu Ruiz

Ecologismo y animalismo. De la tensión entre espectros ético-políticos al  
entrecruzamiento de enfoques críticos, Ernesto Cabrera García

Kangaroos and Dragons: how much is a species worth?, Luis David Reyez  
Sáenz

La animalidad en el pensamiento de Nick Land o la línea de fuga hacia lo  
desconocido, Camilo Andrés Vargas Guevara

Ensayo animal, Rodolfo Bernal Escalante

Ser responsable por la naturaleza según Hans Jonas: el futuro de la humanidad  
inseparable de la preservación de la naturaleza, Ericbert Tambou

Kamgue

Máximas éticas frente al desplazamiento y la inmovilidad climáticos, Bernardo  
Bolaños Guerra

TRADUCCIÓN

Malcolm, Norman. "La concebibilidad del mecanicismo", Marc Jiménez-Rolland



ISSN 2683-2518

**Malcolm, Norman. (1968). The  
Conceivability of Mechanism. *The  
Philosophical Review*, 77 (1): 45-72.**

**Sobre el texto y la traducción**

En el núcleo de nuestra imagen de sentido común se encuentra la idea de que los seres humanos somos agentes, es decir, orientamos nuestras acciones a partir de pensamientos, intenciones y deseos. Esta idea parece avalada por nuestra práctica cotidiana de aludir a tales estados mentales en la explicación del comportamiento humano. De este modo, tanto nuestra imagen de sentido común de los seres humanos como nuestra práctica explicativa para dar cuenta de su comportamiento parecen asumir que los estados mentales pueden ser causalmente responsable del movimiento corporal, el cual tiene efectos adicionales sobre el mundo físico. En este texto, Norman Malcolm examina si disponer de una teoría neurofisiológica completa podría llevarnos a abandonar esta suposición. Aunque “dicha teoría no existe en la actualidad” (§2), muchos esfuerzos científicos se encaminan a su consecución. Si tal teoría llegase a existir, ¿ofrecería bases científicas para una dramática revisión de nuestra imagen de sentido común de los seres humanos?, ¿pondría en entredicho nuestra práctica

explicativa del comportamiento humano? Malcolm argumenta que nuestras explicaciones cotidianas del comportamiento humano dependen de principios a priori que conectan a lo mental con la acción. Tales principios enuncian, por ejemplo, que si una persona desea lograr algo y cree que se requiere realizar cierta acción para lograrlo, entonces realizará dicha acción en caso de que nada se lo impida. Cuando no se presentan impedimentos, equiparamos el hecho de que no se realice la acción con que en realidad la persona no creía que dicha acción fuese requerida para alcanzar sus propósitos o no deseaba la consecución de dichos propósitos. Esta equiparación sugiere que los principios que conectan a ciertos estados mentales con la acción son verdades conceptuales. Si es así, tales principios no podrían ser falsos. En consecuencia, ni siquiera una teoría neurofisiológica completa podría mostrar que los principios en la base de nuestras explicaciones psicológicas cotidianas son falsos. ¿Significa esto que una teoría así sería compatible con nuestra imagen de sentido común? Malcolm sostiene que no lo sería. Puesto que las explicaciones proporcionadas por una teoría neurofisiológica completa dependerían de leyes contingentes, no de verdades conceptuales, tales explicaciones estarían en competencia con las explicaciones psicológicas. Para respaldar esta afirmación, Malcolm introduce un supuesto que se ha vuelto extremadamente influyente en el debate sobre la causación mental: el 'principio de exclusión causal'. Si las leyes neurofisiológicas fuesen suficientes para explicar la acción, entonces podría prescindirse por completo de las explicaciones psicológicas. Puesto que en las explicaciones neurofisiológicas solo los sucesos físicos son causalmente responsables de la acción, los estados mentales no desempeñarían ningún papel en la producción (y, por ende, en la explicación) de

la acción. Si a la incompatibilidad entre explicaciones psicológicas y explicaciones neurofisiológicas se suma que las primeras involucran verdades conceptuales, pareceríamos estar frente a la refutación a priori de la posibilidad de una teoría neurofisiológica completa. No obstante, la conclusión de Malcolm es más moderada. Él afirma que tal teoría es inconcebible, es decir, que no podemos siquiera vislumbrar que sea verdadera. Si lo fuese, no podría haber acciones intencionales y “una vida que estuviese totalmente desprovista de actividades ciertamente no podría ser una vida humana” (§20). Pese a ser inconcebible, la verdad de una teoría así sigue siendo una posibilidad lógica. Pero “el acto humano de aseverarla [como] doctrina” (§22), el mecanicismo, es absurdo. El mecanicismo tiene la peculiaridad de que, si es verdadero, nadie puede afirmarlo: “presenta una antinomia hostil, y quizá insoluble, al pensamiento humano” (§27). Además de ofrecer una estimulante revitalización de la perenne discusión filosófica en torno al determinismo, este artículo de Norman Malcolm hizo contribuciones novedosas a los debates sobre la explicación de la acción y la causación mental que avalan la pertinencia de disponer de una versión del texto en nuestra lengua. Se solicitó autorización a Diane Grossé, encargada de derechos y permisos de la Duke University Press, para publicar la presente traducción. Ella indicó que, cuando el artículo fue publicado en *The Philosophical Review* –actualmente a cargo de dicha casa editorial–, no se tenía registro de derechos de autor, por lo que el texto se considera de dominio público. Agradecemos la información oportuna por parte de esta institución. La traducción y el párrafo introductorio fueron elaborados por Marc Jiménez-Rolland.

## **La concebibilidad del mecanicismo**

Norman Malcolm

1. Por “mecanicismo” entenderé una aplicación especial del determinismo físico –a saber, a todos los organismos con sistemas neurológicos, incluyendo a los seres humanos. La versión del mecanicismo que deseo estudiar asume una teoría neurofisiológica, la cual es adecuada para explicar y predecir todos los movimientos de los cuerpos humanos excepto aquellos que son causados por fuerzas externas. Se asume que el cuerpo humano es un sistema causal tan completo como lo es un motor de gasolina. Los estados y procesos neurológicos son concebidos como correlacionados por leyes generales con los mecanismos que producen movimientos. Se asume que los cambios eléctricos y químicos en el tejido nervioso del cuerpo causan contracciones musculares, las cuales a su vez causan movimientos tales como el parpadeo, la respiración, el apretar los labios, así como los movimientos de los dedos, las extremidades y la cabeza. Dichos movimientos en ocasiones son producidos por fuerzas (empujones y tirones) aplicadas externamente al cuerpo. Si alguien forzara mi brazo a moverse encima de mi cabeza, la teoría no podría explicar el movimiento de mi brazo. Pero podría explicar cualquier movimiento que no se deba a un empujón o tirón externo. Podría explicar, y predecir, los movimientos que ocurren cuando una

persona le hace señas a un taxi, juega ajedrez, escribe un ensayo o entra a una tienda.<sup>1</sup>

Se asume que el sistema neurofisiológico del cuerpo humano está sujeto a varias clases de estimulación. Cambios de temperatura o presión en el entorno; sonidos, olores; la ingesta de alimentos o líquidos: todo esto tendrá un efecto en los impulsos nerviosos que ponen en marcha los mecanismos productores de movimiento del cuerpo.

2. La teoría neurofisiológica que imaginamos sería, como se dijo, lo suficientemente fértil para proporcionar explicaciones causales sistemáticas de todos los movimientos corporales que no se deban a causas físicas externas. Estas explicaciones deberían entenderse como enunciando condiciones *suficientes* del movimiento y no meramente condiciones necesarias. Emplearían leyes que conectaran estados o procesos neurofisiológicos con movimientos. Las leyes serían proposiciones universales de la siguiente forma: siempre que un organismo de la estructura  $E$  se encuentre en el estado  $q$  emitirá un movimiento  $m$ . Al determinar que un organismo específico tiene la estructura  $E$  y se encuentra en el estado  $q$ , uno podría deducir la incidencia del movimiento  $m$ .

Debe enfatizarse que esta teoría no contempla deseos, objetivos, metas, propósitos, motivos o intenciones. Al explicar acontecimientos tales como el caminar de un hombre a través de una habitación, será indiferente para la teoría si el propósito, la intención o el deseo del hombre era abrir una ventana, o incluso si

---

<sup>1</sup> Si dijese "¡Levántate!" y me levantara, la teoría explicaría mis movimientos en términos de sucesos neurofisiológicos producidos por el impacto de las ondas sonoras en mis órganos auditivos.

su caminar a través de la habitación era intencional. Este aspecto de la teoría puede ser indicado al decir que es un sistema de explicación “carente de propósitos”.

El punto de vista del mecanicismo, de este modo, asume una teoría que proporcionaría explicaciones causales sistemáticas, completas y carentes de propósitos de todos los movimientos humanos que no sean producidos por fuerzas externas. Dicha teoría no existe en la actualidad. Pero en nuestros días cada vez son más quienes sostienen que en un futuro no muy distante tal teoría existirá –y que será verdadera. Plantearé la pregunta de si esto es concebible. El tema pertenece a una antigua controversia. No sería realista que esperara hacer alguna contribución sobresaliente a su solución. Pero el problema es en sí mismo de gran interés humano y digno de reiterado estudio.

3. Para apreciar la importancia del mecanicismo, uno debe estar consciente de la medida en la cual una teoría neurofisiológica comprehensiva de la conducta humana diferiría de las explicaciones cotidianas de la conducta con las que todos nosotros estamos familiarizados. Estas explicaciones refieren a propósitos, deseos, metas e intenciones. “Él corre para tomar el autobús”. “Él sube la escalera con el fin de revisar el techo”. “Él se detiene en la tienda porque quiere unos cigarrillos”. Nuestro discurso diario está lleno de explicaciones de la conducta en términos de los propósitos o intenciones del agente. Se afirma que la conducta ocurre con el fin de que algún estado de cosas sea provocado o evitado –que se tome el camión, se revise el techo, se compren los cigarrillos. Digamos que éstas son explicaciones “mediante propósitos”.

Podemos notar varias diferencias entre estas explicaciones comunes mediante propósitos y las explicaciones neurofisiológicas imaginadas. Primera, las últimas fueron concebidas por nosotros como sistemáticas –esto es, pertenecientes a una teoría comprensiva–, mientras que las explicaciones familiares mediante propósitos no están organizadas en una teoría. Segunda, las explicaciones neurofisiológicas no emplean el concepto de propósito o intención. Tercera, las explicaciones neurofisiológicas plasman leyes contingentes, pero las explicaciones mediante propósitos no lo hacen.

Detengámonos en este tercer punto. Se asume que una explicación neurofisiológica de cierta conducta que ha ocurrido tiene la siguiente forma:

Siempre que un organismo  $O$  de estructura  $E$  se encuentra en el estado neurofisiológico  $q$  producirá el movimiento  $m$ .

El organismo  $O$  de estructura  $E$  estaba en el estado neurofisiológico  $q$ .

Por tanto,  $O$  produjo  $m$ .<sup>2</sup>

La forma general de la explicación mediante propósitos es la siguiente:

Siempre que un organismo  $O$  tiene la meta  $M$  y cree que la conducta  $C$  se requiere para provocar  $M$ ,  $O$  producirá  $C$ .

$O$  tuvo  $M$  y creyó que  $C$  se requería para  $M$ .

Por tanto,  $O$  produjo  $C$ .

---

<sup>2</sup> Una *predicción* neurofisiológica tendría la misma forma, con las siguientes diferencias: la segunda premisa diría que  $O$  está o estará en el estado  $q$  (en lugar de que *estaba*), y la conclusión diría que  $O$  *producirá*  $m$  (en lugar de que *produjo*).

Comparemos la primera premisa de una explicación neurofisiológica con la primera premisa de una explicación mediante propósitos. La primera premisa de una explicación neurofisiológica es una proposición contingente, pero la primera premisa de una explicación mediante propósitos no es una proposición contingente. Esta diferencia se presentará más claramente si consideramos cómo, en ambos casos, la primera premisa tendría que ser limitada para ser de hecho verdadera. En ambos casos debe añadirse una cláusula *ceteris paribus* a la primera premisa, o al menos debe entenderse implícitamente. (Será más claro traducir "*ceteris paribus*" como "siempre y cuando no haya factores que lo impidan" en lugar de como "siendo todo lo demás igual".)

Consideremos qué significará "*ceteris paribus*" en términos concretos. Supóngase que un hombre subió una escalera que conducía a un techo. Se desea una explicación. El hecho es que el viento voló su sombrero al techo y él lo quiere de vuelta. La explicación sería presentada en detalle como sigue:

Si un hombre quiere recuperar su sombrero y cree que esto requiere que suba una escalera, él lo hará siempre y cuando no haya factores que lo impidan.

Este hombre quería recuperar su sombrero y creía que esto requería que subiera una escalera, y no había factores que lo impidieran.

Por tanto, él subió la escalera.

¿Qué clases de cosas podrían incluirse como "factores que lo impidan" en un caso como éste? Que no haya una escalera disponible, el miedo de subir a una, la creencia de que alguien quitaría la escalera mientras uno está en el techo, y así

sucesivamente. (Que el hombre no suba una escalera *no* sería un factor que lo impidiera.)

Un punto importante que surge aquí es que añadir la cláusula *ceteris paribus* a la primera premisa convierte a esta premisa en una proposición *a priori*. Si no hubiese cualesquiera factores que lo impidieran (si el hombre supiese que la escalera estaba disponible, si no temiera a las escaleras o los lugares altos, ni creyera que podría ser abandonado en el techo, y así sucesivamente); si no hubiese obstáculos o peligros, reales o imaginados, físicos o psicológicos; entonces si el hombre no subiera una escalera no sería verdadero que él *quería* su sombrero de vuelta o *tenía la intención* de recuperarlo.<sup>3</sup>

En su importante libro reciente, *The Explanation of Behaviour*, Charles Taylor presenta el punto como sigue:

Esto es parte de lo que queremos decir con “tener la intención de X”, que, en ausencia de factores que interfirieran, es seguido por hacer X. No podría decirse que tengo la intención de X si, incluso sin obstáculos u otros factores que lo impidan, aun así no lo hiciera (Taylor, 1964: 33).

Este rasgo del significado de “tener la intención” también es verdadero de “querer”, “propósito” y “meta”.

---

<sup>3</sup> El diagnóstico correcto de una falla como ésta no será evidente en todos los casos. Supóngase que un joven quiere ser un trapeartista en un circo, y cree que esto requiere ejercicio diario en las barras paralelas. Pero es holgazán y frecuentemente no se ejercita. ¿Acaso no tiene realmente la meta que dice tener?, ¿es sólo habladoría? ¿O acaso no cree realmente en la necesidad del ejercicio diario? ¿O es que tiene la meta y la creencia, pero su holgazanería es un genuino impedimento? Uno tendría que conocerlo muy bien con el fin de dar la respuesta correcta. En algunos casos podría no haber una respuesta correcta evidente.

De este modo, la premisa universal de una explicación mediante propósitos es un principio *a priori*, no una ley contingente. Algunos filósofos han tomado esto como una base para decir que una explicación mediante propósitos no es una explicación causal.<sup>4</sup> Pero ésta es una estipulación (quizá una útil), más que una descripción de cómo se usa de hecho la palabra “causa” en el lenguaje ordinario.

Consideremos el efecto de añadir una cláusula *ceteris paribus* a la premisa universal de una explicación neurológica de la conducta. ¿Sería una premisa de esta forma verdadera *a priori*? Ciertamente no. Supóngase que se creyera que siempre que un ser humano se encuentra en el estado neuronal  $q$  su mano derecha se levantará por encima de su cabeza, siempre y cuando no haya factores que lo impidan. ¿Qué factores podrían impedirlo? Que el brazo derecho del sujeto esté roto o que esté atado a su costado, y situaciones similares. Pero la exclusión de tales factores que lo impidan no tendería a hacer a la premisa verdadera *a priori*. No hay conexión de significado, explícita o implícita, entre la descripción de cualquier estado neuronal y la descripción de cualquier movimiento de la mano. Sin importar cuántos factores que lo impidan sean excluidos, la proposición no perderá su carácter de una ley contingente (a menos, por supuesto, que contemos al hecho mismo de que la mano no se mueva como un factor que lo impidiera, caso en el que la premisa se convierte en una tautología).

---

<sup>4</sup> E.g., Taylor (1964) dice que la intención del agente no es un “antecedente causal” de su conducta, pues la intención y la conducta “no están contingentemente conectadas de la forma normal”.

4. Hacer explícitas las condiciones *ceteris paribus* destaca las distintas naturalezas lógicas de las premisas universales de las dos clases de explicación. Premisas de un tipo expresan correlaciones contingentes entre los procesos neurológicos y la conducta. Premisas del otro tipo expresan conexiones *a priori* entre las intenciones (propósitos, deseos, metas) y la conducta.

Esta diferencia es de la mayor importancia. Algunos estudiosos de la conducta han creído que se descubrirá que las explicaciones mediante propósitos de la conducta son menos básicas que las explicaciones que surgirán de una teoría neurofisiológica. Piensan que los principios de la explicación mediante propósitos resultarán ser dependientes de las leyes neurofisiológicas. Desde esta perspectiva nuestras explicaciones ordinarias de la conducta a menudo serán verdaderas: pero las explicaciones neuronales también serán verdaderas –y serán *más fundamentales*. De modo que podríamos, teóricamente, *eludir* las explicaciones de la conducta en términos de propósitos, y podría llegar el día en el que simplemente cayeran en desuso.

Me gustaría mostrar que las leyes neurofisiológicas no podrían ser más básicas que los principios mediante propósitos. Entenderé el enunciado de que una ley  $L_2$  es “más básica” que una ley  $L_1$  de modo que signifique que  $L_1$  depende de  $L_2$ , pero  $L_2$  no depende de  $L_1$ . Por dar un ejemplo, supongamos que hay una conexión uniforme entre la abstinencia de alimentos y el hambre: esto es, pasar  $n$  horas sin alimentos siempre tiene como resultado el hambre. Ésta es  $L_1$ . Otra ley  $L_2$  se descubre –a saber, una conexión uniforme entre cierta condición química del tejido corporal (llamada “inanición celular”) y el hambre. Siempre que ocurre la inanición celular, tiene como consecuencia el hambre. También se descubre que  $L_2$  es más básica que  $L_1$ . Esto equivaldría al siguiente

hecho: la abstinencia de alimentos durante  $n$  horas no tendrá como consecuencia el hambre *a menos que* ocurra la inanición celular; si esta última ocurre, el hambre resultará *sin importar si* ocurre la abstinencia de alimentos. De modo que la regularidad  $L_1$  depende contingentemente de la regularidad  $L_2$ , y la converso no es verdadera. Nuestro conocimiento de esta dependencia nos revelaría las condiciones bajo las cuales la regularidad  $L_1$  no se sostendría.

Nuestra comparación de las distintas naturalezas lógicas de los principios mediante propósitos y las leyes neurofisiológicas nos permite ver que los primeros no pueden depender de las últimas. La conexión *a priori* entre intención o propósito y conducta no puede dejar de sostenerse. No puede ser dependiente contingentemente de ninguna regularidad contingente. Las explicaciones neurofisiológicas de la conducta no podrían, en el sentido explicado, resultar ser más básicas que nuestras explicaciones mediante propósitos cotidianas.<sup>5</sup>

5. Hay una segunda consecuencia importante de la diferencia lógica entre leyes neurofisiológicas y principios mediante

---

<sup>5</sup> Taylor presenta el punto como sigue:

Puesto que la explicación por intenciones o propósitos es como una explicación por un "antecedente" que está conectado de manera no contingente con su consecuente, *i.e.*, debido a que el hecho de que la conducta se sigue de la intención siendo las demás cosas igual no es un hecho contingente, no podemos dar cuenta de este hecho mediante leyes más básicas. Pues explicar un hecho mediante leyes más básicas es proporcionar las regularidades de las que depende causalmente este hecho. Pero al no ser contingente, el que la conducta dependa de la intención no es contingente sobre nada, y por ende no depende de ninguna de tales regularidades (1964: 44).

propósitos. Alguien podría suponer que, aunque las explicaciones mediante propósitos no pueden depender de explicaciones carentes de propósitos, serían refutadas por medio de la verificación de una teoría neurofisiológica comprensiva de la conducta. Creo que este punto de vista es correcto: pero es necesario comprender lo que *no puede* significar. No puede significar que se demostraría que los principios (las premisas universales) de las explicaciones mediante propósitos son falsos. No puede demostrarse que son falsos. No puede no ser verdadero que si una persona quería *X* y creía que *Y* era necesaria para *X*, y no hubo ningún factor que lo impidiera, él lo haría.<sup>6</sup> Este principio mediante propósitos es verdadero *a priori*, no debido a su forma sino debido a su significado –esto es, debido a la conexión de significado entre las palabras “Él quería *X* y se percató de que *Y* era necesario para *X*” y las palabras “Él hizo *Y*”. El principio mediante propósitos no es una ley de la naturaleza, sino una verdad conceptual. No puede ser confirmado o refutado por la experiencia. Puesto que la verificación de una teoría neurofisiológica nunca podría *refutar* ningún principio mediante propósitos, el único posible resultado de tal verificación, lógicamente hablando, sería demostrar que los principios mediante propósitos no tienen aplicación en el mundo. Volveré sobre este punto más adelante.

6. Debemos comprender con más precisión la relación lógica exacta entre las explicaciones neuronales de la conducta y las que

---

<sup>6</sup> Esto es verdadero si usamos “quiere *X*” para decir “apunta hacia *X*”. Pero a veces podemos no querer decir más que “le gustaría tener *X*”, lo que puede representar un mero anhelo.

emplean propósitos. ¿Pueden ser verdaderos ambos tipos de explicaciones del mismo fragmento de conducta en una y la misma ocasión? ¿Hay alguna rivalidad entre ellas? Algunos filósofos dirían que no. Dirían eso, por una parte, porque las dos clases de explicación explican cosas distintas. Las explicaciones mediante propósitos explican acciones. Las explicaciones neurofisiológicas explican movimientos. Por otra parte, podría sostenerse que las dos clases de explicaciones pertenecen a diferentes “cuerpos de discurso” o diferentes “juegos de lenguaje”. Emplean diferentes conceptos y suposiciones. Una clase de explicación relaciona a la conducta con leyes causales y con conceptos de la bioquímica y la fisiología, con impulsos nerviosos y reacciones químicas. La otra clase de explicación relaciona a la conducta con los deseos, intenciones, metas y razones de las personas. Las dos formas de explicación pueden coexistir, puesto que son irrelevantes la una hacia la otra.<sup>7</sup>

Es verdad que las dos clases de explicación emplean diferentes conceptos y, en un sentido, explican cosas distintas: pero, ¿son realmente independientes una de otra? Considérese el

---

<sup>7</sup> Las siguientes observaciones de A. I. Melden presentan ambos puntos:

Donde nos ocupamos de explicaciones causales, de sucesos de los cuales los acontecimientos en cuestión son efectos con alguna ley de causalidad, en esa medida no nos ocupamos de las acciones humanas en absoluto, sino, en el mejor de los casos, de movimientos u acontecimientos corporales; y donde nos ocupamos de explicaciones de la acción humana, ahí los factores causales y las leyes causales en el sentido en el que, por ejemplo, estos términos son empleados en las ciencias biológicas, son completamente irrelevantes para la comprensión que buscamos. La razón es simple, a saber, las características lógicas radicalmente diferentes de los dos cuerpos de discurso que empleamos en estos distintos casos –los diferentes conceptos que son aplicables a estos distintos órdenes de investigación (1961: 184).

ejemplo de un hombre que sube una escalera con el fin de recuperar su sombrero del techo. Esta explicación relaciona su subir con su intención. Una explicación neurofisiológica del que subiera no diría nada acerca de su intención, sino que conectaría sus movimientos sobre la escalera con cambios químicos en el tejido corporal o con la activación de neuronas. ¿Interfieren las dos explicaciones una con otra?

7. Creo que *habría* una colisión entre las dos explicaciones si fuesen ofrecidas como explicaciones de uno y el mismo acontecimiento de un hombre subiendo una escalera. Recordaremos que se suponía que la teoría neurofisiológica imaginada proporcionaba explicaciones causales *suficientes* de la conducta. De este modo, los movimientos del hombre en la escalera serían *completamente* explicados en términos de los procesos eléctricos, químicos y mecánicos en su cuerpo. Esto seguramente implicaría que su deseo o intención de recuperar su sombrero no tuvieron nada que ver con su movimiento de ascender la escalera. Implicaría que en esta misma ocasión él habría ascendido la escalera exactamente de la misma manera incluso si no hubiese tenido la intención de recuperar su sombrero, o incluso sin tener la intención de subir la escalera. Dados los estados neurológicos antecedentes de su sistema corporal junto con leyes generales que correlacionan esos estados con las contracciones de los músculos y los movimientos de las extremidades, él se habría movido como lo hizo sin importar su deseo o intención. Si cada uno de sus movimientos fuese completamente explicado por sus estados neurofisiológicos antecedentes (su “programación”), entonces no era verdad que esos movimientos ocurrieron *debido* a que él quería o tenía la intención de ir por su sombrero.

8. Consideraré brevemente tres posibles objeciones a mi afirmación de que si el mecanicismo fuese verdadero el hombre habría subido la escalera tal como lo hizo incluso si no hubiese tenido ninguna intención de subir la escalera. La primera objeción proviene de un filósofo que apoya actualmente la popular tesis de la identidad psicofísica. Él sostiene que hay una condición neuronal que causa los movimientos del hombre de ascender la escalera, y además sostiene que la intención del hombre de subir la escalera (o, posiblemente, el que tenga la intención) es contingentemente idéntico con la condición neuronal que causa los movimientos. De este modo, si el hombre no tuviese la intención de subir la escalera, la causa de sus movimientos no habría existido, por lo que esos movimientos no habrían ocurrido. Mi respuesta sería que la concepción de que puede haber una identidad contingente (y no meramente una equivalencia extensional) entre una intención (o el tener la intención) y una condición neuronal no es una hipótesis con significado. Una versión de la tesis de la identidad es que la intención de *A* de subir la escalera es contingentemente idéntica con algún proceso en el cerebro de *A*. Verificar esta identidad requeriría el paso carente de significado de intentar descubrir si la intención de *A* está ubicada en su cerebro. Uno podría darle significado a la noción de la ubicación de la intención de *A* en su cerebro al estipular que tiene la misma ubicación que el proceso neuronal correlacionado. Pero la identidad que surgió de esta estipulación no sería contingente.<sup>8</sup> Otra versión de la tesis de la identidad es que el suceso de que

---

<sup>8</sup> Este punto es defendido en mi (1964); también en mi monografía (sec. 18).

Smith tuviera la intención *I* es idéntico con el suceso de que Smith se encontrara en la condición neuronal *N*. Esta versión evita el “problema de la ubicación” arriba planteado: pero debe emprender la tarea (que parece desesperada) de explicar cómo la propiedad de “tener la intención *I*” y la propiedad de “estar en la condición neuronal *N*” podrían ser contingentemente idénticas y no meramente coextensivas.<sup>9</sup>

La segunda objeción proviene de un epifenomenalista. Él sostiene que la condición neurofisiológica que causa contingentemente la conducta en la escalera también causa contingentemente la intención de subir la escalera, pero que la intención no se encuentra en ninguna relación causal con la conducta. Si la intención no hubiese existido, su causa y la de la conducta no habrían existido, de modo que la conducta no habría ocurrido. Una objeción decisiva al epifenomenalismo es que, de acuerdo con él, la relación entre la intención y la conducta sería puramente contingente. Sería concebible que la condición neurofisiológica que siempre causa movimientos de subir escaleras siempre causara también la intención de *no* subir una escalera. El epifenomenalismo permitiría que fuese universalmente verdadero que siempre que cualquier persona tuviese la intención de *no* realizar alguna acción, la realizara, y que siempre que cualquier persona tuviese la intención de realizar alguna acción, no la realizara. Esto es un absurdo conceptual.

La tercera objeción surge de un filósofo que combina el mecanicismo con el conductismo lógico. Él sostiene que alguna condición de los sistemas neurofisiológicos causa los

---

<sup>9</sup> Para una exposición de este problema véase el artículo de Jaewgon Kim (1966).

movimientos, gestos y preferencias preparatorios que son expresiones de la intención del hombre de subir la escalera; y también causa sus movimientos de ascender la escalera. El componente del conductismo lógico en la totalidad de su concepción es éste: sostiene que el que el hombre tenga la intención de subir la escalera es simplemente una construcción lógica de la ocurrencia de las expresiones de la intención junto con la ocurrencia de movimientos de ascender la escalera. Tener la intención no es nada más que la conducta expresiva aunada al subsiguiente ascender la escalera. Tener la intención es definido en términos de sucesos-conducta que son contingentemente causados por una condición neurofisiológica. La suposición de que el hombre no tenía la intención de subir la escalera sería idéntica con la suposición de que ya sea la conducta expresiva o la conducta de subir, o ambas, no ocurrieron. Si cualquiera de las dos no ocurre, entonces ninguna ocurrió, puesto que por hipótesis ambas tienen la misma causa. De este modo sería falso que el hombre hubiera ascendido la escalera como lo hizo incluso si no hubiera tenido una intención de subir la escalera.

Creo que esta tercera posición ofrece una explicación insatisfactoria de la naturaleza de la intención. De hecho, subir la escalera no es una condición necesaria *simpliciter* para la existencia de la intención de subir la escalera. Es una condición necesaria *siempre y cuando* no haya factores que lo impidan. Pero no hay un número definido de factores que lo impidan, y, por ende, no pueden ser exhaustivamente enumerados. Aunado a esto, algunos de ellos involucrarán en sí los conceptos de deseo, creencia o propósito. Por ejemplo: un hombre tiene la intención de subir la escalera, pero además no quiere verse ridículo; cuando está a punto de subir le sorprende la idea de que lucirá ridículo; de modo

que no sube la escalera, aunque tenía la intención de hacerlo. Un conductismo lógico adecuado tendría que analizar no sólo la referencia inicial a la intención, sino también la referencia al deseo, la creencia, el propósito y todos los otros conceptos psicológicos, que aparecerían al enlistar los posibles factores que lo impedirían. No hay razón para pensar que tal programa de análisis pueda emprenderse.

De esta forma, un mecanicista puede esperar que se evite la consecuencia de que el hombre habría subido la escalera como lo hizo incluso si no hubiese tenido la intención de subir la escalera, al combinar su doctrina mecanicista con la tesis de la identidad psicofísica, o con el epifenomenalismo, o con el conductismo lógico. Pero estas posiciones suplementarias son tan objetables o inverosímiles que no salvan realmente al mecanicista de la consecuencia anterior.

9. Recordemos que la teoría neurofisiológica postulada es comprensiva. Se asume que proporciona explicaciones causales completas para todos los movimientos corporales que no son producidos por fuerzas físicas externas. Es un sistema cerrado en el sentido de que no admite, como condiciones antecedentes, nada que no sea estados y procesos neurofisiológicos. Los deseos e intenciones no tienen lugar en ella.

Si la teoría neurofisiológica fuese verdadera, entonces en ningún caso los deseos, intenciones y propósitos serían condiciones necesarias para ningún movimiento humano. Nunca sería verdadero que un hombre *no* se habría movido como lo hizo si *no* hubiera tenido tal y cual intención. Tampoco sería verdadero que cierto movimiento suyo se debiera a, fuese producido o causado por, el que él tuviera una cierta intención o propósito. Las

explicaciones mediante propósitos de los movimientos corporales humanos *nunca* serían verdaderas. Los deseos y las intenciones no serían incluso causas potenciales de los movimientos humanos en el mundo real (por oposición con algún mundo posible en el cual la teoría neurofisiológica no fuese verdadera).

Podría pensarse que existiesen dos sistemas distintos de explicaciones causales de los movimientos humanos: un sistema mediante propósitos y un sistema neurofisiológico. Las condiciones antecedentes en uno de los sistemas serían los deseos y las intenciones de los seres humanos; en el otro serían los estados y procesos neurofisiológicos de esos mismos seres humanos. Cada sistema proporcionaría explicaciones causales adecuadas de los mismos movimientos.

Hablando de manera general, es posible que haya una pluralidad de condiciones causales suficientes simultáneas para un suceso. Pero si tenemos presente el aspecto comprensivo de la teoría neurofisiológica –esto es, el hecho de que proporciona condiciones causales suficientes para todos los movimientos– veremos que los deseos e intenciones no podrían ser causas de los movimientos. Como se ha notado con frecuencia, decir que *B causa C* no significa meramente que siempre que *B* ocurre, *C* ocurre. La causalidad tiene también implicaciones subjuntivas y contrafácticas: si *B ocurriese*, *C ocurriría*; y si *B no* hubiese ocurrido, *C no* habría ocurrido. Pero la teoría neurofisiológica proporcionaría condiciones causales suficientes para cualquier movimiento humano, y así no habría casos en los cuales cierto movimiento no habría ocurrido si la persona no hubiese tendido ese deseo o intención. Puesto que el contrafáctico sería falso en todos los casos, los deseos e intenciones no serían causas de los

movimientos humanos. Nunca serían condiciones causales suficientes ni serían jamás condiciones causales necesarias.

10. Afrontemos este punto inmensamente importante desde un ángulo distinto. Muchas descripciones de la conducta adscriben acciones a personas: dicen que alguien *hizo* algo –por ejemplo, “Él firmó el cheque”, “Tú levantaste la mesa”, “Ella rompió el jarro”. Una adscripción de una “acción” a una persona implica dos cosas:<sup>10</sup> primera, que cierto estado de cosas llegó a existir (el que su firma esté presente en el cheque, el que la mesa sea levantada, el que el jarro esté roto); segunda, que la persona tuvo la intención de que este estado de cosas ocurriera. Si el estado de cosas mencionado no llegó a existir (por ejemplo, el jarro no se rompió), entonces la adscripción de la acción (“Ella rompió el jarro”) debe ser retractada. Si en efecto llegó a existir, pero sin la intención de la persona, entonces la adscripción de la acción a la persona debe ser disminuida por alguna cualificación como “no intencionalmente” o “accidentalmente” o “por error” o “inadvertidamente”, siendo cuestión de las circunstancias cuál cualificación es más apropiada. Una adscripción cualificada como acción aún implica que la persona desempeñó algún papel en el surgimiento del estado de cosas –por ejemplo, su mano golpeó el jarro. Si no desempeñó ningún papel en absoluto, entonces no puede decirse correctamente, incluso con cualificaciones, que ella rompió el jarro.

Supóngase que un hombre pretende abrir la puerta que está frente a él. Gira la perilla y la puerta se abre. Puesto que girar la

---

<sup>10</sup> Sigo aquí a Charles Taylor (1964: 27-32).

perilla es lo que normalmente causa que la puerta se abra, deberíamos pensar que es correcto decir que *él* abrió la puerta. Luego descubrimos que hay un mecanismo eléctrico oculto en la puerta que causó que la puerta se abriera al momento en que *él* giró la perilla, y además que no hay ninguna conexión causal entre el girar la perilla y la operación del mecanismo. De modo que su acto de girar la perilla no tuvo nada que ver con que se abriera la puerta. Ya no podemos decir que *él* abrió la puerta: nada de lo que hizo tuvo ninguna influencia causal sobre ese resultado. Podríamos poner el asunto de esta manera: debido a la operación del mecanismo eléctrico *él* no tuvo oportunidad de abrir la puerta.

El hombre de nuestro ejemplo podría decir que al menos *él* giró la perilla. Debería renunciar a esta afirmación, no obstante, si saliera a la luz que otro mecanismo eléctrico causó que la perilla girara cuando lo hizo, independientemente del movimiento de su mano. El hombre podría afirmar que, en todo caso, *él* movió su mano. Pero ahora la teoría neurofisiológica entra en escena, proporcionando una explicación causal completa del movimiento de su mano, sin considerar su intención.

El problema de qué decir se vuelve agudo. ¿Deberíamos negar que *él* movió su mano? ¿Deberíamos admitir que *él* movió su mano, pero con alguna cualificación? ¿O deberíamos decir, sin cualificación, que *él* movió su mano?

11. Hay una similitud y una diferencia importantes entre nuestros tres ejemplos. La similitud es que en los tres casos un mecanismo produjo los estados de cosas deseados, y nada de lo que el agente hizo tuvo influencia alguna sobre la operación del mecanismo. Pero hay una diferencia entre los casos. En cada uno de los dos primeros, podemos especificar algo que el hombre hizo (una

acción) que normalmente causa que el resultado deseado ocurra, pero que no tuvo ese efecto en esa ocasión. La acción en el primer caso fue girar la perilla, y en el segundo fue sujetar la perilla y hacer un movimiento de girar con la mano. En cada uno de estos casos hubo una acción, cuya eficacia causal fue nulificada por la operación de un mecanismo. En consecuencia, podemos decir correctamente que la acción del hombre *no logró* hacer una contribución a la ocurrencia deseada, y así podemos negar que *él* abriera la puerta o girara la perilla.

¿Hay algo que hizo el hombre en el tercer caso que normalmente causa el movimiento de la mano? ¿Qué fue? ¿Hay algo más que *yo hago* que cause que mi mano se mueva cuando muevo mi mano de la manera normal? No. Ocurren varios sucesos en mi cuerpo (por ejemplo, impulsos nerviosos), pero no puede decirse que sean mis *acciones*. No son cosas que yo hago.

Pero en el tercer caso el hombre *se proponía* hacer un movimiento de girar con su mano. ¿Es ésta una base de similitud entre el tercer caso y los dos primeros? ¿Puede decirse que la intención de mover la propia mano es normalmente una causa del movimiento de la mano, pero que en nuestro tercer caso la eficacia causal de la intención fue nulificada por la operación del mecanismo neurofisiológico?

Sobre la cuestión de si las intenciones son causas de las acciones, Taylor dice algo que es tanto interesante como intrigante. Señala que llamar a algo una acción, en un sentido no cualificado, “significa no sólo que el hombre que exhibió esta conducta se había formado la intención relevante o había tenido este propósito, sino

también que *su proponérselo lo produjo*" (1964: 33. *Cursivas más*).<sup>11</sup> Ahora bien, decir que A "produjo" B es usar el lenguaje de la causación. "Produce" es, en efecto, un sinónimo de "causa".

12. ¿Hay algún sentido en el que la intención de un hombre de hacer algo puede ser la causa de que lo haga? Al lidiar con este punto usaré la palabra "causa" en el sentido más amplio, de acuerdo con el cual lo que sea que explique, o parcialmente explique, la ocurrencia de alguna conducta es la causa, o parte de la causa, de la conducta. Saber que un hombre se proponía subir una escalera no explicaría, en muchos casos, por qué la subió. No explicaría para qué la subió, cuál era su razón o propósito para subirla, mientras que decir cuál era su propósito daría, en el sentido más amplio, la causa o parte de la causa de que la subiera.

Al considerar la intención como una causa de la conducta X, es importante distinguir entre la intención de hacer X (llamemos a esta *intención simple*) y la intención de hacer algo más al hacer X (llamemos a esta *intención adicional*). Decir que un hombre se proponía subir una escalera usualmente no daría una causa de que la subiera; pero enunciar su propósito al subirla usualmente sería dar la (o una) causa de la acción. Es un uso natural del lenguaje preguntar "¿Qué causó que subieras la escalera?"; y es una respuesta apropiada decir "Quería recuperar mi sombrero". (*Pregunta*: "Por todos los cielos, ¿qué causó que votaras sólo por

---

<sup>11</sup> Taylor dice que una intención *no* es "un antecedente causal" de la conducta deseada, puesto que intención y conducta no están *conectadas contingentemente*. Creo que podría representársele con justicia como sosteniendo que una intención sí *causa* la conducta deseada, aunque no en el sentido de "causa" en el que una causa y un efecto están contingentemente correlacionados.

candidatos Republicanos?” *Respuesta*: “Quería restaurar el sistema bipartidista.”) Nuestro uso del lenguaje de la causación no se restringe a casos en los que se asume que causa y efecto están relacionados contingentemente.

13. ¿La intención simple de hacer X puede en algún caso ser una causa de que se haga X? ¿Puede decirse que la intención de una persona de subir una escalera causó que la subiera, o produjo su acción de subirla? Ciertamente es verdad que el hecho de que un hombre se proponga o no se proponga hacer X producirá una diferencia en si hará X. Esto se pone fuertemente de manifiesto cuando nos preocupa predecir si hará X; obviamente, sería importante descubrir si se propone hacerlo. ¿No implica esto que su intención tiene “un efecto sobre su conducta”?<sup>12</sup>

Habitualmente pensamos en las disposiciones como causas de la conducta. Si ante la misma provocación un hombre pierde la calma y otro no, esta diferencia en sus reacciones podría ser explicada por el hecho de que el primero, pero no el segundo, se encuentra en una disposición irritable. Si las disposiciones son causas, difícilmente podemos negarles el mismo papel a las intenciones. Ambas son útiles al predecir la conducta. Si intento estimar la probabilidad de que este hombre haga esto-y-aquello, la información de que tiene una disposición a hacerlo en circunstancias como éstas será una consideración afirmativa.

---

<sup>12</sup> La frase es de Taylor (1964: 34). En mi reseña de su libro (Malcolm, 1967) señalé que se equivoca al sostener que una intención simple *produce* la conducta correspondiente. Ahora considero que en parte está en lo correcto y en parte está equivocado: tiene razón sobre intenciones simples formadas previamente, está equivocado sobre intenciones simples meramente simultáneas.

Tengo derecho a darle igual o posiblemente mayor peso a la información de que él se propone hacerlo.

Las intenciones simples no sólo tienen peso al predecir acciones, sino que también aparecen en la explicación de acciones que ya ocurrieron. Si un hombre que acaba de salir de prisión sube rápidamente a una asta bandera, puede que yo quiera una explicación de por qué ocurrió eso. Si descubro que se había decidido previamente a hacerlo, pero que su encarcelamiento lo había prevenido, obtengo una explicación parcial de por qué está subiendo a la asta bandera, incluso si aún no sé cuál es su intención adicional, si la hay, al subirla. En general, si una acción me sorprende, descubrir que el agente previamente había decidido efectuarla, pero se lo impidió un obstáculo que acaba de retirarse me ayudará a comprender por qué ocurrió.

14. Las intenciones simples consideradas hasta ahora se formaron antes de la acción correspondiente. Pero muchas intenciones simples no se forman antes de la acción correspondiente. Al conducir un auto, uno presiona abruptamente (e intencionalmente) el pedal del freno: pero no hubo un momento antes de esta acción en el que uno se propusiera hacerla. La intención existió sólo en el momento de la acción, o sólo *en* la acción. Llamemos a ésta una intención simple meramente concurrente. ¿Puede una intención de esta clase ser un factor causal en la acción correspondiente?

Aquí tenemos que recordar que, si el conductor no presionó el freno intencionalmente, su presionar el freno no fue una acción sin cualificación. La presencia de intención simple en la acción (esto es, el que sea intencional) es una condición analíticamente necesaria para que sea una acción sin cualificación. Esta condición

no es una causa sino una condición definitoria de la acción sin cualificación. Si no se satisface esta condición, tendríamos que usar alguna frase mitigante –por ejemplo, que el conductor presionó el freno por equivocación. De este modo, una intención simple que es meramente concurrente no puede ser una causa de la acción correspondiente.

15. ¿Acaso podemos evitar comprometernos con la suposición de que la presión del pie del conductor sobre el freno fue ya sea intencional ya sea no intencional? ¿No podemos pensar en ella, de manera neutral, meramente como conducta? Sí, podemos hacerlo. Pero *fue* ya sea intencional, ya sea no intencional. Si fue esto último, no hubo intención simple que fungiera como una causa de la conducta. Si fue lo primero, entonces la conducta fue acción, y la intención simple meramente concurrente del conductor fue una condición definitoria y no una causa de la conducta. La “manera neutral” de pensar acerca de la conducta sería meramente incompleta. Se debería a la ignorancia, no a la existencia de una tercera alternativa. Es imposible, por la definición de “acción”, que la conducta de presionar el freno sea una acción y aun así no sea intencional. De modo que es imposible que una intención simple meramente concurrente haya causado la conducta de presionar el freno, ya sea que la conducta fuese o no una acción.

Para resumir esta discusión sobre las intenciones como causas: necesitamos distinguir entre intenciones simples e intenciones adicionales. Si un agente hace *X* con la intención adicional *Y*, entonces es apropiado hablar de su intención adicional como la (o una) causa del hacer *X*. Las intenciones simples pueden dividirse en aquellas que se forman antes de la acción correspondiente y aquellas que son meramente

concurrentes con las acciones. En virtud de haberse formado previamente, una intención simple puede ser una causa de la acción. Pero en tanto es meramente concurrente, una intención simple no puede ser una causa de la acción correspondiente.

16. Intentemos ahora evaluar el punto de vista de Taylor sobre el papel causal de la intención en la conducta. Él sostiene que no sería verdad, sin cualificación, que una persona apuñaló a otra a menos que su intención de apuñalarla “produjera” el apuñalamiento (1964: 33). Se supone que el ejemplo es de una intención formada previamente –pues Taylor habla de la *decisión* del agente de apuñalar a alguien. Pero una gran cantidad de acciones no incorporan intenciones formadas con anticipación. Incorporan intenciones meramente concurrentes. No puede decirse que estas últimas causen (produzcan) las acciones correspondientes. Posiblemente debido a que ha fijado su atención demasiado estrechamente en casos de decisión, Taylor se equivoca al sostener que, en general, el concepto de acción requiere que la intención del agente deba producir la conducta. Cuando la acción es meramente intencional (sin intenciones previas), no puede decirse que la intención del agente produce su conducta. En tales casos su intención le da a su conducta el carácter de *acción*, pero lo hace en virtud de ser una condición definitoria de la acción, no en virtud de ser una causa de ya sea conducta o acción.

17. Nuestras reflexiones sobre la relación de la intención con la conducta surgieron de una consideración de tres ejemplos de supuesta acción –abrir una puerta, girar una perilla, hacer el movimiento de girar con la mano. En los primeros dos casos imaginamos mecanismos que producían los resultados esperados

independientemente de la intervención del agente. En consecuencia, tuvimos que negar que *él* abriera la puerta o girara la perilla. Luego imaginamos una causa neurofisiológica del movimiento de su mano, y nos preguntamos si esto implicaría, a su vez, que *él* no movió su mano.

¿Es el movimiento de su mano independiente de su “intervención” en virtud de ser independiente de su intención? Vimos previamente (Sección 8) que una teoría neurofisiológica comprensiva no dejaría lugar para los deseos e intenciones como factores causales. En consecuencia, ni la intención simple previamente formada del hombre de mover su mano ni su intención adicional (de abrir la puerta) podrían ser causas del movimiento de su mano.

18. Notamos antes que es verdadero *a priori* que, si un hombre quiere *Y*, o tiene a *Y* como meta, y cree que *X* es requerido para *Y*, entonces en ausencia de factores que lo impidan él hará *Y*. También es verdadero *a priori* que si un hombre se forma la intención de (por ejemplo, decide) hacer *X*, entonces en ausencia de factores que interfieran él hará *X*. Estos principios *a priori* de la acción son asumidos en nuestras explicaciones cotidianas de la conducta.

Vimos que las explicaciones mecanicistas no podrán ser más básicas que las explicaciones en términos de intenciones o propósitos.

Vimos que la verificación de leyes mecanicistas no podría refutar los principios *a priori* de la acción.

Aun así, una explicación mecanicista de la conducta descarta cualquier explicación de ella en términos de las intenciones del agente. Si una teoría neurofisiológica comprensiva es verdadera,

entonces las intenciones de las personas nunca son factores causales en la conducta.

19. Así, si el mecanicismo es verdadero, los principios *a priori* de la acción no se aplican al mundo. Esto tendría que significar una u otra de entre dos alternativas. La primera sería que la gente no tiene intenciones, propósitos o deseos, o que no tienen creencias sobre qué conducta es requerida para la satisfacción de sus deseos y propósitos. La segunda alternativa sería que, aunque tienen intenciones, creencias y demás, siempre hay factores que intervienen –esto es, factores que interfieren con la operación de intenciones, deseos y decisiones.

La segunda alternativa no puede tomarse en serio. Si un hombre quiere estar en el banco opuesto de un río y cree que nadar es la única cosa que lo llevará ahí, él nadará a menos que haya factores que lo impidan, tales como una incapacidad de nadar o un miedo a ahogarse o una fuerte aversión a mojarse. En este sentido no es verdad que factores que intervienen están presentes *siempre* que alguien tiene una meta. No *siempre* hay obstáculos a la satisfacción de cualquier propósito o deseo.

Podría objetarse que la causación mecánica es un factor universal de interferencia. Ahora bien, si esto fuera así implicaría que los propósitos, intenciones y deseos nunca han tenido efecto alguno sobre la conducta. Pero no es una posición coherente sostener que algunas criaturas tienen propósitos y demás, y aun así éstos no tienen ningún efecto sobre su conducta. Los propósitos y las intenciones están, en concepto, tan estrechamente ligados a efectos conductuales que la completa ausencia de efectos conductuales significaría ausencia total de propósitos e intenciones. Así, la única posición abierta al defensor del

mecanicismo es la primera alternativa –a saber, que las personas no tienen intenciones, propósitos o creencias.

Lo que he llamado “un principio de la acción” es una proposición condicional, que tiene un antecedente y un consecuente. El condicional entero es verdadero *a priori*, y por ende si el antecedente se sostiene en un caso particular, el consecuente debe sostenerse en ese caso. Decir que el antecedente se sostiene en un caso particular significa que es verdadero de alguna persona (o animal). Significa que la persona tiene algún deseo o intención, y también tiene la creencia requerida. Si esto fuese así, y si no hubiese factores que interfirieran, se seguiría que la persona actuaría de una manera apropiada. Su intención o deseo sería, en nuestro amplio sentido, una causa de su acción –esto es, sería un factor en la explicación de la ocurrencia de la acción.

Pero esto es incompatible con el mecanicismo. Un mecanicista debe sostener, por tanto, que los principios de la acción no tienen aplicación a la realidad, en el sentido de que nadie tiene intenciones o deseos o creencias.

Algunos filósofos considerarían a este resultado como una refutación adecuada del mecanicismo. Pero otros no. Ellos dirían que la confirmación de una teoría neurofisiológica de la conducta es una posibilidad lógica, y por ende es lógicamente posible que no haya deseos, intenciones y demás, y que negar estas posibilidades lógicas es ser dogmático y acientífico. Evitaré adoptar esta posición “dogmática” y “acientífica”, y formularé una crítica del mecanicismo desde un punto de vista más “interno”.

20. Deseo hacer una aproximación más cercana a la cuestión de la concebibilidad del mecanicismo. Hemos visto que el mecanicismo

es incompatible con la conducta dirigida a propósitos, pero aún no hemos establecido que es incompatible con la existencia de conducta meramente intencional. Un hombre puede hacer algo intencionalmente, pero sin ninguna intención adicional: su conducta es intencional pero no de propósitos. Una posibilidad es que su conducta encarne una intención simple meramente concurrente. Puesto que tales intenciones no son causas de la conducta a la que pertenecen, su existencia no parece estar en conflicto con la causación mecánica. La incompatibilidad del mecanicismo con la conducta dirigida a propósitos aún no ha mostrado que sea incompatible con la conducta intencional como tal.

Pero, ¿podría ser verdad que a veces las personas actuaron intencionalmente aunque nunca fue verdad que actuaron por algún propósito? ¿Podrían hacer cosas intencionalmente pero nunca hacerlas con alguna intención adicional?

Si algunas acciones intencionales carecen de propósito, no se sigue que todas ellas podrían carecer de propósito. Y no creo que esto sea realmente una posibilidad. No intentaré lidiar con todas las clases de acciones. Pero considérese aquella subclase de acciones que son actividades. Cualquier actividad física es analizable en componentes. Si un hombre está pintando una pared, él está sosteniendo una brocha, sumergiendo la brocha en la pintura, moviendo su brazo de un lado a otro. Él hace todas estas cosas al pintar. Son partes de su actividad de pintar. Si alguien se mece en una silla, se empuja contra el piso con los pies y presiona su espalda contra el respaldo de la silla. Éstas son actividades subordinadas en la actividad de mecerse. Si al que está pintando se le pregunta por qué sumerge la brocha en la pintura, él puede responder “Estoy pintando esta pared”. Ésta es una explicación de

lo que está haciendo al sumergir la brocha, y también de *para qué* la está sumergiendo. Es una explicación mediante propósitos. Una persona puede poner pintura sobre una pared, o mecarse en una silla, o caminar de aquí para allá, sin tener ningún propósito al hacerlo. Aun así, estas actividades podrían ser intencionales, aunque sin propósito alguno.

Sean intencionales o no, estas actividades serían analizables en partes componentes. Si la actividad es intencional, entonces al menos alguno de sus componentes será intencional. Si ninguno de ellos lo fuera, el todo al cual pertenecen no sería intencional. Un hombre no podría poner pintura intencionalmente sobre una pared si no sostuviera intencionalmente una brocha. Ahora bien, esto no es estrictamente verdadero puesto que él podría no percatarse de que está sosteniendo una *brocha*, sino más bien un rodillo o un trapo. Pero tendría que haber *alguna* descripción de lo que estaba sosteniendo de acuerdo con la cual sea verdadero que estaba intencionalmente sosteniéndolo e intencionalmente sumergiéndolo en la pintura.

De modo que una actividad intencional debe tener componentes intencionales. Los componentes estarán dirigidos a un propósito en relación con la actividad completa. Si  $X$  es un componente intencional de  $Y$ , puede decirse con igual verdad que al hacer  $X$  se hace  $Y$ , o que se hace  $X$  con el fin de  $Y$ . Al mover el lápiz sobre el papel se dibuja una figura: pero también se mueve el lápiz con el fin de dibujar una figura.

Concluyo que, si no pudiese haber conducta dirigida a propósitos, no podría haber actividades intencionales. Estrictamente hablando, esto no demuestra que no podría haber acción intencional, puesto que muchas acciones no son actividades

(por ejemplo, atrapar un balón o ganar una carrera, en contraste con jugar a la pelota o correr una carrera). Pero muchas de las acciones que no son actividades son etapas en, o terminaciones de, actividades y no podría existir si no lo hicieran las actividades. Aunque no sé cómo demostrar el punto para todos los casos, me parece bastante plausible que, si no pudiese haber actividades intencionales, no podría haber conducta intencional de ninguna clase –me parece tan plausible que asumiré que es el caso. Una vida que estuviese totalmente desprovista de actividades ciertamente no podría ser una vida humana. Mi conclusión es que, puesto que el mecanicismo es incompatible con la conducta dirigida a propósitos, es incompatible con las actividades intencionales y, en consecuencia, es incompatible con *toda* conducta intencional.

21. La muy aplazada pregunta de si el hombre de nuestro ejemplo movió su mano sobre la perilla será respondida como sigue. La acción de mover su mano no puede serle adscrita correctamente a él. Ni siquiera debería adscribirsele con cualificaciones. No debería adscribirsele con alguna cualificación como “de manera no intencional” o “accidentalmente”, pues el uso de estas cualificaciones implica que hay casos en los que es correcto decir que un hombre hizo algo “intencionalmente” o “con cierto propósito”. Pero el mecanicismo descarta esto. Por otra parte, decir “Él no movió su mano” sería confuso, no sólo por la razón recién enunciada, sino también por la razón ulterior de que este enunciado normalmente conllevaría la implicación de que su mano no se movió –lo cual es falso. Ni la oración “Él movió su mano” ni la oración “Él no movió su mano” serían apropiadas. Por supuesto, diríamos “Él movió su mano” si entendiéramos esto como meramente equivalente a “Su mano se movió”. (Es

interesante que usemos estas dos oraciones de manera intercambiable cuando observamos a alguien de quien sabemos que está dormido o inconsciente: estamos igualmente dispuestos a decir ya sea “Él movió su mano” o “Su mano se movió”.) Pero si llegáramos a creer en el mecanicismo, para ser consistentes deberíamos renunciar a adscribir acciones, incluso de manera cualificada.

22. Podemos ahora proceder directamente a la pregunta de si el mecanicismo es concebible. A veces cuando los filósofos preguntan si una proposición es concebible, se proponen preguntar si es auto-contradictoria. Nada en nuestro examen ha indicado que el mecanicismo sea una teoría auto-contradictoria, y yo estoy seguro de que no lo es. Lógicamente hablando, la Tierra y el universo entero podrían haber sido habitados únicamente por organismos de tal naturaleza que todos sus movimientos pudiesen ser completamente explicados en términos de la teoría neurofisiológica que hemos imaginado. Podemos concebir que el mundo hubiese sido tal que el mecanicismo fuese verdadero. En este sentido, el mecanicismo es concebible.

Pero hay un aspecto en el que el mecanicismo no es concebible. Ésta es una consecuencia del hecho de que el mecanicismo es incompatible con la existencia de cualquier conducta intencional. El habla de los seres humanos es, en su mayor parte, conducta intencional. En particular, enunciar, aseverar o decir que tal-y-cual es verdadero requiere la preferencia intencional de alguna oración. Si el mecanicismo es verdadero, por tanto, nadie puede enunciar o aseverar nada. En un sentido, nadie puede *decir* nada. Específicamente, nadie puede aseverar o enunciar que el mecanicismo es verdadero. Si alguien aseverara

esto, la ocurrencia de su “acto de habla” intencional implicaría que el mecanicismo es falso.

De este modo, hay un absurdo lógico en aseverar que el mecanicismo es verdadero. No es que la doctrina del mecanicismo sea auto-contradictoria. El absurdo yace en el acto humano de aseverar la doctrina. La ocurrencia de este acto de aserción es inconsistente con el contenido de la aserción. La mera proposición de que el mecanicismo es verdadero no es auto-contradictoria. Pero la proposición conjuntiva “El mecanicismo es verdadero y alguien asevera que es verdadero” sí es auto-contradictoria. Así que la aseveración de cualquiera de que el mecanicismo es verdadero es necesariamente falsa. La aseveración implica su propia falsedad en virtud de que proporciona un contra-ejemplo a lo que es aseverado.

23. Un defensor del mecanicismo podría afirmar que, puesto que el absurdo que hemos estado describiendo es una mera “paradoja pragmática” y no una auto-contradicción en la doctrina del mecanicismo, no proporciona un sentido en el que el mecanicismo sea inconcebible. Puede decir que la paradoja es similar a la de un hombre aseverando que él mismo está inconsciente. Hay una inconsistencia entre el acto de este hombre de enunciar que está inconsciente y lo que enuncia. Su acto de enunciarlo implica que lo que enuncia es falso. Pero esta paradoja no establece que un hombre no pueda estar inconsciente, o que no podamos concebir que un hombre esté inconsciente. Ahora bien, hay alguna similitud entre la paradoja de enunciar que uno está inconsciente y la paradoja de enunciar que el mecanicismo es verdadero. Pero hay una diferencia importante. No puedo enunciar, sin absurdo, que yo estoy inconsciente. Pero cualquier otra persona puede, sin

absurdo, enunciar que yo estoy inconsciente. Sólo hay una persona (yo) cuyo acto de enunciar esta proposición es inconsistente con la proposición. Pero una aseveración del mecanicismo por parte de cualquier persona es inconsistente con el mecanicismo. Que yo estoy inconsciente no es (consistentemente) enunciable para mí. La no enunciabilidad es relativa sólo a una persona. Pero la no enunciabilidad del mecanicismo es absoluta.

Además, nadie puede consistentemente aseverar que, aunque el mecanicismo no es enunciable, puede ser verdadero. Pues esta aseveración, también, requeriría una preferencia intencional (acto de habla) y así sería incompatible con el mecanicismo.

Hemos elucidado un sentido en el que puede decirse propiamente que el mecanicismo es inconcebible. El sentido es que nadie puede consistentemente aseverar (o enunciar o decir) que el mecanicismo es, o puede ser, verdadero.

Si alguien insistiese en aseverar que el mecanicismo es o puede ser verdadero, su único recurso (si fuese consistente) sería adoptar una forma de solipsismo. Afirmaría que el mecanicismo es verdadero para otros organismos, pero no para él. De esta manera liberaría a su aseveración de inconsistencia, pero al costo de aceptar el bochorno y las dificultades lógicas del solipsismo. También repudiaría la respetabilidad científica del mecanicismo al negar la universalidad de las leyes neurofisiológicas imaginadas.

24. Nuestra crítica de que el mecanicismo no es una doctrina consistentemente enunciable es, por supuesto, de naturaleza puramente lógica. Consiste en deducir una consecuencia del mecanicismo. Ahora uno podría sentir que esta consecuencia no puede refutar al mecanicismo o poner en peligro su estatus

científico. Parecería corresponderle sólo a la ciencia determinar si hay una teoría neurofisiológica comprensiva que explique todos los movimientos corporales de acuerdo con leyes universales. Si la investigación científica confirmara una teoría tal, ¡entonces que así sea! Confirmarla sería confirmar sus consecuencias. Si confirmar la teoría demostrara que la gente no tiene deseos, propósitos o metas, entonces este resultado tendría que ser aceptado, sin importar cuán terrible sería no sólo para nuestras creencias ordinarias sino también para nuestros conceptos ordinarios.

Casi cualquiera sentirá alguna fuerza persuasiva en esta doctrina. El determinismo es un penoso problema debido a que genera una severa tensión entre dos puntos de vista, cada uno de los cuales es considerablemente atractivo: uno es que los conceptos de propósito, intención y deseo, de nuestro lenguaje ordinario, no pueden volverse vacíos por el avance científico; el otro es que esos conceptos no pueden prescribir límites a lo que es posible que logre la ciencia empírica.

Veamos cuál sería el efecto de una confirmación científica del mecanicismo sobre nuestro pensamiento. Supóngase que estoy jugando a atrapar la pelota con un niño pequeño. La bola escapa a su agarre y corre tras ella. Cualquier observador estaría de acuerdo en que el niño está corriendo tras la pelota. Esta descripción implica que el propósito del correr del niño es atrapar la pelota, o que él está corriendo porque quiere capturar la pelota.

Ahora supongamos que un técnico neurólogo podría explicar y predecir cualquier movimiento de las extremidades del niño sin considerar la ubicación de la pelota, sólo en términos de los estados cambiantes del sistema neurofisiológico del niño. O, lo que es peor, supóngase que el técnico podría controlar los movimientos del niño alterando los estados de su sistema nervioso

central a voluntad –esto es, mediante “programación”. Podemos imaginar que sería imposible para nosotros distinguir en un caso particular, mediante la observación de la conducta externa del niño y de las circunstancias, si las extremidades del niño responden a programación o si él corre con el fin de recuperar la pelota. Y supóngase que en muchos casos en los que creíamos que la conducta era intencional, se nos demostrara subsecuentemente que ocurrían exactamente los mismos procesos fisiológicos internos que en aquellas ocasiones en las que el técnico controlaba los movimientos del niño. También podemos suponer que las predicciones de conducta del neurólogo serían más confiables y más precisas que las predicciones basadas en suposiciones sobre propósitos.

Si tales demostraciones ocurrieran a escala masiva, estaríamos aprendiendo que los principios de la explicación mediante propósitos tienen una aplicación mucho más estrecha de lo que habíamos pensado. En más y más ocasiones nos veríamos (cada uno de nosotros) forzados a considerar a otros seres humanos como mecanismos. El resultado final de este desarrollo sería que dejáramos de pensar en la conducta de otros como siendo influenciada por deseos e intenciones.

25. Habiéndonos convertido en creyentes de las explicaciones mecanicistas de la conducta de otros, ¿podría cada uno de nosotros también llegar a creer que la causación mecánica es la doctrina verdadera para su propio caso? No si nos diéramos cuenta de lo que esto implicaría, pues cada uno de nosotros vería que no podría incluirse a sí mismo dentro del alcance de la doctrina. Decir o hacer algo *por una razón* (tanto en el sentido de fundamento como en el sentido de propósito) implica que el decir o hacer es intencional.

Puesto que el mecanicismo es incompatible con la intencionalidad de la conducta, mi aceptación del mecanicismo como verdadero para mí mismo implicaría que soy incapaz de decir o hacer algo por una razón. Podría haber *una* razón (esto es, una causa), pero no podría haber algo como *mi* razón. No podría haber, por ejemplo, algo como mi razón para enunciar que el mecanicismo es verdadero. Así, mi aseveración del mecanicismo involucraría una segunda paradoja. No solo la aseveración sería inconsistente, en el sentido previamente explicado, sino que también implicaría que soy incapaz de tener fundamentos racionales para aseverar algo, incluyendo el mecanicismo.

Nuevamente vemos que el mecanicismo engendra una forma de solipsismo. Al aseverar el mecanicismo debo negar su aplicación en mi propio caso: pues de otra manera mi aseveración implicaría que no podría estar aseverando el mecanicismo sobre fundamentos racionales.

26. Algunos filósofos sostienen que, si el mecanicismo es verdadero, entonces se requiere una revisión radical de nuestros conceptos. Necesitamos deshacernos de todos los términos tales como “intencionalmente”, “no intencionalmente”, “a propósito”, “por equivocación”, “deliberadamente”, “accidentalmente”, y la lista sigue. La clasificación de preferencias como “aseverar”, “repetir”, “citar”, “arremedar”, “traducir”, y demás, tendría que ser abandonada. Necesitaríamos un repertorio enteramente nuevo de descripciones de una clase que fuese compatible con el punto de vista del mecanicismo.

Creo que estos filósofos no han captado toda la severidad del predicamento. Si el mecanicismo es verdadero, no sólo deberíamos renunciar a hablar de “aseverar”, sino también de “describir” o

incluso de “hablar”. Ni siquiera sería correcto decir que una persona *quiso decir* algo mediante el ruido que salió de ella. Ninguna marca o sonido significaría nada. No podría haber *lenguaje*.

Un defensor del mecanicismo no debería pensar que en la actualidad estamos usando los conceptos erróneos y que hace falta una revisión. Si está en lo correcto, no usamos conceptos en absoluto. No hay nada que revisar –y nada que decir. El lema de un mecanicista habría de ser: No se puede hablar, por tanto, se debe callar.

27. Para concluir: hemos puesto al descubierto dos aspectos en los que el mecanicismo no es una doctrina concebible. El primero es que la ocurrencia de un acto de aseverar el mecanicismo es inconsistente con que el mecanicismo sea verdadero. El segundo es que la aseveración del mecanicismo implica que quien la hace no puede estarla haciendo sobre fundamentos racionales. Con el fin de evitar estas paradojas, debe negarse que el mecanicismo sea universalmente verdadero. Se puede sostener que es verdadero para otros, pero no para uno mismo. Es bastante irónico que la afirmación del mecanicismo requiera que uno afirme su opuesto metafísico y metodológico –el solipsismo. La inconcebibilidad del mecanicismo, en los dos aspectos que hemos elucidado, no establece que el mecanicismo sea falso. Parecería, lógicamente hablando, que una teoría neurofisiológica comprensiva de la conducta humana debería ser corroborable mediante investigación científica. Aun así, la aseveración de que se ha logrado esta corroboración involucraría las dos paradojas que hemos elucidado. De este modo, el mecanicismo presenta una antinomia hostil, y quizá insoluble, al pensamiento humano.

*Post-scriptum no científico a manera de conclusión:* Debo confesar que no estoy completamente convencido de que la posición que he adoptado con respecto al punto crucial de este ensayo sea correcta –a saber, el problema de si es posible que haya tanto una explicación neurofisiológica completa como una explicación mediante propósitos completa de una y la misma secuencia de movimientos. No creo que realmente haya demostrado que esto sea imposible. Por otra parte, para mí (y para otros también) es verdad que una secuencia de sonidos tiende a perder el aspecto de habla (lenguaje) cuando concebimos a esos sonidos como siendo causados neurofisiológicamente (en especial si imaginamos que un técnico está controlando la producción de los sonidos). De manera análoga, una secuencia de movimientos pierde el aspecto de acción. ¿Se debe esta tendencia a una falsa imagen o a una analogía desorientadora? Es posible; pero también es posible que no. Quizá la publicación de este ensayo se justifique si incita a ofrecer una defensa realmente convincente de la compatibilidad de las dos formas de explicación.<sup>13</sup>

## Referencias

Kim, J. (1966). On the Psycho-Physical Identity Theory. *American Philosophical Quarterly* 3: 227-235.

---

<sup>13</sup> Varias personas leyeron distintas versiones de este ensayo y me he beneficiado de sus críticas. Estoy especialmente en deuda con Elizabeth Anscombe, Keith Donnellan, Philipa Foot, G. H. von Wright y Ann Wilbur. No son responsables de los errores que persisten en el texto.

- Malcolm, N. (1964). Scientific Materialism and the Identity Theory. *Dialogue* 3 (2): 15-25.
- Malcolm, N. (1967). Review of Explaining Behaviour. *Philosophical Review*, LXXVI: 97-104.
- Malcolm, N. (1971). *Problems of Mind: Descartes to Wittgenstein*. New York: Harper & Row.
- Melden, A.I. (1961). *Free Action*. New York: Routledge.
- Taylor, C. (1964). *The Explanation of Behaviour*. New York: Routledge.

Marc Jiménez-Rolland  
Universidad Autónoma Metropolitana, Unidad Iztapalapa  
marcji2121@yahoo.com