

¿LA TERCERA HUMILLACIÓN? (SOBRE NEUROCIENCIA, FILOSOFÍA Y LIBRE ALBEDRÍO) *

Daniel González Lagier

Universidad de Alicante

Para Ernesto Garzón Valdés

RESUMEN. En este trabajo se describen los experimentos realizados en la década de los cincuenta del siglo pasado por el neurocientífico B. LIBET, cuyos resultados se han interpretado como una demostración empírica de la falsedad del libre albedrío. Frente a las interpretaciones deterministas de estos experimentos (y frente a la idea de que la neurociencia puede desplazar a la filosofía en la discusión sobre estos temas) se esgrimen dos argumentos de G. H. VON WRIGHT que tratan de mostrar la imposibilidad (por razones conceptuales y filosóficas) de probar empíricamente el determinismo universal (que es un presupuesto de las tesis de quienes niegan el libre albedrío con este tipo de argumentos) y la imposibilidad de reducir lo mental a lo neuronal.

Palabras clave: Neurociencia, determinismo, libre albedrío, responsabilidad, causalidad.

ABSTRACT. This paper describes the experiment that the neuroscientist B. LIBET carried out during the 1950s. The conclusions of these experiments have been interpreted as an empirical demonstration of the falseness of free will thesis. Against the determinist interpretation of these experiments (and against the idea that neuroscience can displace philosophy in the debate of these questions), the paper presents two arguments of G. H. VON WRIGHT which attempt to show that it is impossible (by conceptual and philosophical reasons) to prove empirically the universal determinism (which is a presumption of the thesis rejecting free will on the basis of arguments of this type) and the impossibility of reducing the mental dimension to the neuronal one.

Keywords: Neuroscience, determinism, free will, responsibility, causality.

* Fecha de recepción: 26 de septiembre de 2012. Fecha de aceptación: 15 de octubre de 2012.

Este trabajo se ha desarrollado en el marco del Proyecto de Investigación sobre «Argumentación y Constitucionalismo» concedido por el Ministerio de Ciencia e Innovación.

La primera, cuando [el hombre] se percató de que nuestra Tierra no es el centro del universo, sino una minúscula e insignificante parte de un sistema cósmico apenas imaginable en su grandeza. Esta humillación se liga para nosotros al nombre de Copérnico, si bien la ciencia alejandrina ya había proclamado algo semejante. La segunda, después, cuando la investigación biológica destruyó los presuntos privilegios del hombre en la creación, desterrándolo a su procedencia del reino animal y a lo inextinguible de su naturaleza animal. Esta transmutación de valores se ha llevado a cabo en nuestros días bajo el influjo de Ch. Darwin, Wallace y sus predecesores, no sin la más vehemente resistencia de los contemporáneos. La tercera y más sentida humillación la ha de sufrir, empero, el ansia de grandeza del hombre en virtud de la actual investigación psicológica, que va a demostrar al yo que ni siquiera es señor en su propia casa, sino que se halla supeditado a las mezquinas noticias de aquello que sucede inconscientemente en su vida anímica.

(Una dificultad del psicoanálisis, S. FREUD)

1. ¿Es el hombre responsable de sus acciones? Desde hace unos años, cuando se intenta responder a esta pregunta nos topamos con un nuevo paradigma o esquema de interpretación (a veces se hace referencia al mismo como «neuroética»¹ o, incluso, «neurocultura»², aunque no es una terminología unívoca) que pretende dar cuenta de aspectos básicos del comportamiento humano, como el libre albedrío, la deliberación, la consciencia e, incluso, la conducta moral, a partir de los recientes conocimientos acerca del funcionamiento del cerebro (conocimientos que pueden agruparse bajo el nombre de «neurociencia»). En sus pretensiones más amplias y radicales, este nuevo paradigma propone replantear estos problemas al margen de la filosofía y de las humanidades, estableciendo para ellos un nuevo marco —un «campo de conocimiento unificado» bajo la dirección de la ciencia— que conducirá, ni más ni menos, a una nueva concepción del hombre: se afirma, en definitiva, que «las humanidades están abocadas a ser estudiadas, y en algún caso reevaluadas desde los conocimientos científicos. Sin ellos, sin los conocimientos científicos, las humanidades quedarán como el agua descrita a nivel poético, es decir, algo así como una descripción superficial de los fenómenos, sin entrar en el conocimiento de los pilares que la sustentan y le dan fundamento»³.

Uno de los problemas que muchos neurocientíficos creen haber desentrañado es el del libre albedrío⁴. De acuerdo con estos autores, el análisis del cerebro ha mostrado que la libertad humana es meramente una ilusión, una ficción generada por nuestro propio cerebro. Determinados experimentos —los famosos experimentos de la década de los ochenta de B. LIBET⁵ y otros en la misma línea⁶— parecen permitir concluir que el cerebro «se pone en marcha» para realizar una determinada acción antes de que seamos conscientes de haber tomado la decisión correspondiente, de lo que suele inferirse (1) que nuestra decisión consciente no tiene efecto causal en la acción realizada y (2) que la sensación o impresión de actuar libremente y a consecuencia de nuestra decisión consciente es un efecto causal de la propia actividad cerebral. Consecuente-

¹ Sobre los distintos sentidos de la denominación «Neuroética» puede verse CORTINA, 2010. También CAPÓ, NADAL, RAMOS, FERNÁNDEZ, CELA CONDE, 2006.

² *Vid.*, por ejemplo, MORA TERUEL, 2007.

³ MORA TERUEL, 1996: 170.

⁴ *Vid.*, por todos, al principal representante de esta tesis en España: RUBIA, 2009.

⁵ Puede encontrarse una descripción de los mismos en LIBET, 1999.

⁶ *Vid.* SOON, BRAAS, HAYNES, 2008. Citado por SOLER GIL, 2009.

mente, al ser la libertad una mera ilusión, o bien se sostiene que la responsabilidad por nuestras acciones no puede encontrar fundamento o bien que éste no puede basarse en la idea de reproche o merecimiento, sino que ha de encontrarse en otros principios. Como ha señalado E. GARZÓN, en realidad el desafío que las neurociencias plantean para nuestra noción de responsabilidad no es nuevo, sino que se trata de un nuevo ropaje para un viejo problema persistente de la filosofía⁷.

Los experimentos de LIBET son bastante sencillos. Cuando tomamos la decisión de realizar un movimiento simple, por ejemplo, flexionar un dedo, el córtex motor del cerebro envía una señal eléctrica a los nervios motores implicados en el movimiento. En unos experimentos anteriores se había descubierto que previamente a ese proceso se puede detectar ya cierta actividad eléctrica en el cerebro, cuya función parece ser preparar el movimiento. A esta actividad se le denomina «potencial de preparación». LIBET diseñó su experimento con el fin de señalar en qué momento de la secuencia de actividad cerebral (que supuestamente se inicia con la toma de decisión y concluye con el movimiento) se produce el potencial de preparación. Para ello reunió una serie de voluntarios, cada uno de ellos frente a un cronómetro, y les pidió que tomaran en un momento determinado la decisión de levantar un dedo e indicaran en qué posición se encontraba la aguja del reloj en el momento en que tomaron la decisión consciente de mover el dedo. A través de unos electrodos se medía el momento en que tenía lugar el potencial preparatorio y se contrastaba con el momento en que cada voluntario databa su decisión consciente de mover el dedo. Sorprendentemente, el potencial preparatorio no se encontraba entre la decisión y el movimiento, sino que era previo a la decisión consciente (unos 350 milisegundos antes), lo que se interpretó como prueba de que el cerebro tomaba la decisión por su cuenta, antes de que fuéramos conscientes de ella. Posteriormente, otros experimentos (con medios más refinados) parecen avalar estos resultados. En el año 2008 J. D. HAYNES y sus colaboradores detectaron hasta siete segundos antes de la decisión consciente pautas cerebrales que permitían predecir cuál sería la acción de un sujeto enfrentado a la elección entre presionar uno u otro botón, obteniendo un 60 por 100 de aciertos en la predicción.

La negación del libre albedrío basada en este tipo de experimentos no es una conclusión unánime entre los neurocientíficos (el propio LIBET no la aceptaba, entendiendo que sus experimentos dejaban espacio para una libertad de la voluntad entendida no como inicio de la acción, pero sí como posibilidad de veto o censura de la misma), pero sí muy extendida. En mi opinión, hay muchos aspectos oscuros y discutibles en la argumentación de los neurocientíficos que pretenden haber demostrado el determinismo a propósito de la conducta humana, como la pretensión de que la respuesta a estos problemas sea exclusivamente empírica (cuando, en realidad, están asumiendo determinados conceptos de libertad, causalidad, sujeto agente, etc., en la mayoría de ocasiones sin mayor reflexión), la identificación de la decisión consciente con el momento en que el sujeto del experimento observa y registra la posición del cronómetro (todo eso requiere tiempo), el extraño desdoblamiento que sugieren entre el cerebro y el sujeto (¿como si mi cerebro y yo fuéramos entidades independientes!) o la posibilidad de interpretaciones distintas o explicaciones alternativas de los mismos experimentos que impedirían una conclusión tan general como la que se pretende (los experimentos de

⁷ E. GARZÓN VALDÉS, 2007: 226 y ss.

los neurocientíficos se han realizado con acciones muy simples, como mover un dedo u otro, en situaciones en que resulta indiferente qué movimiento realizar, por lo que lo único que parecen demostrar es que las conductas simples que no requieren una deliberación basada en un balance de razones son, en cierto sentido, «mecánicas» o automáticas, lo que en realidad no es sorprendente)⁸.

En general, este tipo de experimentos no son suficientemente sólidos como para minar la creencia de sentido común de que las acciones tienen lugar porque hemos decidido realizarlas. Parece razonable exigir para que una hipótesis quede bien establecida que ésta se apoye en datos empíricos obtenidos de manera fiable, que provea una adecuada explicación de los mismos, que elimine otras hipótesis alternativas igualmente plausibles con tales datos empíricos y que se inserte de manera coherente con el resto de conocimientos. Pues bien, al parecer ni los experimentos «tipo LIBET» son altamente fiables (persisten problemas de medición del tiempo y no se ha logrado un alto porcentaje de predicciones acertadas), ni explican adecuadamente todos los datos empíricos (la consciencia de la toma de decisión queda como un epifenómeno al que es difícil encontrar un sentido), ni eliminan hipótesis alternativas plausibles (como que los resultados sólo sean válidos para acciones sencillas que no requieren deliberación); y desde luego chocan frontalmente con creencias muy arraigadas. En un breve pero esclarecedor artículo, el filósofo y físico F. J. SOLER GIL resume las críticas a los experimentos de este tipo de la siguiente manera: «El porcentaje de aciertos en el experimento de HAYNES es muy bajo»; «La libertad de decisión no está reñida con el hecho de que la mayoría de las acciones estén dirigidas inconscientemente» y «Los experimentos de LIBET y HAYNES descartan la deliberación, por lo que no estudian acciones potencialmente libres»⁹. Consiguientemente, de acuerdo con este autor, los experimentos deberían diseñarse de manera que se cumplan las siguientes condiciones:

— En primer lugar, el experimento debería realizarse en situaciones en las que la persona se encuentra ante alternativas que no son indiferentes. Es preciso que se requiera una deliberación para resolver entre ellas. Y, cuanto más serio sea el asunto en juego, mejor (podría tratarse, por ejemplo, de una decisión vital clave, como la elección de los estudios a seguir).

— En segundo lugar, debería de tratarse de situaciones en las que la persona no se encuentra obligada a tomar su decisión inmediatamente, sino que puede dilatar su reflexión tanto tiempo como considere oportuno (pues, de otro modo, y ante la urgencia y la imposibilidad de decidir lo mejor, podría ponerse en marcha un proceso inconsciente que concluyera en una «corazonada»).

— En tercer lugar, y si se pretende descartar la conjetura de LIBET sobre la libertad como instancia supervisora con derecho de veto, sería necesario llegar a un nivel muy alto de predicción de las decisiones tomadas.

— Y finalmente, si se quiere descartar la posibilidad de que los propios procesos inconscientes en los que se basa la predicción hayan sido, de algún modo, «encargados» por la conciencia —a la manera en que un gobierno marca las directrices generales de su política, pero luego delega en distintas instancias la elaboración de los detalles concretos de las leyes y actuaciones a realizar—, sería preciso mostrar que la actividad cerebral que ha tenido lugar en dichos procesos, venía determinada por estados cerebrales previos al inicio de los procesos conscientes deliberativos¹⁰.

⁸ Para una accesible valoración de los experimentos «tipo LIBET», *vid.* SOLER GIL, 2009.

⁹ SOLER GIL, 2009: 544 y ss.

¹⁰ *Ibid.*, 546.

Obviamente, la neurociencia está todavía muy lejos de poder satisfacer estas condiciones. Pero, en todo caso, se trata de objeciones que dejan abierta la puerta para que en un futuro sea posible establecer empíricamente la negación del libre albedrío. Sin embargo, al margen de este tipo de objeciones, quizá es posible dar razones basadas en consideraciones conceptuales y filosóficas para negar que el libre albedrío pueda ser refutado empíricamente —ahora o en el futuro— de una manera concluyente. En esta dirección me parecen pertinentes dos argumentos de G. H. VON WRIGHT en contra de la posibilidad de demostrar empíricamente el determinismo universal entendido como una ley irrestricta (válida también para las acciones humanas) y contra la reducción de lo mental a lo neuronal. Incluso aunque no se acepte que son argumentos concluyentes contra el ataque al libre albedrío, al menos son un ejemplo de que la filosofía tiene mucho que decir antes de ser desplazada por la «neurocultura». Veámoslos.

2. En mi opinión, las pretensiones de quienes niegan el libre albedrío presuponen la corrección de una tesis más general, según la cual todo lo que ocurre (no sólo las acciones humanas) está determinado. Llamaré a esta tesis «determinismo universal». Lo que se afirma a partir de los experimentos de LIBET es que es el cerebro, y no nuestra voluntad consciente, el que causa nuestras acciones. Sin embargo, no parece aceptable la idea de que el cerebro estimula unas u otras conductas de manera azarosa. Si así fuera, no podría explicarse la supervivencia humana a lo largo de la evolución. Esto es, el cerebro debe estar aproximadamente programado para actuar de una manera instrumentalmente adecuada para el fin de la supervivencia. El cerebro debe actuar de acuerdo con determinadas leyes causales (ante tal o cual estímulo externo en tal o cual circunstancia, el cerebro reacciona de determinada manera). Esto inserta al cerebro en la red de cadenas causales que recorre el mundo. De hecho, las conclusiones de los experimentos «tipo LIBET» suelen reforzarse aduciendo que no es creíble que el hombre sea la única isla de libertad dentro de un mundo causalmente determinado. Dos ejemplos cercanos, ambos desde el ámbito de la ciencia, pueden ilustrar esta idea: Uno de los más acérrimos defensores en España de la tesis de que la libertad es una ilusión generada por el cerebro, F. RUBIA, catedrático de medicina de la Universidad Complutense, escribe: «Los físicos nos dicen que todo el universo está sometido a leyes deterministas de la naturaleza, de manera que sería extraño que el ser humano, su cerebro/mente no lo estuviese»¹¹. Y en un debate reciente en la revista *Thémata*, M. LÓPEZ CORREDOIRA, del Instituto Astrofísico de Canarias, afirma: «Somos fragmentos de Naturaleza sometidos a sus leyes, y ciertos movimientos mecánicos de la materia en nuestros cerebros en interacción con su entorno producen (= “son la causa de”) nuestros pensamientos y decisiones»¹².

Para ser realmente concluyentes, por tanto, los experimentos deberían señalar cuál fue realmente la causa de la acción: qué actividad cerebral no consciente produce las acciones complejas e insertar esa actividad en un determinismo universal que se retrotrae al pasado más remoto y se proyecta al futuro más lejano. El primer argumento de VON WRIGHT que vamos a examinar trata de mostrar precisamente la imposibilidad de probar empíricamente la hipótesis del determinismo universal.

Para desarrollar su argumento debemos introducir las nociones de necesidad ontológica (lo que es necesario que ocurra y, por tanto, está determinado) y de contingencia

¹¹ F. RUBIA, 2009: 11.

¹² M. LÓPEZ CORREDOIRA, 2008: 266.

ontológica (lo que puede o no ocurrir). Pues bien, el argumento de VON WRIGHT consiste en mostrar que para establecer que un acontecimiento es una necesidad ontológica necesitamos presumir que otro acontecimiento (su causa) es una contingencia ontológica.

Debemos distinguir también entre una mera regularidad accidental (aunque pueda ser universal) en la sucesión entre dos acontecimientos y una genuina relación causal (como es sabido, HUME reducía las últimas a las primeras). Cuando afirmamos que existe una relación causal entre dos acontecimientos «*p*» y «*q*» queremos decir algo más fuerte que el simple hecho de que en el mundo ambos acontecimientos se suceden uno al otro una y otra vez. Si se trata de una genuina relación causal, pensamos que el primer acontecimiento genera el segundo, es «responsable» de que ocurra el segundo. Un criterio para saber si realmente «*p*» es causa de «*q*» consiste en preguntarnos si estamos dispuestos a afirmar un «contrafáctico causal» del siguiente tipo: «Si “*p*” hubiera ocurrido en un momento en que, de hecho, no tuvo lugar, “*q*” hubiera ocurrido también». Supongamos que se nos muestra un trozo de cobre que luego se destruye: si estamos convencidos de que entre la aplicación de calor a un metal y su dilatación existe una relación causal, entonces podremos afirmar el enunciado contrafáctico causal «si ese trozo de cobre se hubiera calentado, se habría dilatado». ¿Cómo podemos verificar un condicional contrafáctico causal?

Como señala VON WRIGHT, no hay una manera directa de comprobar qué hubiera sucedido si en el pasado hubiera ocurrido algo que no llegó a ocurrir. Pero sí podemos manipular el futuro con fines experimentales¹³. Podemos producir un acontecimiento *p* y observar si le sigue *q*. Después, en una situación similar, podemos abstenernos de producir (o evitar que ocurra) *p* y observar de nuevo qué sucede con *q*. Si en nuestro primer experimento *q* apareció y en el segundo no, entonces hemos verificado la conexión entre los dos acontecimientos de la manera más fiable posible (lo cual no implica que sea posible una verificación concluyente) y, por tanto, podremos afirmar con cierta seguridad que «si *p* hubiera tenido lugar en la segunda ocasión, *q* le habría seguido»¹⁴.

¿Por qué estos experimentos aumentan nuestra confianza en el contrafáctico causal? Una regularidad constante entre dos acontecimientos *p* y *q* puede deberse *a*) a que existe una relación causal entre ellos; *b*) a que ambos tienen una causa común (la posición de la aguja en un barómetro y el mal tiempo ocurren sucesiva o simultáneamente porque tienen una causa común, no porque el barómetro cause el mal tiempo), o *c*) a que uno de ellos se dé siempre, con independencia de que el otro esté presente o ausente. Estos experimentos pretenden descartar las posibilidades *b*) y *c*).

Si lo anterior es así, para confirmar una relación causal no basta con observar pasivamente la evolución de un sistema de estados de cosas. Hay que «poner a prueba» esta relación completando la observación pasiva con la experimentación activa. Es decir, es necesario realizar varias acciones y omisiones: producir *p* y omitir (o impedir) *p*. Ahora bien, toda acción involucra también un contrafáctico: si afirmo que he abierto la ventana, esta afirmación implica el contrafáctico siguiente: «Si yo no hubiera abierto la ventana en el momento *t*, ésta no se hubiera abierto en el momento *t*». De la misma manera, si afirmo que he producido *p* para comprobar si le sigue o no otro acontecimiento, en mi

¹³ G. H. VON WRIGHT, 1972: 39.

¹⁴ *Ibid.*, 45.

afirmación está presupuesto el contrafáctico: «Si yo no hubiera intervenido, *p* no hubiera tenido lugar». Es decir, en este tipo de experimentos presumimos que operamos con un «sistema cerrado» en el que las relaciones causales las «ponemos en marcha» nosotros mismos. Si el sistema no está cerrado, si «*p*» es causado por alguna relación causal previa con un acontecimiento externo al sistema, no podremos descartar lo que hemos llamado las posibilidades *b*) (que «*p*» y «*q*» tengan una causa común sin que medie otra relación entre ellos) y *c*) (que «*q*» siempre esté presente, con independencia de que lo esté «*p*»).

De este razonamiento VON WRIGHT concluye que la verificación de un contrafáctico causal (y, por tanto, de que estamos ante una genuina relación causal y no ante una regularidad accidental) presupone un contrafáctico relativo a la acción. Ambos son de un tipo distinto: mientras el primero, con el que pretendemos comprobar una relación causal, se basa en la confianza de que la *presencia* de una causa producirá un efecto, el segundo —el relativo a la acción— se basa en la confianza de que en el momento en que actuamos estaba *ausente* cualquier otra causa del resultado de la acción. Es decir, el primer contrafáctico se basa en nuestra creencia de que dado el primer acontecimiento, el segundo es una necesidad ontológica. El segundo contrafáctico se basa en nuestra creencia de que el primer acontecimiento es una contingencia ontológica.

Supongamos que descubrimos que «*p*» (la supuesta causa) no era una contingencia ontológica (estábamos equivocados y «*p*» estaba a su vez determinado, era necesario que tuviera lugar). El margen de acontecimientos del mundo que consideramos indeterminados se reduce, pero para mostrar cuál era la causa de «*p*» (esto es, para mostrar que «*p*» está determinado por un hecho anterior) hay que recurrir de nuevo a la experimentación, para lo cual de nuevo debemos presumir que el nuevo experimento se realiza en un sistema cerrado y, por tanto, que en algún punto es una contingencia ontológica.

Veámoslo con un ejemplo: supongamos que queremos aumentar la temperatura en un sistema químico para observar si se produce determinada reacción. Debemos aumentar la temperatura y observar qué ocurre, si se produce o no la reacción. Después debemos observar qué ocurre si la temperatura no aumenta. Al hacerlo estamos presumiendo que el aumento o no de la temperatura es una contingencia que depende de nuestra intervención. Establecer que la reacción química es una necesidad causalmente determinada en un mundo en el que se produce el aumento de temperatura requiere asumir que el aumento de temperatura no está determinado (ocurrirá si lo producimos; en caso contrario, no ocurrirá). Supongamos que antes de que hayamos podido interferir la temperatura aumenta. Algo ha causado el aumento de temperatura y presumimos que este aumento también estaba determinado. Pero ahora necesitamos comprobarlo y para ello buscamos un hecho previo que pueda explicar el aumento de temperatura. ¿Será el aumento de la presión que se produjo en el sistema? Debemos manipular el contexto de manera que aumente la presión del sistema por nuestra intervención y observemos qué sucede, y después debemos abstenernos de aumentar la presión (o evitarlo) y de nuevo observar. Es decir, estamos nuevamente asumiendo que el aumento de la presión es contingente, que no ocurre necesariamente, sino sólo si lo provocamos.

El argumento de VON WRIGHT depende de su concepción de la causalidad: para VON WRIGHT el concepto de causa depende del de acción. Si no pudiéramos actuar, no comprenderíamos qué es una relación causal. Aprendemos la existencia de relaciones causales porque podemos actuar para producir cambios en el mundo. ¿No podríamos

descubrir la existencia de una relación causal entre dos acontecimientos sólo mediante la observación pasiva y reiterada (sin intervenir por medio de acciones) de que cuando la causa está presente se sigue el efecto y cuando la causa está ausente no se sigue el efecto? Por supuesto que sí, pero cuando hacemos esto, dice VON WRIGHT, asumimos que *si pudiéramos eliminar la causa, el efecto no tendría lugar*. Esto es, asumimos que la causa es una contingencia ontológica. En palabras de VON WRIGHT:

La idea de que la causalidad puede representar una «amenaza» para la libertad contiene una buena dosis de verdad empírica, a saber: aquélla de la que la impotencia y la inaptitud dan fe. Pero metafísicamente no deja de ser una ilusión. La ilusión se ha venido alimentando de nuestra propensión a creer, con espíritu humano, cabría decir, que el hombre en un estado de pura pasividad, simplemente observando secuencias regulares, puede registrar conexiones causales y cadenas de acontecimientos causalmente eslabonados que él, por extrapolación, se figura entonces que colman el universo desde un pasado infinitamente remoto hasta un futuro infinitamente lejano. Este punto de vista no alcanza a advertir que las relaciones causales son relativas a fragmentos de la historia del mundo que tienen el carácter de lo que hemos venido llamando sistemas cerrados¹⁵.

Una manera sencilla de formular el argumento consiste en decir que no sólo la noción de acción presupone que el mundo no está completamente determinado (que lo que producimos con nuestra acción no ocurriría sin ella), sino también (dada su vinculación con la noción de acción, a través de la experimentación) la de causalidad. Y la idea de determinismo (en el sentido que estamos discutiendo aquí) descansa en la causalidad.

Proyectemos ahora el anterior argumento sobre los experimentos de LIBET. Al sujeto voluntario que colabora en el experimento se le pide que tome la decisión de levantar el dedo y lo levanta. Los experimentadores observan que el potencial preparatorio se activa antes de que el sujeto haya tomado la decisión, de lo que deducen que hay una causa previa que dispara el potencial preparatorio. Lo que creían una contingencia (el dedo se levantaría o no dependiendo de lo que decidiera el voluntario) se muestra como una necesidad ontológicamente determinada (estaba determinado: el dedo se levanta en todo caso). ¿Qué es lo que hizo que el dedo se levantara? Hay que buscar una causa de la acción. El experimentador debería ser capaz de identificar cierta actividad cerebral como causa del levantamiento del dedo. Puede que tenga la hipótesis de que estimular ciertas inervaciones en una determinada zona del cerebro hace que el sujeto voluntario «sienta» que ha tomado una decisión y que su dedo se levante. Para contrastar esta hipótesis debe diseñar un nuevo experimento, y en este experimento tendrá que estimular las inervaciones y ver qué sucede y después no estimularlas y ver qué sucede. Es decir, tiene que asumir que la estimulación de las inervaciones es una contingencia ontológica, algo que no está determinado. Porque si la estimulación de las inervaciones ocurriera en todo caso, no podría comprobar que éstas son la causa de la acción. Supongamos ahora que llega a la conclusión de que el cerebro del voluntario se pone en marcha reaccionando ante algún estímulo externo, por ejemplo, las instrucciones que él mismo ha dado al voluntario. Si el experimento ha de tener sentido, ha de asumir que si él no hubiera dado las instrucciones al voluntario, éste (o su cerebro) no hubiera levantado el dedo. Y, de nuevo, si cree que son sus instrucciones las que causaron todo eso, ha de asumir que estas instrucciones son una contingencia ontoló-

¹⁵ VON WRIGHT, 1987: 106.

gica. De otra manera no podrá contrastar la presencia de relaciones causales entre sus instrucciones y la respuesta del voluntario.

Para concluir con este argumento, recordemos una última precisión: VON WRIGHT no pretende negar con el anterior razonamiento que el determinismo universal sea una hipótesis plausible. Lo que pretende es mostrar que, por razones conceptuales (dados nuestros conceptos de causa y de acción), no puede ser demostrado. Si —como sostengo— los argumentos de los neurocientíficos para negar el libre albedrío presuponen la corrección del determinismo universal, tampoco pueden ser demostrados.

3. El segundo argumento de VON WRIGHT que quiero traer a colación tiene que ver con la imposibilidad de reducir los fenómenos mentales (intenciones, emociones, creencias...) a fenómenos neuronales y físicos (impulsos eléctricos, reacciones químicas, etc.). Lo que los neurocientíficos deterministas parecen decir muchas veces es que hay correlatos empíricos (neuronales) de los estados mentales y que una vez conocida esta correlación la conducta humana se podrá explicar completamente en términos del sustrato neuronal. Lo que VON WRIGHT pretende, por el contrario, es mostrar que no se pueden eliminar los estados mentales de las explicaciones de la conducta¹⁶.

De acuerdo con VON WRIGHT el estudio de los fenómenos mentales, esto es, la psicología científica, puede realizarse desde tres enfoques distintos: El primer enfoque consiste en analizar lo mental desde el punto de vista de la introspección o «auto-observación». Los estados mentales se caracterizan porque no son intersubjetivamente observables y, en ese sentido, no son objetivos; son «propiedad privada del sujeto que los tiene», que dispone de un acceso directo a los mismos (sabemos, sin necesidad de observarnos «desde fuera», qué intenciones tenemos, qué creencias sostenemos, qué emociones experimentamos, etc.). A este enfoque lo podemos llamar la «psicología de la consciencia» (a veces se la ha llamado también, en un sentido despectivo, «psicología popular»). Pero lo mental tiene un sustrato físico o corporal, que a su vez presenta otros dos aspectos: lo mental, por un lado, se expresa en la conducta, en los movimientos corporales del sujeto y los cambios que producen en el mundo; y, al mismo tiempo, también está constituido por sucesos intracorporales que el sujeto no puede, en principio, observar en sí mismo: los fenómenos neuronales que ocurren en el cerebro del sujeto. Al primer aspecto, VON WRIGHT lo llama el aspecto conductual de lo mental; al segundo, el aspecto neuronal. La psicología que se desarrolla estudiando el primer aspecto es la psicología de la conducta o conductista; la que se desarrolla analizando el segundo aspecto es la neuropsicología.

Como señala VON WRIGHT, las tres ramas de la psicología «no se llevan bien» y la psicología de la conducta y la neuropsicología sospechan de la psicología de la consciencia, hasta el punto de que han pretendido eliminarla. El behaviorismo clásico de J. B. WATSON intentó reducir la psicología de la consciencia a la conductista, sosteniendo que la consciencia y los estados mentales no existen y son sólo fenómenos

¹⁶ Sigo su argumento tal como fue formulado en la Lección de doctorado «honoris causa» en la Universidad de Leipzig el 21 de mayo de 1996, con el título de *La posición de la psicología entre las ciencias*. Agradezco a E. GARZÓN VALDÉS y a F. LAPORTA el haberme llamado la atención sobre esta conferencia y haberme proporcionado el texto. Las citas entrecomilladas, salvo que se indique otra cosa, pertenecen al texto de esa conferencia. La explicación de las tres prioridades se encuentra también en «Valor, norma y acción en mis escritos filosóficos. Con un epílogo cartesiano», en *Doxa*, núm. 26, 2003 (trad. de L. MANRIQUE y P. NAVARRO) y en *In the Shadow of Descartes*, Springer, 1998. También se apuntan algunos aspectos de este argumento de VON WRIGHT (aunque menos definidos) en su libro *Sobre la libertad humana*, Paidós, 2002.

conductuales complejos. La neuropsicología o neurociencia (al menos sus defensores más radicales), por su parte, pretende reducir los estados mentales a los fenómenos neuronales. Esta reducción no es posible, según VON WRIGHT, dadas las relaciones peculiares que lo mental, lo conductual y lo neuronal, mantienen entre sí. Su argumento consiste en mostrar el juego de relaciones y prioridades que se dan entre estas tres dimensiones de los estados de consciencia. Estas relaciones son la *prioridad causal* de lo neuronal frente a lo conductual, la *prioridad epistemológica* de lo mental sobre lo neuronal y la *prioridad semántica* de lo conductual frente a lo mental. Veámoslas:

Supongamos que se produce de pronto un ruido y el sujeto A vuelve la cabeza inmediatamente hacia el lugar de donde éste procede. Podemos explicar el movimiento corporal aludiendo a ciertos procesos fisiológicos. En palabras de VON WRIGHT: «Las ondas sonoras se introducen en el oído interno y provocan allí procesos sensoriales centrípetos que se propagan hacia el centro de la audición en el cerebro. Desde allí son transmitidos de nuevo a un centro motor y originan impulsos motores centrífugos que se propagan hasta los músculos y, finalmente, dan como resultado movimientos del cuerpo». Ésta es una explicación fisiológica del movimiento corporal como reacción a un estímulo. Si queremos explicar causalmente cómo se produjo el movimiento corporal, debemos recurrir a este tipo de explicaciones, en las que lo neuronal aparece como causa de la conducta (del movimiento corporal). En esto consiste la *prioridad causal* de lo neuronal sobre lo conductual.

Supongamos ahora que le preguntamos al sujeto por qué ha vuelto la cabeza y éste responde que lo ha hecho porque ha oído un ruido y, dado que estaba esperando a alguien, quería saber si ya había llegado. Ahora tenemos, junto con la explicación fisiológica, una explicación racional de la conducta. La explicación neuronal y la explicación racional se sitúan en niveles distintos, pero en cierto sentido hay una correlación entre ellos: debe haber un sustrato neuronal de «percibir un sonido» y un sustrato neuronal de la razón aducida para volver la cabeza («querer averiguar si es la persona que esperaba»). De manera que sería posible, a partir de observaciones del sistema nervioso, descubrir si una persona ha oído un sonido o, incluso, si tiene cierto deseo, o una creencia, etc. Por ejemplo, cuando detectáramos en el sujeto A la actividad cerebral x podríamos decir que el sujeto ha oído un ruido. Ahora bien, para llegar a esto previamente hemos tenido que establecer una correspondencia entre «oír ruidos» y la actividad cerebral x, y para establecer inicialmente esta correspondencia necesitamos criterios distintos de los neuronales para identificar que el sujeto está oyendo algo. Y lo mismo ocurre con el correlato neuronal de las razones para hacer u omitir algo. En palabras de VON WRIGHT: «El hecho, por ejemplo, de que ciertas alteraciones hormonales sean indicativas de un estado de miedo o cansancio es algo que se ha podido establecer sobre la base de investigaciones anatómico-fisiológicas en seres vivos de los que ya se sabía que estaban asustados o cansados. Y para saber esto debemos saber ya qué significa estar asustado o cansado y saber cómo se puede comprobar eso sin tener que apelar a criterios intracorporales». Lo mental tiene, por tanto, *prioridad epistemológica* sobre lo neuronal: para establecer cuál es el correlato neuronal de estados mentales como «oír un sonido», «tener miedo», «estar sediento», etc., necesitamos previamente tener ya identificados estos estados mentales. Sólo si ya los tenemos identificados, podemos descubrir cuáles son sus correlatos neuronales. Por el contrario, observando sólo los continuos procesos neuronales, no podremos saber a qué estado mental se corresponden.

Surge ahora una nueva pregunta: ¿qué criterios usamos para identificar lo mental? La respuesta de VON WRIGHT, siguiendo a WITTGENSTEIN, es que descubrimos que un sujeto tiene una u otra razón o está en uno u otro estado mental a través de su conducta externa. «¿Cómo sabemos si un animal, por ejemplo un perro, ha oído un sonido? Normalmente porque vuelve la cabeza en dirección al sonido o de otro modo porque adopta una actitud de atención, o echa a correr si está asustado, o corre por el contrario en dirección al sonido si quiere averiguar lo que pasa». En el caso de los seres humanos, la conducta que usamos como criterio es frecuentemente —pero no siempre— verbal. Cuando le preguntamos al sujeto de nuestro ejemplo por qué ha vuelto la cabeza y nos responde que porque quería comprobar si ya ha llegado la persona que esperaba, su declaración es un tipo de conducta —conducta verbal— que nos indica qué razón tenía. El resto de su comportamiento (no verbal) nos ayuda a confirmar que ésta es realmente su razón. Usando la distinción wittgensteniana entre síntoma y criterio podríamos decir que lo neuronal es síntoma de lo mental, pero lo conductual es algo más fuerte: es el criterio que usamos para determinar que un sujeto está bajo uno u otro estado mental¹⁷. VON WRIGHT sugiere que la relación entre la conducta y lo mental debe entenderse como una relación semántica: «Aquéllos [los criterios conductuales] nos dicen qué quiere decir o significa que, por ejemplo, un sujeto oiga un ruido o esté asustado por algo o esté cansado». En esto consiste la *prioridad semántica* de lo comportamental frente a lo psíquico o mental. Sin conducta externa, no podríamos entender qué quiere decir estar en uno u otro estado mental, ni identificar que un sujeto está bajo uno u otro estado mental.

Lo neuronal, por tanto, causa la conducta; lo mental es necesario para poder identificar lo neuronal (y dota de sentido a la conducta) y lo conductual es el criterio que usamos para comprobar lo mental. Dadas estas tres prioridades, ninguna de estas dimensiones puede reducirse a otra. Lo conductual nos es necesario para entender los estados mentales de los demás (y los nuestros) y como criterio de identificación de los mismos, y tener identificados los estados mentales es necesario para encontrar sus correlatos neuronales. Todo intento de reduccionismo conlleva una pérdida importante en nuestra capacidad de comprendernos a nosotros mismos. Cuando la neurociencia pretende reducir lo mental a lo neuronal arguyendo que lo neuronal tiene prioridad causal frente a lo mental, se olvida del resto de dimensiones de la relación, respecto de las cuales lo neuronal es *secundario*.

* * *

Los dos argumentos que hemos examinado se dirigen contra las pretensiones de probar empíricamente la falsedad del libre albedrío y de reescribir las humanidades en términos neurocientíficos, pero no son una refutación de las tesis deterministas. De hecho, el propio VON WRIGHT considera que no existen genuinas relaciones causales entre fenómenos mentales y fenómenos físicos (lo que parece ser uno de los presupuestos del libre albedrío, aunque a veces se sustituya la idea de relación causal por la de «emergencia» para explicar la relación entre los dos niveles). La relación entre los fenómenos mentales y los físicos puede, en su opinión, describirse de la siguiente forma: «Siempre que pueda decirse correctamente que un evento en la naturaleza E causa un

¹⁷ VON WRIGHT, 2002: 113.

evento mental M, existe otro evento físico F tal que E causa F y M coincide temporalmente con (en todo o en parte) la duración de F. Y cuando, a la inversa, puede decirse correctamente que un evento mental M causa E en el mundo físico, existe otro evento físico F tal que F causa E y la duración temporal de M es incluida en F»¹⁸. Existen, por tanto, dos cadenas de acontecimientos (los mentales y los físicos) paralelas (cuyo ajuste se ha ido realizando a través de la educación). La secuencia física que culmina en un movimiento corporal (desde cierta actividad cerebral hasta la actividad muscular) es paralela a una secuencia mental (desde los deseos, creencias e intenciones hasta mi consciencia de haber actuado), sin que pueda decirse haya genuinas relaciones causales «de abajo hacia arriba» o «de arriba hacia abajo». Esto sitúa a VON WRIGHT cerca de posturas que presentan lo mental como un «epifenómeno» (algo, por tanto, cercano a la «ilusión» o el «engaño» de los deterministas más duros). Pero sin caer en el reduccionismo ni en la ingenuidad de creer que sólo con experimentos neurocientíficos podemos eliminar el mundo de lo mental o resolver el delicado problema de la libertad.

BIBLIOGRAFÍA

- CAPÓ, M.; NADAL, M.; RAMOS, C.; FERNÁNDEZ, A., y CELA CONDE, C. J., 2006: «Neuroética. Derecho y neurociencia», *Ludus Vitalis*, vol. XIV, núm. 25.
- CORTINA, A., 2010: «Neuroética: ¿Las bases cerebrales de una ética universal con relevancia política?», *Isegoría*, núm. 42.
- GARZÓN VALDÉS, E., 2007: «30 minutos de filosofía del derecho. Viejos y nuevos problemas», *Doxa. Cuadernos de Filosofía del Derecho*, 30.
- LIBET, B., 1999: «Do we have free will», *Journal of Consciousness Studies*, 6, núms. 8-9.
- LÓPEZ CORREDOIRA, M., 2008: «Naturaleza vs. libertad», *Thémata. Revista de Filosofía*, núm. 40.
- MORA TERUEL, F., 1996: «Neurociencias. ¿Hacia una nueva concepción del hombre?», *Árbor*, núm. 153/602.
- *Neurocultura. Una cultura basada en el cerebro*, Alianza Editorial.
- RUBIA, F., 2009: *El fantasma de la libertad. Datos de la revolución neurocientífica*, Crítica.
- SOLER GIL, F. J., 2009: «Relevancia de los experimentos de Benjamin Libet y de John-Dylan Haynes para el debate en torno a la libertad humana en los procesos de decisión», *Thémata. Revista de Filosofía*, núm. 41.
- SOON, Ch. S.; BRAAS, M., y HAYNES, J.-D.: «Unconscious determinants of free decisions in the human brain», *Nature Neuroscience*, 11.
- VON WRIGHT, G. H., 1972: *Causality and Determinism*, Columbia University Press.
- 1987: *Explicación y comprensión*, Madrid, Alianza Editorial (trad. de L. VEGA).
- 2002: *Sobre la libertad humana*, Paidós.
- 2003: «Valor, norma y acción en mis escritos filosóficos. Con un epílogo cartesiano», *Doxa*, núm. 26 (trad. de L. MANRIQUE y P. NAVARRO).

¹⁸ VON WRIGHT, 2006: 75.

