

Neurofilosofía y libre albedrío

Neurophilosophy and free will

*JOSÉ MANUEL MUÑOZ**

Resumen: Analizo la trascendencia de la neurociencia para el estudio de la relación entre determinismo y libre albedrío. Diversos trabajos muestran el vínculo entre la actividad de ciertas áreas nerviosas y el desempeño de las funciones volitivas, el trabajo de Benjamin Libet y de Daniel Wegner otorga gran importancia al inconsciente en nuestros actos, y hay pruebas de la influencia causal del entorno sociocultural sobre la voluntad y las vías neurales del individuo. Todo esto parece apuntar a una concepción determinista de nuestra voluntad, pero sostengo que ni esta ni la indeterminista están absolutamente claras actualmente y que la investigación futura debe profundizar en la neuroplasticidad, las decisiones distales y la deliberación.

Palabras clave: libre albedrío, neurofilosofía, córtex prefrontal, potencial preparatorio, voluntad consciente, causalidad descendente.

Abstract: I analyze the impact of neuroscience on the study of the relation between determinism and free will. Several papers show the link between activity on certain nervous areas and performance of volitional functions, Benjamin Libet and Daniel Wegner's work gives great importance to the unconscious in our acts, and there is evidence of causal influence of sociocultural environment on the individual will and neural pathways. All that seems to point to a deterministic conception of our will, but I hold that neither this one nor the indeterministic one are absolutely clear at present and that future research must go deep into neuroplasticity, distal decisions and deliberation.

Key words: free will, neurophilosophy, prefrontal cortex, readiness potential, conscious will, downward causation.

1. Introducción

En el presente trabajo me aproximo a la relación entre el libre albedrío y el determinismo desde el prisma de la neurofilosofía, disciplina que analiza la repercusión filosófica de los descubrimientos de la neurociencia. No parto para ello de la aceptación de una posición metafísica concreta, entre las muchas posibles, acerca del clásico problema mente-cuerpo¹. En la sección 2 abordo las bases biológicas de la volición centrándome en estudios de

Fecha de recepción: 08/01/2013. Fecha de aceptación: 12/06/2013.

* Departamento de Lógica, Historia y Filosofía de la Ciencia; Universidad Nacional de Educación a Distancia (UNED). Línea de investigación actual: neurociencia y libre albedrío. Correo electrónico: jomuor@gmail.com. El autor agradece a los editores y los revisores el esfuerzo invertido en la lectura y la consideración de este trabajo.

1 Este tipo de aproximación se inspira en un método denominado «neurofilosofía mínima», propuesto por Walter (2001, pp. 127- 34).

lesiones, de neuroimagen funcional y de electrofisiología. En la sección 3 presento los influyentes experimentos de Benjamin Libet acerca de la génesis temporal de actos voluntarios, así como opiniones que creo relevantes en el debate relacionado con dichos experimentos. La posición de Daniel Wegner, defendiendo que la voluntad consciente es una ilusión, es el tema tratado en la sección 4. Tras ello, en la sección 5, reclamo la relevancia de la causalidad descendente a la hora de tratar la neurofilosofía del libre albedrío. El trabajo finaliza con una breve conclusión que pretende orientar al lector hacia posibles vías de profundización. Para un mejor aprovechamiento de la lectura sería recomendable un conocimiento básico de neuroanatomía, así como de las posturas que existen sobre la relación entre determinismo y libre albedrío (compatibilismo, libertarismo, determinismo duro, incompatibilismo duro, etc.)², ambos temas de gran calado que, por razones de espacio, no han sido desarrollados aquí.

2. Bases biológicas de la volición

Existen diversos trabajos que muestran la relación entre las lesiones en ciertas áreas nerviosas y la modificación de funciones volitivas. Damasio y van Hoesen (1983) han vinculado una lesión en el cíngulo anterior con el trastorno conocido como «mutismo acinético», en el cual los afectados son incapaces de moverse y comunicarse de forma normal a pesar de estar conscientes. El síndrome de la mano extraña, consistente en la realización de movimientos con la mano en contra de la voluntad del afectado, y descrito, por ejemplo, por Gasquoin (1993), también ha sido asociado con lesiones en el cíngulo anterior, así como en sus alrededores (Walter 2002, p. 571). Damasio *et al.* (1994), por su parte, revisitan el conocido caso de un capataz de construcción ferroviaria llamado Phineas Gage, quien en un accidente laboral sufrió una seria lesión en su córtex prefrontal (de ahora en adelante, etiquetado como CPF). Gage sobrevivió al accidente y, en muchos sentidos, su comportamiento siguió siendo perfectamente normal: tenía la misma inteligencia que antes del accidente, no había perdido la memoria, podía moverse y hablar sin ningún problema, adquirir nuevos conocimientos... Sin embargo, se convirtió en una persona soez, irreverente e incapaz de realizar tareas que requirieran planificación y organización, cuando antes del accidente se le tenía por una persona responsable y bien adaptada a las convenciones sociales. Por otro lado, Lhermitte (1983) describe cinco casos diferentes de lesiones en las áreas laterales del córtex frontal (de ahora en adelante, etiquetado como CF). Todos ellos guardaban en común la manifestación de la denominada «conducta de utilización», en la cual el afectado utiliza automáticamente un objeto por el simple hecho de serle presentado mediante estímulos visuales y/o táctiles. Lhermitte sugiere que estas lesiones tienen el efecto de eliminar la inhibición que el lóbulo frontal es capaz de ejercer sobre el parietal, desencadenando la acción de este último y haciendo que el paciente dependa enormemente de los mencionados estímulos.

Las técnicas de neuroimagen funcional nos dan la oportunidad de observar qué cambios ocurren en el cerebro mientras tiene lugar una actividad, permitiéndonos detectar qué áreas manifiestan activación y por tanto participan en ella. Es decir, nos proporcionan un seguimiento de la actividad cerebral. Entre estas técnicas destacan especialmente dos: la tomo-

2 En un artículo reciente publicado en esta misma revista (Muñoz 2012) realicé una aproximación sistemática a este asunto.

grafía por emisión de positrones, o PET (*Positron Emission Tomography*), y la imagen por resonancia magnética funcional, o fMRI (*functional Magnetic Resonance Imaging*). Dolan *et al.* (1993), por ejemplo, presentan un estudio realizado con treinta pacientes de esquizofrenia y cuarenta de depresión. Detectaron cómo se reducía la actividad en el CPF dorsolateral izquierdo de los sujetos con pobreza del habla (dificultad para hablar con espontaneidad). Frith *et al.* (1991) y Hyder *et al.* (1997) describen una activación en esa misma zona durante el desempeño de actos volitivos. Además, Spence *et al.* (1997) hallaron una especial actividad tanto en el córtex parietal como en el córtex del cíngulo en sujetos con delirios de control ajeno (un síntoma común entre los esquizofrénicos que consiste en la convicción por parte del afectado de que alguien externo a él posee el control de sus pensamientos y sus acciones).

La electrofisiología es la disciplina que se ocupa de investigar la actividad eléctrica de las células y los tejidos del organismo. Los estudios de esta especialidad han contribuido al avance en el conocimiento acerca de los mecanismos nerviosos que subyacen al proceso volitivo. Por ejemplo, Schultz (1999) ha observado en el macaco cangrejero (*Macaca fascicularis*) que el cuerpo estriado, puente de información entre los ganglios basales y el córtex cerebral, cumple un importante papel en el procesamiento de movimientos voluntarios y en la regulación del comportamiento dirigido a objetivos. Hay más autores que han desarrollado su investigación en torno a la electrofisiología de la volición. Entre ellos destaca Benjamin Libet, cuyos trabajos no se centran en las estructuras implicadas en el proceso volitivo sino más bien en su dimensión temporal. Por su gran repercusión, dedicaremos a estos trabajos la sección 3.

Parece que, como podemos ver, existe un vínculo innegable entre la actividad de ciertas áreas nerviosas y la génesis de decisiones de carácter voluntario. El CF juega un papel central en conjunción con otras estructuras estrechamente relacionadas con él, como el lóbulo parietal, la circunvolución del cíngulo (en especial el cíngulo anterior) y los ganglios basales (entre los que destaca el cuerpo estriado). Tal y como explica Walter (2001, p. 253), a grandes rasgos «la función del córtex frontal es *organizar el comportamiento a lo largo del tiempo*. Mientras que el área motora se preocupa de organizar y ejecutar movimientos, es trabajo del córtex prefrontal “controlar” procesos cognitivos que garanticen que movimientos apropiados son seleccionados en el momento correcto para el lugar correcto». El CPF desempeña un rol especialmente importante. Esta área guarda una serie de conexiones de distinta índole (Squire *et al.* 2008, pp. 1201-3) con sistemas de tipo sensorial y motor, con el sistema límbico (involucrado en las respuestas emocionales y formado por estructuras como el hipotálamo, el hipocampo y la amígdala, entre otras) y con los ganglios basales. Esta variedad de interconexiones convierte al CPF en un sistema multimodal y se considera un requisito indispensable para ejercer un control de la cognición (*ibid.*, pp. 1211-2). A la multimodalidad se unen propiedades de sus neuronas tales como la capacidad para mantener su actividad y conservar información, el procesamiento de reglas relacionadas con tareas, la flexibilidad para modificar su actividad en respuesta a cambios en las necesidades, o la codificación de información relacionada con recompensas. Todo ello hace que el CPF sea una estructura fundamental para ejercer el control de la cognición y desempeñar un papel central en el procesamiento de comportamientos complejos (pp. 1211-7). El CPF, en palabras de Spence y Frith (1999, p. 27), «parece estar involucrado en mantener posibles acciones en mente antes de que sean ejecutadas, y en seleccionar cuál será llevada a cabo» (si bien

estos autores se refieren en concreto a la región dorsolateral del CPF). Es posible que la selección tenga lugar mediante la inhibición de todas las acciones posibles a excepción de una, que será la que finalmente se produzca (Goldman-Rakic 1987). Esta posibilidad estaría en sintonía «con la observación de que los pacientes con daño en el córtex prefrontal son a menudo incapaces de inhibir respuestas inapropiadas para su ambiente» (Spence y Frith 1999, p. 19), tal y como atestiguan casos como el de Phineas Gage, mencionado anteriormente. También podría ser compatible con la opción, propuesta por Benjamin Libet y tratada en la sección 3, de que el inconsciente constituya una fuente de opciones distintas y nuestra voluntad consciente efectúe una selección en forma de veto.

Si bien es indudable la conexión entre la actividad de las áreas mencionadas en el párrafo anterior y el desarrollo del proceso de volición, hay matices que sugieren evitar una excesiva simplificación de la situación. En primer lugar, hay evidencias de un fuerte vínculo entre la volición y la emoción. Koenigs *et al.* (2007), por ejemplo, presentan un estudio en el cual los individuos con lesiones en su CPF ventromedial, región que se considera fundamental para la aparición de emociones de tipo social, manifestaban una alta tendencia a esgrimir opiniones de tipo utilitario en situaciones morales con fuerte contenido emocional, mientras que sus opiniones en situaciones menos potentes emocionalmente se ajustaban más a la tendencia habitual. Además, parece cada vez más claro que «la memoria y la emoción están íntimamente conectadas en el cerebro, y alterar nuestra experiencia de una puede afectar drásticamente a nuestra experiencia de la otra» (Boleyn-Fitzgerald 2010, p. 90). Teniendo en cuenta esto, podemos pensar en una especie de triángulo sistémico que conecta de modo inextricable la memoria, la emoción y la volición. En segundo lugar, encontramos, según Walter (2001, p. 256), un problema de «regresión infinita» hacia los precursores de la voluntad en el cerebro. Fuster (1995, p. 296) nos avisa sobre ello de esta forma: «asignar la voluntad a cualquier región frontal del cerebro nos lleva obviamente a preguntarnos por una orden previa sobre esa región desde otra estructura; la misma cuestión puede ser planteada sobre esa otra estructura, cualquiera que esta fuera, y entonces sobre su precursora, y así sucesivamente». En la misma obra, Fuster plantea atender a un ciclo que él bautiza como «ciclo percepción-acción», en el cual existe un bucle de influencia mutua entre los córtex frontal y posterior y las estructuras subcorticales del cerebro, de manera que no podemos pensar en un «verdadero origen» (p. 296). En tercer lugar, la propia metodología de los estudios podría ser responsable de que extraigamos conclusiones que nos conducen a localizar determinados procesos neurales en regiones concretas del CPF (Walter 2011, p. 524). De hecho, hay teorías que presentan la actividad del CPF de un modo integrante (*e.g.*, Miller y Cohen 2001). En cuarto y último lugar, debemos tener en cuenta importantes datos proporcionados por la fisiología y la parasitología que guardan relación con la volición. Parece ser que ciertas moléculas como la dopamina o la serotonina, entre otras, ejercen funciones de modulación en el sistema nervioso influyendo en procesos como la atención, la motivación, la impulsividad, etc. (Fellous 1999; Walter 2011, pp. 525-6). Además, la infección por el parásito *Toxoplasma gondii* repercute en factores como el nivel de dopamina, el tiempo de reacción e, incluso, las posiciones morales (Flegr *et al.* 1996; Flegr 2007, 2013).

Teniendo en cuenta los matices expuestos, podemos concluir que resultaría un tanto precipitado situar la volición, el libre albedrío o la conciencia en compartimentos estancos de nuestro sistema nervioso (*cf.* Boleyn-Fitzgerald 2010, p. 113). Probablemente sería más

prudente pensar que, si bien ciertas áreas resultan decisivas en la generación de nuestras decisiones voluntarias (en especial el CPF), dichas áreas no podrían ejercer esta función sin estar integradas en un amplio sistema formado por diversas estructuras, en el cual la cognición, la emoción y la memoria se encuentran fuertemente interrelacionadas. La capacidad de un área nerviosa para funcionar de forma independiente y para, a su vez, integrarse de forma coordinada en un conjunto es conocida como «metaestabilidad» (Kelso y Tognoli 2009). El mencionado sistema, además, no puede ser aislado completamente del individuo en su conjunto, sujeto a su vez a influencias del entorno en el cual se desenvuelve (la infección por *Toxoplasma gondii* sería un ejemplo de ello, aunque también debemos pensar en factores de tipo social, cultural, etc.). Por lo tanto, la actividad de las áreas nerviosas mencionadas es una *condición necesaria* pero no forzosamente *suficiente* para que tenga lugar el proceso volitivo. También hay que tener en cuenta que la correlación entre dos fenómenos (en este caso, la actividad nerviosa y los procesos volitivos) no indica por sí misma que uno de ellos sea la única causa necesaria para que el otro tenga lugar. Por todas estas razones, sería inconveniente posicionarse a favor de ciertas posiciones en el problema mente-cuerpo, como el materialismo eliminativo o la teoría de la identidad, basándonos *exclusivamente* en la correlación actividad nerviosa – volición.

3. Benjamin Libet y la génesis temporal de actos voluntarios

El interés de Libet en la electrofisiología de la volición se centra en la relación existente entre la aparición de un cambio eléctrico cerebral denominado «potencial preparatorio» (*readiness potential*, de ahora en adelante RP) y el origen de los actos voluntarios. En Libet (1999) encontramos un análisis integrador de su investigación que, si bien menciona trabajos previos (e.g.: Libet, Wright y Gleason 1982, 1983; Libet *et al.* 1983; Libet 1985), nos será útil aquí como guía para exponer sus aportaciones de manera global.

La definición de libre albedrío con la que Libet trabajó requería dos condiciones (1999, p. 47): (1) que el acto fuera endógeno en el sentido de que no hubiera constricciones que pudieran afectar al desempeño del acto voluntario, y (2) que el sujeto de estudio sintiera que poseía el control sobre dicho acto. En la mayoría de las pruebas no había restricción temporal, así que «los sujetos efectuaban un simple golpe o flexión de la muñeca [o los dedos] en cualquier momento que sentían el impulso o deseo de hacerlo» (p. 49). El RP aparecía con anterioridad a la actividad de los músculos que participaban en la ejecución del acto (pp. 48-9), pero tan importante como registrar dicha aparición era averiguar en qué momento eran los individuos del experimento conscientes de su decisión. Para ello se utilizó un reloj modificado en el cual un punto de luz efectuaba un giro completo cada 2.56 segundos, y no cada 60 segundos como sucede en los relojes normales, con el resultado de que cada segundo imaginario del reloj adaptado se correspondía con 43 milisegundos reales (pp. 49-50). Tras completar la misión de flexionar la muñeca o los dedos en el instante deseado, cada sujeto tenía que comunicar en qué posición del reloj modificado se encontraba el punto de luz en el momento en que era consciente de su deseo de efectuar dicha flexión (momento etiquetado por Libet con la letra W). Asimismo, la precisión con la que cada uno realizaba la mencionada comunicación fue debidamente calibrada, encontrando un error medio razonablemente escaso consistente en -50 milisegundos (*id.*). Todo el procedimiento de registro de los RP

y de los tiempos en el reloj modificado fue realizado cuarenta veces para cada individuo, y se obtuvo un promedio de los resultados.

Libet descubrió dos tipos de RP (p. 48). Por un lado estaban los que aparecían cuando el sujeto afirmaba haber tenido un plan para efectuar el movimiento dentro del plazo de un segundo (RP I). Por otro, aquellos que tenían lugar de forma espontánea a juicio del sujeto y, en consecuencia, sin haber efectuado ningún plan previo para ejecutar la acción (RP II). Los experimentos mostraron (pp. 50-1) que en las flexiones espontáneas el RP II aparecía a los -550 milisegundos, previamente a la actividad de los músculos involucrados en el movimiento (considerada como tiempo 0). En el caso del RP I, registrado en las flexiones planificadas, el registro tenía lugar a los -1050 milisegundos. Además, se detectó la aparición de W a los -150 milisegundos (tras corregir su desviación de -50 segundos) independientemente de si la flexión era espontánea o planeada. Esquemmatizando:

RP I [-1050 ms] ----- W [-150 ms] --- Actividad muscular [0 ms]

RP II [-550ms] ----- W [-150 ms] --- Actividad muscular [0 ms]

Se observa una diferencia de 900 milisegundos entre RP I y W, y una de 400 milisegundos entre RP II y W. Esta última, apunta Libet, podría ser mayor porque «el proceso iniciador real en el cerebro probablemente comienza antes que nuestro RP registrado, en un área desconocida que entonces activa el área motora suplementaria en el córtex cerebral» (p. 51). El RP, añade el autor, podría surgir precisamente del área motora suplementaria.

A la vista de los resultados obtenidos, Libet sostiene lo siguiente:

(1) La iniciación de los actos voluntarios tiene lugar de manera inconsciente, ya que se produce antes de que el sujeto sea consciente de su decisión (*id.*).

(2) La consciencia, a pesar de la conclusión anterior, podría intervenir en el proceso volitivo (pp. 51-2). El sujeto dispone en principio de 150 milisegundos para intervenir de manera consciente. Este es el tiempo que media entre la aparición de W y la actividad muscular. No obstante, el córtex motor primario, mediante un proceso ya irreversible, se comunica con la región motora espinal durante los últimos 50 milisegundos previos a la actividad muscular. Por esta razón el plazo real queda reducido a 100 milisegundos. Durante este breve espacio de tiempo, el individuo podría vetar de forma consciente el proceso de movimiento muscular que había sido iniciado inconscientemente. Debido al diseño de los experimentos, en los que no se efectuaba registro eléctrico si no se realizaba un movimiento, no fue posible detectar el veto a pesar de que los participantes afirmaron en alguna ocasión haberlo llevado a cabo. Sí se pudo detectar, con otro tipo de diseño experimental, procesos de veto en movimientos previstos para ser realizados en un momento ya fijado. En estos casos sucedió que, tras la aparición de un RP prediciendo un movimiento, los sujetos vetaron este con un registro temporal que se encontró entre los 100 y los 200 milisegundos antes del mencionado momento.

(3) No parece que la voluntad consciente pueda hacer las veces de «gatillo» que facilite que un proceso volitivo evolucione hasta el punto en el que tiene lugar una acción (p. 52). Si se piensa en los actos voluntarios que realizamos de forma automática, se puede observar que pueden producirse sin que se detecte una intención preliminar consciente de ejecutarlos.

Además, con anterioridad a este tipo de actos se detecta un RP que tiene una amplitud muy pequeña y una muy escasa duración (*id.*).

(4) Cabe preguntarse si la aparición del veto puede tener un origen inconsciente, siendo el sujeto más tarde consciente de dicho veto, de un modo similar a lo que sucede con la actividad muscular estudiada (pp. 52-3). Sin embargo, no sería necesario que esto sucediera. La razón es que *el veto constituye un mecanismo de control*, contrariamente a lo que sucede cuando se es consciente de querer hacer un movimiento, y «No hay imperativo lógico en ninguna teoría mente-cerebro [...] que requiera actividad neural específica que preceda y determine la naturaleza de una función de control consciente» (p. 53). Tampoco se ha demostrado experimentalmente que sean imprescindibles mecanismos inconscientes para que tenga lugar el mecanismo de control (*id.*). En cualquier caso, no podemos descartar que dichos mecanismos den lugar a factores sobre los cuales se fundamente la decisión de vetar y que sean necesarios para que el sujeto tenga conocimiento del veto que realiza, «pero el *contenido* de ese conocimiento (la decisión real de vetar) es una característica separada que no necesita tener el mismo requisito» (*id.*).

(5) A pesar de que los actos investigados son eminentemente simples, los resultados tienen un gran impacto a la hora de considerar todos los tipos de actos voluntarios, independientemente de su grado de complejidad o planificación (pp. 53-4). Distinguiendo la intención «de actuar ahora» [*sic*] de un eventual proceso de deliberación previa, y teniendo en cuenta que tanto el RP I (en casos de cierta planificación para actuar) como el RP II (en casos de acciones espontáneas) aparecen con anterioridad a W (deseo consciente de efectuar el movimiento), se puede decir que el orden temporal de sucesos hallado para la intención «de actuar ahora» sería extrapolable a todos los actos voluntarios.

En conclusión, Libet considera que el libre albedrío tiene cabida en el ser humano, pero no como creador sino más bien como censor: los mecanismos inconscientes nos proporcionarían una fuente de opciones variadas, y entre ellas nuestra voluntad consciente, mediante el uso del veto, efectuaría la selección de cuáles terminarían finalmente desembocando en actos volitivos (p. 54).

Los hallazgos de Libet han dado lugar a una amplia discusión. Sirva como muestra lo que sigue. Walter (2011, pp. 519-20), por ejemplo, es muy crítico con el planteamiento metodológico de los experimentos de Libet. Gomes (1999, p. 65), por su parte, sostiene que el término «intención» puede dar lugar a diversas interpretaciones, y otorga relevancia al hecho de que en algunas de las pruebas diseñadas por Libet se hallara un RP precediendo a un veto en un movimiento concebido para su ejecución en un momento previsto, ya que demuestra que puede generarse un RP sin necesidad de que se termine efectuando un acto, y por esta razón no se debe considerar que el RP equivale a «la decisión irrevocable de actuar» (p. 68). En cuanto a Mele (2011, p. 501) efectúa una distinción entre las intenciones y decisiones de tipo «proximal» (relacionadas con acciones que se llevarán a cabo de forma inminente) y de tipo «distal» (relacionadas con acciones que se efectuarán más adelante)³. Obsérvese que en los experimentos de Libet podríamos estar tratando con ejemplos del primer tipo, y así lo interpreta también Mele, aunque sostiene que en dichos experimentos cabe la posibilidad de que los RP I y RP II registrados en los casos de movimientos ejecutados no correspondan a

3 Nótese que el autor se refiere a intenciones/decisiones inmediatas y previas (futuras), respectivamente.

intenciones proximales de efectuar un movimiento, sino a sus causas potenciales (pp. 506-7). Las intenciones proximales, en ambos casos, harían su aparición más adelante (p. 507).

4. Daniel Wegner y la ilusión de la voluntad consciente

En las discusiones recientes acerca del libre albedrío ha penetrado con fuerza la obra *The Illusion of Conscious Will* (Wegner 2002), que, dicho sea de paso, otorga una gran relevancia a los experimentos de Libet. Su autor, el psicólogo Daniel Wegner, sostiene en ella la idea de que las intenciones conscientes no juegan un rol causal en la ejecución de las acciones que las subsiguen. Ofrece argumentos que apoyan la idea de que los mecanismos que provocan que las acciones sean producidas por la mente se diferencian en anatomía y psicología de los mecanismos que conducen a que se produzca la voluntad consciente (*ibid.*, pp. 29-61), y afirma que esta última «es una ilusión en el sentido de que *la experiencia de querer conscientemente una acción no es una indicación directa de que el pensamiento consciente ha causado la acción*» (p. 2). Tanto nuestra intención consciente de actuar como nuestra acción voluntaria tienen un origen inconsciente, pero cuando asociamos ambos causalmente, mediante un vínculo que realmente no es cierto, se produce la experiencia de voluntad consciente (pp. 63-98). Esta ilusión está protegida por las personas «porque tienen un *ideal de agencia consciente* que guía sus inferencias sobre qué deben haber conocido y querido incluso cuando efectúan acciones que no pretendían» (p. 146), así que «aspiran a ser agentes ideales que conocen todas sus acciones por adelantado» (p. 145). Wegner, además, equipara la voluntad consciente al libre albedrío (2004, p. 656): «nuestra discusión ha sido realmente *sobre* la experiencia de libre albedrío, examinando en detalle cuándo la gente lo siente y cuándo no. La idea especial que hemos estado explorando es explicar la experiencia de libre albedrío en términos de procesos deterministas o mecanicistas».

Numerosas observaciones parecen jugar a favor de la postura de Wegner. Un ejemplo puede ser la conducta de utilización descrita por Lhermitte (1983) y mencionada en la sección 2. Otro ejemplo podemos encontrarlo en el trabajo de Soon *et al.* (2008), quienes, mediante fMRI, observaron indicios de toma de decisiones, en forma de activación en el córtex prefrontal y parietal, con varios segundos de antelación al registro de la actividad consciente. Sin embargo, Wegner no está exento de críticos. Nahmias (2002, p. 536), por ejemplo, destaca la escasa importancia que este autor otorga a las intenciones conscientes vinculadas a acciones planeadas, deliberadas o prolongadas. También critica (p. 533) que las «excepciones a la regla de que nuestras experiencias conscientes de nuestras acciones se ajustan a aquellas acciones» sean convertidas por Wegner en una norma general, ya que el hecho de que «la consciencia de querer una acción *pueda* ser separada de la acción no significa que cuando no están separadas, la consciencia no sea causalmente relevante». Hay trabajos, en cambio, que llegan a conclusiones similares a las de Wegner, pero con argumentos distintos. Es el caso de Lau, Rogers y Passingham (2007), quienes, mediante un estudio realizado con una técnica estimuladora del córtex conocida como «técnica de estimulación magnética transcraneal», o TMS (*Transcranial Magnetic Stimulation*), hallaron que el momento en que el sujeto experimentaba tener la intención de ejecutar una acción puede tener lugar con posterioridad a dicha acción. A juicio de los autores (*ibid.*, p. 81), esto podría apoyar la idea de que la voluntad consciente es una ilusión de un modo más convincente

que el de Wegner. De todos modos, como ellos mismos reconocen, los resultados obtenidos no permiten descartar que el momento en que se experimenta la intención de actuar pueda estar presente también antes de la acción y desempeñar un papel causal en el proceso (p. 89). Lo que sí permiten es contradecir la idea de que «la experiencia de intención [...] está completamente determinada antes de una acción» (*id.*)⁴.

5. Neurofilosofía y causalidad descendente

Schwartz (1999) ha demostrado que es posible modificar la sintomatología y los circuitos cerebrales asociados al trastorno obsesivo-compulsivo entrenando la atención del afectado mediante terapia cognitivo-conductual. Otros estudios han puesto de manifiesto que, con un debido entrenamiento, se puede controlar voluntariamente la activación del córtex del cíngulo anterior rostral hasta lograr una disminución en la percepción del dolor (deCharms *et al.* 2005). También hay pruebas de que, mediante estrategias cognitivas de regulación emocional, es posible modificar la manera en que se conecta la amígdala con el CPF y, así, disminuir el efecto diferencial que la genética provoca en el nivel de respuesta de la amígdala de distintos individuos a estímulos de carácter indeseable (Schardt *et al.* 2010). Estos ejemplos atestiguan la capacidad del ser humano para modificar el funcionamiento de su sistema nervioso, denominada «neuroplasticidad».

Frith (2009) subraya el control que el contexto social y cultural ejerce en la voluntad del individuo, y considera que es en dicho contexto donde debemos enmarcar la interacción que se produce entre experimentadores y sujetos de estudio en los experimentos relativos a la ejecución de actos volitivos. Podríamos decir que dicha interacción es extensible a los tres ejemplos mencionados, pues las prácticas llevadas a cabo por los participantes difícilmente se pueden concebir sin tener en cuenta a los experimentadores. En este sentido, podemos considerar que tiene lugar un influjo causal procedente del ambiente que rodea a dichos participantes; con el resultado de modificar mecanismos neurales. Por otra parte, el control del entorno sociocultural queda patente también en trabajos que no son propiamente de neurociencia. Por ejemplo, Kurzban, DeScioli y O'Brien (2007) observaron que en sus experimentos el castigo de carácter moral era mayor si era infligido bajo observación.

La realidad y las relaciones causales dentro de su seno pueden ser descritas en forma de jerarquía. Ellis (2009, p. 64) enumera ocho niveles en función de las disciplinas que los estudian: (1) física de partículas, (2) física atómica, (3) química, (4) bioquímica, (5) biología celular, (6) fisiología, (7) psicología, y (8) sociología, economía, política. Si pensamos en un universo organizado de forma jerárquica, podemos decir que la aproximación que habíamos realizado hasta ahora al problema del libre albedrío se había basado tanto en un sentido ascendente de la causalidad, desde nuestro cerebro hacia nuestro comportamiento, como en una causalidad dentro de un mismo nivel, que se manifiesta en la relación entre distintas regiones nerviosas (*cf.* sección 2). Sin embargo, si tenemos en cuenta la neuroplasticidad

4 Conviene aclarar que es controvertida la concepción de la intención como una experiencia. Tal y como destaca Moya (2012, p. 112), Wittgenstein sostenía que las intenciones (entendidas como propósitos) no son experiencias o sensaciones, ya que no poseen duración temporal.

y la influencia causal del entorno sociocultural sobre la voluntad del individuo apreciaremos un tipo de aproximación que afronta el problema desde una «causalidad hacia abajo» (*downward causation*), concepto introducido por Campbell (1974).

Si bien se trata de un concepto estelar entre los autores partidarios de la emergencia, la causalidad descendente no tiene por qué ser entendida solamente dentro de un ámbito emergentista. Por otra parte, la organización jerárquica del universo puede servir como argumento para opinar en contra de la propia emergencia o, en general, en contra de la causalidad hacia abajo. Esto se puede observar en el trabajo de Moya (2011), quien ha analizado el reto que supone para la causalidad mental la concepción jerárquica según la cual hay una dependencia unidireccional de los niveles superiores respecto de los inferiores. Por lo tanto, podemos comprobar que la jerarquía del universo no tiene por qué ajustarse a una posición ontológica concreta sino a una descripción abarcable de la realidad: diferentes disciplinas del saber estudian realidades de distintas características. Pero si tenemos en cuenta los hallazgos mencionados anteriormente, que parecen mostrar la existencia de fenómenos de causalidad descendente, tendremos que reconocer que el reduccionismo no basta para afrontar la neurofilosofía del libre albedrío: no todo puede explicarse desde abajo hacia arriba. La causalidad en sentido ascendente es, no cabe duda, imprescindible para comprender la volición. Pero es solo una parte de la historia. Si pretendemos describir la causalidad en el fenómeno volitivo de una forma lo más completa posible, debemos adoptar una triple perspectiva: ascendente, intranivel y descendente (*cf.* Ellis 2009, pp. 78-9).

Es muy importante recalcar que el hecho de negar la suficiencia de la reducción en la neurofilosofía del libre albedrío no significa que necesariamente debamos descartar la verdad del determinismo. Un determinista podría alegar que la influencia del contexto cultural y social sobre la voluntad del individuo y sus vías neurales puede estar determinada desde un origen externo al agente volitivo. Si bien es cierto que la reducción es empleada a menudo para defender el determinismo, no debemos caer en el error de confundir ambos conceptos o de enlazarlos de modo inseparable.

6. Conclusión

Las técnicas de la neurociencia han permitido ahondar en los mecanismos cerebrales relacionados con la volición. Sin embargo, aproximándonos al debate sobre el libre albedrío sin efectuar posicionamiento metafísico previo acerca del problema mente-cuerpo, la conclusión general que podemos extraer es que la neurociencia actual no nos permite alcanzar una solución a dicho debate (*cf.* Mele 2011, p. 513-4), y que por ello (y sin menoscabo de la necesaria ayuda de la reflexión filosófica) debe ahondar en sus avances, especialmente en lo relativo a *la neuroplasticidad, las decisiones distales y la deliberación*. Esta conclusión parte principalmente de tres argumentos:

[I] Algunas estructuras nerviosas, entre las que el córtex prefrontal es especialmente importante, parecen ser responsables de la función volitiva. También hay que considerar la estrecha conexión de la voluntad con la emoción. Si a esto le unimos el hecho de que el trabajo de Libet y Wegner apunta hacia un gran peso de los procesos inconscientes en nuestros actos, y además consideramos el influjo causal descendente del entorno sociocultural sobre la voluntad y las vías neurales del individuo, una primera impresión podría apuntar a que

nuestra voluntad sería explicable mediante teorías deterministas. Sin embargo, hay razones para no abandonar el indeterminismo:

(a) Resulta problemático hallar un origen causal último de nuestra voluntad a nivel anatómico y, además, la actividad de las áreas nerviosas implicadas es condición necesaria pero no obligatoriamente suficiente para el proceso volitivo, tal y como expliqué en la sección 2.

(b) Las posturas de Libet y Wegner han recibido diversas críticas, como dejé patente en las secciones 3 y 4.

(c) La causalidad descendente, cuya existencia es motivo de discusión, no es incompatible con mecanismos indeterministas. No podemos descartar en absoluto que, por ejemplo, indeterminaciones de tipo cuántico intervengan en fenómenos ecológicos o climatológicos que influyen en la configuración del entorno sociocultural en el que nos desenvolvemos, o que, dado que la sociedad y la cultura son creaciones humanas, una conjunción de indeterminaciones en diversos agentes intervenga en la producción de dichas creaciones.

[II] Libet no extrae de sus experimentos una vinculación del indeterminismo con la consciencia en nuestros actos (Libet 1999, p. 55), pero sí piensa que decantarse por un libre albedrío indeterminista es igualmente válido, incluso preferible por ciertas razones, que rechazarlo (*ibid.*, pp. 56-7). Por lo tanto, su opinión supone un guiño al libertarismo. Teniendo en cuenta que deja la puerta abierta a la posibilidad de que ejerzamos nuestra libertad mediante el veto consciente de algunas opciones de acción frente a otras, los libertaristas podrían, en principio, utilizar dicha posibilidad para apoyar sus teorías, pero se encontrarían con las siguientes dificultades:

(a) Es necesario explicar cómo es posible que se pueda seleccionar conscientemente entre opciones de naturaleza inconsciente. Parece existir un vacío explicativo entre ambas fases, con dificultades de inteligibilidad y no aclarado por Libet.

(b) Cabe la posibilidad de que el mecanismo de veto tenga lugar sin que exista un control último, capacidad que consiste en ser la fuente, el autor, el origen último de las propias decisiones y/o acciones. Nótese que, a simple vista, las opciones de acción y la selección en forma de veto de los que habla Libet podrían considerarse, respectivamente, como posibilidades alternativas y como control último, ambas condiciones necesarias para la libertad según el libertarismo. Sin embargo, un libertarista consideraría que un veto consciente determinista no constituye en absoluto un control último, y en el caso de un veto consciente indeterminista tendría que lidiar con el argumento de *Mind*, según el cual no hay control último si la voluntad surge del azar.

(c) Hay que aclarar cómo podría el indeterminismo originar tanto mecanismos inconscientes (surgimiento de opciones de acción) como conscientes (proceso de veto).

[III] Aunque tomáramos la obra de Libet y de Wegner como un apoyo para defender el determinismo de nuestras decisiones y nuestras acciones, hay razones para sostener que la neurociencia no nos permite descartar la existencia de un libre albedrío incompatible con el determinismo, a pesar de las dificultades para aceptarla. A saber:

(a) La neuroplasticidad que nuestro cerebro es capaz de mostrar en algunas ocasiones abre la posibilidad de que mecanismos cerebrales indeterministas aún no descubiertos puedan moldear nuestro cerebro y modificar nuestras capacidades al ejercer nuestra libertad (si bien estos mecanismos necesitarían de una descripción complementaria que explicara coherentemente la causalidad mental, de tipo descendente, y el control último por parte del agente).

(b) No podemos dejar a un lado el riesgo que supone explicar decisiones relevantes de tipo distal de la misma forma en que explicamos decisiones proximales irrelevantes, como las que aparecen en los experimentos de Libet (Mele 2011, p. 507).

(c) No debemos olvidar la importancia de la deliberación para nuestras decisiones y nuestros actos, tratada secundariamente por Wegner (*cf.* Nahmias 2002, p. 536) y por Libet en beneficio de la espontaneidad.

Bibliografía

- BOLEYN-FITZGERALD, M. 2010. *Pictures of the mind: What the new neuroscience tells us about who we are*. Upper Saddle River, New Jersey: FT Press.
- CAMPBELL, D.T. 1974. 'Downward causation' in hierarchically organised biological systems. En *Studies in the philosophy of biology: Reduction and related problems*, eds. F.J. Ayala & T. Dobzhansky, 179-86. Berkeley & Los Ángeles: University of California Press.
- DAMASIO, A.R., & G.W. VAN HOESEN. 1983. Emotional disturbances associated with focal lesions of the limbic frontal lobe. En *Neuropsychology of human emotion*, eds. K.M. Heilman & P. Satz, 87-103. Nueva York: Guilford.
- DAMASIO, H., T. GRABOWSKI, R. FRANK, A.M. GALABURDA, & A.R. DAMASIO. 1994. The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science* 264: 1102-5.
- DECHARMS, R.C., F. MAEDA, G.H. GLOVER, D. LUDLOW, J.M. PAULY, D. SONEJI, J.D.E. GABRIELI, & S.C. MACKEY. 2005. Control over brain activation and pain learned by using real-time functional MRI. *Proceedings of the National Academy of Sciences of the United States of America* 102: 18626-31.
- DOLAN, R.J., C.J. BENCH, P.F. LIDDLE, K.J. FRISTON, C.D. FRITH, P.M. GRASBY, & R.S.J. FRACKOWIAK. 1993. Dorsolateral prefrontal cortex dysfunction in the major psychoses; symptom or disease specificity? *Journal of Neurology, Neurosurgery, and Psychiatry* 56: 1290-4.
- ELLIS, G.F.R. 2009. Top-down causation and the human brain. En *Downward causation and the neurobiology of free will*, eds. N. Murphy, G.F.R. Ellis & T. O'Connor, 63-81. Berlín: Springer.
- FELLOUS, J-M. 1999. Neuromodulatory basis of emotion. *Neuroscientist* 5: 283-94.
- FLEGR, J. 2013. Influence of latent *Toxoplasma* infection on human personality, physiology and morphology: Pros and cons of the *Toxoplasma*-human model in studying the manipulation hypothesis. *The Journal of Experimental Biology* 216: 127-33.
- 2007. Effects of *Toxoplasma* on human behavior. *Schizophrenia Bulletin* 33: 757-60.
- FLEGR, J., Š. ZITKOVÁ, P. KODYM, & D. FRYNTA. 1996. Induction of changes in human behaviour by the parasitic protozoan *Toxoplasma gondii*. *Parasitology* 113: 49-54.
- FRITH, C.D. 2009. Free will and top-down control in the brain. En *Downward causation and the neurobiology of free will*, eds. N. Murphy, G.F.R. Ellis & T. O'Connor, 199-209. Berlín: Springer.
- FRITH, C.D., K. FRISTON, P.F. LIDDLE, & R.S.J. FRACKOWIAK. 1991. Willed action and the prefrontal cortex in man: A study with PET. *Proceedings of the Royal Society B: Biological Sciences* 244: 241-6.

- FUSTER, J.M. 1995. *Memory in the cerebral cortex: An empirical approach to neural networks in the human and nonhuman primate*. Cambridge, Massachusetts: MIT Press.
- GASQUOINE, P.G. 1993. Alien hand sign. *Journal of Clinical and Experimental Neuropsychology* 15: 653-67.
- GOLDMAN-RAKIC, P.S. 1987. Motor control function of the prefrontal cortex. En *Motor areas of the cerebral cortex: Ciba Foundation Symposium 132*, ed. Ciba Foundation Symposium 132, 187-200. Chichester (Reino Unido): Wiley.
- GOMES, G. 1999. Volition and the readiness potential. En *The volitional brain: Towards a neuroscience of free will*, eds. B. Libet, A. Freeman & K. Sutherland, 59-76. Exeter (Reino Unido): Imprint Academic.
- HYDER, F., E.A. PHELPS, C.J. WIGGINS, K.S. LABAR, A.M. BLAMIRE, & R.G. SHULMAN. 1997. «Willed action»: A functional MRI study of the human prefrontal cortex during a sensorimotor task. *Proceedings of the National Academy of Sciences of the United States of America* 94: 6989-94.
- KELSO, J.A.S., & E. TOGNOLI. 2009. Toward a complementary neuroscience: Metastable coordination dynamics of the brain. En *Downward causation and the neurobiology of free will*, eds. N. Murphy, G.F.R. Ellis & T. O'Connor, 103-24. Berlín: Springer.
- KOENIGS, M., L. YOUNG, R. ADOLPHS, D. TRANEL, F. CUSHMAN, M. HAUSER, & A. DAMASIO. 2007. Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446: 908-11.
- KURZBAN, R., P. DESCIOLO, & E. O'BRIEN. 2007. Audience effects on moralistic punishment. *Evolution & Human Behavior* 28: 75-84.
- LAU, H.C., R.D. ROGERS, & R.E. PASSINGHAM. 2007. Manipulating the experienced onset of intention after action execution. *Journal of Cognitive Neuroscience* 19: 81-90.
- LHERMITTE, F. 1983. 'Utilization behaviour' and its relation to lesions of the frontal lobes. *Brain* 106: 237-55.
- LIBET, B. 1999. Do we have free will? En *The volitional brain: Towards a neuroscience of free will*, eds. B. Libet, A. Freeman & K. Sutherland, 47-57. Exeter (Reino Unido): Imprint Academic.
- 1985. Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences* 8: 529-66.
- LIBET, B., C.A. GLEASON, E.W. WRIGHT, & D.K. PEARL. 1983. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain* 106: 623-42.
- LIBET, B., E.W. WRIGHT, & C.A. GLEASON. 1983. Preparation- or intention-to-act, in relation to pre-event potentials recorded at the vertex. *Electroencephalography and Clinical Neurophysiology* 56: 367-72.
- 1982. Readiness-potentials preceding unrestricted 'spontaneous' vs. pre-planned voluntary acts. *Electroencephalography and Clinical Neurophysiology* 54: 322-35.
- MELE, A.R. 2011. Free will and science. En *The oxford handbook of free will*, ed. R. Kane, 2ª ed., 499-514. Nueva York: Oxford University Press.
- MILLER, E.K., & J.D. COHEN. 2001. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* 24: 167-202.

- MOYA, C.J. 2012. Libertad, responsabilidad moral y neurociencia. *Pasajes: Revista de Pensamiento Contemporáneo* 37: 109-18.
- 2011. Mind, brain, and downward causation. En *Moral behavior and free will: A neurobiological and philosophical approach*, eds. J.J. Sanguinetti, A. Acerbi & J.A. Lombo, 185-200. Morolo (Italia): IF Press.
- MUÑOZ, J.M. 2012. Hacia una sistematización de la relación entre determinismo y libertad. *Daímon* 56: 5-19.
- NAHMIAS, E. 2002. When consciousness matters: A critical review of Daniel Wegner's *The illusion of conscious will*. *Philosophical Psychology* 15: 527-41.
- SCHARDT, D.M., S. ERK, C. NÜSSER, M.M. NÖTHEN, S. CICHON, M. RIETSCHEL, J. TREUTLEIN, T. GOSCHKE, & H. WALTER. 2010. Volition diminishes genetically mediated amygdala hyperreactivity. *NeuroImage* 53: 943-51.
- SCHULTZ, W. 1999. The primate basal ganglia and the voluntary control of behaviour. En *The volitional brain: Towards a neuroscience of free will*, eds. B. Libet, A. Freeman & K. Sutherland, 31-45. Exeter (Reino Unido): Imprint Academic.
- SCHWARTZ, J.M. 1999. A role for volition and attention in the generation of new brain circuitry: Toward a neurobiology of mental force. En *The volitional brain: Towards a neuroscience of free will*, eds. B. Libet, A. Freeman & K. Sutherland, 115-42. Exeter (Reino Unido): Imprint Academic.
- SOON, C.S., M. BRASS, H-J HEINZE, & J-D HAYNES. 2008. Unconscious determinants of free decisions in the human brain. *Nature Neuroscience* 11: 543-5.
- SPENCE, S.A., D.J. BROOKS, S.R. HIRSCH, P.F. LIDDLE, J. MEEHAN, & P.M. GRASBY. 1997. A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain* 120: 1997-2011.
- SPENCE, S.A., & C.D. FRITH. 1999. Towards a functional anatomy of volition. En *The volitional brain: Towards a neuroscience of free will*, eds. B. Libet, A. Freeman & K. Sutherland, 11-29. Exeter (Reino Unido): Imprint Academic.
- SQUIRE, L.R., D. BERG, F.E. BLOOM, S. DU LAC, A. GHOSH, & N.C. SPITZER (eds.). 2008. *Fundamental neuroscience*, 3ª ed. Burlington, Massachusetts: Academic Press.
- WALTER, H. 2011. Contributions of neuroscience to the free will debate: From random movement to intelligible action. En *The oxford handbook of free will*, ed. R. Kane, 2ª ed., 515-29. Nueva York: Oxford University Press.
- 2002. Neurophilosophy of free will. En *The oxford handbook of free will*, ed. R. Kane, 565-76. Nueva York: Oxford University Press.
- 2001. *Neurophilosophy of free will: From libertarian illusions to a concept of natural autonomy*. Cambridge, Massachusetts: MIT Press.
- WEGNER, D.M. 2004. Précis of *The illusion of conscious will*. *Behavioral and Brain Sciences* 27: 649-59.
- 2002. *The illusion of conscious will*. Cambridge, Massachusetts: MIT Press.