

Sobre los presupuestos filosóficos de las pruebas neurocientíficas.
(A propósito de “Proceso y neurociencia. Aspectos generales”, de M. Taruffo).

Daniel González Lagier
Universidad de Alicante

“Como el entomólogo a la caza de mariposas de vistosos matices, mi atención perseguía, en el vergel de la sustancia gris, células de formas delicadas y elegantes, las misteriosas mariposas del alma, cuyo batir de alas quién sabe si esclarecerá algún día el secreto de la vida mental”

(Santiago Ramón y Cajal, *Recuerdos de mi vida*).

1. La naturalización de la responsabilidad.

La filosofía en torno a la responsabilidad, la filosofía práctica, está experimentando un proceso de “naturalización”, que es una consecuencia o manifestación de lo que se ha llamado “la naturalización de la mente” o “naturalización de la intencionalidad” y, en última instancia, de una tendencia a la naturalización de la filosofía¹ en general. No hay un claro acuerdo acerca de qué entender por “naturalización” de la filosofía, pero se podría coincidir en que implica la reconstrucción de la filosofía a partir de conceptos admitidos o, al menos, admisibles por las ciencias empíricas (frecuentemente identificadas, en un sentido estricto de “ciencia”, con la física y la química)². Esta es una caracterización amplia (y mínima) que incluye tanto a las posturas más científicas, que pretenden la reducción de la filosofía (o alguna parte de la misma) a alguna ciencia empírica (como propuso Quine a propósito de la epistemología, a la que consideraba psicología cognitiva³) como las posturas menos radicales que entienden que hay cierto continuo entre filosofía y ciencia, aunque ambas puedan tener dominios distinguibles.

La naturalización de la mente es, por consiguiente, el intento de reconstruir los conceptos mentales -como creencia, decisión, intención, deseos, emociones, dolor, etc.- de manera que puedan ser aceptados por las ciencias de la naturaleza. Es una tarea de

¹ Para un análisis de este fenómeno puede verse T. Grimaldos y J. Pacho (eds.), “La naturalización de la filosofía: problemas y límites”, en Pre-Textos, 2005.

² Carlos Moya, “La naturalización de la responsabilidad moral”, en T. Grimaldos y J. Pacho (eds.), *La naturalización de la filosofía: problemas y límites*, pág. 59.

³ W.V. Quine, “Naturalización de la epistemología”, en *La relatividad ontológica y otros ensayos*, Ed. Tecnos, 2002. Este artículo puede verse como uno de los impulsos más importantes de la naturalización de la mente desde el lado de la filosofía.

inspiración claramente materialista y antimetafísica, pero no es ni mucho menos una tarea fácil. Los estados mentales tienen ciertos rasgos (en lo que sigue me referiré a ellos como “las propiedades de lo mental”) que parecen encajar mal en lo que las ciencias empíricas nos dicen del mundo: En primer lugar, nuestros estados mentales nos son accesibles a nosotros mismos, por introspección, de una manera directa, al margen de la evidencia empírica y de inferencias a partir de ella (es el rasgo de la *conciencia*). En segundo lugar, y relacionado con lo anterior, hay cierta diferencia cualitativa en cómo emergen a mi conciencia los dolores, deseos o preocupaciones, o en la manera en que experimentamos escuchar un concierto de Bach o el sabor de un trozo de chocolate. Puede decirse, entonces, que estos estados mentales se corresponden con diferentes sensaciones internas⁴, que los filósofos -en analogía con el término *quanto* usado en física- llaman *qualia*. En tercer lugar, muchos estados mentales (las creencias, las intenciones,...) poseen un contenido, un significado, versan sobre otros hechos (es el rasgo de la “intencionalidad” o “contenido mental”). En cuarto lugar, los estados mentales se relacionan con nuestra conducta externa, pareciendo tener eficacia causal sobre ella (es el fenómeno de la *causación mental*)⁵. Conciencia, carácter cualitativo, intencionalidad y eficacia causal son cuatro características de la mente que no es evidente que puedan explicarse por referencia exclusiva a procesos físico-químicos y leyes empíricas. Supongamos que la mente se identifica con el cerebro y que los estados mentales son conexiones eléctricas, físicas o químicas entre neuronas: ¿cómo es posible que algo material, físico, nos resulte accesible sin observación externa y posea un contenido semántico?⁶ Y si la mente no forma parte del mundo físico, ¿cómo puede interactuar causalmente con él?

La adscripción de responsabilidad, por su parte, tal como la concebimos, presupone tanto la existencia de ciertos estados mentales como estos rasgos básicos de la mente. Somos responsables de nuestras acciones en la medida en que éstas sean causadas por deseos y creencias respecto del mundo sobre los que tenemos cierto control. De manera que la posibilidad de naturalizar la responsabilidad depende (al menos, en parte) del éxito de una teoría naturalizada de la mente.

Con el desarrollo de la neurociencia, la idea de una teoría naturalizada de la mente cobró un nuevo impulso. Suele situarse el nacimiento de la neurociencia moderna en

⁴ T. Nagel, “Qué se siente ser un murciélago”, Ed. Fondo de Cultura Económica, 2000.

⁵ Una explicación más detallada de estos rasgos puede verse en C. Moya, *Filosofía de la mente*, Capítulo I (“Algunos rasgos de nuestro conceptopreteórico de lo mental”), Universidad de Valencia, 2006.

⁶ Una presentación de una estrategia para la naturalización de la intencionalidad puede verse en Carlos Moya, “Las emociones y la naturalización de la intencionalidad”, en *Anales del Seminario de Metafísica*, núm. 28, Ed. Complutense, 1994.

1888, cuando Santiago Ramón y Cajal descubre que el cerebro es una red de células, las neuronas. Posteriormente, Charles Sherrington analizó las conexiones entre ellas y Edgard Adrian registró actividad bioeléctrica en todo el sistema nervioso. Los avances en psicofarmacología a mediados del siglo XX y, sobre todo, la aparición de las técnicas de neuroimagen en los noventa y principios del siglo XXI, han dado un impulso decisivo a la neurociencia. La gran mayoría de noticias espectaculares que trascienden a la opinión pública tienen que ver con estas técnicas de neuroimagen, que han dado lugar a lo que se ha llamado la neurociencia cognitiva, esto es, el estudio del funcionamiento del cerebro en los procesos de adquisición del conocimiento y de la formación de estados mentales.⁷ Del tronco central de la neurociencia se han ramificado nuevas disciplinas que pueden verse como aplicaciones de la misma a distintos ámbitos. En 2002 se celebró en San Francisco el congreso *Neuroética: esbozando un mapa del terreno* que consagraba esta nueva disciplina, dedicada a la aplicación de la neurociencia a los temas tradicionales de la ética, y en 2007 la MacArthur Foundation creó el proyecto “Derecho y Neurociencia” para reunir a varios conocidos neurocientíficos, filósofos y juristas de diversos países con el fin de profundizar en la intersección entre las neurociencias y el Derecho. De esta manera, la “neuroética” y el “neuroderecho” se convierten en el principal intento de naturalizar la filosofía práctica y la responsabilidad⁸.

Para adscribir responsabilidad jurídico-penal o moral a un agente es necesario (1) que exista un sistema de reglas (con el que juzgamos la conducta del individuo), (2) que el agente haya actuado libremente y (3) que podamos probar que los estados mentales requeridos para la adscripción de la responsabilidad están presentes en ese agente. Estos tres requisitos constituyen tres grandes desafíos para la naturalización de la responsabilidad: ¿Puede una filosofía práctica naturalizada dar cuenta de la normatividad? ¿Qué respuesta da la neurociencia al problema del libre albedrío? ¿Qué aporta la neurociencia a la prueba de los estados mentales necesarios para la atribución de responsabilidad? En este trabajo no me voy a ocupar de la primera de estas cuestiones. Mi objetivo es la tercera cuestión, pero -en la medida en que, como veremos, pueden estar vinculadas- diré algo también de la segunda.

⁷ Una descripción detallada de la evolución del conocimiento del cerebro desde el Antiguo Egipto puede verse en J.P. Changeux, *El hombre neuronal*, Ed. Espasa-Calpe, 1985. Véase también, J. M. Giménez Anaya y J. I. Murillo, “Mente y cerebro en la neurociencia contemporánea. Una aproximación a su estudio interdisciplinar”, en *Scripta Theologica*, 39, 2007, págs. 607 y ss.

⁸ Una reflexión acerca del impacto que la neurociencia puede tener en el Derecho puede verse en M. Narváez, “El impacto de la neurociencia en el Derecho: El caso de la responsabilidad subjetiva”, en *Revista Telemática de Filosofía del Derecho*, núm. 15, 2012. Una muestra de los temas tratados bajo el rótulo de neuroderecho puede verse en S. Zeki y O. Goodenough, *Law & the Brain*, Oxford University Press, 2006.

En realidad, el problema de la prueba no es menor. Las dificultades para probar los estados mentales de un sujeto tienen que ver con su carácter subjetivo: Aunque, como hemos visto, podemos tener información acerca de nuestros propios estados mentales sin necesidad de observación empírica, respecto de los estados mentales de terceros sólo podemos tratar de inferirlos a partir de su conducta externa y ciertos datos contextuales. Existen varios argumentos que debilitan nuestra confianza en esas inferencias: Se ha señalado, por ejemplo, que no existen leyes o generalizaciones fiables que nos permitan correlacionar la conducta de los demás con sus estados mentales (y así, por ejemplo, un estado mental como una intención se puede tratar de llevar a cabo por medio de cursos de acción distintos, o es posible que dos personas reaccionen de manera distinta ante un mismo tipo de emoción); que al atribuir a otro un estado mental determinado puede que sólo estemos proyectando sobre él lo que nosotros sentiríamos, pensaríamos o querríamos en su situación; o que no hay ningún tipo de hecho que permita verificar la corrección de una atribución de un estado mental, más allá de los propios criterios que hemos usado para hacer la atribución. Todo ello explica que muchos juristas -e, incluso, nuestros tribunales- hayan asumido que los estados mentales relevantes para la adscripción de responsabilidad jurídica deben entenderse como hechos normativos cuya prueba es en realidad una atribución o adscripción a partir de criterios fijados normativamente y no la determinación de algún hecho empírico⁹. En este clima, la neurociencia parece poder ofrecer técnicas que permiten superar los problemas de la prueba de los hechos psicológicos, tratándolos como hechos empíricos y no como ficciones normativas y “levantando” su velo de subjetividad: por ejemplo, la resonancia magnética funcional permite detectar los cambios en el flujo sanguíneo en el cerebro en el momento en que el individuo realiza determinadas tareas motoras o está en ciertos estados cognitivos o emocionales, lo que parece permitir correlacionar estados mentales con estados cerebrales, de manera que ante la presencia de un estado cerebral determinado se podría suponer que el sujeto tiene el estado mental correspondiente.

2. La neurociencia y la prueba de los estados mentales.

En su artículo “Proceso y Neurociencia. Aspectos generales”, Michele Taruffo trata de evaluar esta aspiración de la neurociencia de servir de prueba de los estados mentales. Para ello, presenta una ordenada panorámica general de las posibilidades y las

⁹ Sobre los problemas de la prueba de los estados mentales, específicamente la prueba de la intención o el dolo, y la discusión sobre su carácter normativo, puede verse Daniel González Lagier, “Buenas razones, malas intenciones”, en *Doxa*, núm. 26 y “La prueba de la intención y la explicación de la acción”, en *Isegoría*, núm. 35, 2006.

dudas que ofrece la aplicación de la neurociencia a la prueba judicial. Taruffo parte del examen de una serie de supuestos, cada vez más complejos, en los que la neurociencia podría aportar información relevante a efectos de prueba:

- 1) Probar la naturaleza y gravedad de una lesión cerebral.
- 2) Determinar si un sujeto padece dolor (y con qué intensidad) a consecuencia de determinada causa.
- 3) Establecer si una lesión cerebral condiciona el comportamiento físico de un sujeto (o si provoca la incapacidad para realizar cierto comportamiento).
- 4) Descubrir si cierta lesión cerebral incide sobre la voluntad del sujeto (y sobre la conducta correspondiente).
- 5) Probar, a partir de ciertos datos sobre el cerebro de un sujeto, si concurría algún vicio en su voluntad en el momento de celebrar algún contrato.
- 6) Establecer si una lesión cerebral puede reducir o eliminar la voluntariedad de un comportamiento supuestamente delictivo, o incluso determinar si el comportamiento delictivo puede ser consecuencia de alguna anomalía cerebral.
- 7) Averiguar la verdad o falsedad de lo que una persona dice.
- 8) Mostrar, con carácter general, la ausencia de voluntariedad de cualquier comportamiento.

En el análisis de estos casos se plantean dos tipos de problemas: El primero es un problema -compartido con otros tipos de pruebas científicas- de fiabilidad de las técnicas empleadas. Aunque el estatus científico de la neurociencia, con carácter general, no es puesto en cuestión, también es cierto que todavía está en un estadio incipiente, por lo que muchas de sus aplicaciones pueden carecer de fiabilidad suficiente. Y así, por ejemplo, quizá la neurociencia cuenta con mecanismos más fiables para probar que el sujeto no puede realizar cierta actividad debido a una lesión cerebral que para establecer la intensidad del dolor de un sujeto o si éste está mintiendo. El segundo problema es más específico: Se trata de determinar, en palabras de Taruffo, “si, cuándo y en qué condiciones, las neurociencias proporcionan conocimientos e informaciones utilizables como *prueba* en juicio”¹⁰. No se trata, por tanto, para Taruffo de una cuestión exclusivamente de fiabilidad, sino de pertinencia. ¿En qué sentido?

Por un lado, Taruffo sugiere que algunos de los conceptos que aquí estamos usando han de entenderse en un sentido jurídico. En su opinión, “la determinación de la incapacidad de entender y de querer es un problema *jurídico*, no un problema científico:

¹⁰ M. Taruffo, “Proceso y neurociencia”, en M. Taruffo y J. Nieva (Dirs.), *Neurociencia y proceso judicial*, Ed. Marcial Pons, 2013, pág. 16.

es el juez quien debe valorar las informaciones sobre el estado del cerebro del sujeto para determinar si éstas sirven para calificar al sujeto como incapaz de entender y de querer sobre la base de las normas que se refieren a esta situación”¹¹. Aunque me parece que, en parte, el argumento es correcto, no estoy seguro de todo su alcance: En otro lugar he sostenido que los conceptos que el Derecho usa relativos a los distintos tipos de hechos (acción, omisión, causalidad, etc.), incluyendo los conceptos correspondientes a hechos psicológicos, deben construirse teniendo en cuenta que se trata de instrumentos para una finalidad práctica (adscribir responsabilidad de una manera que nos parezca correcta y ajustada a nuestras intuiciones); pero eso no implica que sean conceptos nuevos *completamente*, con un significado totalmente ajeno al del lenguaje popular¹². En caso contrario, si el sentido de los términos que se refieren a hechos psicológicos fuera radicalmente distinto al de la “psicología popular”, muchas normas resultarían incomprensibles y su aplicación podría verse como un sinsentido. Dworkin se enfrenta a un argumento similar en su artículo “¿Deben nuestros jueces ser filósofos? ¿Pueden ser filósofos?”, sosteniendo que “Los hombres de Estado y los jueces que estipularon que nadie debe ser castigado si no es responsable de sus actos, o que las personas deben ser tratadas como iguales ante la ley, quisieron llevar a la práctica jurídica los juicios morales y los principios bien conocidos y, por lo tanto, emplearon los conceptos ordinarios en los cuales se expresan estos juicios y principios. Si, por el contrario, supusiéramos que los legisladores estuvieron exhibiendo conceptos jurídicos especiales y completamente diferentes, para los cuales usaron las mismas palabras para designarlos (“responsable” e “igual”), creeríamos entonces que lo hicieron de manera perversa o infundada”¹³. Nosotros podemos sostener que, en la medida en que los conceptos relativos a hechos psicológicos usados por el Derecho mantienen un sentido usual (y debe ser una medida importante), y en la medida en que ese sentido no sea alterado por los científicos, todavía queda abierta la posibilidad de que la neurociencia sea pertinente para la prueba de los estados mentales.

Pero, por otro lado, en opinión de Taruffo la pertinencia de la neurociencia para la prueba jurídica queda condicionada por la cuestión, mucho más compleja, de si puede

¹¹ M. Taruffo, “Proceso y neurociencia”, pág. 18.

¹² D. González Lagier, *Quaestio Facti. Ensayos sobre prueba, causalidad y acción*, Ed. Fontamara, pág. 83 y ss.

¹³ R. Dworkin, “¿Deben nuestros jueces ser filósofos? ¿Pueden ser filósofos?”, en *Isonomía*, núm. 32, 2010.

reducirse la mente al cerebro, identificando las actividades mentales con las actividades cerebrales¹⁴. Para Taruffo una respuesta afirmativa tiene dos problemas:

En primer lugar, advierte que responder afirmativamente a la pregunta por la posibilidad de reducir la mente al cerebro lleva a aceptar los postulados del determinismo. En palabras suyas: “Se trataría, de hecho, de adoptar una perspectiva teórica estrictamente materialista y determinista, según la cual toda la vida mental, espiritual, volitiva y cognitiva se reduciría al funcionamiento de las células cerebrales y sus conexiones. Siguiendo esta perspectiva, se acabaría descartando que exista la voluntad, que se pueda elegir y tomar decisiones, que se pueda hablar de responsabilidad por los propios comportamientos, que exista conciencia de estos comportamientos y que estos puedan referirse a conocimientos de cualquier naturaleza. En esencia, el hombre se vería reducido a una especie de máquina del cerebro, y toda su vida estaría determinada mecánicamente por el funcionamiento sin control de esta máquina”¹⁵. Las consecuencias son tan distorsionadoras de nuestra visión del mundo que podrían verse como un buen argumento para rechazar la identificación entre mente y cerebro, si realmente se siguieran de esta identificación. De hecho, algunos autores, como Peter Strawson, han tratado de mostrar la imposibilidad de aceptar el determinismo a partir de la constatación de que ciertas actitudes espontáneas y naturales hacia las acciones de los demás, como el resentimiento, el agradecimiento, el perdón, el enfado, etc., que no tenemos cuando aceptamos que el sujeto no podía haber actuado de otro modo, están en la base de nuestras relaciones interpersonales y, en definitiva, de la vida social¹⁶.

El segundo problema es que al identificar mente y cerebro se estaría incurriendo en una falacia. Al identificar cierta actividad cerebral detectada por alguna técnica neurocientífica con una intención, una mentira, un recuerdo o una emoción, se están atribuyendo estos hechos al cerebro, o a una parte del mismo, cuando en realidad son estados sólo atribuibles al sujeto como un todo. Por ejemplo, señala Taruffo que “aun admitiendo que una determinada actividad cerebral sea necesaria para mentir (al igual que para decir la verdad), sin embargo resulta conceptualmente imposible identificar la mentira con esa actividad (...) Las mentiras no suceden en el cerebro”¹⁷; igualmente, “las intenciones no son procesos cerebrales, y del mismo modo la intencionalidad no se

¹⁴ M. Taruffo, “Proceso y neurociencia”, pág. 23.

¹⁵ M. Taruffo, “Proceso y neurociencia”, pág. 23.

¹⁶ P. Strawson, “Libertad y resentimiento”, Ed. Paidós, 1995.

¹⁷ M. Taruffo, “Proceso y neurociencia”, pág. 22.

'coloca' en una zona del cerebro y no se reduce a un estado cerebral"¹⁸. Asimismo, "la conciencia no es algo que 'sucede' en el cerebro, como la digestión se produce en el estómago, sino que implica el contacto con el mundo exterior, en una compleja interacción de cerebro, cuerpo y mundo"¹⁹. Toda estas identificaciones, en definitiva, "incurren en una falacia fundamental, que consiste en atribuir a una parte del sujeto humano, es decir, al cerebro, habilidades y funciones que son propias del sujeto en su totalidad, con toda la complejidad que se manifiesta en la vida"²⁰.

Taruffo toma esta objeción de los trabajos de Michael S.Pardo y Dennis Patterson, quienes, a su vez, como veremos más adelante, aplican al ámbito jurídico una crítica que Peter Hacker y Maxwell Bennet, un filósofo y un neurocientífico, dirigen a los fundamentos filosóficos de la neurociencia (y a destacados filósofos, como John Searle y Daniel Dennet, que siguen de cerca en sus explicaciones de la mente y la conciencia las aportaciones de la neurociencia), considerando que ésta incurre en lo que ellos llaman la "falacia mereológica", al atribuir a una parte (el cerebro) lo que debería ser atribuido al todo (la persona).

De manera que tenemos dos problemas de honda complejidad filosófica -la conexión entre una teoría materialista de la mente y la aceptación del determinismo, por un lado, y la falacia mereológica, por otro – que se erigen como serios obstáculos o limitaciones al uso probatorio de la neurociencia. Lo que Michele Taruffo hace es, por tanto, cuestionar los fundamentos filosóficos de la neurociencia y, con ella, de las pruebas neurofisiológicas. ¿Tiene razón?

3. La búsqueda de la mente en la materia: Un mapa de teorías de la mente.

Para evaluar los argumentos de Taruffo necesitamos tener en cuenta qué opciones se han propuesto para explicar la naturaleza de la mente. Trataré de presentar un mapa de las teorías de la mente más extendidas y conocidas, clasificándolas a partir de sus respuestas a una serie de preguntas encadenadas:

PRIMERA PREGUNTA: ¿Está la realidad constituida por entidades físicas y mentales, entendidas como sustancias separadas, o sólo existe un tipo de entidades? El *dualismo*, como el sostenido por Descartes, sostiene lo primero; el *monismo*, lo segundo. El dualismo ha sido la concepción de la mente dominante en la historia de occidente al

¹⁸ M. Taruffo, "Proceso y neurociencia", pág. 20

¹⁹ M. Taruffo, "Proceso y neurociencia", pág. 21

²⁰ M. Taruffo, "Proceso y neurociencia", pág. 23.

menos desde Descartes y ha “empapado” el lenguaje común y lo que podríamos llamar la “psicología popular” (esto es, la teoría que comúnmente usamos en las explicaciones de la conducta y las atribuciones de estados mentales). De acuerdo con el dualismo, el pensamiento y el mundo físico son sustancias distintas, cada una de ellas con propiedades distintas (para Descartes, por ejemplo, la mente es consciente, se conoce directamente, es libre, indivisible e indestructible, mientras que el cuerpo posee extensión, se conoce indirectamente, es determinado, infinitamente divisible y destructible). Hoy en día las propiedades que los dualistas asignarían a cada sustancia no son las mismas, pero la idea de dos tipos de sustancias con propiedades irreducibles permanece, por lo que la mente no puede ser estudiada por las ciencias de la naturaleza.

SEGUNDA PREGUNTA: Si se asume el monismo, ¿qué tipo de sustancia compone la realidad? El monismo puede ser *idealista* o *materialista*²¹. El monismo idealista se inclina del lado de la mente: por ejemplo, al sostener, como Berkeley, que las cosas físicas son ideas en la mente de Dios, o al analizar la realidad, como el positivismo lógico, como una construcción a partir de datos de los sentidos. El monismo materialista es el intento de decantar la distinción hacia el lado de lo físico, entendiendo que lo físico es lo único que es irreduciblemente real y que, por lo tanto, la mente debe ser explicada a partir de lo material (o eliminada). Así como el dualismo es la concepción que está detrás de la psicología popular, el materialismo es la concepción dominante hoy entre científicos y filósofos de la mente.

TERCERA PREGUNTA: Si se asume el monismo materialista, ¿tiene sentido hablar de estados mentales? Para el dualismo, los *estados* mentales existen genuinamente en el mundo mental, por lo que los *conceptos* mentales forman legítimamente parte de la psicología, que es una ciencia distinta de las ciencias físicas. Pero si se acepta el monismo materialista (dejaré de lado al idealismo, que es una postura prácticamente abandonada hoy día), entonces se plantea la cuestión de si es posible acomodarlos en una concepción fisicalista del mundo. Surge, entonces, la alternativa entre el eliminacionismo, el reductivismo y el no reductivismo. El *materialismo eliminacionista* cree que la tarea de encontrar los correlatos neuronales de los estados mentales está

²¹ Dejo de lado el monismo neutral de autores como Bertrand Russell, para quien el componente último del mundo es una sustancia que no es ni material ni espiritual (éstas serían formas de observar esa sustancia neutral). Puede verse, por ejemplo, B. Russell, *Análisis de la materia*, 1969 o *El conocimiento humano*, 1993. Para una explicación actual (y defensa) del monismo neutral puede verse Rafael A. Alemañ Berenguer, “La relación mente-materia y el monismo neutral”, en *Naturaleza y libertad. Revista de estudios interdisciplinarios*, núm. 1, 2012.

condenada al fracaso, porque no hay estados mentales, tal como los considera la psicología y el sentido común. Presuponer que la neurociencia permitirá establecer una correlación entre estados cerebrales y mentales es dar carta de existencia a estos últimos. Pero la psicología está equivocada, trata de dar explicaciones de la conducta que la investigación científica mostrará que no logran decir nada verdadero sobre la realidad. Como la teoría del flogisto (una supuesta sustancia, en realidad inexistente, que se creía que se liberaba en los procesos de combustión u oxidación), las teorías y los conceptos mentales están condenados a desaparecer, simplemente porque hacen referencia a cosas que no existen. A medida que avance la neurociencia, podremos sustituir el lenguaje de los estados mentales por el lenguaje de los estados cerebrales, de igual manera que el descubrimiento de los virus y las bacterias permitió rechazar las explicaciones de las enfermedades en términos de demonios, maldiciones y espíritus. Éste sería, simplemente, un ámbito más donde la extensión de la ciencia permitiría la expulsión de la superstición. El *materialismo reduccionista* no impugna el sentido de los conceptos mentales (mantiene la psicología como una disciplina epistemológicamente adecuada o conveniente, al menos en cierto grado), pero cree que las teorías mentales ganarán en precisión y poder explicativo si son traducidas o reducidas a un lenguaje en última instancia no mental. Una teoría es reducida a otra, de acuerdo con Paul Churchland, cuando “una teoría nueva y eficaz logra abarcar un conjunto de proposiciones y principios que reflejan perfectamente (o casi perfectamente) las proposiciones y principios de una teoría o marco conceptual anteriores”²². Si el ejemplo preferido de los eliminacionistas es el flogisto, el de los reduccionistas es la explicación del calor en términos de promedio de energía cinética de las moléculas (a diferencia de lo que ocurría con el flogisto, las teorías o leyes que usan el concepto de calor, como la que afirma que el calor dilata los metales, son útiles, pero su reducción a la nueva teoría proporciona una descripción más precisa del fenómeno). El *materialismo no reduccionista* sostiene que la traducción del lenguaje mental al lenguaje físico es imposible sin pérdidas significativas: aunque haya una sola sustancia, lo físico, podemos hablar de estados o propiedades cerebrales y de estados o propiedades mentales. El materialismo no reduccionista puede verse como un monismo respecto de las sustancias y, simultáneamente, un dualismo respecto de las propiedades. Su desafío, por tanto, es explicar cómo de lo físico pueden surgir las propiedades de lo mental.

²²

P. Churchland, *Materia y conciencia*, pág. 51.

CUARTA PREGUNTA: Si se asume que hablar de estados mentales no es un sinsentido, esto es, si se asume un materialismo no eliminacionista (reductivista o no) ¿qué son realmente los estados mentales? Se han propuesto varias alternativas:

1) La primera opción consiste en identificar lo mental con la conducta o las disposiciones de comportamiento. A esta estrategia se le ha llamado *conductismo lógico o conceptual*. Lo que tienen en común las distintas concepciones conductistas (como las de Carnap, Ryle o Wittgenstein) es la idea de que existe algún tipo de relación constitutiva e interna entre la mente y el comportamiento²³. Esta conexión surge por el modo de construcción de los conceptos mentales y el aprendizaje de los términos correspondientes. De acuerdo con el conductismo este proceso de aprendizaje no va de lo “interno” a lo “externo”: no reconocemos un dolor en nosotros mismos y luego aplicamos el concepto a otros (si así fuera, como no es posible saber que lo que sienten los otros es exactamente lo mismo, no habría criterios de corrección en el uso de las expresiones mentales, esto es, no tendrían significado), sino que aprendemos a usar el concepto a partir de las manifestaciones externas del dolor que observamos en los demás y en nosotros mismos. El conductismo es una concepción reductivista: el lenguaje mentalista debe ser traducido a la descripción del comportamiento y a la expresión de disposiciones conductuales. “En su forma más fuerte y más directa -escribe Paul Churchland- el conductismo filosófico postula que toda oración acerca de un estado mental se puede parafrasear, sin pérdida de significado, por una oración larga y compleja acerca de cuál sería la conducta observable que se produciría si una determinada persona se encontrara en esta, o aquella o cualquier otra circunstancia observable”²⁴. Volveremos más adelante sobre el conductismo filosófico, puesto que la objeción de Taruffo contra la neurociencia parece conducirle a esta postura o a una similar.

2) La segunda estrategia consiste en identificar los estados mentales de un sujeto con sus estados cerebrales. Esto es lo que hace el *materialismo de la identidad de tipos*, un materialismo reduccionista que sostiene que cada *tipo* de estado mental es una y la misma cosa que un *tipo* de estado cerebral, o que cada propiedad mental es idéntica a una propiedad cerebral. Esta teoría no niega que existan estados mentales: lo que niega es que sean algo distinto de estados del cerebro. De la misma manera que la luz son ondas electromagnéticas, el agua es H₂O y la temperatura es promedio de energía

²³ C. Moya, *Filosofía de la mente*, pág. 43.

²⁴ P. Churchland, *Materia y conciencia*, pág. 46.

cinética molecular, las creencias son un determinado estado cerebral, los deseos otro, etc. Se trata de “identidades interteóricas”²⁵. Para el materialismo de la identidad esta identificación entre estados mentales y estados cerebrales tiene el carácter de una hipótesis científica, que cuenta con algunos indicios a su favor y que se espera que llegue a ser confirmada en el futuro permitiendo la reducción de los conceptos mentales a la neurociencia. Un problema de esta teoría es que si los estados cerebrales y los estados mentales son la misma cosa, entonces todas las propiedades de los primeros han de ser también propiedades de los segundos, y viceversa. Esto es, los estados mentales deben tener propiedades espaciales (deben estar en algún sitio) y los estados cerebrales deben tener propiedades semánticas²⁶. Las dificultades para entender esto suelen llevar a negar las propiedades de lo mental, o alguna de ellas.

3) Existe una versión distinta del materialismo de la identidad. Ante la desconfianza de que sea posible encontrar los correlatos neuronales de los estados mentales, algunos autores han defendido un tipo de materialismo que postula no una identidad entre tipos de estados mentales y tipos de estados cerebrales, sino sólo entre *casos* o *instancias* de estados mentales y *casos* o *instancias* de estados cerebrales. Son las *teorías de la identidad de instancias*. De esta forma, cada estado mental se corresponde con un evento físico, pero dos sujetos pueden tener el mismo estado mental y estar en estados cerebrales distintos, o tener estados mentales diferentes pese a estar en estados cerebrales idénticos. Esto implica que no puede hacerse una reducción teórica -esto es, general- de los estados mentales a los estados cerebrales (aunque puede hacerse, como sugiere la siguiente estrategia, el funcionalismo, a otro tipo de entidades). Una de estas teorías es el *monismo anómalo* de Donald Davidson²⁷. Se trata de una teoría relevante, porque pretende compaginar el determinismo del mundo físico con el indeterminismo (al menos, epistemológico) de la conducta humana. Davidson parte de la asunción de la tesis de la identidad de instancias y de tres principios básicos: (a) El principio de interacción causal, de acuerdo con el cual al menos algunos eventos mentales interactúan causalmente con eventos físicos, (b) el principio del carácter nomológico de la causalidad, de acuerdo con el cual donde hay causalidad, debe haber una ley que correlacione tipos de eventos y (c) el principio del carácter anómalo de lo mental, según el cual no hay leyes estrictamente causales a partir de las cuales puedan predecirse los eventos mentales, dado que las interpretaciones de los estados mentales de otros sujetos -a diferencia de lo

²⁵ P. Churchland, *Materia y conciencia*, pág. 52.

²⁶ P. Churchland, *Materia y conciencia*, pág. 55.

²⁷ D. Davidson, *Sucesos mentales*, en *Ensayos sobre acciones y sucesos*, Crítica, 1995.

que ocurre con las relaciones causales entre eventos físicos- están sujetas al principio rector de la racionalidad.

4) La estrategia que tiene mayor difusión actualmente es el *funcionalismo*, que en realidad puede verse como una concepción que funde cierta herencia del conductismo con la tesis de la identidad de instancias. Para los funcionalistas los estados mentales se identifican con roles, papeles o funciones causales. De la misma manera que un reloj se caracteriza por su función para medir el tiempo con exactitud, un estado mental se caracteriza por la función que cumple. Dicha función se describe a partir de ciertas relaciones con el contexto (*inputs*) y con la conducta (*outputs*). Para el funcionalismo “lo esencial para que un determinado sujeto tenga la creencia de que *p* no es que en su cerebro haya una configuración neuronal de un tipo determinado, sino que se halle en un estado que guarde las relaciones adecuadas con otros estados y procesos y cumpla el papel causal adecuado en la organización de dicho sujeto”²⁸. El funcionalismo es compatible con el materialismo, porque tener una propiedad funcional requiere una estructura subyacente física, pero difiere del materialismo de la identidad de tipos porque admite que las mismas funciones podrían ser realizadas por estructuras físicas organizadas de manera distinta. Así, un ser (o una máquina) que tuviera una organización funcional suficientemente semejante a la de un hombre podría tener estados mentales. Por tanto, aunque todo estado mental se corresponde con un estado físico, no hay una correspondencia entre tipos de estados mentales y tipos de estados físicos. El funcionalismo puede verse, además, como una concepción reductivista, en el sentido de que asume la posibilidad de una descripción de las propiedades mentales que no haga referencia a otras propiedades mentales, pero a diferencia del conductismo y del materialismo de la identidad no reduce los estados mentales a estados cerebrales, sino a propiedades funcionales²⁹.

5) La última alternativa que analizaremos es la del materialismo emergentista. El *emergentismo* es una postura materialista que identifica los estados mentales con propiedades emergentes o sistémicas, es decir, propiedades que surgen a partir de cierto grado de complejidad de un sistema. Son propiedades del conjunto del sistema, pero no de cada una de sus partes. Para explicar qué es una propiedad emergente John Searle - cuya concepción, que él mismo llama *naturalismo biológico*, es una teoría emergentista- recurre al siguiente ejemplo: “Supongamos que tenemos un sistema S, compuesto de los

²⁸ C. Moya, *Filosofía de la mente*, pág. 101.

²⁹ C. Moya, *Filosofía de la mente*, pág. 104. En contra, P. Churchland, *Materia y conciencia*, pág. 65.

elementos *a*, *b*, *c*... Por ejemplo, *S* podría ser una piedra y los elementos podrían ser moléculas. En general, habrá rasgos de *S* que no son, o no necesariamente, rasgos de *a*, *b*, *c*... Por ejemplo, *S* podría pesar 20 kilogramos, sin que las moléculas individualmente pesen 20 kilogramos. Denominemos a estos rasgos 'rasgos del sistema'. Algunos rasgos del sistema pueden ser deducidos, o determinados, o calculados a partir de los rasgos *a*, *b*, *c*..., simplemente por la forma en que se componen y ordenan (y, a veces, por sus relaciones con el entorno). Ejemplos de ellos serían la forma, el peso y la velocidad. Pero algunos otros rasgos del sistema no pueden ser determinados sólo a partir de los elementos que los componen y de las relaciones con el entorno: han de ser explicados a partir de las relaciones entre los elementos. Llamémosles 'rasgos del sistema causalmente emergentes'. La solidez, la liquidez y la transparencia son ejemplos de rasgos del sistema causalmente emergentes³⁰. De manera que, para Searle, una propiedad emergente es una propiedad que surge a partir de los elementos que componen el sistema, de las relaciones del sistema con el entorno y -esto parece lo decisivo- de las relaciones de esos elementos entre sí. Los estados mentales se identifican con propiedades emergentes de los cerebros. Los cerebros son sistemas físicos que han ido adquiriendo complejidad a lo largo de su historia evolutiva. El naturalismo biológico es descrito por Searle a partir de cuatro tesis³¹:

- 1) Los estados mentales, con su ontología subjetiva de primera persona, son fenómenos reales del mundo real.
- 2) Los estados mentales son causados en su totalidad por procesos neurobiológicos de nivel inferior localizados en el cerebro.
- 3) Los estados mentales son rasgos del sistema cerebral en su conjunto y existen, por tanto, en un nivel superior al de las neuronas y sus sinapsis ("Por sí misma -dice Searle-, una neurona no es consciente, pero las partes del sistema cerebral compuestas por ella sí lo son").
- 4) Los estados mentales, en la medida en que son reales, tienen eficacia causal ("Mi sed consciente, por ejemplo, me lleva a tomar agua").

También ha sostenido una postura emergentista Mario Bunge (presentándola como una prolongación del programa de Santiago Ramón y Cajal de estudiar el espíritu en la materia), para quien (1) todos los estados, sucesos y procesos mentales son estados, sucesos y procesos en los cerebros de vertebrados superiores, (2) estos estados, sucesos y procesos emergen de los estados, sucesos y procesos de los componentes

³⁰ J. Searle, *El redescubrimiento de la mente*, Ed. Crítica, 1996, pág. 121.

³¹ J. Searle, *La mente. Una breve introducción*, Ed. Norma, 2006, págs. 147 y 148.

celulares del cerebro y (3) las relaciones entre la mente y el cerebro son relaciones entre subsistemas diferentes del cerebro o entre algunos de ellos y otros componentes del organismo³².

El materialismo emergentista no es reductivista. De acuerdo con Searle, aunque los estados mentales pueden ser reducidos causalmente a -en el sentido de explicados por- los componentes del cerebro y sus relaciones, esta explicación no es suficiente, porque lo característico de lo mental es que es un fenómeno en primera persona, requiere un lenguaje subjetivo que no puede ser descrito con el lenguaje objetivo de la neurociencia. En el mismo sentido, para Bunge la neurofisiología es necesaria, pero no suficiente para dar cuenta de los estados mentales, dado que olvida conceptos como pensamiento, intención, etc.; y lo mismo puede decirse de la psicología, puesto que no presta atención al cerebro o Sistema Nervioso Central.

Teorías de la mente

1. Dualismo (dos sustancias)
2. Monismo (una sustancia).
 - 2.1. Monismo idealista (todo es espíritu).
 - 2.2. Monismo materialista (todo es materia).
 - 2.2.1. Materialismo eliminacionista (no hay estados mentales, es una descripción falsa).
 - 2.2.2. Materialismo reductivista (los estados mentales son estados de un tipo distinto).
 - 2.2.2.1. Tesis de la identidad de tipos (reducción a eventos neurofisiológicos)
 - 2.2.2.2. Conductismo (reducción a conducta).
 - 2.2.2.3. Funcionalismo (reducción a roles causales).
 - 2.2.3. Materialismo no reductivista.
 - 2.2.3.1. Tesis de la identidad de instancias (los estados mentales son estados físicos, pero no hay correlación entre tipos).
 - 2.2.3.2. Emergentismo (los estados mentales son propiedades emergentes. Monismo de sustancias, dualismo de propiedades).

No puedo señalar aquí los problemas y las ventajas de cada una de estos enfoques distintos sobre la mente. Pero, en mi opinión, sólo algún tipo de emergentismo causal puede satisfacer tres presupuestos que me parecen básicos y que espero que se acepten, a efectos de este trabajo, sin necesidad de mayor discusión:

a) El primer presupuesto sería el "materialismo ontológico", esto es, la idea de que el mundo en el que vivimos es en última instancia el mundo físico o material. Es posible que tenga sentido hablar de un mundo de instituciones y de un mundo de ideas y conceptos (y es probable que en el análisis de estos mundos no haya que aceptar un "materialismo

³²

M. Bunge, *El problema mente-cerebro. Un enfoque psicobiológico*, Ed. Tecnos, 2002, pág. 42.

metodológico", esto es, la idea de que todo debe ser explicado de acuerdo con el modelo de las ciencias de la naturaleza), pero estos presuponen el mundo material (donde "presuponer" no significa que sean sin más "reducibles" a él). Este presupuesto nos compromete con el monismo materialista³³.

b) El segundo presupuesto sería el de la aceptación de la existencia de los estados mentales. Eso es algo de lo que todos tenemos experiencia directa. Es un dato (obtenido por introspección) innegable. Como ha insistido una y otra vez John Searle, el materialismo eliminativista "parece una locura. Parece una locura afirmar que yo nunca he sentido sed o deseos, que nunca he tenido dolor, que, en realidad, nunca he creído nada, o que mis creencias y mis deseos no juegan ningún papel en mi conducta"³⁴. Este presupuesto nos obliga a rechazar el materialismo eliminacionista.

c) El tercer presupuesto es la aceptación de que la mente (los estados mentales) tienen cuatro propiedades básicas: Son conscientes, poseen (al menos, algunos estados mentales) *qualias*, pueden tener contenido semántico y pueden causar la conducta. Este presupuesto nos obliga a rechazar el reductivismo, pues las distintas formas de reductivismo plantean el mismo problema: no pueden dar cuenta del aspecto subjetivo o de primera persona de los estados mentales, al sustituirlo por un discurso en tercera persona. Las concepciones reductivistas, en palabras de Thomas Nagel, "omiten algo esencial que descansa más allá de los fundamentos externamente observables para atribuir estados mentales a otros, a saber, el aspecto del fenómeno mental que es evidente desde la primera persona, el punto de vista interior del sujeto consciente: por ejemplo, la forma en que el azúcar te sabe o la forma en que ves el rojo o sientes la ira"³⁵. Este rechazo del reductivismo nos lleva a buscar una teoría que mantenga las propiedades de lo mental.

El materialismo emergentista trata de conciliar estos tres presupuestos, por lo que creo que debe ser asumido como un programa de investigación, aunque ciertamente todavía no como una teoría plenamente satisfactoria. En todo caso, el repaso de estas

³³ Las propias limitaciones del dualismo de sustancias nos conducen al materialismo, pues, en palabras de Searle: "1) Nadie ha logrado proporcionar jamás una descripción inteligible de las relaciones entre esos dos reinos. 2) la postulación es innecesaria. Es posible explicar todos los hechos de primera y todos los hechos de tercera persona sin postular reinos separados. 3) La postulación genera dificultades intolerables. De acuerdo con esta concepción, se hace imposible explicar de qué manera los estados y sucesos mentales pueden causar estados y sucesos físicos. En síntesis, es imposible evitar el epifenomenalismo" (J. Searle, *La mente. Una breve introducción*, pág. 168. También puede encontrarse una clara crítica al dualismo en P. Churchland, *Materia y conciencia*, págs. 40 y ss.

³⁴ J. Searle, *El redescubrimiento de la mente*, pág. 62.

³⁵ T. Nagel, *Mente y cosmos. Por qué la concepción neo-darwinista materialista de la naturaleza es, casi con certeza, falsa*, Capítulo II (*La conciencia*), Ed. Biblioteca Nueva, 2004.

concepciones nos ha llevado a la conclusión de que hay varios tipos de materialismo y que éste no tiene por qué ser eliminacionista o reductivista.

4. Teorías de la mente y determinismo.

Volvamos a nuestro problema original: la pertinencia de la neurociencia para la prueba de los estados mentales. De acuerdo con el primer argumento de Taruffo, si se acepta la identificación de las actividades mentales con las actividades cerebrales, "se acabaría descartando que exista la voluntad (...). En esencia, el hombre se vería reducido a una especie de máquina del cerebro, y toda su vida estaría determinada mecánicamente por el funcionamiento sin control de esta máquina". ¿Es esto necesariamente así? Es cierto que el determinismo se encuentra más cómodo entre las posturas materialistas y la tesis del libre albedrío parece encontrar mejor acomodo en el dualismo, pero antes de afirmar que una cosa lleva a la otra es necesario tener en cuenta tres matizaciones: La primera es la discusión entre compatibilismo e incompatibilismo; la segunda es la distinción entre determinismo ontológico y determinismo epistemológico, y la tercera es el carácter gradual de la identificación entre mente y cerebro que podemos apreciar en las distintas teorías materialistas de la mente.

a) La compatibilidad o no entre el determinismo y el libre albedrío (y la responsabilidad moral) constituye ya una venerable cuestión filosófica. Taruffo asume una postura *incompatibilista*: si el determinismo fuera cierto, no podrían sostenerse ni el libre albedrío ni la responsabilidad. Para los *compatibilistas*, en cambio, sería posible sostener que nuestras acciones están determinadas sin que, sin embargo, se siguieran las consecuencias para la responsabilidad o para la disolución de la vida social que señala Taruffo. Si el compatibilismo tuviera razón, la aceptación de una concepción materialista de la mente no llevaría necesariamente al determinismo.

Me parece, sin embargo, que en este punto no le falta razón a Taruffo. En mi opinión, siempre hay algo insatisfactorio en las propuestas compatibilistas, porque su objetivo lo alcanzan, en el mejor de los casos, redefiniendo explícitamente lo que se entiende por determinismo o por libre albedrío, y otras veces simplemente deslizándose inadvertidamente de un sentido a otro de estos términos. Hobbes, por ejemplo, un compatibilista clásico, sostenía que la libertad consiste en la ausencia de obstáculos para actuar conforme a nuestros deseos. Esto es compatible con que nuestros deseos estén determinados. Pero si aceptamos que un sujeto es realmente libre cuando hubiera podido actuar de una manera distinta a como lo hizo, entonces el libre albedrío requiere no sólo

que la acción sea libre (*libertad de acción*), en el sentido de que no haya obstáculos que impidan al agente actuar de acuerdo con sus planes y objetivos, sino también que la voluntad sea libre (*libertad de voluntad*), en el sentido de la elección por esos planes y objetivos no esté determinada por condiciones suficientes previas. Muchos compatibilistas suelen identificar la libertad con algo (la libertad de acción sin libertad de la voluntad, libertad como noúmeno pero no fenomenológica, deseos de segundo orden, etc.) que no satisface plenamente nuestra percepción de ser libres, por lo que la conciliación entre libertad y determinismo se logra en la mayor parte de concepciones compatibilistas asumiendo un sentido fuerte de determinismo, pero débil de libertad. Creo que tiene razón Searle cuando afirma que "el compatibilismo hace un planteamiento lógico sobre los conceptos de 'libre' y 'determinado' y señala, con acierto, que hay un uso de estos según el cual decir que una acción es libre no significa, hasta aquí, formular interrogante alguno acerca de si está determinada o no, en el sentido de la existencia de condiciones causales previas suficientes. Pero una vez aceptado este planteamiento lógico, todavía queda abierta una cuestión empírica fáctica. *¿Es verdad que toda acción humana ocurrida en el pasado, que ocurre ahora o que ocurrirá alguna vez fue, es y será causada por condiciones previamente suficientes?*"³⁶.

b) En segundo lugar, debe tenerse en cuenta la distinción entre el determinismo como tesis ontológica (todo está causalmente determinado) y el determinismo como tesis epistemológica (conociendo el estado del mundo en un momento determinado y las leyes causales adecuadas, todo suceso es predecible). El primero requiere que las causas sean siempre condiciones suficientes de los efectos. Pero precisamente esto es algo que se niega en ocasiones. La mecánica cuántica, como es sabido, parece introducir un elemento de azar e indeterminación en el nivel sub-atómico del mundo físico, de manera que sus leyes sólo pueden ser estadísticas. Esta indeterminación no afecta en condiciones normales a niveles superiores. ¿Es posible que en el cerebro se filtre algún grado de indeterminación por influencia de la mecánica cuántica? No han faltado autores que lo han sugerido así (Eddington, R. Penrose, J. Eccles, Robert Kane, etc.). Esta estrategia tiene dos enormes problemas (por no decir, como señala Searle, que trata de resolver un misterio con otro): ¿cómo es posible que la indeterminación de la mecánica cuántica afecte al nivel neuronal (superior al sub-atómico), a diferencia de lo que ocurre en el resto de contextos físicos? Y, más grave aún: ¿No es el libre albedrío también

³⁶ J. Searle, *La mente. Una breve introducción*, pág. 278.

incompatible con el ingrediente azaroso que la indeterminación introduciría en nuestra conducta?³⁷ Tan desastroso es para el libre albedrío el determinismo rígido como la ausencia de toda determinación: El libre albedrío requiere del tipo adecuado de determinación. Esto ha llevado a otros autores a tratar de encontrar alguna vía intermedia entre el azar y la determinación necesaria. Kathinka Evers, en su libro *Neuroética. Cuando la materia se despierta*, distingue entre una causalidad necesaria, que produciría el mismo efecto en todos los mundos posibles, y una causalidad contingente, para la que habría al menos un mundo posible en el que esa misma causa ha producido un efecto distinto³⁸. Desde el lado de la neurociencia se han presentado algunas investigaciones que sugieren que en el cerebro (al menos, para algunas de sus funciones) las relaciones causales son de esta segunda manera (Evers se centra sobre todo en los trabajos del neurocientífico J.P. Changeux tratando de explicar la creatividad a partir de la plasticidad neuronal, esto es, la capacidad del cerebro de modificarse físicamente en función del aprendizaje, incluso del proveniente del medio social). Si esto es así, "es teóricamente posible *influir* deliberadamente en el resultado para orientarlo en cierta dirección"³⁹. Creo que, respecto del determinismo ontológico, hay suficiente grado de discusión entre neurocientíficos como para poder dejar todavía la cuestión abierta, sin identificar necesariamente reducción de la mente al cerebro con aceptación del determinismo.

El determinismo epistemológico, por su parte, añade al requisito de la existencia de condiciones causalmente suficientes previas de la conducta la idea de una correlación entre tipos de estados cerebrales y tipos de estados mentales. Si esta correlación no existe o no puede establecerse, entonces la conducta humana, aunque pueda estar causalmente determinada, es impredecible. "Leyendo" los estados cerebrales no podríamos saber qué estados mentales tendría el sujeto y, por tanto, cómo va a actuar. El materialismo de la identidad de instancias, por tanto, no puede sustentar un determinismo neurobiológico epistemológico.

c) En tercer lugar, como hemos visto al exponer las distintas teorías materialistas, cabe una graduación en el grado de identificación entre mente y cerebro, que tiene consecuencias para las implicaciones deterministas. De mayor a menor identificación:

³⁷ Para una clara exposición de algunos filósofos que han pretendido fundamentar el libre albedrío en la mecánica cuántica y los problemas que esto plantea, véase C. G. Patarroyo, "Indeterminación cuántica, libertad y responsabilidad", en *Ideas y valores*, núm. 136, 2008, Bogotá. Para una exposición de los intentos del lado de algunos neurocientíficos, véase J. Pastor-Gómez, "Mecánica cuántica y cerebro. Una revisión crítica", en *Revista de Neurología*, núm. 35, 2002.

³⁸ K. Evers, *Neuroética. Cuando la materia se despierta*, Ed. Katz, 2010, pág. 86 y ss.

³⁹ K. Evers, *Neuroética. Cuando la materia se despierta*, pág. 96.

1) El materialismo eliminacionista y el materialismo reduccionista proponen una identificación fuerte entre cerebro y mente: las explicaciones de la conducta basadas en creencias, deseos, emociones, etc. podrán ser eliminadas o al menos traducidas a explicaciones neurofisiológicas una vez que se hayan encontrado los correlatos neuronales de los estados mentales. Sin embargo, no llevarían necesariamente a sostener que todas las conductas están determinadas si se encuentra alguna explicación neurofisiológica para ello (como las que destacan la plasticidad del cerebro y la flexibilidad de algunas relaciones causales en su funcionamiento).

2) Un grado menor de identificación se da en el materialismo de la identidad de instancias o casos: es una identificación en el nivel de los casos, pero sin una correlación de tipos. De esta manera, como hemos visto, se abre la puerta a aceptar una noción de libertad como imposibilidad de predecir la conducta, esto es, como indeterminismo epistémico.

3) Menor identificación aún se da en las posturas que sostienen un monismo de sustancias pero un dualismo de propiedades. Es el caso del materialismo emergentista. Si la relación de emergencia es algún tipo de relación causal entre el sustrato físico y lo mental, el libre albedrío podría explicarse desde posturas emergentistas si pudiera haber influencia causal de las propiedades sistémicas sobre las propiedades de nivel inferior. Mario Bunge, por ejemplo, trata de ofrecer una explicación del libre albedrío, entendido como la capacidad del pensamiento de influir en la conducta, desde los postulados del materialismo emergentista.

4) El funcionalismo, por último, establece una conexión muy débil entre el cerebro y las propiedades mentales (que, recordémoslo, se identifican con propiedades funcionales). Puede ser compatible con algún tipo de determinismo, pero en la medida en que define la mente al margen de su sustrato neuronal (hasta el punto de que las mismas funciones mentales pueden realizarse en soportes físicos distintos), se aleja de un determinismo neurobiológico.

En definitiva, si se tiene en cuenta los distintos sentidos de determinismo y los distintos sentidos de materialismo, no parece que haya necesariamente una contradicción entre ser materialista y aceptar el libre albedrío. Es decir, siempre hay un argumento disponible para cada tipo de teoría que permite rechazar el determinismo.

En todo caso, como he sostenido en otro lugar, hay argumentos conceptuales que permiten sostener que no cabe probar empíricamente la tesis determinista. Precisamente en esta idea cabe sustentar la responsabilidad: no en el hecho de que el determinismo sea una hipótesis que no ha sido todavía probada, sino en el hecho de que, por razones

conceptuales, no puede serlo. No puede serlo porque las categorías conceptuales implicadas, como la de relación causal, están construidas a partir del concepto de acción, que a su vez presupone fragmentos del mundo no determinados⁴⁰. Probablemente, éste y otros argumentos conceptuales semejantes descansan en la imposibilidad de construir un lenguaje lo suficientemente objetivizante como para demostrar que carecemos de algo de lo que tenemos una convicción subjetiva tan intensa. Como ha escrito Adela Cortina: “Si quiero comprobar experimentalmente si tenía posibilidad de actuar de otro modo o si tengo posibilidad de decidir, entonces he de asumir una perspectiva impersonal u objetiva para poder diseñar experimentos que me permitan comprobar la verdad o falsedad de las afirmaciones. En este caso puede muy bien ocurrir que no podamos traducir el lenguaje de primera persona al de tercera persona”⁴¹

5. ¿Es realmente falaz la falacia mereológica? De nuevo sobre conductismo lógico y funcionalismo.

La objeción hacia las pruebas neurofisiológicas a la que Taruffo presta más atención es la de que sus defensores incurren en una falacia grave al atribuir al cerebro, o a alguna de sus partes, estados mentales que sólo son atribuibles al sujeto en su conjunto. De esta manera, Taruffo dirige contra las pruebas neurofisiológicas la objeción que Pardo y Patterson dirigen contra la neuroética y el neuroderecho, quienes, a su vez, la toman de la objeción que Hacker y Bennet dirigen a la neurociencia -al menos a la neurociencia cognitiva-.

Los argumentos centrales en Pardo y Patterson son, por un lado, la distinción entre criterios e indicios y, conectada con esta distinción, la identificación de los estados mentales con "habilidades" o con la conducta del sujeto. Esta definición de estado mental les lleva a negar que la mente (y los estados mentales) pueda ser localizada en algún lugar del cuerpo.

El punto de partida de su argumentación es la distinción wittgensteniana entre pruebas que son criterios y pruebas empíricas o inductivas para la adscripción de predicados psicológicos, como creer, percibir, conocer, etc. Las primeras tienen una relación constitutiva o conceptual con dicho predicado psicológico, mientras que las segundas sólo prestan apoyo inductivo, en la medida en que pueden estar empíricamente

⁴⁰ D. González Lagier, *¿La tercera humillación? Sobre neurociencia, filosofía y libre albedrío*, en M. Taruffo y J. Nieva, *Neurociencia y proceso judicial*.

⁴¹ A. Cortina, *Neuroética y neuropolítica*, Ed. Tecnos, 2012, pág. 157.

correlacionadas con el estado mental⁴². De acuerdo con los autores, las pruebas "criteriales" de los atributos psicológicos son "varios tipos de comportamiento": "Comportarse de un modo determinado es lógicamente una buena prueba y, por eso, parte constitutiva de estos conceptos. Para la percepción visual esto incluye, por ejemplo, que uno de los ojos siga el fenómeno que se persigue, que uno reporte ajustadamente lo que se observa, y así respectivamente. El creer incluye, por ejemplo, que uno asevere o refrende lo que uno cree, que uno actúe en direcciones consistentes con lo que se cree, que uno no crea proposiciones directamente contradictorias y así sucesivamente"⁴³. La conexión entre el comportamiento y los estados psicológicos, según los autores, es conceptual, de manera que "el comportamiento no sólo aporta pruebas de que alguien en una ocasión dada está percibiendo o tiene una creencia, sino que también determina parcialmente lo que significa percibir o creer"⁴⁴. Por el contrario, la actividad neuronal, aunque pudiera establecerse una correlación entre ella y los estados mentales, sólo sería prueba inductiva o empírica de los mismos. Lo neuronal no podría ser criterio de los estados mentales porque para saber que cierto tipo de evento neuronal se corresponde con un determinado tipo de estado mental, previamente hemos de haber identificado el estado mental⁴⁵ (a través del comportamiento).

La conexión interna o constitutiva entre la conducta y los estados mentales les lleva a definir la mente como "un conjunto diverso de habilidades ejercidas por una persona"⁴⁶. "La mente -afirman- no es en absoluto una entidad o sustancia (ni física ni no física). Tener una mente implica poseer una serie de ciertos poderes racionales que se manifiestan con el pensamiento, con los sentimientos y las acciones"⁴⁷. Pardo y Patterson contraponen esta concepción al materialismo (por el que entienden la identificación de mente y cerebro) y al dualismo, presentándola como una tercera vía.

El último paso en su argumentación es el siguiente: si la mente no es el cerebro, ni ningún otra parte del sujeto, no tiene sentido atribuir a un órgano los estados mentales. Hacerlo es incurrir en la falacia que -siguiendo a Hacker y Bennet- llaman falacia mereológica.

Esta argumentación sigue punto por punto la de Peter Hacker y Maxwell Bennet en sus *Fundamentos filosóficos de la neurociencia* y en *La naturaleza de la conciencia* (que

⁴² M. S. Pardo y D. Patterson, "Fundamentos filosóficos del Derecho y la Neurociencia", en *In Dret*, 2/2011 pág. 13.

⁴³ M. S. Pardo y D. Patterson, "Fundamentos...", pág. 14.

⁴⁴ M. S. Pardo y D. Patterson, "Fundamentos...", pág. 14.

⁴⁵ M. S. Pardo y D. Patterson, "Fundamentos...", pág. 15.

⁴⁶ M. S. Pardo y D. Patterson, "Fundamentos...", pág. 7.

⁴⁷ M. S. Pardo y D. Patterson, "Fundamentos...", pág. 43.

reproduce parcialmente el primer libro): De acuerdo con estos autores, los neurocientíficos han heredado en cierta medida los errores del dualismo cartesiano: han pasado de adscribir atributos psicológicos a la mente a adscribir los mismos atributos al cerebro, pero “suponer que el cerebro es el portador de los atributos psicológicos no supone avance alguno sobre esa concepción falsa [el dualismo]”⁴⁸. Esta confusión les lleva a decir que el cerebro *sabe cosas, razona, construye hipótesis, decide*, o que las neuronas son *inteligentes, transmiten mensajes, calculan probabilidades, etc.*⁴⁹. Se trata en todos los casos, opinan, de un grave error conceptual que lleva a sinsentidos, y no una simple forma metafórica de hablar. Y es un grave error porque la atribución de estados mentales no es una cuestión empírica, sino conceptual. Para poder adscribir estos atributos al cerebro necesitaríamos algún criterio conceptual. En el caso del ser humano, el criterio de atribución es su comportamiento; pero en el caso del cerebro “ni siquiera sabemos qué es lo que demostraría que (...) posee tales atributos”⁵⁰. El comportamiento está, por tanto, constitutiva y conceptualmente vinculado con los estados mentales: es parte de lo que significa tener esos estados mentales. Por tanto, los estados mentales sólo son atribuibles al agente que actúa como un todo.

¿Cómo debemos valorar estos argumentos? Me parece que John Searle, en su contestación a las críticas de Hacker y Bennet, muestra algunas debilidades de estas objeciones a la neurociencia. Como es evidente, la concepción de la mente de estos autores se basa en el conductismo lógico de Wittgenstein⁵¹. Uno de los rasgos distintivos del conductismo lógico es la existencia de esta relación lógica o conceptual entre la conducta (o las disposiciones conductuales, o los poderes causales) y la mente que señalan tanto Pardo y Patterson como Hacker y Bennet. El argumento para establecer esta conexión, como hemos visto, parte de las consideraciones de Wittgenstein sobre la imposibilidad de un lenguaje privado: Si no hay criterios públicos de corrección del uso de

⁴⁸ P. Hacker y M. Bennet, “Fundamentos filosóficos de la neurociencia”, en M. Bennet, P. Hacker, D. Dennet y J.Searle, *La naturaleza de la conciencia. Cerebro, mente y lenguaje*, Ed. Paidós, 2008, pág. 19.

⁴⁹ P. Hacker y M. Bennet, “Fundamentos filosóficos de la neurociencia”, pág. 31.

⁵⁰ P. Hacker y M. Bennet, “Fundamentos filosóficos de la neurociencia”, pág. 35.

⁵¹ El conductismo lógico no niega necesariamente la existencia de un elemento subjetivo en los estados mentales (al menos, al parecer, Wittgenstein no lo hacía), pero en la medida en que no lo introduce en la definición de los mismos se le presentan algunas paradojas. Así, por ejemplo, tendría que aceptar que un sujeto superespartano que nunca expresara su dolor, ni siquiera en situaciones extremas –por ejemplo, al extraerle una muela sin anestesia– no tendría realmente dolor, y que un actor tan magnífico que pudiera fingir expresiones externas de dolor que engañaran a todos tendría, sin embargo, dolor. Otra objeción importante es que analizar determinados estados mentales nos lleva necesariamente a otros estados mentales, lo que convierte a las definiciones de los conductistas en circulares. Así, por ejemplo, mi creencia de que llueve se manifestará en que yo me llevo un paraguas sólo si tengo también el deseo de no mojarme. Y mi deseo de no mojarme se manifestará en esta conducta de coger el paraguas sólo si tengo la creencia de que el paraguas me mantendrá seco. Véase W. Betchel, *Filosofía de la mente: Una panorámica para la ciencia cognitiva*, Ed. Tecnos, 1991.

un lenguaje, éste carece de sentido. Atribuir a un sujeto un estado mental -dolor, por ejemplo- es algo que hacemos porque le vemos actuar de determinada manera. Luego determinada conducta, o cierta disposición a ella, es parte de lo que significa tener un dolor. Frente a este argumento, Searle señala que Hacker y Bennet confunden los criterios de atribución de estados mentales con los estados mentales atribuidos: “La falacia -escribe Searle-, en pocas palabras, reside en confundir las reglas de uso de las palabras con la ontología. Del mismo modo que el antiguo conductismo confundía las pruebas de estados mentales con la ontología de los estados mentales, este conductismo lógico wittgensteniano confunde las bases para hacer la atribución con el hecho que se atribuye. Es una falacia decir que las condiciones para el correcto funcionamiento del juego del lenguaje son condiciones para la existencia del fenómeno en cuestión”⁵². Thomas Nagel hace una objeción similar al conductismo lógico cuando señala que “todas estas estrategias son verificacionistas. Suponen que todo lo que se necesita decir sobre el contenido de un estado mental es lo que verificaría o confirmaría o justificaría su aseveración desde el punto de vista de un observador. De una forma u otra reducen los atributos mentales a las condiciones de lo externamente observable con el criterio de que atribuimos estados mentales a los otros”⁵³ (Esta misma crítica, por cierto, puede hacerse a los juristas que han sostenido que, para el Derecho, los fenómenos psicológicos no son genuinos hechos empíricos que haya que probar, sino hechos normativos, dado que el Derecho necesita fijar criterios normativos –reglas de presunción- para su atribución⁵⁴).

Con esta reducción de la mente a la conducta, a habilidades o a roles causales, el conductismo (y su heredero sofisticado, el funcionalismo) deja fuera la perspectiva subjetiva o de primera persona, que en opinión de autores como Searle o Nagel es precisamente lo característico de lo mental. El conductismo asume respecto de los estados mentales un punto de vista externo, desplazando el punto de vista interno característico de los estados mentales. Es concebible un estado mental que no se exterioriza en una conducta, pero no lo es un estado mental que no se perciba de una manera característica por el sujeto que lo tiene (esa percepción no tiene por qué ser una sensación; las creencias, por ejemplo, no tienen por qué llevar aparejada una sensación. Pero hay algo que nos hace saber que tenemos una creencia, sin necesidad de observarnos “desde fuera”). Los *qualia*, en definitiva, son negados explícitamente por Hacker y Bennet.

⁵² J. Searle, “Réplica a Philosophical Foundations of Neuroscience”, en *La naturaleza de la conciencia*, pág. 130.

⁵³ T. Nagel, *Mente y cosmos*, Capítulo II.

⁵⁴ D. González Lagier, “Buenas razones, malas intenciones”, pág. 59.

Tampoco otorgan la debida importancia estos autores al hecho de que la neurociencia ha establecido que existe alguna relación causal entre los estados cerebrales y los estados mentales. Por supuesto, no es algo que nieguen. Pero si se toma en serio que el funcionamiento del cerebro es causa de la existencia de las experiencias mentales, entonces la atribución de estados mentales al cerebro o componentes del mismo puede verse como una manera aceptable de hablar, no muy distinta de cuando decimos que el estómago hace la digestión. Se trata sólo de localizar la causa de un proceso, el encargado de una función. No parece haber aquí una falacia grave, ni por ello se está identificando al cerebro con el sujeto, ni se está reduciendo el concepto de estado mental al concepto de estado cerebral (pues un concepto completo de estado mental ha de incluir las “propiedades de lo mental”): No se está diciendo –por ejemplo- que las intenciones son estados cerebrales, sino que son causadas por estados cerebrales. El argumento de Hacker, Bennet, Pardo y Patterson para considerar esta atribución como una falacia es que los estados mentales son conductas o disposiciones a actuar de una u otra manera, por lo que sólo se pueden atribuir a quien puede actuar. Este es un concepto que tampoco da cuenta de las propiedades de lo mental. Si se rechaza esta definición reductivista, la acusación de falacia a estas atribuciones cae por sí sola. Pueden a veces resultar extrañas desde el punto de vista del lenguaje común, pero probablemente ello se explique por la carga de dualismo cartesiano que persiste en ese lenguaje.

6. Para terminar: Teorías de la mente y pruebas neurofisiológicas.

¿Qué consecuencias tiene todo lo anterior para las pruebas neurofisiológicas de los estados mentales? Creo que Michele Taruffo tiene razón cuando vincula la admisibilidad de las pruebas neurocientíficas con cuestiones filosóficas y conceptuales. Me parece que esta apertura a los argumentos filosóficos es uno de los grandes méritos no sólo del artículo que he comentado, sino de toda su obra. La ciencia tiene necesariamente presupuestos conceptuales que la filosofía debe revisar, por lo que la aceptación y comprensión de las hipótesis científicas debe hacerse teniendo en cuenta precauciones conceptuales que eviten afirmaciones absurdas y sin sentido. En esto, tanto Taruffo como Pardo y Patterson, por un lado, y Hacker y Bennet, por otro, tienen razón. Pero dentro de una “naturalización no reductivista” de la filosofía, debe aceptarse cierta influencia de los descubrimientos empíricos en la filosofía. La filosofía no puede dar la espalda a la ciencia, al menos en el sentido de que sus conceptos no deben contradecir, en última instancia, lo establecido empíricamente. Y una de las cosas que parece que la ciencia ha mostrado es la relación causal de los estados mentales respecto del funcionamiento del cerebro.

Las distintas teorías de la mente dan diferentes respuestas acerca de la relación entre neurociencia y psicología, esto es, acerca de la relevancia de la neurociencia para el conocimiento de los estados mentales y, por tanto, para la prueba de los mismos:

(1) Para el materialismo eliminacionista, la psicología debe ser eliminada y la neurociencia tiene la única palabra sobre nuestros estados internos. La neurociencia, por tanto, puede probar que el sujeto está en determinado estado cerebral, pero no puede probar ningún estado mental, porque tal cosa no existe. Por tanto, uno de los presupuestos de las pruebas neurocientíficas de los estados mentales es precisamente el rechazo de este tipo de materialismo.

(2) Para el materialismo de la identidad de tipos, que postula una correlación entre tipos de estados mentales y tipos de estados cerebrales, la psicología proporciona una explicación incompleta, imprecisa, de nuestra conducta, por lo que debe ser reducida a la neurociencia. Ésta tiene un valor probatorio pleno de los estados mentales, que no son otra cosa que estados cerebrales.

(3) El materialismo de la identidad de sustancias, de acuerdo con el cual cada estado mental se corresponde con un estado cerebral, pero no siempre el mismo, no puede reducir la psicología a la neurociencia y, por tanto, no puede conceder a la neurociencia ningún valor probatorio (o sólo en la medida en que se acepte algún grado de probabilidad en la coincidencia entre estados de uno y otro tipo).

(4) Para el conductismo (y para el funcionalismo) la psicología no puede ser reducida a la neurociencia, sino que ha de ser reconstruida como una psicología conductista (o, en el caso del funcionalismo, a una ciencia cognitiva que estudie los procesos mentales como modos de procesar información). Desde este punto de vista, no existe una conexión intrínseca entre el cerebro y la mente, pero, como reconocen Pardo y Patterson, la neurociencia puede constituir prueba indiciaria de los estados mentales si se acepta cierta correlación entre estados de uno y otro tipo. No obstante, estas pruebas cederían frente a la conducta (o la presencia de las relaciones causales adecuadas con la conducta y otros estados mentales), que, al ser constitutiva del estado mental, hace prueba plena del mismo. Estos autores sugieren que, en caso de conflicto entre lo que indican las pruebas neurofisiológicas y lo que muestra la conducta, es ésta la que debe prevalecer (lo que plantea enormes problemas dada la posibilidad de la mentira y el fingimiento). En mi opinión, aunque puede aceptarse que la conducta tiene -por usar la expresión de G. H.

von Wright⁵⁵ - una *prioridad semántica o conceptual* sobre lo mental, lo neuronal tiene una prioridad ontológica sobre lo mental, que por supuesto es relevante a efectos de prueba.

(5) Para el materialismo emergentista no cabe una reducción de la psicología a la neurociencia, porque los estados mentales tienen propiedades que no quedan reflejadas en el estudio de los procesos neuronales. La psicología, por tanto, puede proporcionar explicaciones válidas de la conducta que pueden ser usadas para la prueba de los estados mentales. Ahora bien, si los estados cerebrales son condiciones causalmente suficientes de los estados mentales, como reconoce el emergentismo, y si la relación causal es entre tipos de estados cerebrales y tipos de estados mentales, técnicas neurocientíficas suficientemente avanzadas, precisas y fiables tendrían un valor probatorio decisivo, porque sería imposible estar en ese estado cerebral y no tener el estado mental correspondiente. En definitiva, tendríamos dos vías complementarias para probar el estado mental, y la vía neurocientífica sería incontestable si alcanzara suficiente fiabilidad. Pero la tarea de correlacionar tipos de estados cerebrales y tipos de estados mentales es todavía incipiente. Mientras tanto, para el emergentismo las pruebas basadas en la psicología pueden ser suficientemente válidas.

La lección que podemos extraer del repaso de las teorías de la mente es que la *pertinencia* de las pruebas neurocientíficas requiere (a) una relación entre estados cerebrales y estados mentales, (b) no eliminar uno de los polos (como hace el materialismo eliminacionista); (c) no eliminar las propiedades de lo mental (como hacen muchos reduccionismos, entre ellos el conductismo); (d) una correlación estable entre tipos de estados cerebrales y tipos de estados mentales (lo cual no depende, obviamente, de la teoría, sino que es una cuestión empírica).

Creo que el materialismo emergentista puede satisfacer o dar cuenta de estas condiciones. Y creo igualmente que por misteriosa que todavía pueda parecer la idea de emergencia, apunta en la buena dirección. Puede que, por el momento, sea poco más que un programa de investigación, un balbuceo, pero en mi opinión es un balbuceo que trata de dar coherencia a los ingredientes adecuados, y hasta un balbuceo puede ser mejor que las extrañas imágenes del mundo mental y de nosotros mismos que ofrecen algunas de las teorías rivales.

⁵⁵ Como una forma de mostrar la imposibilidad de reducir una a otra la psicología popular o psicología de la conciencia, el conductismo y la neurociencia, G.H. Von Wright ha destacado que lo neuronal tiene una *prioridad causal* sobre la conducta, lo mental tiene *prioridad epistémica* sobre lo neuronal (pues es necesario tener identificado el estado mental al margen del estado cerebral para poder establecer la correlación entre estados de uno y otro tipo) y la conducta tiene una *prioridad semántica* (o conceptual, dirían los conductistas) sobre lo mental (dado el argumento wittgensteniano de cómo adquieren significado los conceptos mentales). G.H. v. Wright, "Valor, norma y acción en mis escritos filosóficos. Con un epílogo cartesiano", en *Doxa*, núm. 26, 2003.

Actitud hacia los estados mentales	Teorías de la mente	Definición de Estado mental	Problema Libre albedrío	Atribución de estados mentales al cerebro	Valor epistemológico de la psicología y de la neurociencia	Objeciones
Eliminacionismo	Materialismo eliminacionista	No existen	[compatibilismo] Posibilidad indeterminismo ontológico	No: no hay estados mentales. No puede cometer la falacia mereológica	La psicología esta equivocada. La neurociencia prueba estados cerebrales, no mentales	¿No existen los estados mentales? ¿No siento dolores ni tengo creencias?
Reduccionismo	Tesis de la identidad de tipos	Los e. mentales son e. cerebrales. Correspondencia tipo-tipo.	Posibilidad indeterminismo ontológico	si	Psicología como aproximación Neurociencia valor epistemológico pleno	Tendencia a eliminar las propiedades de lo mental al hacer la reducción. Elimina punto de vista interno.
Reduccionismo	Conductismo	Los e. mentales son acciones o disposiciones conductuales. Cualquier entidad que exhiba la conducta adecuada tiene estados mentales (con indep.de si tiene o no cerebro como órgano)	No comprometido con determinismo neurofisiológico	No	Conducta como "prueba criterial" de los e. mentales. Neurociencia puede ser prueba indiciaria, pero cede en caso de conflicto.	No explica propiedades de lo mental. Elimina punto de vista interno. No posibilidad causación mental. Objeción superespartano. Confunde cuestión epistemológica y semántica con cuestión ontológica
Reduccionismo	Funcionalismo	Los e.mentales son propiedades funcionales o roles causales, que se definen por correlacionar ciertos inputs con ciertos outputs.	No comprometido con determinismo neurofisiológico	No	Roles causales como "prueba criterial". Neurociencia como prueba indiciaria, pero cede en caso de conflicto.	No explica propiedades de lo mental. Elimina punto de vista interno.
No reduccionismo	Tesis de la identidad de instancias	Los estados mentales son estados cerebrales. Sólo correlación caso-caso.	Posibilidad indeterminismo ontológico. Indeterminismo epistemológico.	Si, pero al cerebro en su conjunto, no a estados cerebrales concretos.	Neurociencia no tiene valor epistemológico de los e.mentales. Psicología es ineliminable e irreductible.	No explica propiedades de lo mental (aunque no las elimine). ¿cómo pueden estados físicos tener propiedades como los qualia?
No reduccionismo	Materialismo Emergentista	Los e.mentales son propiedades emergentes o sistémicas. Monismo de sustancias, dualismo de propiedades.	Posibilidad de causación "de arriba a abajo"	Si. Reduccion causal, no ontológica. Los e.cerebrales son causa suficiente de los e.mentales.	Neurociencia tiene valor pleno de la prueba de la existencia del e.mental (si se demuestran correlaciones tipo-tipo), pero éste no puede describirse completamente sin referencia a conceptos psicológicos.	Relación de emergencia es un concepto oscuro.

