

Domingo Juan José Díaz

Estudio de la relación memoria
autobiográfica-sociabilidad: hacia
una neurofilosofía de la
sociabilidad

Director/es
Pérez Chico, David

<http://zaguan.unizar.es/collection/Tesis>

ISSN 2254-7606



Prensas de la Universidad
Universidad Zaragoza



Universidad de Zaragoza
Servicio de Publicaciones

ISSN 2254-7606

Tesis Doctoral

**ESTUDIO DE LA RELACIÓN MEMORIA
AUTOBIOGRÁFICA-SOCIABILIDAD: HACIA UNA
NEUROFILOSOFÍA DE LA SOCIABILIDAD**

Autor

Domingo Juan José Díaz

Director/es

Pérez Chico, David

**UNIVERSIDAD DE ZARAGOZA
Escuela de Doctorado**

Programa de Doctorado en Filosofía

2024



Universidad
Zaragoza

Tesis Doctoral

ESTUDIO DE LA RELACIÓN MEMORIA AUTOBIOGRÁFICA-SOCIABILIDAD: HACIA UNA NEUROFILOSOFÍA DE LA SOCIABILIDAD

Autor

Domingo Juan José Díaz

Director

David Pérez Chico

Facultad de Filosofía
2024

A Pilita

Había aprendido sin esfuerzo el inglés, el francés, el portugués, el latín. Sospecho, sin embargo, que no era muy capaz de pensar. Pensar es olvidar diferencias, es generalizar, abstraer. En el abarrotado mundo de Funes no había sino detalles, casi inmediatos.

Jorge Luis Borges
Funes el memorioso

La razón de que el hombre sea un ser social, más que cualquier abeja y que cualquier otro animal gregario, es clara. La naturaleza, pues, como decimos, no hace nada en vano. Sólo el hombre, entre los animales, posee la palabra.

Aristóteles
La Política

Agradecimientos

Quiero comenzar mis agradecimientos con unas palabras para mi director de tesis, David Pérez Chico, por sus recomendaciones y lecciones impartidas, por haberme escuchado en los momentos de dudas y por haber alentado mis motivaciones a lo largo de estos años. Todo este proyecto hubiese sido imposible sin su ayuda y paciencia infinitas.

Asimismo, deseo agradecer a todas las personas de la Universidad de Zaragoza que me han prestado su colaboración, en especial, al personal del Departamento de Filosofía, Biblioteca de la Facultad de Filosofía y Letras María Moliner y Escuela de Doctorado. También, vaya un muy especial agradecimiento a los profesores Manuel Liz, Javier García Campayo, Manuel Bedia y Cristina Andreu.

De la misma manera, quiero agradecer a Marta García Alonso del Departamento de Filosofía y Filosofía Moral y Política de la UNED, quien fue guía y directora del TFM que dio origen a esta investigación doctoral, motivando mi interés en las relaciones entre la filosofía y las neurociencias.

En un trabajo centrado en la memoria autobiográfica y la sociabilidad no quiero dejar de lado, mucho menos olvidar, los momentos complicados que hemos atravesado. Afortunadamente, el amor, la compañía y las palabras de los seres queridos siempre dan fuerzas extras para superar los obstáculos.

Por último, agradezco a mi familia todo su apoyo y preocupación, a mi hermana María Guadalupe, a mi madre María Águeda y, muy especialmente, a mi esposa María Pilar por todo su amor, comprensión y ayuda.

Resumen

El presente trabajo es una indagación sobre la relación entre la memoria autobiográfica y la sociabilidad, donde se aprecia una vía de comunicación bidireccional entre ambos conceptos. El carácter peculiar de la memoria autobiográfica reside en que es un tipo de memoria específicamente humana. Está apoyada sobre tres estructuras: narrativa, experiencia y recuerdos. La sociabilidad humana, a diferencia de otras especies del reino animal, integra elementos cognitivos, culturales y valores sociales que son únicos. La relación se ha analizado desde la perspectiva de la neurofilosofía, considerándose, además, aspectos clave como la influencia de determinadas patologías sociales y mentales. En cuanto a aspectos filosóficos, su importancia está focalizada en lo que se considera como la «antesala de la moral» donde conviven elementos genéticos, sociales, culturales y medioambientales. Asimismo, el estudio incorpora una propuesta para abrir una línea investigación interdisciplinar desde la neurofilosofía con el propósito de indagar el impacto de la inteligencia artificial, la robótica y la realidad extendida en la sociabilidad humana, y analizar sus efectos sobre la identidad.

Abstract

The present work is an inquiry into the relationship between autobiographical memory and sociability, where a bidirectional communication between both concepts is appreciated. The peculiar character of autobiographical memory lies in the fact that it is a specifically human type of memory. It is supported by three structures: narrative, experience and memories. Human sociability, unlike other species in the animal kingdom, integrates cognitive, cultural, and social values that are unique. The relationship has been analysed from the perspective of neurophilosophy, and key aspects such as the influence of certain social and mental pathologies have also been considered. Philosophically, its importance is focused on what is considered the «threshold of morality», where genetic, social, cultural and environmental elements coexist. Likewise, the study incorporates a proposal to open an interdisciplinary research line from neurophilosophy with the aim of investigating the impact of artificial intelligence, robotics and extended reality in human sociability and analyse its effects on identity.

Prefacio

« [...] Emprendí este proyecto porque quería entender qué es lo que tienen los cerebros de los mamíferos altamente sociables que permite esa sociabilidad» (Churchland P. S., 2012, pág. 20).

Las palabras precedentes que encabezan este Prefacio están allí no por azar ni por relleno intelectual, todo lo contrario. Su importancia es doble, pues ellas contienen el germen que ha despertado en mí una curiosidad movilizadora y creciente. Asimismo, son las portadoras de la luz que ha guiado los objetivos de este trabajo: indagar los misteriosos caminos de la sociabilidad humana.

Quienes nos dedicamos a la tarea de filosofar sabemos muy bien que nuestra labor no se traduce en encontrar respuestas claras, evidentes y complacientes, sino en construir preguntas que incomodan, que cuestionan lo establecido, que dan pie a nuevas preguntas, en definitiva, preguntas que movilizan el pensamiento y la acción. De esta manera, la senda recorrida me ha conducido por extraños territorios fronterizos donde conviven, y no siempre en feliz armonía, disciplinas filosóficas y científicas en medio de un espacio claroscuro de actividad interdisciplinar.

El trabajo no está concluido, sino que forma parte de un proyecto más grande, por lo menos en intenciones, para avanzar hacia nuevos planteamientos que dejo esbozados en la presente tesis. En el origen, se halla el TFM dirigido por la Dra. Marta García Alonso del Departamento de Filosofía y Filosofía Moral y Política de la UNED titulado *Neuroética: una aproximación a la ética contemporánea basada en el cerebro*. Este trabajo fue el primer acercamiento a unas disciplinas y cuestiones que comenzaban eclosionar en el mundo académico, sobre todo, anglosajón. Casi una década después, las investigaciones relacionadas con el campo de la neuroética y las ciencias cognitivas se han multiplicado en todo el mundo, últimamente, alentadas por el desarrollo de la inteligencia artificial (IA), la realidad extendida, la relación humano-máquina o los planteamientos que surgen desde la corriente del Transhumanismo. Echando la vista atrás,

todo ello me indica que el camino emprendido, a pesar de los escollos, ha sido el correcto, cuando menos, ilusionante.

De aquellas lecturas en neuroética surgieron las primeras preguntas sobre nuestra naturaleza social y cómo los humanos somos capaces de construir un espacio vital denominado «*ethos social*», un lugar en el que considero reside la antesala de nuestra moralidad en el sentido de hábitos y costumbres, esas que luego nos permiten analizar y construir todo el edificio de la ética. Asimismo, me han sido de gran ayuda las investigaciones que desde la neurociencia se hacen sobre las patologías que repercuten en nuestras funciones cognitivas y conducta, tales como el *Alzheimer* o los diversos tipos de demencias, a las cuales he considerado como una «investigación inversa», pues nos permiten reconstruir las hipótesis sobre cómo es posible la construcción social a partir de la destrucción de la identidad.

Como he comentado, este trabajo no es más que una pequeña porción de una pregunta gigante que me gustaría poder responder, pero sé que es una tarea imposible, pues el tiempo la construye y destruye con el devenir, y es allí donde nuestro trabajo se transforma en un minúsculo aporte. Parafraseando a Albert Camus, «imaginemos a Sísifo feliz».

Índice

Agradecimientos.....	7
Resumen	9
Abstract	10
Prefacio	12
Capítulo 1. Introducción.....	17
Capítulo 2. Hipótesis y objetivos de la investigación	22
Capítulo 3. Aspectos metodológicos y objetivos particulares	25
Capítulo 4. Escrutando la memoria: entre ciencia y filosofía.....	28
4.1. Antecedentes	28
4.2. La memoria humana: entre filosofía, conductismo y cognitivismo	33
4.3. Aclaraciones terminológicas: plasticidad cerebral, cerebro, encéfalo y neuronas	39
4.4. La memoria y sus sistemas.....	50
Capítulo 5. La memoria autobiográfica.....	57
Capítulo 6. Sociales por naturaleza.....	68
6.1. La base de la sociabilidad: los valores morales	70
6.2. Neuronas espejo, empatía y desarrollo del cerebro social.....	73
6.3 Imitación y habilidades sociales.....	79
6.4. Altruismo, cooperación y mutualismo.....	82
Capítulo 7. Neurofilosofía	89
7.1. Entre filosofía y neurociencia	91
7.2 El problema del dualismo mente-cuerpo.....	95
7.3 El conductismo lógico	100
7.4 Materialismo eliminativo.....	105
7.5 Neurofilosofía y la interfaz filosofía, psicología y neurociencia	111
7.6 Patricia Churchland y la neurofilosofía.....	120
7.7 De la neurofilosofía a la neuroética.....	126
7.8 Sumario de diapositivas.....	133
Capítulo 8. Propuesta: hacia una neurofilosofía de la sociabilidad	153

Capítulo 9. Conclusiones.....	162
Bibliografía	167
Diapositivas.....	178
Anexos	209

Capítulo 1. Introducción

El señor X, por razones que no vienen al caso, pierde su memoria durante el sueño. A las 7,30 suena el despertador y el señor X despierta: no entiende lo que significa el pitido de un artefacto con agujas y números que hay en su cabeza. El señor X siente hambre y se incorpora. El despertador sigue sonando. Al lado de la cama hay unas zapatillas y ropas. El señor X lo único que sabe es que necesita comer. Descalzo y en pijama, sale del dormitorio y comienza a deambular por las distintas habitaciones en busca de comida. Casualmente, al fin, acaba en la cocina, donde el color rojo de una manzana le invita a morderla. El señor X comienza a sentirse mejor, pero hay algo que le molesta: un pitido cada vez más insoportable que viene del dormitorio. El señor X se dirige a la habitación del pitido y toma con brusquedad el reloj. Lo mira, lo golpea, hasta que por azar presiona una palanquita y el ruido cesa. El señor X tiene más hambre. Vuelve a la cocina y come desordenadamente distintos alimentos. Siente que el suelo está cada vez más frío y que todo su cuerpo comienza a tiritar. Corre junto a la cama. En un sillón cuelgan pantalones, unas camisas, un jersey, una chaqueta..., pero nada de aquello significa nada para él. Instintivamente, se vuelve a la cama y se cobija entre sábanas y mantas. (Ruiz-Vargas J. M., 1998, pág. 28)

La cita describe la vivencia de un ser en la más absoluta soledad: la desconexión del mundo tanto interior como exterior. A decir verdad, su lectura me deja perplejo, pues el intento de experimentar la situación vivida por el «señor X» me sumerge en la insoportable angustia que puede ser la vida humana privada de recuerdos, o lo que es lo mismo, vivir recluido en la prisión del olvido perpetuo.

Camouflada bajo el título se esconde una de las preguntas, quizás la principal, que dio origen de este trabajo de investigación: ¿cómo los mamíferos humanos hemos llegado a convertirnos en «animales sociales y políticos»? (Aristóteles, 2007a, págs. I-7, 58) y (Aristóteles, 2007b, págs. I-2, 47). Considero que la definición aristotélica es del todo acertada, pues nos revela el carácter

biológico y social de la naturaleza humana. Es evidente que una respuesta definitiva que desvele los misterios de esta pregunta excede los propósitos del trabajo. Sin embargo, no he querido renunciar al intento de hallar al menos una aproximación a partir de un análisis interdisciplinar, pues la naturaleza de los conceptos que intervienen así lo requiere. Por esta razón, he diseñado un itinerario de investigación transversal que recorre tres ejes temáticos: (1) la memoria autobiográfica (MA); (2) la sociabilidad humana; y (3) el análisis de un marco conceptual desde la neurofilosofía.

En esta investigación partiré de la posición que determina la existencia de una estrecha relación entre los conceptos de «MA» y «sociabilidad», los cuales están unidos por verdaderos puentes de mutua influencia e interacción. Asimismo, entiendo que esta relación es la base que permite dotar de contenido y significado a dichos conceptos, constituyendo la estructura esencial para el desarrollo y manifestación de las experiencias humanas que, en su organización más elemental, necesita de un sustrato biológico (nivel interno) y de un ámbito socio-cultural (nivel externo).

Los avances en genética, epigenética, biología evolutiva, medio ambiente, antropología, neurociencias y neurotecnología entre otras, muestran un buen número de evidencias sobre la composición y funcionamiento de la MA y la sociabilidad que desvelan la existencia de una red de comunicación entre procesos biológicos y sociales mucho más compleja y estrecha de lo que se pensaba hace tan solo algunas décadas atrás.

En cuanto al marco metodológico elegido, la neurofilosofía, como «interfaz entre filosofía, neurociencias y psicología» (Churchland P. S., 1986), ofrece una síntesis crítica sobre determinados conceptos que atraviesan la historia de la filosofía y que ahora están siendo empleados en el desarrollo de diversas investigaciones neurocientíficas, por ejemplo: «consciencia» (Churchland P. M., 1992), (Koch, 2006) y (Bennett, Dennett, Hacker, & Searle, 2008), «empatía» (Iacoboni, 2011), «altruismo» (Tomasello, Marengo, & Dweck, 2010), «libre albedrío» (Churchland P. S., 1986), «libertad» (Mele, 2009) o «responsabilidad» (Morse, 2011).

La importancia de este marco metodológico está en el análisis de los resultados de disciplinas experimentales y cómo estas utilizan conceptos complejos que atañen a disciplinas filosóficas (no experimentales). Estoy convencido de que una depuración conceptual nos permitirá construir sobre terreno firme y evitar conflictos entre disciplinas.

El modelo de neurofilosofía planteado por Patricia Churchland descansa sobre los pilares del materialismo eliminativo, una concepción que deja a un lado toda psicología popular (*folk*), puesto que «constituye no solamente una mala representación de nuestra naturaleza interna, sino que directamente constituye una mala representación de nuestros estados y actividades internos» (Churchland P. M., 1992, pág. 76). Una vez que se ha eliminado cualquier vestigio de psicología popular, será la neurobiología quien nos guíe por el laberinto fisiológico en donde «el cerebro, más que algunas cosas no físicas, es el que siente, piensa y decide» (Churchland P. S., 2012, págs. 1-2). De esta manera, el antiguo marco popular de interpretación abre el paso a las explicaciones desde una neurociencia más desarrollada (Churchland P. M., 1992, págs. 75-76).

Esta perspectiva deja en claro que para entender cualquier proceso mental «primero debemos entender el cerebro» (Churchland & Di Francesco, 2007, pág. 179), pues la mente es definida como un «nivel de actividad cerebral» (Churchland P. S., 2007, pág. 185).

Según Churchland (2008, pág. 409), «el peso de la evidencia indica que los procesos mentales son en realidad los procesos del cerebro», de este modo avanzamos hacia la solución, en parte, del problema de la dualidad mente-cerebro planteado por Descartes. Por lo tanto, quedan excluidas las interpretaciones dualistas de la relación mente-cerebro y el acceso al conocimiento por medio de las «intuiciones populares» (Churchland P. S., 2007, pág. 187).

No se trata de declarar *a priori* la absoluta inutilidad del conocimiento intuitivo acerca del cerebro, sino de revisar hasta qué punto este órgano está preparado para abordar directamente una investigación sobre sí mismo. Cuestiones tales como «¿por qué nos deprimimos?», «¿qué es el amor?», «¿por qué soña-

mos?» o «¿qué niveles de serotonina influyen en nuestras decisiones?» requieren de un análisis interdisciplinar, sobre todo de aquellas evidencias sólidas provenientes de las ciencias experimentales (Churchland P. S., 2008, pág. 409).

Los estudios sobre la memoria humana no son ajenos a las dificultades que plantea la elección de una metodología adecuada. Por ejemplo, el neurocientífico Howard Eichenbaum reconoce que «con la aparición de métodos modernos de neurociencia y ciencia cognitiva, así como su combinación, han surgido nuevas y profundas percepciones sobre los mecanismos de la memoria» (2003, págs. 19-20), las cuales han renovado y fortalecido los estudios a partir de nuevas evidencias.

La neurociencia describe al cerebro como una construcción biológica y social, es decir, una materia plástica cuya fisiología es modelada en gran parte por el entorno social y el medioambiente (Gazzaniga M. S., 1993). Ahora bien, si nuestra memoria necesita un soporte físico ubicado en el cerebro, es evidente que esta plasticidad bio-social de la que hablamos no le puede ser ajena. Las patologías neurodegenerativas muestran que cualquier alteración o modificación grave en la materia-soporte tiene consecuencias directas en la vida interna y social del individuo. Sin memoria, el mundo deja de ser un lugar conocido, un espacio vacío de contenido.

La memoria, en cuanto «experiencia subjetiva», no puede ser abordada exclusivamente desde una perspectiva objetiva y experimental (Redolar Ripoll, 2014, pág. 3). Por ello, considero que el marco de análisis teórico-experimental que nos ofrece la neurofilosofía, como interfaz entre las neurociencias, la psicología y la filosofía, es el más adecuado para tender un puente de comunicación entre la MA y la sociabilidad humana.

En las primeras décadas del siglo XXI, la irrupción de la inteligencia artificial y el desarrollo de la realidad extendida están configurando significativamente la casi totalidad de los ámbitos de la vida humana, la sociedad y hasta nuestra forma de conocer y relacionarnos con el medio ambiente, una relación mediada y transformada por el software y las pantallas.

La experiencia cognitiva en ningún caso es ajena a todo este torbellino, pues mezcla sensaciones de escasas certidumbres con piélagos de dudas difíciles de atravesar. Desde el actual escenario, podemos visualizar, allá a lo lejos y entre densos bancos de niebla, el faro que nos sugiere que tampoco será igual la relación entre nuestra memoria autobiográfica y la construcción de la sociabilidad, y para ello tendremos que prepararnos.

El desafío de la fabricación de mundos virtuales nos enfrenta de lleno con aquellas preguntas kantianas que dormían en el viejo arcón filosófico: ¿qué puedo saber?; ¿qué debo hacer?; ¿qué me cabe esperar?; y ¿qué es el hombre? De tal magnitud está siendo esta especie de temblor total que se nos hace casi imposible la misión de imaginar tan siquiera el horizonte del siglo venidero. Desde la incertidumbre de nuestro presente podemos agregar una cuestión y preguntarnos ¿cuál será nuestro lugar?

Capítulo 2. Hipótesis y objetivos de la investigación

La hipótesis que anima esta investigación señala que entre la memoria autobiográfica (MA) y la sociabilidad existe una influencia recíproca que permite el desarrollo bio-sociocultural de nuestra especie. Ambos elementos han sido modelados por siglos de evolución, creando una comunicación que es esencial para la adaptación al medio y la supervivencia, lo cual constituye una ventaja evolutiva.

Este trabajo propone dos objetivos que podrían considerarse en un mismo nivel de interés. Por un lado, señalar la importancia que un estudio sobre la relación entre la MA y la sociabilidad no es otra cosa que la búsqueda de aquellos elementos que conforman el sustrato fundamental y necesario para la construcción de la moralidad humana, pues nos encontramos en la antesala de moralidad como *ethos* social de nuestras habitudes y carácter.

Por otro lado, el segundo objetivo principal que persigue esta investigación es poner de manifiesto la riqueza conceptual y metodológica que podría llegar a aportar la creación de una línea de investigación interdisciplinar a través de lo que he llamado *neurofilosofía de la sociabilidad*. En tal sentido y como ejemplo de ello he desarrollado la relación entre la memoria autobiográfica (específica de nuestra especie) y el proceso de construcción de la sociabilidad humana. A partir de esta relación se abre un abanico de cuestiones y planteamientos que desbordan, por supuesto, los límites de lo estrictamente filosófico o lo estrictamente científico.

Todo ello me ha animado a construir puentes entre territorios con límites bien demarcados, pero a la vez, fronterizos, lo cual les otorga un tratamiento especial. Considero que esta es la natural esencia del conocimiento humano, siempre en desarrollo, entre territorios vecinos.

La sociabilidad (*ethos* o carácter social) comprende un sofisticado sistema cognitivo basado en el aprendizaje, la atención, la percepción y, sobre todo, la

memoria. Con respecto a esta última, hoy sabemos que además de almacenar, procesar y recuperar, también es capaz de proyectar información de las experiencias hacia el futuro, por ejemplo, en la toma de decisiones o en la evaluación de circunstancias.

La MA es un sistema de memoria auténticamente humano que comprende los recuerdos (hechos acontecidos, fantasías, ilusiones y proyecciones), la construcción de la narración de tales recuerdos y, por supuesto, la experiencia. Según Ruiz Vargas, «constituye el punto crítico en el que convergen los sentimientos, las motivaciones y los deseos, las metas y los logros, los valores, las creencias y los significados de cualquier persona» (2004, pág. 1). Para la psicología este tipo de memoria es fundamental para la construcción del Yo. De aquí extraigo su importancia para la filosofía, ya sea si hablamos de un «sujeto cognitivo X» o un «agente moral Y», cualquiera sea el concepto que busquemos construir, no será posible pensarla aislada de esta estructura de memoria, pues:

La información que genera sirve a importantes constructos de autodefinición del individuo, como la identidad, que le aporta unidad, propósito y sentido a la vida de una persona, cuando recuerda el pasado o proyecta el futuro, algo tan necesario en una sociedad globalizada que busca cada vez más homogeneizar. (Beltrán-Jaimes, Moreno-López, Polo-Díaz, Zapata-Zabala, & Acosta-Barreto, 2012, págs. 109-110)

En relación a los objetivos particulares derivados de los capítulos, los mismos están centrados en revisar la actualidad del proyecto neurofilosófico originario propuesto por Patricia S. Churchland; realizar una revisión documental de la bibliografía sobre la MA; y exponer las posibles evidencias que confirmen el nexo entre la sociabilidad de los mamíferos humanos —ni más ni menos que nuestra sociabilidad— y la MA. Estos objetivos nos permitirán acercarnos a lo que considero constituye el origen de la conceptualización aristotélica sobre cómo los mamíferos humanos nos constituimos en animales políticos y sociales o zoón politikón (*ζῶον πολιτικόν*).

Considero que este estudio, si bien teórico, posee un interés eminentemente práctico para comprender las bases de la vida humana, puesto que recorre un camino de ida y vuelta entre el individuo y la sociedad en sus estructuras fundamentales, pues nos encontramos en lo que podemos considerar la *antesala de la moralidad* entendida como el ámbito natural de las hábitudes que dan forma a nuestra peculiar forma de ser y estar en el mundo.

Para concluir, considero que abrir una línea de investigación interdisciplinar como la que se propone en este trabajo nos puede permitir elaborar nuevas preguntas para explorar los desafíos que plantean a la sociedad el desarrollo de la inteligencia artificial (IA), la neurotecnología (en la relación cuerpo-máquina/invasivo-no invasivo), la realidad extendida y los múltiples universos explorados a través de la tecnología del metaverso.

Tales cuestiones han puesto en revisión conceptos que atraviesan toda la historia de la filosofía y de la ciencia, y que necesitaremos volver a revisar su alcance y significado, por ejemplo, *experiencia, mundo, realidad, libre albedrío o identidad humana* entre otros. Asimismo, se hace perentoria una investigación a gran escala sobre las causas que están generando diversas socio-patologías atribuidas a las nuevas tecnologías y cómo estas han colonizado todos los ámbitos sociales, dejando apenas reductos libres de virtualidad.

Llegados a este punto, no cabe más que preguntarnos si estaremos a las puertas de una nueva sociabilidad y cómo hemos de habitar este mundo para dejar un resquicio de libertad para las generaciones futuras.

Capítulo 3. Aspectos metodológicos y objetivos particulares

La complejidad del tema exige una investigación que debe abordarse desde diversas disciplinas. Por tal motivo, como metodología propongo una revisión de los conceptos más relevantes en diversas áreas de las neurociencias, la psicología, la antropología, la sociología y la filosofía, sobre cuáles son los aspectos generales y específicos de la memoria autobiográfica y la sociabilidad. El objetivo de ello es encontrar un puente de comunicación entre ambos conceptos que ayude a esclarecer la hipótesis fundamental del trabajo: la influencia recíproca.

Si bien algunos capítulos incorporan un breve planteamiento histórico de la situación, no es objetivo de la investigación indagar más allá en las líneas históricas. Por ello, se pondrá énfasis en las discusiones científicas y filosóficas más recientes. Desde una perspectiva estrictamente filosófica, se revisarán algunos de los elementos del modelo original de la neurofilosofía planteado por Patricia Churchland.

Asimismo, se realizará una lectura crítica de las fuentes bibliográficas específicas sobre la memoria autobiográfica y la neurofilosofía, las cuales están asentadas en las neurociencias, la neurotecnología, la neurociencia social y cognitiva, las ciencias sociales, la psiquiatría, la psicología y la filosofía. Para ello se consultarán bases de datos bibliográficas y otros recursos electrónicos.

Con el propósito de hacer más cómodo el seguimiento de la lectura, el trabajo está estructurado solo en capítulos y subcapítulos. Con respecto a las citas bibliográficas y formato del texto, he optado por el estilo autor-fecha de la norma APA (*American Psychological Association*) 7^a edición. Aquellas citas textuales que no estuviesen traducidas al castellano, serán traducidas mí y el texto original transrito como nota a pie de página para su confrontación. En cuanto a la bibliografía, esta aparecerá ordenada alfabéticamente en un solo cuerpo al

final del trabajo. A continuación de la Bibliografía, he agregado una sección de Anexos que incorpora diapositivas y una breve explicación de las mismas con el propósito de mostrar algunos de los hitos que marcan el desarrollo de la neurofilosofía y neuroética.

Con respecto a los objetivos particulares marcados para cada capítulo, el trabajo propone los siguientes puntos. Para el Capítulo 4 y sus apartados, el objetivo está centrado en mostrar cuáles han sido las líneas de investigación más significativas que han recorrido la historia de la memoria dentro de la propia *psicología de la memoria y las ciencias cognitivas*. En tal sentido, se abordarán aquellos aspectos que se consideran dentro de la bibliografía especializada han servido para fundamentar la estructura de la propia disciplina. Además, se analizarán en líneas generales las principales cuestiones epistemológicas surgidas desde el conductismo y el cognitivismo, algunos términos clave y los diversos sistemas de memoria.

El objetivo del Capítulo 5 será desarrollar y explicar el concepto de memoria autobiográfica, mostrar su importancia dentro de la actual filosofía de la memoria y cómo interviene en la construcción de la identidad.

Para el Capítulo 6 y sus apartados estarán centrados en el estudio de la sociabilidad humana y su relación con la memoria autobiográfica. El objetivo planteado es mostrar cómo se construye el proceso de sociabilidad y su fundamental tarea cognitiva determinada por las neuronas espejo y el desarrollo del cerebro social. Además, mostrar aquellos elementos que he considerado son centrales y que muestran la naturaleza social de nuestra especie.

El Capítulo 7 engloba los temas principales para el tercer eje que se propone en el trabajo, siendo los otros dos (a) el eje relacionado con la memoria autobiográfica y (b) el eje que recorre el estudio de la sociabilidad. Dichos temas hilvanan un recorrido histórico por algunos de los principales debates suscitados en torno al denominado «problema de la dualidad mente-cuerpo» que llega hasta nuestros días. Los temas seleccionados permiten, por un lado, mostrar la complejidad del tratamiento del problema de la dualidad, por otro lado, arribar a las

puertas de la neurofilosofía. La elección no ha sido fácil, mucho menos la exclusión deliberada de temas de la filosofía de la mente, tales como *el funcionalismo*, *la concepción fenomenológica*, *el emergentismo* o *el enactivismo y mente encarnada* entre otros.

Para finalizar, el Capítulo 8 recogerá, a modo de intención personal, una propuesta que considero novedosa para la apertura de una línea de investigación interdisciplinar dedicada a estudiar los desafíos actuales que nos presentan la relación cuerpo-máquina, la inteligencia artificial y la robótica, y su impacto directo en el desarrollo de la relación memoria autobiográfica – sociabilidad, es decir, la cimentación de la identidad humana.

De forma complementaria, el Anexo ubicado al final del trabajo incorpora las figuras que ilustran los capítulos. Asimismo, he considerado pertinente agregar una serie de Diapositivas explicativas, a modo de resumen visual, con algunos de los hitos y nombres más relevantes de los campos de la neurofilosofía y la neuroética. Ello permitirá situar, en un rápido recorrido histórico, sus problemáticas y desafíos actuales.

Este trabajo está relacionado con tres objetivos ODS: ODS 4 Educación de Calidad; ODS 9 Industria, Innovación e Infraestructura; y ODS 16 Paz, Justicia e Instituciones Sólidas.

El autor de esta tesis suscribe y manifiesta su compromiso con el Código de Buenas Prácticas adoptado por la Escuela de Doctorado de la Universidad de Zaragoza.

Capítulo 4. Escrutando la memoria: entre ciencia y filosofía

Este capítulo está dedicado a indagar la memoria humana desde sus conceptos más generales, para luego ir avanzando y revisando su compleja estructura. En la actualidad, la ciencia empírica (neurología, neurociencias y ciencias cognitivas) que está centrada en la investigación de la memoria ha realizado grandes avances, aunque no definitivos, motivados por la necesidad de desvelar los orígenes de determinadas patologías que inciden sobre ella y mitigar el sufrimiento de las personas que las padecen y, claro está, su entorno más cercano. Por otro lado, no hay que olvidar que este tipo de patologías tienen un impacto económico incalculable en los sistemas sanitarios de los países que colaboran en la financiación de los tratamientos y cuidados paliativos, asimismo, en las economías domésticas de las familias.

La filosofía contemporánea se ha hecho eco de estos avances desarrollando disciplinas específicas como, por ejemplo, la denominada filosofía de la memoria en el *Centre for Philosophy of Memory* de la Universidad de Grenoble en Francia. No obstante, el interés por la memoria recorre toda la historia de la filosofía, desde la antigüedad hasta nuestros días.

En lo que respecta a este trabajo, el estudio de la memoria autobiográfica es un eje clave, ya que se considera a su estructura cognitiva como un elemento necesario para la construcción del proceso de sociabilidad, de la identidad y el reconocimiento de los otros.

4.1. Antecedentes

En las últimas tres décadas, los estudios sobre la memoria humana han alcanzado progresos considerables. Sin embargo, las evidencias obtenidas no

han podido derribar determinadas creencias populares, muy arraigadas en la sociedad, que la definen como un «simple contenedor» en el que se van acumulando de forma más o menos ordenada los recuerdos de las experiencias vividas, tanto individuales como colectivas. Si bien esta descripción contiene algo de verdad, no deja de ser limitada, pues concibe a la memoria desde una función primaria que ha quedado superada con las recientes investigaciones: como un mero almacén de recuerdos a los que podemos acceder por medio de la evocación y recuperarlos a voluntad en su forma y contenido.

Las investigaciones en psicología de la memoria y neurociencia cognitiva de la memoria describen a la memoria humana como un sistema activo que permite la construcción de las experiencias presentes y proyecciones de experiencias hacia el futuro. Estas nuevas perspectivas cambian la visión clásica que se tenía de la memoria, esto es, como un sistema anclado solo en el pasado.

En tal sentido, podemos considerar a la memoria «como un complicado sistema de procesamiento de la información que opera a través de procesos de almacenamiento, codificación, construcción, reconstrucción y recuperación de la información» (Ruiz-Vargas J. M., Psicología de la memoria, 1998, pág. 30).

El interés por la naturaleza y función de la memoria ha recorrido toda la historia humana, suscitando diversas especulaciones y representaciones simbólicas y conceptuales. Esto quedó reflejado, por ejemplo, en el arte, la mitología (la mitología griega centraba a Mnemósine como madre de las nueve musas), la religión, la filosofía y, más recientemente, en los estudios científicos sobre la memoria.

El avance en las exploraciones científicas no tardó en dejar al descubierto que las meras «intuiciones subjetivas», es decir, «la introspección y la argumentación lógica, únicos métodos disponibles antes de la exploración científica sistemática de la mente, son insuficientes para desentrañar los misterios de un sistema tan complejo» (Koch, 2006, pág. 205). Por esta razón, la neurociencia contemporánea en sus diversas especialidades —neurociencia molecular, neurociencia celular, neurociencia de sistemas, neurociencia conductual y neurociencia cognitiva— ha asumido, podríamos decir con cierto éxito, la mayor parte de

las investigaciones sobre las cuestiones de la mente y el cerebro, especialmente aquellas relacionadas con temas de importancia para la salud. Un ejemplo de lo que acabamos de mencionar son los estudios que buscan frenar los efectos negativos que tienen las patologías neurodegenerativas sobre los propios pacientes, sus familias y los sistemas sanitarios (Bear, Connors, & Paradiso, 2008, pág. 13).

Las investigaciones científicas sobre la memoria se bifurcan históricamente en dos grandes líneas, una neurobiológica y otra psicológica, que han tenido un «impacto duradero en la neurobiología de la memoria» (Eichenbaum, 2003, pág. 23). Partiendo de estas líneas, H. Eichenbaum (2003, págs. 20-22) configura un esquema histórico-temático basado en «cuatro cuestiones» que, desde su punto de vista, han guiado estas investigaciones a lo largo de aproximadamente un siglo, a saber: «conexión», «cognición», «compartimentalización» y «consolidación». Menciona, asimismo, una «Edad de Oro» de la neurociencia —desde mediados del siglo XIX hasta principios del siglo XX— basada en los trabajos de Willis en 1664 sobre anatomía, y en la descripción realizada por Galvani en 1791 acerca de los mecanismos de conducción de la electricidad a través del sistema nervioso.

E. Kandel (2007) destaca la importancia del estudio de las señales eléctricas y la transmisión de la información, pues según sus investigaciones, son «el fundamento de toda vida mental (...) el lenguaje de la mente, la herramienta que emplean las células nerviosas —piezas constitutivas del cerebro— para comunicarse entre sí a grandes distancias» (pág. 97). Este investigador hace hincapié en «cuatro etapas» del desarrollo de la «biología de las señales eléctricas», desde los experimentos de Galvani en el siglo XVIII hasta las investigaciones sobre los potenciales de acción realizadas Alan Hodgkin y Andrew Huxley. A continuación, expongo un resumen de estos ocho hitos mencionados, las «cuatro etapas» de Kandel y las «cuatro cuestiones» de Eichenbaum, que han sido fundamentales para avanzar en las investigaciones sobre las bases nerviosas, conductuales y cognitivas de la memoria.

En la *primera etapa*, Kandel menciona los experimentos de Galvani en 1791 y las investigaciones en física de Hermann von Helmholtz, quien descubrió

que la electricidad transmitida de un lugar a otro por los axones llevaba en realidad información sensorial del exterior hacia la médula espinal y el cerebro y luego el camino inverso, es decir, desde el cerebro y la médula hacia los músculos (Kandel E. R., 2007, pág. 98). Las señales de von Helmholtz se conocerían más tarde con el nombre de «potenciales de acción».

La *segunda etapa* hace referencia a las investigaciones realizadas en la década de 1920 por Edgar Douglas Adrian, quien prosiguió los estudios sobre las «señales» (potenciales de acción) de H. von Helmholtz. Kandel (2007) destaca que su aporte sobre «la intensidad» como producto de la emisión de los potenciales de acción fue esencial para la historia de la fisiología (págs. 101-102).

La *tercera etapa* en la historia de las señales nerviosas corresponde a los estudios de electrofisiología realizados por Julius Bernstein sobre los mecanismos que intervienen en los potenciales de acción, especialmente la «hipótesis de la membrana» (Kandel E. R., 2007, págs. 103-104).

Por último, en la *cuarta etapa*, Kandel (2007) destaca los trabajos de Hodgkin, Huxley y Katz sobre la generación de los potenciales de acción, «señales clave para la transmisión de información acerca de las sensaciones, los pensamientos, las emociones y la memoria desde una región del cerebro a otra» (pág. 109).

Por su parte, H. Eichenbaum (2003) describe una línea histórica basada en cuatro aspectos esenciales que han recorrido las investigaciones sobre el sistema nervioso y las funciones cerebrales que atañen a la memoria: conexión, cognición, compartimentalización y consolidación.

La «conexión» corresponde al nivel más elemental de la función de la memoria a escala celular y molecular. Los antecedentes más destacados se encuentran en los trabajos de Santiago Ramón y Cajal (doctrina neuronal), Charles Sherrington (arco reflejo) e Iván Pavlov (reflejo condicionado). Siguiendo el análisis de Eichenbaum (2003, pág. 20), «el énfasis en la conexión refleja la principal conclusión que ha surgido en estas investigaciones: que la memoria se codifica

en la dinámica —es decir, la variabilidad o plasticidad— de las conexiones entre células nerviosas».

La «cognición» aborda la «naturaleza de los recuerdos en el nivel máximo de análisis: el psicológico» (Eichenbaum, 2003, pág. 21). En la mencionada «Edad de Oro», aparecieron dos escuelas antagónicas que rivalizaron sus teorías psicológicas: el «conductismo» y el «cognitivismo». Los trabajos de H. Ebbinghaus y W. James, como representantes principales de cada una de estas escuelas, marcaron toda la investigación posterior en la psicología de la memoria.

La «compartimentalización» indica la localización de la memoria y la organización de la corteza cerebral en la que destacan las investigaciones de J. Gall, P. Broca, C. Wernicke, P. Flourens, G. Fritsch, E. Hitzig y D. Ferrier.

La «consolidación» remite fundamentalmente a la función memorística (amnesia o pérdida de la memoria) en «pacientes neurológico con trastornos de la memoria así como en individuos humanos normales» (Eichenbaum, 2003, pág. 38). Los trabajos de T. Ribot, A. Alzheimer, S. Korsakoff, G. Muller, A. Pilzecker y W. Burham están en la base de la consolidación, «un fenómeno hipotético derivado de la observación de que al principio los recuerdos son inestables y más adelante se vuelven resistentes a la pérdida, lo que sugiere un proceso prolongado durante el cual adoptan una forma permanente» (Eichenbaum, 2003, pág. 43).

Eichenbaum afirma que las bases celulares-moleculares y los mecanismos cognitivos-conductuales contribuyen a perfeccionar el conocimiento sobre procesos de consolidación de la memoria (2003, pág. 41).

Con esta idea como base, resulta evidente que cualquier planteamiento filosófico serio sobre el estudio de la memoria humana no puede pasar por alto estas investigaciones experimentales, y depender solo de las intuiciones y del análisis introspectivo.

4.2. La memoria humana: entre filosofía, conductismo y cognitivismo

Parte de lo que hoy conocemos acerca de la estructura y el funcionamiento de la memoria humana se lo debemos, en buena medida, a las discusiones metodológicas en el siglo pasado entre el conductismo y el cognitivismo. Sin embargo, en lo que va del siglo XXI, aquellas acaloradas discusiones han quedado relegadas a un segundo plano por el avance imparable de la neurotecnología, la genética, la informática y la neurociencia cognitiva.

Según E. Kandel (2001, pág. 382), es la neurociencia cognitiva (disciplina que incluye estudios sobre la percepción, la acción, la emoción, el lenguaje y la memoria) quien cuenta con los recursos más completos y complejos para describir un enfoque funcional del cerebro, combinando principalmente la biología celular, la ciencia de los sistemas nerviosos, las diversas técnicas de imagen cerebral, la psicología cognitiva, la neurología conductual y la informática.

Si nos remontamos hasta mediados del siglo XIX, las principales investigaciones sobre la mente, los estados mentales y la memoria provenían de una filosofía que estaba ceñida al método introspectivo. El nuevo siglo trajo consigo un importante desarrollo de las ciencias experimentales, especialmente las ciencias naturales, las cuales eclipsaron de forma gradual los intentos filosóficos de una explicación sólida. En esta nueva etapa, las indagaciones sobre la memoria humana partirán de la psicología experimental. Kandel (2001) afirma:

Se ocupó fundamentalmente del estudio de la sensación: la secuencia de sucesos por los cuales un estímulo originaba una respuesta conductual. Para el cambio de siglo los intereses de los psicólogos se habían dirigido a las propias conductas: aprendizaje, memoria, atención, percepción y acción voluntaria. (pág. 382)

Una figura destacada fue el filósofo y psicólogo alemán Hermann Ebbinghaus (1850-1909), quien realizó los primeros trabajos experimentales de la

escuela conductista relacionados con el aprendizaje y el funcionamiento la memoria en humanos. A él pertenece el célebre experimento de la «sílaba sin sentido», una lista de letras («ket», «poc», «baf», por ejemplo) que buscaba eliminar cualquier sesgo estético o interés especial del agente al memorizarlas en un tiempo determinado (Eichenbaum, 2003, págs. 30-31). Con respecto a estos estudios, Baddeley (1999) comenta:

Decidió aplicar los métodos experimentales que se acababan de desarrollar para el estudio de la percepción a investigaciones más ambiciosas sobre los «procesos mentales superiores» y, más específicamente, a la memoria humana. Prefirió evitar la riqueza y complejidad de la memoria en la vida cotidiana, estudiando el aprendizaje y olvido de materiales artificiales por un único sujeto, él mismo, bajo condiciones de aprendizaje y recuerdo rígidamente controladas. Por medio de esta implacable simplificación, pudo demostrar características importantes de la memoria humana desconocidas por investigadores anteriores. Sin embargo, la verdadera importancia de su trabajo radica menos en sus descubrimientos que en su demostración de que el método experimental podía utilizarse para investigar algo tan complicado como el aprendizaje y la memoria humanos. Este asunto, que incluso las complicadas funciones mentales podían ser estudiadas dadas unas condiciones suficientemente simplificadas y controladas, ha denominado el estudio científico de la memoria humana desde entonces. (pág. 1)

Luego, cumpliendo con rigurosos métodos de laboratorio, vinieron los estudios de la memoria y la conducta animal realizados por Iván Pavlov, Edgard Thorndike (caja puzzle) o Small (laberinto de Small). Para otros conductistas, como es el caso de J. B. Watson (aprendizaje de ratas en el laberinto) y B. F. Skinner (conductas aprendidas), los procesos mentales no observables como la experiencia consciente quedarían fuera de cualquier consideración científica.

Para Kandel, la validez de los trabajos estará en la evaluación de la conducta observable, es decir, «la relación entre estímulos físicos específicos y las respuestas observables en animales intactos» (Kandel E. R., 2001, pág. 382). De ello podemos derivar que para la escuela conductista la vida mental se reduce

a la mera conducta observable, constituyendo su objetivo fundamental el «proporcionar una ciencia rigurosa de la memoria que concordara con los hallazgos de la neurofisiología de los reflejos condicionados, y trató de explicar toda conducta aprendida basándose en los elementos de la asociación y las respuestas condicionadas» (Eichenbaum, 2003, págs. 30-31). No obstante, podemos preguntarnos si esta reducción conductista es capaz de asegurarnos una investigación más sólida. Ello aún es confuso.

Las respuestas de George Miller, Noam Chomsky, Ulric Neisser o Herbert Simon, entre otros cognitivistas, no tardaron en llegar, generando amplias repercusiones dentro y fuera de la psicología.

Para el cognitivismo no solo era necesario estudiar la conducta observable sino, además, cómo se procesaba la información que va desde los receptores sensitivos hasta el cerebro, transformándose luego en memoria y acción. De ese modo se produce una «representación interna», es decir, una equiparación entre representaciones mentales y representaciones nerviosas conductuales. Kandel (2001) señala aquí una importante contradicción metodológica que tanto los cognitivistas como los conductistas tuvieron que aceptar, pues «los procesos mentales siguen siendo en gran parte inaccesibles al análisis experimental. Sin acceso directo a los sustratos nerviosos de las representaciones internas es difícil, si no imposible, diferenciar entre teorías rivales» (pág. 383).

La neurociencia cognitiva ofrece una respuesta alternativa para salvar la brecha abierta entre conductistas y cognitivistas, es decir, entre los estudios de la actividad mental y los avances técnicos y conceptuales a partir de una perspectiva integradora. Kandel (2001) señala cinco puntos fundamentales que describen esta integración:

Primero, en las décadas de los años 60 y 70, Ed Evarts y Vernon Mountcastle desarrollaron técnicas para estudiar la actividad de células aisladas de los cerebros de animales intactos con comportamientos activos. Pronto se emplearon estas técnicas para estudiar la correlación entre las acciones de las células individuales en condiciones de conducta controladas. También resultó posible estimular pequeños grupos de células y aumentar su actividad o lesionarlas de forma que disminuyesen su actividad (...). El

resultado es que sabemos que los mecanismos de la percepción son muy similares en el mono, y en otros animales más simples, y en el ser humano. Segundo, los estudios celulares en monos proporcionaron también la posibilidad de correlacionar patrones de activación de células individuales en regiones cerebrales específicas con procesos cognitivos superiores como son la atención y la toma de decisiones. Esto cambió la forma de estudiar la conducta, tanto en animales de experimentación como en seres humanos. A diferencia de los conductistas, ya no nos centramos solo en las propiedades estímulo respuesta de la conducta; por el contrario, dedicamos la atención al procesamiento de la información en el cerebro que se traduce en conducta. Tercero, el desarrollo de la neurociencia de los sistemas y la psicología cognitiva despertó un renovado interés por el análisis de la conducta de pacientes con lesiones cerebrales que interfirían en el funcionamiento mental (...). Los pacientes con lesiones en regiones específicas del cerebro muestran déficits cognitivos bastante específicos. Las consecuencias conductuales de las lesiones cerebrales nos dicen mucho sobre la función de áreas y vías nerviosas específicas. Los estudios de lesiones han demostrado que la cognición no es un proceso unitario, sino que existen varios sistemas cognitivos, cada uno de los cuales posee muchos módulos independientes de procesamiento de la información. Cuarto, las nuevas técnicas de imagen radiológica: tomografía de emisión de positrones (PET), imagen por resonancia magnética (RM), magnetoencefalografía y colorantes sensibles al voltaje, han permitido relacionar variaciones de la actividad de poblaciones enteras de neuronas con actos mentales específicos en seres humanos vivos. Finalmente, la informática ha hecho una contribución fundamental a las neurociencias cognitivas. Los ordenadores han permitido crear modelos de la actividad de grandes poblaciones de neuronas y empezar a poner sobre el papel las ideas respecto a componentes específicos del cerebro en conductas concretas. (págs. 383-384)

Según Eichenbaum (2003, pág. 101), fue William James, a principios del siglo XX (*The Principles of Psychology*), quien mejor entendió las diferencias entre conductistas y cognitivistas de la memoria. Sin embargo, el método experimental desarrollado por el conductismo terminó por imponerse, pues ofrecía a una mayor solidez al paradigma científico imperante en aquella época (por ejem-

plo, los experimentos de Pavlov). La hegemonía conductista comenzó a resentirse con las investigaciones de Yerkes (1876-1956), Edward Tolman (1886-1959) y Frederic Bartlett (1886-1969).

El problema fundamental que Bartlett veía en los conductistas eran las situaciones artificiales y simplificadas que estos recreaban en los laboratorios y que poco tenían que ver con el funcionamiento de la vida cotidiana (Ruiz-Vargas J. M., 1998, pág. 38). Si bien ambos enfoques —conductismo y cognitivismo— dominaron las principales teorías en el estudio de la memoria, los conductistas tuvieron una mayor acogida, ya que las investigaciones cognitivistas resultan «imposibles de comprobar» (Ruiz-Vargas J. M., Psicología de la memoria, 1998, pág. 41).

A mediados de 1960 se publica *Psicología cognitiva* de Ulric Neisser, una obra que describe cómo la gente adquiere, almacena, manipula y utiliza la información (Ruiz-Vargas J. M., 1998, pág. 43). Dicho enfoque estará centrado en la relación entre las «representaciones internas y el modo como son utilizadas en las actividades mentales». De esta manera, los estudios científicos de la memoria hallaron «un vasto y fértil marco de trabajo», pues «se hará hincapié en la interdependencia de la memoria y el resto de los procesos mentales» (Ruiz-Vargas J. M., 1998, pág. 43).

En el campo de la memoria, las primeras investigaciones de la psicología cognitiva estuvieron influenciadas por los ordenadores y la creación de nuevos conceptos y nuevos lenguajes «tales como “buffer de memoria”, “retroalimentación”, “codificación” y “recuperación”» (Baddeley, 1999, pág. 2).

La crítica elaborada por Neisser sobre el exceso de control en el laboratorio que impulsaba el conductismo no fue suficiente para erradicar de forma definitiva el paradigma de investigación. Si bien es «imposible captar en el laboratorio muchos de los más ricos e intrigantes aspectos de la memoria humana» (Baddeley, 1999, págs. 1-3), el conductismo de Ebbinghaus continuó dominando la escena de la psicología norteamericana. La principal preocupación de U. Neisser es la «validez ecológica», es decir, el mundo exterior frente al control del laboratorio.

E. Kandel destaca el nacimiento de la llamada «biología mental», una unión entre la filosofía del espíritu, la psicología conductista y la psicología cognitiva. Esta unificación dio paso a la psicología cognitiva moderna (2007, págs. 25-26) que, unida a la neurociencia, conforma lo que hoy conocemos como neurociencia cognitiva, una «rama de la ciencia que aportó a la moderna psicología cognitiva métodos biológicos para estudiar los procesos mentales», incorporando técnicas de imagen y biología molecular (Kandel E. R., 2007, pág. 26). A decir del propio Kandel (2007):

La nueva biología mental es, en potencia, más perturbadora aun, pues sugiere que no sólo el cuerpo, sino la mente y las moléculas específicas que intervienen en los procesos mentales superiores –la conciencia de sí y de los otros, del pasado y del futuro– evolucionaron a su vez desde la época de nuestros antepasados. Además, esta nueva biología postula que la conciencia es un proceso biológico que, a su debido tiempo, podrá explicarse en términos de vías de señalización moleculares utilizadas por poblaciones de células nerviosas que interactúan entre sí (...). Al haber unificado la psicología conductista y la cognitiva, la neurociencia y la biología molecular, esta nueva ciencia de la mente puede abordar cuestiones filosóficas con las que los pensadores más eminentes han lidiado durante milenios. ¿Cómo adquiere la mente el conocimiento sobre el mundo? ¿Qué proporción de ella se hereda? ¿Nos imponen las funciones mentales innatas una manera fija de experimentar el mundo? ¿Qué cambios físicos se producen en el cerebro cuando aprendemos y recordamos? ¿Cómo es que una experiencia que dura unos minutos se transforma en un recuerdo que dura toda la vida? Estos interrogantes ya no son terreno de especulaciones metafísicas sino fértiles áreas de investigación experimental. (págs. 27-28)

Las palabras de Kandel ilustran las cuestiones fundamentales que aborda la neurofilosofía contemporánea. El interrogante que se abre para otras disciplinas filosóficas, sean estas teóricas o prácticas, es saber hasta dónde pueden avanzar sus investigaciones al margen del caudal de informativo que nos ofrece, por ejemplo, la neurociencia social cognitiva o la neurobiología. Sin embargo, para nuestro itinerario de investigación estas disciplinas científicas constituyen una fuente de información imprescindible.

El siguiente apartado está dedicado a algunos conceptos que son clave en la neurociencia contemporánea. Además, constituyen la base para cualquier investigación sobre de la memoria humana.

4.3. Aclaraciones terminológicas: plasticidad cerebral, cerebro, encéfalo y neuronas

Algunas de las principales investigaciones en epigenética revelan que los cambios en el medioambiente, descrito este como el medio bio-físico y socio-cultural en el que se desarrolla la vida, influyen en la conducta y la ontogenia (proceso que describe el desarrollo de un organismo desde la fecundación del cigoto hasta la senescencia).

Sabemos que la evolución ha dotado al ser humano para absorber estos cambios y modificar su conducta a través de la adaptación al medio, lo cual es una ventaja para la supervivencia del individuo y la especie. Siguiendo este razonamiento, podemos decir que los seres humanos estamos «evolucionando».

El cerebro humano es un órgano biológico compuesto por cien mil millones de neuronas aproximadamente más sus sinapsis, y pesa alrededor de 1,4 kg. Siguiendo la descripción de Butcher, Pacheco y Tirado (2006):

La mayoría de las neuronas no se reproducen. Cuando se llega a la tercera edad, el cerebro ha perdido cerca de 100 g de peso. Miles de neuronas habrán desaparecido. Dentro de esa desventaja de las neuronas respecto a las demás células (de no poderse reproducir), hay una gran ventaja: su plasticidad, es decir, esa maravillosa capacidad de poder crecer y hacer múltiples conexiones con otras neuronas por medio de las dendritas y los axones, de manera que se pueden formar circuitos de neuronas haciendo posible el almacenamiento de información y habilidades. Las neuronas son células «informáticas» que registran, almacenan y transmiten información, y la plasticidad es inherente a ellas. (pág. 11)

Darwin entendió a la conducta, reflejo de la actividad del sistema nervioso, como un rasgo que se hereda a través de la adaptación. Un ejemplo de ello son las manifestaciones corporales de una reacción de pánico: «pelos erizados y pupilas dilatadas» (Bear, Connors, & Paradiso, 2008, pág. 11). Esta conexión es importante ya que tanto la conducta como la adaptación están mediadas por la «plasticidad cerebral», una ventaja evolutiva fundamental para el aprendizaje y la memoria (Butcher López, Pachecho Railey, & Tirado Medina, 2006, págs. 36-38) ligada a la experiencia del individuo:

La interacción del individuo con su medio modifica los procesos metabólicos a nivel de neurotransmisores y a nivel de la estructura de los axones y dendritas. Los cambios cerebrales posteriores al nacimiento son determinados en gran medida por la estimulación del ambiente. (Butcher López, Pachecho Railey, & Tirado Medina, 2006, pág. 37)

Por su parte Kandel (2001), apoyándose en la «hipótesis de la plasticidad» de J. Konorski (1948), realiza la siguiente afirmación:

La aplicación de un estímulo produce cambios de dos clases en el sistema nervioso (...). A la primera propiedad, por la cual las células nerviosas *reaccionan* con el impulso que llega (...), la llamamos *excitabilidad* (...). Los cambios que se producen (...) a causa de esta propiedad, los denominaremos *alteraciones debidas a la excitabilidad*. La segunda propiedad, merced a la cual se producen ciertas transformaciones funcionales permanentes en determinados sistemas de neuronas como resultado de estímulos apropiados o combinaciones de los mismos, la denominaremos *plasticidad*, y *alteraciones plásticas* a las modificaciones correspondientes. (págs. 34-35)

La evolución de la conducta desde antepasados comunes encuentra una fundamentación sólida basada en los modelos de investigación con animales, por ejemplo, en las ratas de laboratorio. M. Bear (2008), siguiendo la línea del desarrollo evolutivo propuesto por Darwin, sostiene:

La idea de que el sistema nervioso de diferentes especies evolucionó desde antepasados comunes y tiene mecanismos comunes es la razón

por la que se relacionan los resultados de los experimentos en animales con los de los experimentos en seres humanos. Así pues, por ejemplo, muchos de los detalles de la conducción del impulso a través de fibras nerviosas se estudiaron primero en el calamar, pero ahora se conoce que pueden aplicarse también a seres humanos. La mayoría de los neurocientíficos actuales utiliza *modelos animales* para examinar el proceso que quieren entender en los seres humanos. Por ejemplo, las ratas muestran claros signos de adicción a la cocaína, si se les da la oportunidad de autoadministrarse cocaína de forma repetida. Por tanto, las ratas son un modelo de animal válido para la investigación sobre el modo en que las drogas psicoactivas ejercen sus efectos sobre el sistema nervioso. (pág. 11)

El concepto de «plasticidad cerebral» marcó el inicio de una nueva senda para entender el funcionamiento del sistema nervioso en relación con el medioambiente y la sociabilidad, pues las evidencias aportadas desde la neurobiología y la neurofisiología contribuyeron a dejar a un lado las viejas concepciones que consideraban al cerebro como un órgano estático e inmutable.

En la actualidad, las fronteras entre lo biológico y lo social comienzan a no ser tan claras como se pensaban hasta hace tan solo unas décadas atrás. No obstante, para comprender la complejidad que encierra el concepto de plasticidad cerebral y su relación con la sociabilidad es necesario que nos detengamos un momento en el análisis de la estructura del cerebro y sus unidades celulares: las neuronas.

Estas células nerviosas componen una sofisticada red celular que recibe y procesa información en un determinado tiempo, soportando, además, la estructura de la memoria. Christof Koch (2006) va un poco más allá y relaciona esta red con la estructura de la conciencia (biológica):

Las neuronas son los átomos de la percepción, la memoria, el pensamiento y la acción, y las conexiones sinápticas entre ellas determinan y orientan el modo en que las células individuales se ensamblan transitoriamente en coaliciones de mayor tamaño que generan percepción. Por

tanto, cualquier teoría que explique las bases neuronales de la conciencia debe describir interacciones específicas entre células nerviosas en una escala temporal de milisegundos. (pág. 42)

Los descubrimientos de Santiago Ramón y Cajal (1852-1934) determinaron que las neuritas —dendritas y axón— de «diferentes neuronas no se continúan con las otras y se deben *comunicar por contacto, no por continuidad*. Esta idea de que la neurona se adhería a la teoría celular se vino a conocer como la doctrina neuronal» (Bear, Connors, & Paradiso, 2008, pág. 27). La historia de la doctrina neuronal propuesta por Ramón y Cajal recién pudo confirmarse con la aparición del microscopio electrónico en la década de 1950.

Nuestro sistema nervioso está compuesto de neuronas y células gliales, «dos categorías amplias, dentro de las cuales hay muchos tipos de células que difieren en estructura, química y función. No obstante, la distinción entre neurona y glía es importante» (Bear, Connors, & Paradiso, 2008, pág. 24). Según explica Bear (2008):

Las células neuronales perciben los cambios del entorno, comunican estos cambios a otras neuronas y ordenan las correspondientes respuestas corporales a estas sensaciones. Se piensa que la glía o células gliales, contribuyen a la función cerebral sobre todo aislando, apoyando y alimentando a las neuronas vecinas. (pág. 24)

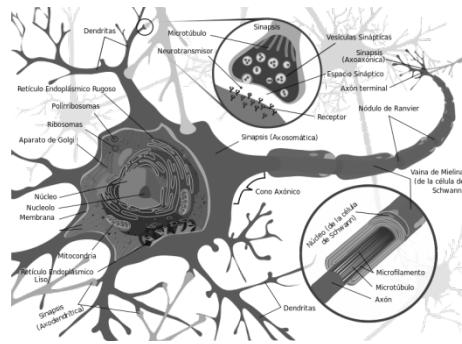


Figura 1. La estructura básica de una neurona. La ilustración fue extraída y modificada de Wikimedia Commons (https://commons.wikimedia.org/wiki/File:Complete_neuron_cell_diagram_es.svg).

La neurona está compuesta por el soma o cuerpo celular, las dendritas, el axón y las terminales presinápticas (Kandel E. R., 2001, pág. 21). El soma que mide unos 20 µm de diámetro y las membranas envuelven unas estructuras, las organelas. El soma de las neuronas contiene las mismas organelas que existen en todas las células animales: núcleo, retículo endoplasmático liso, retículo endoplasmático rugoso, aparato de Golgi y mitocondrias (Bear, Connors, & Paradiso, 2008, pág. 28). Todo el interior de la membrana celular, incluidas las organelas con excepción del núcleo, se conoce como citoplasma. En el centro de la célula se encuentra una estructura esférica, el núcleo, que mide aproximadamente entre 5-10 µm de anchura y contiene a los cromosomas, los portadores de todo nuestro material genético o ADN (ácido desoxirribonucleico). Según comenta Bear (2008):

El ADN de cada una de nuestras neuronas es idéntico y es el mismo que el ADN de las células de nuestro hígado y nuestro riñón. Lo que distingue a una neurona de una célula hepática son los segmentos concretos de ADN que se utilizan para configurar la célula. Estos segmentos de ADN son los llamados genes. (pág. 30)

El axón es una estructura «presente solo en las neuronas que está altamente especializada en la transferencia de información a distancia en el sistema nervioso» (Bear, Connors, & Paradiso, 2008, pág. 38). Las dendritas, cubiertas de miles de sinapsis, «funcionan a modo de antenas de las neuronas» (Bear,

Connors, & Paradiso, 2008, pág. 41), esto es lo que permite a las neuronas comunicarse entre sí, transformando las señales eléctricas en señales químicas (neurotransmisores).

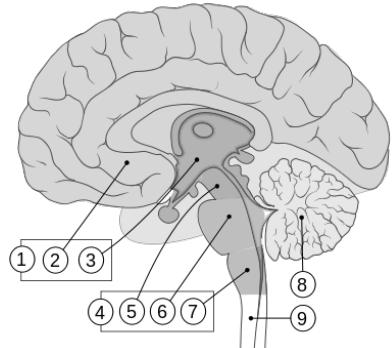


Figura 2. Corte sagital del encéfalo humano. (1) cerebro anterior; (2) telencéfalo; (3) diencéfalo; (4) tronco del encéfalo; (5) mesencéfalo; (6) protuberancia; (7) bulbo raquídeo; (8) cerebelo; y (9) médula. Extraído y modificado de Wikimedia Commons, trabajo de Jordi March i Nogué, CC BY-SA 3.0 (<<https://commons.wikimedia.org/w/index.php?curid=8405926>>).

Kandel disecciona el sistema nervioso central (SNC) en siete partes principales: (1) la médula espinal que se divide en las regiones cervical, dorsal, lumbar y sacra; (2) el bulbo raquídeo; (3) la protuberancia; (4) el cerebelo; (5) el mesencéfalo; (6) el diencéfalo; (7) los hemisferios cerebrales que son «una capa externa muy plegada —la corteza cerebral— y tres estructuras situadas en la profundidad: los ganglios basales, el hipocampo y los núcleos amigdalinos (...). La corteza cerebral se divide en cuatro lóbulos: frontal, parietal, temporal y occipital» (Kandel E. R., 2001, pág. 8). En la Figura 2 se muestran las localizaciones precisas de algunas de las partes que componen el sistema nervioso central (SNC).

Describir la estructura y fisiología del encéfalo remite a la localización de las funciones cognitivas en la corteza cerebral. En tal sentido, Kandel (2001) afirma:

Las funciones cerebrales responsables de nuestras capacidades cognitivas ocurren fundamentalmente en la *corteza cerebral*, la sustancia gris plegada que cubre los hemisferios cerebrales. En cada uno de los dos hemisferios cerebrales la corteza que los reviste está dividida en cuatro

lóbulos diferenciados anatómicamente: *frontal*, *parietal*, *temporal* y *occipital*, denominados así por los huesos del cráneo que los alojan. El lóbulo frontal se ocupa en gran medida de la planificación de acciones futuras y del control del movimiento; el lóbulo parietal de la sensación somática, con la formación del esquema corporal, y de la relación del propio esquema corporal con el espacio extrapersonal; el lóbulo occipital con la visión; el lóbulo temporal con la audición; y a través de sus estructuras profundas —el hipocampo y los núcleos amigdalinos— con aspectos del aprendizaje, la memoria y las emociones. (pág. 9)

La corteza cerebral es la «responsable de gran parte del planteamiento y la ejecución de acciones en la vida diaria» (Amaral, 2001, pág. 324). Se cree que su forma plegada constituida por cisuras, surcos y circunvoluciones es el resultado de la evolución para dar cabida a una mayor cantidad de neuronas, siendo uno de los «determinantes esenciales de su capacidad de procesamiento de información» (Amaral, 2001, pág. 324).

La corteza cerebral humana es mayor a la de cualquier otro primate superior y posee un espesor de entre 2 a 4 milímetros. Su «arquitectura» es compleja y muy parecida visualmente en toda su extensión, lo cual requiere un análisis reduccionista de cada una sus áreas para llegar a comprender el funcionamiento del conjunto (Koch, 2006, pág. 137). No obstante, llegar a unas descripciones anatómicas tan precisas solo ha sido posible gracias al desarrollo tecnológico del microscopio electrónico y, más recientemente, a la tecnología de imagen de alta resolución. Para Koch (2006):

La exploración exhaustiva de la corteza tuvo que aguardar la llegada de los microscopios modernos, las tinciones químicas y los colorantes que se ligaban a los componentes celulares, como la capa de mielina que envuelve los axones o al ácido ribonucleico del soma. Con esta creciente capacidad para seleccionar constituyentes moleculares específicos de neuronas ha florecido el estudio de la arquitectura cerebral basado en variaciones locales sutiles pero acusadas. (pág. 137)

De la cita anterior extraigo dos cosas que considero relevantes en las investigaciones neurocientíficas. Por un lado, que el desarrollo tecnológico es intrínseco a la evolución de las neurociencias, por otro, y esto quizás se deba a la juventud de las disciplinas neurocientíficas, podemos observar que estas han importado de otros campos científicos términos y conceptos no del todo adecuados para explicar mecanismos, funciones o simples descripciones, tales como «arquitectura de la corteza cerebral», «regiones geográficas de la corteza cerebral» o «mapas cerebrales».

También se da el caso inverso y muchas disciplinas que beben de las aguas de las neurociencias confunden y mal difunden conceptos básicos tan distintos como, por ejemplo, «cerebro» y «encéfalo».

Por esta razón es necesario revisar y discutir una eventual corrección terminológica y su adecuación al terreno experimental. La discusión recogida en el libro *La naturaleza de la conciencia: cerebro, mente y lenguaje* (Bennett, Dennett, Hacker, & Searle, 2008) es un excelente trabajo que hace hincapié en esta cuestión, en donde filósofos y neurocientíficos reflexionan acerca de la pertinencia o no del uso de determinados conceptos por parte de los neurocientíficos. En la réplica a los críticos, representados por Dennett y Searle, los aludidos Bennett y Hacker responden (2008):

El objetivo de la investigación filosófica sobre los fundamentos conceptuales de la neurociencia es desvelar y esclarecer verdades conceptuales, verdades que las descripciones plausibles de los descubrimientos y las teorías neurocientíficas y sin las cuales no tienen sentido. Si esa investigación se lleva a cabo correctamente, esclarecerá los experimentos neurocientíficos y su descripción, además de las inferencias que puedan hacerse a partir de ellos. (pág. 160)

Avanzando con la anatomía de la corteza cerebral, observamos que está dividida en cuatro lóbulos principales (Figura 3): (1) frontal, (2) occipital, (3) parietal, y (4) temporal, llamados así por los huesos del cráneo que los albergan. Además, existen otras dos regiones, la corteza del cíngulo y la corteza insular (Amaral, 2001, pág. 325). Continuando con la descripción de Amaral (2001):

Los cuatro lóbulos se definen de forma notoria por llamativos surcos de la corteza cuya posición en los diferentes cerebros humanos es relativamente constante. Una de las muescas más acusadas de la corteza cerebral, la cisura de Silvio o surco lateral, separa el lóbulo temporal de los lóbulos frontal y parietal. La corteza de la ínsula forma el límite medial del surco lateral. Otra muesca importante, la cisura central o de Rolando, discurre por la parte medial y lateral sobre la superficie dorsal del hemisferio y separa los lóbulos frontal y parietal. (pág. 325)

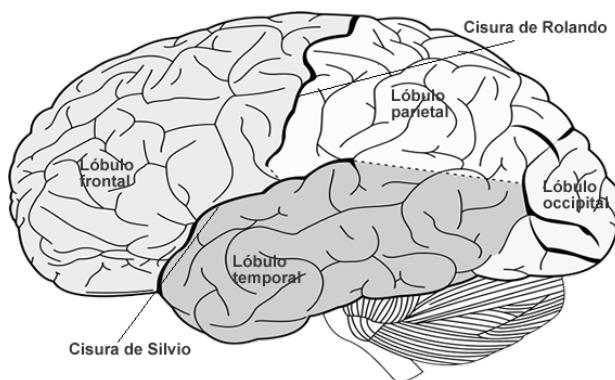


Figura 3. Lóbulos de la corteza cerebral. Imagen extraída y modificada de Wikimedia Commons (<<https://commons.wikimedia.org/w/index.php?curid=1676555>>).

C. Koch (2006) describe a la neocorteza situada detrás del surco central como la encargada de la percepción y el input sensorial, y los lóbulos frontales serán los responsables de la acción:

Mientras que, dicho en términos aproximados, toda la neocorteza de detrás del surco central se ocupa de la percepción y el input sensorial, los lóbulos frontales —la gran extensión de la neocorteza situada en posición anterior respecto al surco central— se encargan de la acción. Las cortezas motoras, premotora, prefrontal y cingulada anterior pertenecen a los lóbulos frontales. Su función es guiar, controlar y ejecutar outputs motores, como movimientos esqueléticos u oculares, y la expresión de las emociones, el habla o los estados mentales internos (como los procesos del pensamiento no consciente). A medida que los organismos evolucionan, aumentan la complejidad de sus acciones y se amplían sus objetivos en el espacio y en el tiempo, llegando a depender menos del impulso instintivo y más de la experiencia previa, la intuición y el razonamiento. Esto precisa planificación, toma de decisiones en entornos inciertos y control cognitivo,

recuerdo del almacenamiento de información en tiempo real, y sensación de autoría. En estas funciones ejecutivas, de alto nivel, nos damos más cuenta que nunca del valor de la corteza *prefrontal* (...). La corteza prefrontal (CPF), la parte más anterior de la corteza cerebral, se define como las zonas corticales receptoras de axones procedentes de neuronas de proyección del núcleo talámico mediodorsal (...). La CPF es la única región neocortical que habla directamente con el hipotálamo, responsable de la liberación de hormonas. Por tanto, las regiones prefrontales se hallan en una posición muy favorable para integrar información derivada de todas las modalidades sensoriales y motoras. Otro de sus papeles es el almacenamiento a corto plazo, en tiempo real, de información importante para el organismo. (pág. 149)

Asimismo, se observa una relación entre los lóbulos frontales y los ganglios basales, pues son regiones que «median en movimientos intencionados, secuencias de acciones o pensamientos motores, y aprendizaje motor. En los vertebrados sin corteza o con la corteza poco desarrollada, los ganglios basales son los centros prosencefálicos más importantes» (Koch, 2006, pág. 150).

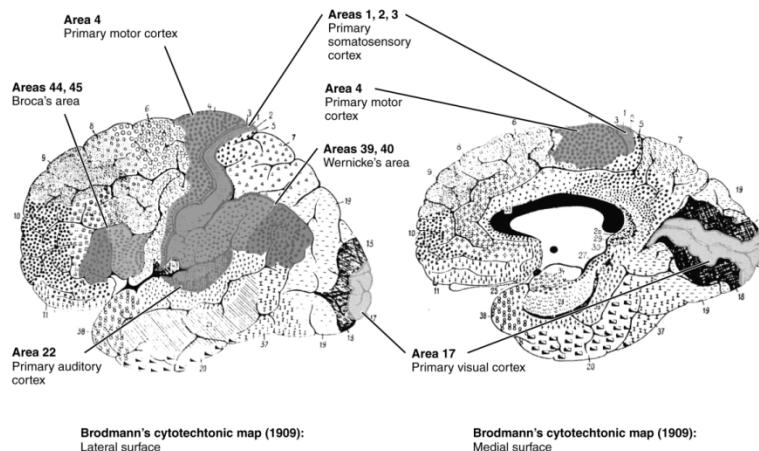


Figura 4. Áreas de Brodmann. Imagen extraída de Wikimedia Commons By OpenStax College - Anatomy & Physiology, Connexions (<<http://cnx.org/content/col11496/1.6/>>), Jun 19, 2013, CC BY 3.0, (<<https://commons.wikimedia.org/w/index.php?curid=30147951>>).

Otras áreas de la corteza cerebral (Figura 1.4) focalizan su actividad en el procesamiento de información sensitiva o emisión de órdenes motoras, diferenciándose unas de otras por sus funciones (Amaral, 2001, págs. 325-326).

Existen, además, otras áreas del encéfalo que no están restringidas a una actividad motora o sensitiva, ocupándose de la regulación. Con respecto a los sistemas reguladores, Amaral (2001) señala:

Son componentes esenciales de los circuitos nerviosos que subyacen a las conductas complejas. Las conductas complejas se orientan a menudo a satisfacer una necesidad primaria como el hambre, la sed o el sueño. Así, los sistemas sensitivos y reguladores del hipotálamo determinan los niveles sanguíneos de glucosa. En cuanto la glucemia desciende por debajo de un nivel crítico, sentimos hambre. Para satisfacer el hambre, primero han de entrar en juego procesos perceptivos y reguladores. Cuando un predador inspecciona el entorno buscando indicios de la presa, como imágenes, sonidos u olores, los sistemas reguladores centran el aparato sensorial sobre estímulos de importancia para la alimentación (...). Una de las cuestiones más interesantes de la neurociencia, fundamental para nuestra comprensión del aprendizaje y la memoria, es cómo interaccionan los sistemas reguladores que se ocupan de la recompensa, la atención, y la motivación de los sistemas sensitivos y motores. (pág. 334)

Como vemos, el sistema nervioso está preparado para recibir información del medioambiente, procesarla y emitir una respuesta conductual adecuada. En las investigaciones podemos seguir patrones bastante claros porque, más allá de la complejidad fisiológica, «la estructura del sistema nervioso es similar en los distintos individuos de una especie» (Amaral, 2001, pág. 335).

La memoria, junto con la atención y el aprendizaje, conforma un conjunto de sistemas de procesamiento y recuperación de información que es clave para la vida individual y social, además de constituir «los mecanismos más importantes a través de los cuales el medio altera la conducta» (Kandel, Kupfermann, & Iversen, 2001, pág. 1227).

4.4. La memoria y sus sistemas

La memoria no sólo es esencial para la continuidad de la identidad sino para la transmisión de la cultura, la evolución y la continuidad de las sociedades a lo largo de las centurias. Aunque el tamaño y la estructura del cerebro humano no se han modificado desde la aparición del *Homo sapiens* en África oriental hace unos 150.000 años, la capacidad de aprendizaje de los individuos y su memoria histórica se han incrementado a lo largo de los siglos en virtud del conocimiento compartido, es decir, mediante la transmisión de la cultura. La evolución cultural, modo de adaptación que no es biológico, obra en paralelo con la evolución biológica como medio de transmisión del conocimiento del pasado y como comportamiento adaptativo a lo largo de las generaciones. Desde la antigüedad hasta nuestros días, todas las hazañas humanas fueron producto de una memoria compartida que se acumuló durante siglos, fuera mediante registros escritos o a través de una tradición oral conservada con esmero. Así como la memoria compartida enriquece nuestra vida en tanto individuos, la pérdida de la memoria destruye la continuidad del yo, corta los lazos con el pasado y con los otros, y puede afligir al niño o al adulto maduro. (Kandel E. R., 2007, págs. 29-30)

Una de las hipótesis que anima este trabajo sostiene que la memoria humana, «proceso de codificación, almacenamiento y recuperación del conocimiento sobre el mundo» (Kandel, Kupfermann, & Iversen, 2001, pág. 1127), constituye una compleja estructura bio-cultural necesaria para la construcción del *ethos* social o sociabilidad, o lo que es lo mismo, sin esta estructura seríamos incapaces de elaborar, almacenar y recuperar información social necesaria para el desarrollo cotidiano del individuo y de la especie. Dicha estructura me hace pensar en la necesidad de la existencia de un proceso de información que fluye en dos direcciones, es decir, información que nace de los individuos y fluye al medio social e información que nace en el medio social y fluye hacia los individuos. Entiendo que con este proceso recíproco de información comienza la construcción de la sociabilidad humana (*ethos* social).

Los individuos almacenamos en nuestra memoria toda la información colectiva generada en el *ethos* social. Una vez procesada esta información, la volcamos nuevamente en la sociedad. La recuperación de esta información nos permite recordar el pasado, comprender el presente y, sobre todo, configurar nuestro futuro.

Antes de entrar de lleno en el tema central de la memoria autobiográfica, debemos presentar una de las clasificaciones, quizás la más básica, sobre los sistemas de memoria que recoge la psicología de la memoria y la neurociencia cognitiva con el propósito de llevar un hilo conductor a modo de guía.

Las principales investigaciones coinciden en la importancia de la función adaptativa de dichos sistemas, tal y como explica Ruiz-Vargas (1994):

Los diferentes sistemas han aparecido en etapas diferentes de la evolución de las especies así como que su emergencia a nivel individual también sigue diferentes etapas en el desarrollo del organismo; es decir, los diferentes sistemas de memoria siguen una secuencia filogenética y ontogenética en su aparición. (pág. 165)

En una primera aproximación, podemos deducir que la memoria humana es el resultado del desarrollo de una capacidad evolutiva del cerebro que nos permite alcanzar aprendizajes complejos, retener y recuperar información y experiencias. Dicha capacidad está compuesta de tres fenómenos esenciales: (1) registro o codificación (estrategias cognitivas que utilizamos para el procesamiento de la información que llega al cerebro); (2) almacenamiento de la información (retención temporal de la información a corto y largo plazo); y (3) evocación o recuperación de la información almacenada (Téllez López, 2006, págs. 105-107).

Por medio de la memoria recuperamos información que obtenemos mediante los sentidos. Con respecto a ello, Baddeley (1999) afirma:

No es un sistema sino muchos. Los sistemas varían en duración de almacenamiento desde fracciones de segundos hasta una vida entre, y en capacidad de almacenamiento desde minutos almacenes momentáneos al

sistema de memoria a largo plazo, que parece exceder largamente la capacidad y flexibilidad al mayor ordenador disponible. (pág. 4)

Para este autor, «una forma de obtener una visión de conjunto de la memoria humana es seguir la pista al modo en que los estímulos visuales y auditivos son procesados y recordados» (Baddeley, 1999, pág. 11).

La «memoria» como concepto genérico de varios sistemas cerebrales, conductuales y cognitivos está rodeado de ciertas controversias dentro del ámbito de la psicología cognitiva (Ruiz-Vargas J. M., 1998, pág. 43). Siguiendo el comentario de Ruiz-Vargas (1998):

Muchos psicólogos cognitivos no comparten, sin embargo, la necesidad de postular formas múltiples de memoria y mantienen una actitud en favor de la idea de que la memoria es una entidad unitaria. En esencia, asumen que las aparentes diferencias que manifiesta el funcionamiento de la memoria no son más que combinaciones y concatenaciones diferentes de determinados procesos cognitivos que operan dentro de los límites de lo que podría llamarse «una gran memoria». (pág. 58)

A pesar de las diferencias a la hora de encontrar una definición adecuada de lo que sea la memoria humana y su funcionamiento, la mayor parte de las investigaciones se decantan por un concepto de memoria como «sistemas múltiples». En este punto se abre otra controversia, pues E. Tulving es partidario de una definición de «corte *estructural*», en cambio Sherry y Schacter optan una definición «*funcional*» (Ruiz-Vargas J. M., 1998, págs. 62-63).

Dejando apartadas las mencionadas controversias, los principales estudios describen dos grandes sistemas de memoria «separables aunque íntimamente relacionados» (Ruiz-Vargas J. M., 1998, pág. 68): memoria a corto plazo (MCP) y memoria a largo plazo (MLP):

Donald Hebb, en su libro *La organización de la conducta* (1949), sugirió la necesidad de asumir dos sistemas de memoria separables: uno a corto plazo, cuya base fisiológica estaría en la actividad de los circuitos neuro-

nales reverberatorios, y otro a largo plazo, que implicaría un cambio estructural permanente en el sistema nervioso. Precisamente porque ese cambio estructural requiere un tiempo considerable para producirse, planteó la necesidad de un sistema de memoria a corto plazo responsable de mantener la información mientras tanto. La propuesta de Hebb no encontró apenas eco hasta finales de 1950 cuando, tras el desarrollo de diferentes paradigmas experimentales de memoria a corto plazo, se reavivó la cuestión que desembocaría, en la década de 1960, en una polémica sobre la naturaleza unitaria o dicotómica de la memoria. Fue, concretamente, a mediados de 1960, cuando empezó a acumularse evidencia a favor de la existencia de dos sistemas de memoria: uno a corto y otro a largo plazo. (Ruiz-Vargas J. M., Psicología de la memoria, 1998, págs. 63-64)

En 2005, la investigación de Tulving y Rosenbaum denominada «paciente K.C.» demostró la hipótesis de que la memoria no es un bloque único y unitario, sino que está compuesta de diversos sistemas.

En tal sentido, E. Tulving diseñó un esquema jerárquico general con cinco tipos de memoria: *procedimental, perceptiva, operativa, semántica y episódica* (Eustache & Desgranges, 2010, pág. 48). Para algunos autores, en la base de la clasificación de la memoria se encuentran dos fuentes que son fundamentales: (a) la filosofía (de Ryle y Bergson), y (b) los trabajos experimentales con adultos, niños, pacientes amnésicos y animales (Ruiz-Vargas J. M., 1998, pág. 68).

A continuación, expondré una breve descripción de cada uno de estos sistemas con sus características principales.

La *memoria procedimental*, también conocida como memoria implícita (diferente de la explícita o declarativa) por su forma de recuperación no consciente, se define como un «sistema de acción» que expresa sus operaciones a través de la conducta y automatismos (García García, 2018, pág. 23) a diferencia de los otros cuatro sistemas que expresan sus operaciones en la cognición o el pensamiento (Ruiz-Vargas J. M., 1994, pág. 156). En tal sentido, Tulving (1985) señala que:

Este sistema permite que los organismos retengan conexiones aprendidas entre estímulos y respuestas, incluyendo aquellos que involucran patrones de estímulo complejos y las cadenas de respuesta, y responder adaptadamente al medio ambiente. (pág. 387)

Eustache & Desgranges (2010) destacan su dependencia de la memoria episódica, sin embargo, al igual que otros investigadores, hacen hincapié en desvincularla de la acción cognitiva que poseen otros sistemas de memoria:

La memoria procedural concierne al aprendizaje y conservación de las competencias: aprender a esquiar, a montar en bicicleta, a tocar un instrumento o incluso a caminar. La memoria procedural es independiente de los otros tipos de memoria (...). La memoria procedural es una parte disociada de la memoria episódica. Podemos conseguir que un amnésico vaya en bicicleta: éste habrá olvidado todo lo acaecido durante los días que aprendió a montar en el vehículo, pero habrá asimilado, mediante su memoria procedural, los gestos necesarios para pedalear y mantenerse en equilibrio. (pág. 50)

Por lo tanto, la memoria procedural es un «sistema de acción» (Ruiz-Vargas J. M., 1994, pág. 156) que se manifiesta «en las habilidades o destrezas perceptivas, motoras, cognitivas adquiridas y solo podemos acceder a ella a través de la acción. Es imposible transmitirla de forma verbal. Su contenido se refiere a “saber cómo”» (Ruiz-Vargas J. M., 1998, págs. 71-72). Según Téllez, este sistema «se refiere a los efectos no conscientes de la experiencia previa en nuestra conducta, cogniciones y emociones» (2006, pág. 113).

Al tener una recuperación implícita, «su uso es automático y su formación y evocación no requieren participación consciente. No es necesario expresar este conocimiento mediante el uso de palabras» (Fernández Ruiz & López García, 1998, págs. 18-25).

La *memoria perceptiva*, también conocida como *Sistema de Representación Perceptual* (priming perceptual de identificación de objetos) (Ruiz-Vargas J. M., 1994, pág. 157), es «el recuerdo de un precepto, incluso antes de que adquiera significado» (Eustache & Desgranges, 2010, pág. 49). Este sistema es

una «forma no consciente de aprendizaje, es decir, es presemántico, pues sus operaciones son independientes de la memoria procedimental, semántica y episódica» (Ruiz-Vargas J. M., 1994, págs. 157-158).

La *memoria a corto plazo*, también llamada *operativa* o *memoria primaria*, es un sistema que «funciona sin descanso en nuestra vida consciente. Es la que mantiene presente las informaciones que necesitamos en tiempo real para hablar, imaginar, reflexionar o calcular» (Eustache & Desgranges, 2010, pág. 49). En tal sentido, Ruiz-Vargas (1994) afirma:

La memoria a corto plazo (MCP) registra y retiene información perceptual y conceptual en un estado altamente accesible por un período corto de tiempo tras el input, lo que permite al individuo tener una impresión persistente del ambiente presente (o de sus pensamientos actuales). (...) La MCP es fácilmente disociable de la MLP (episódica y semántica) en el laboratorio. (pág. 158)

Los sistemas de *memoria semántica* y *memoria episódica*, si bien diferentes, poseen una estrecha relación, ya que forman parte de la memoria *declarativa* o *explícita*, «un tipo de memoria de hechos y eventos que se adquiere y se evoca conscientemente y mediatizada por el uso del lenguaje» (Téllez López, 2006, pág. 112).

A través de la memoria semántica «conocemos el mundo». Ella «organiza la información en hechos, conceptos y vocabulario (Ruiz-Vargas J. M., 1998, pág. 71), es decir, nos permite la representación subjetiva de nuestro entorno y del mundo a través de la construcción de modelos mentales (Téllez López, 2006, págs. 112-113).

Sin embargo, y a diferencia de la memoria episódica, la memoria semántica no está estructurada en parámetros espacio-temporales (Ruiz-Vargas J. M., 1998, pág. 71). Según las descripciones de Eustache & Desgranges (2010):

La memoria semántica nos es familiar. Cuando aprendemos que Moscú es la capital de Rusia, no memorizamos un precepto, sino un sentido, un

conocimiento. Este tipo de memorización reposa sobre la memoria semántica. Todos los conocimientos generales que vamos acumulando, incluidos los relativos a nosotros mismos, se deben a este tipo de memoria, que no implica acordarse de un acontecimiento preciso, ni tampoco de un precepto. (pág. 49)

El sistema de memoria episódica nos permite recordar de forma consciente parte de nuestras experiencias pasadas (Ruiz-Vargas J. M., 1994, págs. 158-159), «instantes únicos, acotados en el tiempo y el espacio» (Eustache & Desgranges, 2010, pág. 49). Remarco aquí la palabra «consciente», una diferencia esencial que sirve para distinguirla de otros sistemas de memoria. Según Téllez (2006), «esta memoria posee nuestra historia personal, nuestra autobiografía, y da continuidad a nuestras acciones» (pág. 113).

Los sistemas mencionados componen en su totalidad aquello que habitualmente conocemos como «la memoria». La complejidad para entender su funcionamiento y estructura hacen de este concepto uno de los retos más grandes que, tanto la ciencia como la filosofía, tienen por delante en este siglo y, posiblemente, el siguiente. Entiendo que para la filosofía el trabajo consistirá en despejar el camino conceptual y encontrar una metodología rigurosa de investigación.

El siguiente capítulo está dedicado a la memoria autobiográfica, uno de los pilares fundamentales de nuestro sistema cognitivo. En la estructura general de esta tesis, la memoria autobiográfica constituye uno de sus ejes principales, pues de ella depende en gran parte la construcción de nuestra identidad subjetiva y el reconocimiento colectivo (de los «otros» y del «nosotros») que podemos adquirir como seres sociales.

Me gustaría destacar que la construcción de la sociabilidad se apoya en una estructura cognitiva compuesta por los sistemas de memoria y el aprendizaje (percepción y atención). Esta transferencia de información es necesaria porque «la memoria implica la adquisición previa de información (aprendizaje) y el aprendizaje requiere la retención de esa información (memoria)» (Téllez López, 2006, pág. 106).

Capítulo 5. La memoria autobiográfica

Freud estaba fascinado por los lapsus y errores de la memoria que tenían en el curso de la vida cotidiana y cómo se relacionaban con la emoción, sobre todo con la emoción inconsciente. Pero también se vio obligado a considerar las distorsiones mucho más burdas de la memoria que exhibían algunos de sus pacientes, sobre todo cuando le relataban que de niños habían sido seducidos o habían sufrido abusos sexuales. Al principio se tomaba estos relatos de manera literal, pero con el tiempo, cuando se topó con varios casos que parecían faltos de pruebas o de plausibilidad, comenzó a preguntarse si esos recuerdos habían sido distorsionados por la fantasía, y si, de hecho, podrían ser fabulaciones totales, construidas de manera inconsciente, pero convincentes hasta tal punto que los pacientes creyeran en ellas a pie juntillas. Las historias que relataban los pacientes y que se contaban a sí mismos, aun cuando fueran falsas, podrían tener un efecto muy poderoso en sus vidas, y a Freud le parecía que su realidad psicológica podía ser la misma ya fueran una experiencia real o una fantasía. (Sacks, 2019, págs. 111-112)

En la actualidad, algunos de los estudios más relevantes sobre la memoria autobiográfica (MA) provienen de la psicología experimental y las neurociencias. Sin embargo, como queda evidenciado en diferentes investigaciones y publicaciones, las mencionadas disciplinas son incapaces de agotar un campo tan vasto y complejo en donde el objeto de estudio muestra una alta resistencia a los límites marcados por los experimentos en el laboratorio. Es por ello que las nuevas perspectivas en las investigaciones están combinando los paradigmas conductistas y cognitivistas con las neurociencias, lo que pone en valor el trabajo interdisciplinar entre estos diversos saberes.

Si bien las investigaciones interdisciplinarias en el campo de la MA están siendo fructíferas, en su camino podemos vislumbrar dos problemas de difícil solución, sobre todo porque su tratamiento excede los límites de las disciplinas mencionadas: (a) la metodología de investigación de la MA y (b) el acuerdo —o

mejor dicho desacuerdo— científico —y quizás también filosófico— en relación a los conceptos fundamentales utilizados en dichas investigaciones.

Entiendo que el mayor inconveniente para resolver estos escollos viene dado por la propia naturaleza experimental de las disciplinas en cuestión, muy alejada de cualquier discusión conceptual. Teniendo en cuenta esta problemática, uno de los ejes principales de esta tesis defiende la incorporación de la neurofilosofía a este conjunto de saberes, una disciplina teórica, pero cercana a las ciencias experimentales, como la encargada del desarrollo de un marco crítico-epistemológico, el cual servirá para poner a prueba tanto la fortaleza de la metodología de investigación de la MA como así también la revisión crítica y unificación de conceptos básicos trabajados por la psicología y las neurociencias.

La importancia de la memoria autobiográfica (MA) reside en la capacidad de recordar las propias experiencias (Ruiz-Vargas, 2004, pág. 1). No obstante, cabe señalar que el éxito del desarrollo evolutivo de este sistema también se halla en la capacidad para olvidar sucesos dolorosos o experiencias traumáticas que ponen en riesgo la propia supervivencia del individuo, son los llamados «procesos inhibitorios». Asimismo, este poderoso sistema de conocimiento nos permite ser autoconscientes de la construcción subjetiva de la identidad, es decir, que podemos «organizar y combinar armónicamente nuestro conocimiento sobre el mundo y nuestro conocimiento sobre nosotros mismos» (Ruiz-Vargas, 2004, pág. 1).

A pesar de los avances en las investigaciones sobre la memoria, la mayoría de las personas desconocen la «verdadera naturaleza de la memoria» (Ruiz-Vargas, 2004, pág. 2) y su importancia tanto en el plano personal como en el social. Con respecto al progreso de las investigaciones en el campo de la memoria, Ruiz-Vargas (2004) señala cuatro factores clave:

(1) el desarrollo de la *Neurociencia Cognitiva*; (2) la aceptación de los sistemas múltiples de memoria; (3) la unificación del *enfoque de laboratorio* y el *enfoque naturalista/ecológico*; y (4) la legitimación de la MA a partir de una metodología rigurosa de investigación. (págs. 2-3)

Cuando hablamos de la MA nos acercarnos de forma consciente a la narración de nuestras experiencias vividas. En otras palabras, los recuerdos autobiográficos contenidos la memoria constituyen un puente imaginario que nos transporta en el tiempo hacia las experiencias que conforman nuestra identidad como individuos pertenecientes a una sociedad determinada. Sin embargo, sumergirnos en el mar de los recuerdos personales está muy lejos de ser una experiencia nítida y veraz de los acontecimientos ocurridos en el pasado. Es probable que una gran mayoría de nosotros solo pueda reconstruir sus recuerdos a partir de una narración caótica de pequeñas imágenes y sensaciones que se suceden una tras otra sin guardar ninguna relación con los hechos reales. Entonces, ¿qué recordamos en realidad?, ¿cómo sabemos que aquello que recordamos, un suceso o experiencia, ha ocurrido en realidad y no es una mera ficción de nuestra imaginación?, ¿cuánto hay del pasado en la construcción de nuestros pensamientos a futuro? Como bien dice el reconocido neurólogo Oliver Sacks (2019), «asusta pensar que nuestros recuerdos máspreciados podrían no haber ocurrido nunca, o podrían haberle ocurrido a otro» (pág. 104).

Si bien la MA es un sistema que almacena, procesa y recupera información sobre nuestras experiencias es importante tener en cuenta que una buena parte de dicha información puede ser una simple ficción generada por nuestra imaginación asociativa. Además, muchos recuerdos dolorosos contenidos en la MA pueden devenir en una patología. Según la definición elaborada por García-Bajos y Migueles (2016):

La memoria autobiográfica nos permite viajar mentalmente al pasado y recrearnos en eventos futuros. Muchas experiencias son triviales, pero las vivencias emocionales dan significado a nuestras vidas. Investigaciones recientes muestran también que estamos motivados a olvidar las experiencias negativas, poniendo énfasis en los procesos inhibitorios. El control inhibitorio puede prevenir que hechos negativos, experiencias que nos generan tristeza y preocupación accedan a la conciencia. La ayuda para desterrar acontecimientos negativos pasados o futuros de nuestra mente puede tener una repercusión importante en las terapias cognitivas (pág. 89).

Como vemos, el control inhibitorio representa un potente mecanismo de autodefensa necesario para lograr un equilibrio emocional frente a experiencias vitales dolorosas o negativas (García-Bajos & Migueles, 2016, pág. 89). En tal sentido, podemos decir que estamos dotados de un potente escudo natural.

La investigación empírica de la MA puede rastrearse de Galton a Freud, ambos contemporáneos Ebbinghaus, pero cada uno con su método particular (Robinson, 1986, pág. 19). Como comenta Robinson (1986), «el acercamiento de Galton a la memoria era el del botánico, mientras que el de Freud era el del biógrafo» (pág. 20). Aquí es importante mencionar dos aspectos clave que centraron las diversas investigaciones: los recuerdos, por un lado, y el sujeto que recuerda, por otro.

Según comenta A. Manzanero (2017), la memoria autobiográfica es «el recuerdo de un suceso vivido en primera persona». Sin embargo, más allá de las diversas definiciones que se pueden encontrar del concepto «memoria autobiográfica», este continúa siendo uno de los campos de la memoria «menos desarrollado» (Brewer, 1986, pág. 25).

Si tenemos en cuenta que uno de los principales métodos de investigación de la memoria autobiográfica corresponde al desarrollado por F. Galton (1822-1911) nos encontramos, según Brewer (1986), ante un verdadero problema, pues se están investigando «las múltiples formas de memorias autobiográficas y no autobiográficas» (pág. 25), es decir, más centrados en la narración del sujeto.

La mayoría de investigadores considera que la MA es parte de la memoria episódica (Manzanero, 2017) que, a diferencia de la memoria semántica, existe un marcado contexto personal. Sin embargo, este punto de vista es confuso y ha traído consigo no pocas discusiones metodológicas y de definición, sobre todo con las conceptualizaciones de E. Tulving (Brewer, 1986, pág. 33).

Brewer (1986, pág. 33) define a la MA como la memoria de información que está relacionada con uno mismo, un campo de aglutina «una intensa activi-

dad investigadora que pretende comprender cómo esta información sobre hechos autobiográficos está organizada en memoria y cómo es recordada» (Cañas & Bajo, 1998, pág. 369).

Como podemos apreciar, los investigadores se encuentran con grandes dificultades a la hora de hallar una definición conceptual unificada de lo que la MA sea. Este problema está relacionado con otro no menos importante en dificultad y alcance: el método de investigación. En tal sentido, Cañas & Bajo (1998) dan cuenta de la situación:

La metodología experimental que venimos utilizando para estudiar la memoria humana requiere que el investigador controle la información que va a ser almacenada para poder, luego, compararla con la información que un sujeto recuerda... [...] no tenemos una forma fácil de comprobar que lo que esta persona recuerda corresponde a los hechos tal y como ocurrieron en la realidad. Por lo tanto, las técnicas que se están diseñando para estudiar MA tienen que tener alguna forma de comprobar la correspondencia de los hechos reales y los hechos tal como son recordados. Entre todas las técnicas propuestas las dos más utilizadas han sido las basadas en los diarios y las de asociación libre. Con ellas se está investigando los aspectos que consideramos más importantes para conocer la estructura y el funcionamiento de la MA: sus parámetros temporales, su organización, los mecanismos de recuperación que trabajan sobre ella y las causas del olvido de hechos autobiográficos. (págs. 369-370)

Si bien el párrafo es claro y describe un cuadro de situación bastante incierto con respecto a los resultados, el recorrido de la investigación de la MA ha pasado a lo largo de su historia por diversos métodos que han ido depurando las técnicas empleadas con el propósito de hacer que los resultados sean menos sesgados y se ajusten a una realidad compleja.

Manzanero (2017) describe las tres técnicas principales que se han empleado para la investigación de la MA: (1) técnica de Galton, (2) técnica de recuperación dirigida, y (3) la confección de diarios. Con respecto a Galton, Manzanero comenta que este utilizaba una técnica:

Consistente en presentar a los sujetos una palabra estímulo y pedirles que recuerden algún suceso personal asociado con ella. Posteriormente se les pide que describan el suceso y lo sitúen en el tiempo, valorando características como el grado de detalle y la vividez con que lo recuerdan. (Manzanero, 2017)

Tanto la «técnica de recuperación dirigida» como la «confección de diarios» ensayadas por Baddeley y Linton respectivamente buscan «superar las limitaciones como la falta de control y la verificación de los recuerdos autobiográficos» (Manzanero, 2017), siendo consideradas como alternativas a la investigación de Galton, quizás más vulnerable a los sesgos interpretativos.

Los estudios de Linton con diarios personales durante trece años arrojaron algunos buenos resultados, otros en cambio son discutibles, por ejemplo: «que los hechos desagradables se olvidaban más rápidamente que los agradables» (Manzanero, 2017) y (Cañas & Bajo, 1998, pág. 370). Es evidente que la investigación con un único sujeto experimental no puede extenderse y considerarse regla general.

Si bien el aporte de Linton al estudio de la MA fue innovador, no es capaz de escapar de ciertos sesgos metodológicos, como señalan Cañas & Bajo (1998):

La técnica utilizada poseía varios problemas: al ser un estudio en el que el investigador es el sujeto de la investigación se podían sesgar datos en favor de las hipótesis de este. Este problema puede solucionarse si el investigador pide a un grupo independiente de sujetos que lleven su propio diario. Sin embargo, esta solución no elimina el problema fundamental de la técnica de los diarios en general. Simplemente por tomar nota de los hechos puede hacer que estos queden marcados en memoria de una forma más persistente. (pág. 370)

Por su parte, Manzanero (2017) destaca otro importante estudio con diarios realizado por Willen A. Wagenaar en las décadas de 1980 y 1990:

Comenzó el estudio cuando tenía 37 años y lo finalizó con 43, apuntando cada día uno o dos sucesos, cuidando de señalar *quién* estaba implicado en el suceso, *qué* ocurrió y *dónde* y *cuándo* tuvo lugar. Además, evaluó en cada hecho su importancia, frecuencia, el grado de implicación emocional y la agradabilidad de su implicación. Durante el estudio apuntó un total de 2.400 sucesos. Posteriormente, trató de recordar los sucesos con un intervalo de retención de al menos 12 meses facilitándose parte de la información como ayuda al recuerdo. El orden de qué, quién, dónde y cuándo fue aleatorio, para comparar la importancia de estos cuatro aspectos, y manipuló el número de indicios que utilizaba como ayuda para recordar. Los resultados obtenidos mostraron que los indicios más importantes fueron la información sobre quién, qué y dónde, que ofrecieron aproximadamente la misma ayuda, mientras que la información sobre cuándo era una ayuda pobre. Este resultado sugiere que, aunque la información temporal se almacena, no es una buena vía de acceso a los recuerdos autobiográficos. No obstante, cuando se combinaba esta información con cualquiera de los restantes indicios, se constataron las mejores ayudas a la recuperación. Respecto a las características de los sucesos, encontró que los hechos inusuales o importantes y los que le implicaban emocionalmente eran los mejor aprendidos y menos olvidados. El efecto de la agradabilidad de los sucesos fue más complicado, ya que los sucesos desagradables eran peor recordados, aunque sólo tras intervalos cortos de tiempo, pues no se detectaron diferencias con intervalos grandes de tiempo (más de cuatro años).

Del párrafo anterior podemos deducir que las diversas metodologías de investigación empleadas no han agotado lo que conocemos, o bien, lo que podemos llegar a decir sobre la MA. En tal sentido, debemos darle la razón a Breuer (1986) quien opina que la MA constituye «una de las áreas menos desarrolladas en el estudio de la memoria humana» (pág. 25).

De las diversas definiciones que podemos encontrar sobre la MA, considero que la más adecuada, quizás por cercanía y utilidad a uno de los ejes de esta investigación, es la que propone R. Fivush (2011) quien describe a la memoria autobiográfica como:

Esa forma de memoria únicamente humana que se mueve más allá del recuerdo de eventos experimentados para integrar la perspectiva, la interpretación y la evaluación a través del yo, el otro y el tiempo para crear una historia personal. Para decirlo de manera sucinta, la memoria autobiográfica es la memoria del yo que interactúa con otros al servicio de objetivos a corto y largo plazo que definen nuestro ser y nuestro propósito en el mundo. (pág. 2)

Esta autora hace un aporte fundamental al demarcar y diferenciar un lugar exclusivo para la memoria episódica y otro para la MA, campos que otros investigadores por cuestiones metodológicas no han diferenciado con claridad o los han mantenido en una unidad. Tal separación, como bien dice Fivush, «permite una comprensión más completa del desarrollo de la memoria episódica y autobiográfica tanto en las especies como en el desarrollo humano» (2011, pág. 18).

El concepto de MA definido Fivush vincula las experiencias del yo pasado con el yo actual en un medio cultural determinado y con actividades sociales específicas. En este sentido, pienso que si algún día se pudiese desentrañar con exactitud el funcionamiento de la MA estaríamos muy cerca de comprender cuáles son los mecanismos cognitivos que intervienen en el desarrollo de lo que conocemos como «nuestro tiempo personal», es decir, cómo los seres humanos hemos llegado a construir el sentido de nuestras vidas y nuestra identidad (Fivush, 2011, pág. 18).

Una investigación de Fivush & Reese (1992), muy cercana a la hipótesis de la interacción entre la MA y la sociabilidad de nuestro trabajo, propone que la construcción de la MA nace en las primeras conversaciones entre padres e hijos sobre el pasado (pág. 115). Estos estudios ponen de manifiesto que las conversaciones mantenidas entre padres e hijos no siguen una estructura estándar o patrón, lo cual genera que los niños aprendan habilidades y valores diferentes (1992, pág. 127).

Otro punto sustancial que destacan dichas investigaciones es que los padres tienden a construir relatos más elaborados y embellecidos según se trate de niños o niñas (Fivush & Reese, 1992, pág. 128).

Entiendo que la MA debe considerarse como una base fundamental del proceso de desarrollo del individuo en un entorno social. Así lo explican Fivush & Reese (1992):

El niño no está construyendo una historia de vida en forma aislada, sino que está involucrado en interacciones sociales que están influenciadas por formas culturales (narrativas) y creencias (estereotipos de género) que conducen a formas particulares de entender y construir el pasado. Para comprender la memoria autobiográfica debemos examinar las formas en que estos tipos de factores socioculturales se muestran en la interacción social y cómo el niño llega a interiorizar estas influencias en el proceso de construcción de una historia de vida¹. (pág. 128)

Asimismo, la perspectiva narrativa que alimenta desde nuestra infancia la MA cobra una vital importancia en el proceso adaptativo de los mamíferos humanos, pues estos relatos nos van esculpiendo como personas integrantes de una sociedad. Más aún, podemos preguntarnos, o quizás imaginar, qué sería de nuestras vidas sin el almacén autobiográfico de estas narraciones. Según comenta Vilarroya (2019):

Narrar surgió de una necesidad adaptativa para la sociabilidad y el relato se convirtió en la herramienta que utilizamos con el fin de aprender a vivir en sociedad. En sus orígenes, narrar se nutrió particularmente de la fuerte predisposición para el juego propia de los mamíferos. Una vez establecida como una facultad propia de nuestra especie, narrar se integró en nuestro desarrollo, en nuestro comportamiento individual y social, en nuestra vida cognitiva y emotiva, y nos permitió enriquecernos a través de la imagina-

¹ «The child is not constructing a life story in isolation, but is engaging in social interactions which are influenced by cultural forms (i.e., narratives) and beliefs (i.e., gender stereotypes) which lead to particular ways of understanding and constructing the past. In order to understand autobiographical memory, we must examine the ways in which these kinds of social-cultural factors are displayed in social interaction, and how the child comes to internalize these influences in the process of constructing a life story».

ción. Nos hizo, por tanto, dependientes de todas aquellas vidas que podíamos imaginarnos para aprender tanto de las experiencias ajenas como de las propias. (págs. 76-77)

La importancia de la memoria recorre toda nuestra existencia. Por ejemplo, para Kandel (2019) es «un adhesivo que aglutina nuestra vida mental», «una parte integrante de todas las funciones cerebrales, desde la percepción hasta la acción» (pág. 117). Como hemos visto a lo largo de este capítulo, la MA contiene el relato de nuestra vida, ella nos ayuda a almacenar, crear y evocar recuerdos de nuestras experiencias individuales y sociales.

Para concluir, me gustaría plantear la cuestión de que por qué considero relevante para la filosofía contemporánea un estudio sobre la MA, un campo dominado por las neurociencias y la psicología. Como hemos comentado a lo largo del capítulo, la MA procesa de forma constante información relevante sobre nuestras experiencias de vida que sirve para consolidar nuestra identidad como individuos pertenecientes a una sociedad determinada (*ethos social*). Sin embargo, la MA no solo almacena los recuerdos de nuestro pasado, sino que además colabora en la construcción y proyección a futuro de nuevas experiencias de vida. Así la describió Williams James en 1890 en sus *Principios de Psicología*:

Un recuerdo es más que un hecho con fecha en el pasado: es un hecho con fecha en mi pasado. [...] Debe aparecerseme envuelto de ese «calor» y de esa «intimidad» [...], que son los criterios con los que la conciencia reconoce y se adueña de cualquier tipo de experiencia. (Viard, 2010, pág. 57)

Si bien la filosofía a lo largo de su historia aportó algunos elementos a la investigación sobre la memoria humana y la mente, no podemos negar que el principal desarrollo ha venido de la mano de la ciencia experimental (Kandel E. R., 2019, págs. 13-16). Sin embargo, como he comentado, los avances científicos no han logrado evitar los importantes escollos conceptuales que no logran poner de acuerdo a sus diversas teorías. En tal sentido, el análisis filosófico puede constituir el punto de partida para discutir, depurar y conectar aquellos

conceptos fundamentales que engloba una investigación sobre la MA, asimismo, poner a prueba una metodología rigurosa de investigación.

Las pinceladas que en este capítulo he ido esbozando sobre la MA y su alcance en el desarrollo de la vida humana dan cuenta del amplio campo que tenemos por delante en nuestra filosofía contemporánea: aspectos antropológicos, cognitivos, sociales, emocionales, cerebrales, mentales, políticos y un largo etcétera justifican el interés. Por ello, entiendo que hablar de la MA no es otra cosa que hablar de nosotros mismos y nuestras experiencias pasadas, presentes y futuras, es decir, nuestra frágil identidad humana.

En el capítulo siguiente abordaré el tema de la sociabilidad humana (*ethos social*), el segundo eje de la investigación. En este punto cabe recordar que la hipótesis central del trabajo supone una estrecha relación entre la MA y la sociabilidad. Desde un punto de vista cognitivo no cuento con fuentes y datos suficientes para determinar si alguno de estos dos elementos de la relación posee una mayor relevancia sobre el otro. En tal sentido, me encuentro en la aporía de no poder determinar cuál de estos elementos es el antecedente, debiendo asumir como punto de partida la coexistencia y realidad de ambos en un mismo nivel.

Capítulo 6. Sociales por naturaleza

Este capítulo y todos sus apartados estarán centrados en el segundo eje de la investigación dedicado al tema de la sociabilidad humana. Como hipótesis de trabajo señalé la existencia de una relación de influencia recíproca entre la memoria autobiográfica (MA) y la sociabilidad, pues considero que ambos elementos son parte de un mismo proceso de desarrollo bio-socio-cultural de los mamíferos humanos que se ha manifestado a lo largo de toda nuestra historia evolutiva.

En la actualidad, las investigaciones provenientes de diversos campos de las neurociencias nos aportan evidencias claras sobre la naturaleza social del cerebro humano y de qué manera el medio bio-social afecta a su plasticidad, cualidad tan necesaria para el aprendizaje y para la supervivencia. Desde mi punto de vista, el proceso de sociabilidad de nuestra especie puede ser considerado como una ventaja evolutiva que nos ha permitido adaptarnos y sobrevivir.

En el capítulo anterior comenté que los recuerdos autobiográficos de nuestras experiencias vividas se hallan almacenados en la memoria autobiográfica y son recuperados a través de una construcción narrativa mediada por el lenguaje, quizás real o ficticia, contada en primera persona.

La información recuperada no solo nos ayuda a evocar situaciones de nuestro pasado, sino también para diseñar e imaginar nuestro presente y futuro. Esta facultad cognitiva nos revela algunas pistas sobre la relación existente entre la MA y la sociabilidad.

Para ilustrar la importancia de la relación entre la memoria autobiográfica y la sociabilidad, podemos tomar como ejemplo el deterioro cognitivo que desarrollan los enfermos de Alzheimer y otras demencias similares. Estas patologías plantean un serio desafío a la ciencia médica contemporánea, pues sus consecuencias destructivas impactan sobre la mente y el cuerpo. En casos extremos podemos afirmar que la enfermedad ha borrado, además de los recuerdos y las

experiencias de vida, todo indicio de autonomía e identidad en la persona que la padece. Esta línea de investigación, a la que he denominado «investigación inversa», puede ayudarnos a construir nuevas hipótesis y preguntas para temas relacionados con la neurofilosofía.

Como parte de la historia natural, el ser humano alberga la necesidad evolutiva de la especie de perdurar en el tiempo: un código natural que trasciende al individuo. En tal sentido, los humanos estamos dotados de capacidades biológicas, cognitivas y culturales que coadyuvan al proceso evolutivo de nuestra especie por medio del altruismo, la cooperación, la convivencia pacífica y la negociación en conflictos. En una palabra, estamos dotados naturalmente para favorecer el desarrollo de la vida en sociedad (Lizón, 2016, pág. 10).

Esta peculiar cualidad social de nuestra especie la defino como «sociabilidad o *ethos* social». Este concepto complejo engloba la constitución biológica y las construcciones culturales necesarias para nacimiento y desarrollo de aquellos valores que constituyen el carácter o naturaleza moral de los individuos en sociedad.

Las corrientes actuales de investigación sociológica y neurociencia social cognitiva profundizan aún más la relación entre biología y cultura, como expone la profesora A. Lizón (2016):

Buena parte de nuestro éxito como especie parece radicar en este hecho fundamental: somos criaturas básicamente sociales. Nuestros cerebros han sido cableados para abrigar motivaciones profundas que nos unen a parientes y amigos; para responder a los retos y compromisos que plantean la reciprocidad mutua y la coordinación dentro de los grandes grupos; para procesar la fruición y el malestar social con la misma intensidad y vehemencia con la que experimentamos el placer o el dolor físico. (pág. 10)

La idea de una «naturaleza moral» no es nueva, sino que data de una larga tradición filosófica y científica como pueden dar cuenta los trabajos de Aristóteles, Hume o Darwin. Según explica Patricia Churchland:

La hipótesis predominante es que lo que nosotros, los humanos, llamamos «ética» o «moralidad» es una estructura de conducta social en cuatro dimensiones que viene determinada por la interrelación de distintos procesos cerebrales: (1) el cuidado o la atención a los demás (enraizado en el apego a nuestros familiares y la preocupación por su bienestar), (2) el reconocimiento de los estados psicológicos de los demás (basado en las ventajas de predecir la conducta de terceros), (3) la resolución de problemas en un contexto social (por ejemplo, cómo deberíamos castigar a los sinvergüenzas) y (4) el aprendizaje de prácticas sociales (mediante un esfuerzo positivo y negativo, por imitación, por ensayo y error, por diversos condicionamientos y por analogía). La sencillez de esta estructura no significa que sus formas, variaciones y mecanismos neuronales sean simples. Al contrario, la vida social es increíblemente compleja, puesto que el cerebro es el órgano que la administra. (2012, pág. 15)

Planteada esta breve introducción, pasaremos a indagar el sustrato y las características que permiten el desarrollo de la moralidad humana, un elemento vinculado al desarrollo biológico y cultural de las personas en sociedad.

6.1. La base de la sociabilidad: los valores morales

El desarrollo de la vida en sociedad está fundamentado en nuestros valores morales, esto es, una moralidad entendida como costumbres o hábitos. Para P. Churchland (2012), los valores «no tienen por qué implicar normas, aunque a veces sea así; no tienen que ser necesariamente explícitas, sino que los niños pueden aprenderlas implícitamente mientras se desenvuelven en su mundo social» (pág. 20). Sin embargo, hay que tener presente que las sociedades humanas son las más complejas en cuanto a variedad, cantidad de individuos, leyes y normas de funcionamiento (Soler, 2009, pág. 282).

En la actualidad, encontramos algunos estudios (biológicos, cognitivos, genéticos, epigenéticos y de redes neuronales) que relacionan la estructura bioquímica-molecular con el desarrollo de valores necesarios para la subsistencia y adaptación al entorno natural y social (por ejemplo, el sistema neuroendocrino

encargado de regular nuestras hormonas, vida sexual, estados de ánimo y emociones, entre otras funciones). Esta relación biología - cultura de los valores morales es sustancial para esta investigación.

El avance de las neurociencias y sus disciplinas transversales nos aportan nuevas perspectivas de estudio sobre las manifestaciones externas de la sociabilidad (convenciones y pactos sociales) y la relación con los sustratos bioquímicos que influyen y modelan nuestro carácter y conducta. En definitiva, cada vez que nos preguntamos por la sociabilidad (*ethos social*) nos adentramos en el oscuro y complejo laberinto de la naturaleza humana.

En un primer acercamiento, podemos decir que la sociabilidad o *ethos* social constituye una cualidad natural ligada a la evolución propia de los mamíferos humanos como parte de nuestro proceso adaptativo. Su desarrollo está cimentado en una estructura que es biológica, cognitiva y cultural a la vez, pues depende de «mecanismos neuronales» que incluyen a la motivación e impulso, recompensa y predicción, percepción y memoria (Churchland P. S., 2012, pág. 18). Destaco estos tres aspectos estructurales de la sociabilidad (biología, cognición y cultura) porque son los mismos que hallamos la base de la memoria autobiográfica. Tales aspectos me hacen pensar que la sociabilidad constituye el espacio-tiempo, como comentamos, propiamente humano, desde donde se despliega toda nuestra memoria autobiográfica.

De momento, no tenemos constancia de que otro ser sea autoconsciente de su propia historia personal, es decir, que tenga un constructo similar a la memoria autobiográfica humana y, además, pueda narrarla.

La estructura biológica nos permite ensayar un camino inverso de investigación y estudiar cómo se destruye la sociabilidad, por ejemplo, cuando las neuropatologías afectan a la memoria, la percepción, la atención y otros procesos cognitivos. Aquí vemos la importancia de la base orgánica (física/material) que acompaña a los procesos cognitivos y culturales.

A medida que avanzamos, vemos que las fronteras que delimitan el nacimiento y desarrollo de los valores morales y la sociabilidad ya no están del todo

claras, es decir, nos hallamos en un espacio compartido donde conviven e interactúan el individuo y la sociedad, la biología y la creación cultural. En tal sentido, entiendo que es clave el concepto de «moralidad» que define P. Churchland (2012):

Es una estructura de conducta social en cuatro dimensiones que viene determinada por la interrelación de distintos procesos cerebrales: (1) el cuidado o la atención a los demás (enraizado en el apego a nuestros familiares y la preocupación por su bienestar), (2) el reconocimiento de los estados psicológicos de los demás (basado en las ventajas de predecir la conducta de terceros), (3) la resolución de problemas en un contexto social (por ejemplo, cómo deberíamos castigar a los sinvergüenzas) y (4) el aprendizaje de prácticas sociales (mediante un refuerzo positivo y negativo, por imitación, por ensayo y error, por diversos condicionamientos y por analogía). La sencillez de esta estructura no significa que sus formas variaciones y mecanismos neuronales sean simples. Al contrario, la vida social es increíblemente compleja, puesto que el cerebro es el órgano que la administra. (pág. 19)

El enfoque propuesto por P. Churchland nos ofrece una fundamentación de la sociabilidad basada en ciertos valores sociales tales como «la atención a las amistades, el apego entre familiares y la necesidad de pertenencia a un grupo» por sobre cualquier tipo de racionalidad moral derivada de nuestras convenciones sociales o culturales (2012, pág. 23).

La autora prioriza los fundamentos biológicos por sobre cualquier otro elemento cultural o de convenciones sociales. Su trabajo aporta un conjunto de evidencias sólidas acerca del sofisticado laberinto que compone la base de la naturaleza social humana con las que podemos estar de acuerdo o no, pero lo cierto es que no podemos mostrarnos indiferentes ante ellas.

La clave del origen de los valores está en la supervivencia y el cuidado de uno mismo, pues «si no tuviera ninguna motivación para preservar su vida, ningún animal sobreviviría mucho tiempo ni tampoco se reproduciría» (Churchland P. S., 2012, pág. 23). Sin embargo, la sociabilidad humana se manifiesta a través de mecanismos de cooperación y cuidados de los otros. Tal conducta, altruista

y cooperativa, debe servir a los individuos involucrados, sino se desecha por representar un coste energético que compromete la supervivencia (Churchland P. S., 2012, pág. 24). Ahora bien, como comenta el biólogo Manuel Soler (2009, págs. 259-260): «vivir juntos en sociedad no solo tiene ventajas, sino también inconvenientes», por ejemplo, las enfermedades y los parásitos (Soler, pág. 262).

Con respecto a la extensión del bienestar más allá del propio individuo, la neuroendocrinología plantea una perspectiva novedosa al asumir una modificación genética-evolutiva en la organización neuronal de los mamíferos a partir de la generación de nuevos valores sociales que incluyen tanto el bienestar de terceros como también de desconocidos, una ampliación de la conducta social que marca el inicio de la moralidad (Churchland P. S., 2012, pág. 25).

El responsable neuroquímico de esta modificación en el cerebro de los mamíferos es una cadena de péptidos muy antigua llamada oxitocina (Churchland P. S., 2012, pág. 25). Los cerebros de los mamíferos han evolucionado creando mecanismos neuronales y corporales que «maternalizan el cerebro de los mamíferos hembra» (Churchland P. S., 2012, pág. 43) para cuidar de su descendencia.

Hasta aquí podemos considerar que el cuidado de las crías representa algo natural y sin cuestionamiento. Sin embargo, los mamíferos humanos creamos sociedades basadas en el cuidado altruista de otros individuos ajenos a nuestro círculo familiar y en la cooperación con otros. ¿Cómo es posible la conducta altruista? ¿Por qué cooperamos? Estas preguntas y las respuestas que de ellas se derivan constituyen una parte de la base de toda acción humana dentro de un marco de sociabilidad.

6.2. Neuronas espejo, empatía y desarrollo del cerebro social

Aristóteles, en sus debates sobre moralidad, hace hincapié en las habilidades sociales como la flexibilidad, la idoneidad y el sentido práctico para

prosperar en el ámbito social. Según esta perspectiva, el ejercicio de las habilidades sociales depende de la adquisición de los hábitos adecuados, y puede verse influenciado por los modelos, las prácticas sociales unas instituciones que encontramos en la vida diaria. Un componente esencial de la sociabilidad normal ataña a nuestra capacidad para atribuir estados mentales a los demás. Sin esta capacidad no podemos sentir empatía hacia el sufrimiento, ni entender sus intenciones, sentimientos, creencias y planes. (Churchland P. S., 2012, pág. 133)

En la actualidad, las neurociencias, dentro de un amplio conjunto de ciencias experimentales, están ayudando a desentrañar parte de la compleja estructura social humana, complementando de esta manera los amplios estudios iniciados por otras disciplinas, tales como la filosofía, la psicología, la mitología, el arte, y la literatura (Churchland P. S., 2012, pág. 135).

En lo que concierne a los ejes de esta investigación, podemos destacar que uno de los hallazgos más relevantes que está permitiendo estudiar más en profundidad la relación entre nuestro cerebro y el medio en que se desarrolla es el concepto de «plasticidad cerebral».

El aumento del volumen de la masa cerebral alcanzado en nuestra historia evolutiva, aproximadamente seis millones de años, como homínidos ha sido un factor determinante para el desarrollo de la plasticidad cerebral, las funciones cognitivas y, fundamentalmente, el aprendizaje (García García, 2008, pág. 70). Una evidencia de ello puede constatarse en la presencia de una mayor cantidad de circunvoluciones y surcos que observamos en la corteza cerebral. Según comenta Vilarroya (2019):

El hecho es que en los últimos 4 millones de años el cerebro de los humanos pasó de 450 centímetros cúbicos a los actuales 1450, es decir, hemos ganado un litro en forma de tejido cerebral. Esto representa un fenómeno biológico extraordinario. Entre otras cosas, el cerebro pasó de tener unos 30 mil millones de neuronas a contar con más de 86 mil millones, y pasó de gastar el 10% de toda la energía que consume un individuo al 25% actual. El hecho de que ocurra algo así requiere, sin ninguna duda, que esa ganancia haya tenido un valor adaptativo brutal. (págs. 56-57)

Como explica García y García (2018), el desarrollo filogenético del *homo sapiens* abrió la puerta al desarrollo cognitivo y al aprendizaje «no solo del otro sino a través del otro», lo que explicaría una Teoría de la Mente (ToM) que «sirve para predecir y modificar el comportamiento» (pág. 71). Con el fin de demarcar a qué nos referimos cuando hablamos de teoría de la mente he escogido dos definiciones que nos pueden ayudar:

La teoría de la mente se refiere a la capacidad que tenemos de «leer» la mente de los demás, esto es: 1) entender que otras personas tienen creencias, deseos e intenciones que no tienen por qué coincidir con los nuestros; 2) determinar el contenido de dichos estados mentales, y 3) predecir cómo se comportarán esas personas en respuesta a esos estados mentales. (Vilarroya, 2019, pág. 70)

El concepto de «teoría de la mente» (ToM) se refiere a la habilidad para comprender y predecir la conducta de otras personas, sus conocimientos, sus intenciones y sus creencias. Desde este punto de vista, este concepto se refiere a una habilidad «heterometacognitiva», ya que hacemos referencia a cómo un sistema cognitivo logra conocer los contenidos de otro sistema cognitivo diferente de aquel con el que se lleva a cabo dicho conocimiento. La terminología asociada a este concepto es variada: se han utilizado diversos conceptos, como «ToM», «cognición social», «mentalización», «psicología popular», «psicología intuitiva» o «conducta intencional» (Tirapu-Ustároz, Pérez-Sayes, Erekatxo-Bilbao, & Pelegrín-Valero, 2007, pág. 479).

La ToM es compleja e implica diversos niveles de profundidad analítica. Uno de estos niveles se utiliza con frecuencia para realizar estudios teóricos sobre la empatía y el juicio moral, por ejemplo, el dilema del prisionero o el dilema del tren, enunciados de la siguiente manera:

Un vagón de tren se dirige sin control hacia un grupo de cinco operarios que realizan obras de mantenimiento en la vía. Todos ellos morirán aplastados por la máquina si no encontramos una solución. Usted tiene la posibilidad de apretar un botón que activará un cambio de agujas y desviará el tren hacia otra vía donde se encuentra un trabajador realizando obras

de reparación. El vagón mataría a este hombre, pero los otros cinco se salvarían. ¿Pulsaría el botón? Ahora vamos a por otra versión del mismo dilema. Usted se halla sobre un paso elevado que cruza sobre la vía y situado en un punto vertical entre la vía y las cinco personas. Un señor con aspecto desaliñado y con pintas de estar bebido se encuentra a su lado. Una manera de frenar el «vagón asesino» consiste en empujar al señor para que caiga sobre la vía y resulte atropellado, lo que provocará que el maquinista reaccione, frene el tren y salve las cinco vidas. ¿Le daría un empujón? (Tirapu-Ustároz, Pérez-Sayes, Erekatxo-Bilbao, & Pelegrín-Valero, 2007, pág. 483)

Estudios con resonancia magnética muestran «la conexión neuronal entre la ToM y la empatía», señalando en una región determinada al llamado «cerebro moral» (Tirapu-Ustároz, Pérez-Sayes, Erekatxo-Bilbao, & Pelegrín-Valero, 2007, pág. 484).

La región prefrontal del córtex (CPF) está vinculada con las funciones cognitivas y las habilidades sociales, unas ventajas evolutivas que aumentan las capacidades físicas y sociales para predecir y capitalizar las predicciones con incidencia en el retraso de la gratificación y el ejercicio del autocontrol (Churchland P. S., 2012, pág. 136).

Las principales teorías neurocientíficas coinciden en que la plasticidad del cerebro, esencial para la vida en sociedad, está vinculada a las neuronas espejo. Este grupo de neuronas nos permiten desarrollar una saludable vida social, convivir en grupo y entender a los demás desde un punto de vista mental y emocional. El hallazgo de estas células ha permitido a los neurocientíficos hallar una explicación a «nuestra capacidad natural de percibir y ser sensible a los estados emocionales de los demás junto a la motivación para cuidarlos» (Decety, Bartal, Uzefovsky, & Knafo-Noam, 2016, pág. 1), es decir, la empatía. Además, por vez primera, se contaba con una explicación neurofisiológica a las formas complejas de cognición social (Iacoboni, 2011, págs. 14-15).

Los primeros estudios destacados sobre neuronas espejo se realizaron en 1995, en el laboratorio de neurobiología de G. Rizzolatti en la ciudad de Parma. En dichos experimentos se investigó una porción de la corteza premotora

(donde se planean e inician los movimientos) denominada «área F5» a través de electrodos implantados en cerebros de macacos, los cuales habían sido entrenados para asir determinados objetos. Los registros evidenciaron que tales grupos de neuronas solo se activaban cuando los macacos observaban tanto la acción como su objetivo (García García, 2008, pág. 73).

Más adelante, los estudios con técnicas de imagen realizados en humanos confirmaron las primeras hipótesis: «comprendemos los estados mentales de los otros simulándolos en el cerebro, y lo logramos por medio de las neuronas espejo» (Iacoboni, 2011, pág. 40). Según los investigadores italianos, estos hallazgos evidenciaban que nuestro cerebro es capaz de elaborar un reflejo espectral en la mente, un mecanismo para «leer intenciones».

Estos descubrimientos fueron la puerta de entrada para que el neurocientífico Vittorio Gallese y filósofo Thomas Metzinger elaboraron una teoría denominada «copia compartida», es decir, que los mamíferos humanos somos capaces de comprender a los demás a partir de la observación de los movimientos, una conexión entre el observador y el actor. La teoría de Gallese y Metzinger venía a confirmar una antigua intuición filosófica que relaciona la empatía con el desarrollo de la sociabilidad humana. En este punto podemos destacar las palabras de S. Ayan (2004):

La suposición de que el activo ponerse en lugar de otro es la base de todo intercambio social tiene una larga tradición filosófica. Bajo el concepto de empatía o capacidad de compenetración, esta capacidad humana ha atraído, sobre todo desde inicios del siglo XX, a grandes pensadores, como Edmund Husserl (1859-1938), que vio en ella la solución del problema de la «intersubjetividad», la cuestión de por qué podemos, pese a ser individuos autónomos, comprendernos unos a otros. La teoría de la «copia compartida» representa, pues, la variante moderna de un pensamiento antiguo. (pág. 79)

A primera vista, el hallazgo de las neuronas especulares alentaba una explicación convincente sobre una «teoría de la mente» y la posibilidad de comprender los estados mentales de los demás.

Luego del impacto causado por las primeras publicaciones, junto a las respuestas que auguraban un futuro prometedor para el campo de las neuronas especulares, aparecieron algunos trabajos de peso que no terminaban de ver con claridad las conclusiones a las que arribaron los neurocientíficos de Parma, ya que las evidencias neurobiológicas no terminaban de ser concluyentes para explicar la relación entre este grupo de neuronas y conducta social de los mamíferos humanos (Churchland P. S., 2012, pág. 138). Con respecto a ello, Churchland (2012) expone la complejidad que subyace a la hora de interpretar con certidumbre origen funcional de la actividad de una neurona espejo:

¿Se relaciona esa actividad con una emoción, con la atención, con un estímulo sensorial, con un dato almacenado en la memoria, con algo esperado en el futuro o con la anticipación a un movimiento?; ¿o bien se la relaciona con alguna combinación de todos estos factores? (pág. 139)

Desconocemos el objetivo último de las cuestiones planteadas por P. Churchland, sin embargo, podemos atisbar que sus intenciones están en pos de una profundización metodológica en la investigación sobre las neuronas especulares antes que una crítica simple y rasa.

Algo similar ocurre con la «teoría de la copia compartida» desarrollada por Gallese y Metzinger la cual, en opinión de S. Ayan, «por fascinante que pueda resultar, pertenece al reino de la especulación» (2004, pág. 81). Si bien el desarrollo de las investigaciones neurobiológicas sobre neuronas espejo y sus derivaciones en otras disciplinas —por ejemplo, la filosofía, las ciencias sociales y las ciencias cognitivas— lleva algunos años de discusión en los círculos académicos y prestigiosas revistas, aún no podemos saber con exactitud cuál es el alcance de la misma y si puede explicar sin fisuras la sociabilidad humana.

Desentrañar la complejidad (Churchland P. S., 2012, pág. 145) del funcionamiento del *ethos* social de los mamíferos humanos, es decir, la manifestación de la sociabilidad o estructura del reino de nuestras costumbres o hábitos —por ejemplo, el saludo entre desconocidos, el respeto a los mayores, la celebración de una cena con amigos, el duelo por la muerte de los seres queridos, las

organizaciones gubernamentales democráticas o los sistemas de servicios sociales en salud y pensiones— nos acerca al despliegue de nuestras capacidades biocognitivas integradas en la inteligencia natural (IN): una evidencia de que estamos dotados naturalmente para el aprendizaje.

6.3 Imitación y habilidades sociales

Las últimas décadas del siglo XX cimentaron las bases de las investigaciones sobre la cognición (memoria, atención y percepción), el carácter y la conducta a partir del desarrollo de la informática, las ciencias biológicas (por ejemplo, la genética) y la psicología experimental. De la conjunción de estas disciplinas emergieron los primeros estudios centrados en las «habilidades sociales».

En la actualidad, alentada por el desarrollo de disciplinas científicas, filosóficas y tecnológicas, esta área multidisciplinar constituye un terreno fértil para las neurociencias, la neurobiología, la IA y diversas corrientes filosóficas enraizadas en las ciencias cognitivas (Churchland P. S., 2012, pág. 179).

Como hemos comentado en otros pasajes, el *ethos* social o sociabilidad aglutina las características de nuestro particular modo ser, es decir, animales sociales. Sin embargo, a diferencia de otras especies animales, la evolución nos ha dotado de facultades cognitivas (lenguaje, memoria, aprendizaje, atención, percepción, entre otras) que dan un formato único a nuestra sociabilidad, pues permiten el modelado de otro elemento diferenciador de nuestra naturaleza: la cultura. Podemos afirmar que la nuestra es una «sociabilidad cultural», una dimensión mediada por valores morales y creaciones propias.

Estas facultades cognitivas moldean una de las características fundamentales que cooperan en la construcción de nuestra sociabilidad: el comportamiento. Como parte del desarrollo evolutivo, el comportamiento posee una base genética importante que, como bien explica Soler (2009):

[...] no es que exista un gen (o grupo de genes) ... [...] lo único que hacen los genes es dirigir la formación de una proteína, no provocar ningún comportamiento. El hecho de que un individuo sea portador del gen o de los genes responsables de un comportamiento, lo único que quiere decir es que tiene la posibilidad de llevar a cabo ese comportamiento, pero no hay seguridad de que vaya a ejecutarlo... [...] El comportamiento es el resultado de la interacción entre los genes y el medio ambiente, y es imposible aislar uno de los componentes como el factor más importante. (págs. 72-73)

Encuentro relevante la definición de Soler (2009, pág. 77) sobre el aprendizaje «como la modificación del comportamiento debida a la experiencia», pues entiendo que abre nuevos interrogantes sobre la relación entre la memoria autobiográfica, almacén de nuestras experiencias, y la base genética que interviene en la movilidad del comportamiento social.

Otro elemento importante que relaciona el aprendizaje, la memoria autobiográfica y la inteligencia social es el peculiar carácter narrativo nuestro lenguaje, el cual está asentado en una estructura biológica y social capaz de almacenar y transmitir diversas experiencias, es decir, podemos contar y hacer partícipe al grupo nuestras vivencias (Vilarroya, 2019, págs. 72-73).

Con respecto al aprendizaje, Soler (2009) deja en claro existen diversas especies animales que poseen elevadas cualidades para el aprendizaje:

Los animales son capaces de aprender comportamientos sorprendentes... [...] pero el aprendizaje necesita que existan unas capacidades previas. Un mito bastante extendido es que los humanos tenemos una capacidad de aprendizaje muy superior a la del resto de animales. Sin embargo, esto no es totalmente cierto. Por ejemplo, las ratas nos superan en la habilidad para evitar un veneno; las palomas mensajeras, en la capacidad para orientarse en grandes espacios; y muy pocas personas serían capaces de igualar a las abejas a la hora de encontrar distintas zonas de alimentación un día después de haber localizado el alimento. De todas formas, es importante destacar que cada especie está adaptada a aprender solo ciertos aspectos, aquellos que le son necesarios y útiles en su medioambiente. La selección natural no permite despilfarros, y menos en

lo que respecta al sistema nervioso, que es muy costoso de mantener, tanto en términos nutricionales como energéticos. Las ratas no podrían aprender a buscar néctar ni las abejas a evitar venenos, simplemente porque esas habilidades no les son beneficiosas en su medio ambiente. (pág. 78)

Los procesos de aprendizaje conllevan un elevado coste energético, tanto para el individuo como para el grupo, que es necesario racionalizar. La transmisión del aprendizaje en sociedad persigue minimizar los costes de la supervivencia que, de otro modo, serían inasumibles, pues estaríamos abocados a aprender y transmitir desde cero y de forma individual cada nueva experiencia. Para compensar esta aparente desventaja, los humanos, entre otras especies animales, hemos sido dotados para el aprendizaje por la imitación, una capacidad innata muy eficaz que nos permite actualizar los conocimientos y transmitirlos a las siguientes generaciones (Churchland P. S., 2012, pág. 145) sin destinar elevados recursos energéticos.

En las últimas décadas, el estudio del «mimetismo inconsciente» es un campo de la psicología que está ganando fuerza, pues considera que este «desempeña un papel significativo en los procesos de afiliación y en el establecimiento de una relación cordial» (Churchland P. S., 2012, pág. 174).

Un aspecto para destacar de esta hipótesis es que considera a la mímica como un «cohesionador social», por ejemplo, de la observación de la conducta y el comportamiento de un recién llegado al grupo podemos deducir que, si actúa como yo, es realmente como yo y no pondrá en peligro el equilibrio social ni la supervivencia del grupo. La ventaja evolutiva nos sitúa en otro peldaño, pues nos ayuda a predecir la conducta (Churchland P. S., 2012, págs. 175-176).

El mimetismo inconsciente revela nuestra capacidad para socializar y adaptarnos al grupo social a través de expresiones faciales, gestos y posturas (Chartrand & Bargh, 1999). La importancia de esta conducta reside en que no existe una intención previa de imitar, sino de que actuamos de forma inconsciente y espontánea, lo cual nos ayuda a crear lazos de filiación.

En la actualidad, este tipo de mimetismo es estudiado bajo el concepto de «efecto camaleón», son investigaciones que buscan arrojar luz a la importancia que los comportamientos no verbales tienen en nuestra historia evolutiva (Lakin, Jefferis, Cheng, & Chartrand, 2003).

6.4. Altruismo, cooperación y mutualismo

En la actualidad, las florecientes investigaciones que conforman el abanico de las ciencias cognitivas nos aportan nuevos elementos que fortalecen las conceptualizaciones teóricas sobre cómo la sociabilidad humana constituye un factor decisivo en la evolución y supervivencia de la especie.

Parte de estos fundamentos están focalizados en el altruismo y la cooperación, sobre todo, en la búsqueda de aquellos elementos comunes que han permitido extender en las diferentes culturas las conductas altruistas y cooperativas más allá de las unidades familiares.

Entender las bases de estos comportamientos ha motivado el interés de muchas disciplinas (biología evolucionista, economía experimental, teoría de juegos, antropología cultural y biológica, psicología cognitiva, social y evolutiva, entre otras) que los han convertido en objeto de estudio (Tomasello, Marengo, & Dweck, 2010, págs. 18-19). Asimismo, son fundamentales para las investigaciones en los campos de la moral y la política, pues permiten comprender la complejidad del funcionamiento social de la naturaleza humana.

La vida en sociedad requiere un alto grado de inteligencia para procesar y organizar toda la información generada, asimismo, y no menos importante, para administrar los recursos disponibles y tomar decisiones eficientes.

La supervivencia del grupo (capacidad vital para conocer amigos y evitar enemigos) y las «presiones selectivas» requieren que la enseñanza de los conocimientos y el flujo de la información trasciendan las fronteras del individuo, de esta manera comienza a formarse lo que podemos considerar como el primer

sustrato de la inteligencia social, la cual nos ayudará a seleccionar nuestros colaboradores y desarrollar habilidades de supervivencia frente al engaño de los rivales (Soler, 2009, pág. 291).

La inteligencia social recorre un camino de ida y vuelta en el desarrollo del altruismo, la cooperación y el mutualismo, los cuales constituyen tipos especiales de conductas que atienden y se preocupan por el cuidado del otro que, en ciertos casos, incluye también a desconocidos externos al núcleo familiar más inmediato, aunque esto conlleve un alto coste energético (Churchland P. S., 2012, pág. 77). ¿Por qué lo hacemos? Según Churchland (2012), la respuesta podemos hallarla en nuestra propia neuroquímica:

En el centro de esta compleja red de conexiones neuronales se encuentra la oxitocina (OXT), un poderoso péptido que en los mamíferos se encarga de organizar el cerebro de modo que el cuidado y la atención de uno mismo se entienda a los bebés, y de ahí a un círculo cada vez más amplio de relaciones de cuidado. La oxitocina se ha relacionado con la confianza, debido en gran parte a que eleva el umbral de tolerancia a los demás y reduce el miedo y las respuestas de evitación. (pág. 77)

Como podemos apreciar, tanto nuestra bioquímica personal como nuestro conocimiento e integración en el medioambiente cultural nos predisponen para la vida en sociedad. Sin embargo, el complejo equilibrio social, siempre en tensión, dependerá en gran medida de las capacidades de organización de las instituciones para la asignación y distribución de los recursos escasos, lo cual requiere la creación de pactos sociales (respeto de derechos y deberes) sobre los que se asientan nuestras acciones colaborativas. Dichos acuerdos, a veces rudimentarios (como dar la palabra o estrechar la mano para cerrar acuerdos comerciales), han permitido alcanzar los niveles básicos de cooperación, entendimiento y respeto requeridos para la supervivencia y continuidad de la especie a través de siglos de evolución.

Estas complejas y primitivas formas de asociación están asentadas sobre habilidades y motivaciones colaborativas que evolucionaron a partir de pequeños grupos, por ejemplo, clanes familiares y tribus. La ampliación del cuidado y la

ayuda más allá de estos núcleos comenzó con el desarrollo de las prácticas agrícolas hace unos diez mil años. Si bien las conductas colaborativas son necesarias para la supervivencia del conjunto, implican un elevado coste que deben asumir todos los individuos (Churchland P. S., 2012, pág. 78).

En este nivel primario de la construcción de las sociedades humanas llamamos que estas fomentan y premian un tipo específico de comportamiento: el altruismo, un rasgo distintivo de nuestra especie (Soler, 2009, pág. 327), un comportamiento «que se ejecuta en beneficio de otros individuos a la vez que supone un coste para el que lo hace» (Soler, 2009, pág. 296). En cambio, si la acción la realiza un grupo de personas en pos de un beneficio mutuo se denomina colaboración mutualista (Tomasello, Marengo, & Dweck, 2010, pág. 70).

Ahora bien, distinguir socialmente qué tipos de acciones podemos considerar como «altruistas puras» es bastante complejo. Sin embargo, en la actualidad, desde diversos sectores de la biología, la etología o la sociología trabajan para ofrecer una definición bastante completa sobre lo que podemos entender como altruismo, tal es el caso de la ofrecida por Soler (2009):

Un comportamiento altruista es una acción que supone un coste para el individuo que la realiza (donante), pero que beneficia a uno o más individuos (receptores). Un verdadero alto altruista sería aquel que no reportara ningún beneficio al donante. Desde un punto de vista evolutivo, esos costes y beneficios se considerarán en términos de supervivencia o éxito reproductor. (pág. 297)

En el comportamiento social podemos identificar dos características propias de nuestra especie (Tomasello, Marengo, & Dweck, 2010, págs. 11-13) que han evolucionado hacia formas avanzadas de asociación y que constituyen el resultado del aprendizaje social: (a) la evolución cultural acumulativa y (b) las instituciones sociales que incluyen normas y prácticas que son aceptadas por el conjunto social, tales como los derechos y las obligaciones, es decir, los pactos contenidos en el contrato social que todos aceptamos en las sociedades modernas.

Según explica Tomasello (2010, pág. 66), las tendencias altruistas para ayudar y cooperar ya se manifiesta en los niños que, con el paso tiempo, aprender a ser selectivos.

Al preguntarnos por la condición de posibilidad de la sociabilidad humana no podemos eludir las características biológicas y culturales que modelan nuestra naturaleza a partir de una conjunción adaptativa entre el medio ambiente natural y el medio socio-cultural, dos elementos indispensables e inseparables que nacen de la convergencia entre la necesidad y la ayuda.

Muchas son las evidencias que sostienen al altruismo como uno de los pilares de nuestra construcción social. Tomasello (2010) remarca que la voluntad de cooperar es parte de nuestra cultura y, sobre todo, nuestra naturaleza humana:

Los *Homo sapiens* están adaptados para actuar y pensar cooperativamente en grupos culturales hasta un grado desconocido en otras especies. De hecho, las hazañas cognitivas más formidables de nuestra especie, sin excepción, no son producto de individuos que obraron solos sino de individuos que interactuaban entre sí, y lo dicho vale para las tecnologías complejas, los símbolos lingüísticos y matemáticos, y las más complicadas instituciones sociales. (pág. 17)

A diferencia de otras especies, los humanos somos altruistas (Soler, 2009, pág. 327), una cualidad fundamental que nos ha permitido desarrollarnos a partir de compartir recursos escasos, cuidados y cooperar entre los diversos individuos que conforman una sociedad.

Encontrar en la vida social actos altruistas puros resulta verdaderamente difícil, pues por definición una auténtica conducta altruista supone un coste sin beneficio a cambio, movilizando la cooperación y castigando las actitudes antisociales (Soler, 2009, pág. 328). Aunque estas cualidades de cooperación y generosidad están presentes en cada uno de nosotros desde la infancia, podemos ver, sin embargo, actitudes selectivas que guían los actos altruistas con el propósito de auto-protegernos del aprovechamiento indiscriminado de otras personas (Tomasello, Marengo, & Dweck, 2010, pág. 67).

La conducta altruista es una cualidad humana aprendida y transmitida que, por factores evolutivos y de supervivencia, ha penetrado en todas las culturas hasta nuestros días. Sin embargo, debido a la imposibilidad de contar con evidencias contrastables, es imposible datar con exactitud su origen. No obstante, algunas investigaciones comparadas de relevancia en el campo de la antropología evolutiva coinciden en ubicar su origen en los asentamientos humanos, más o menos estables, que desarrollaron las primeras técnicas de agricultura como su forma de subsistencia regular.

Un ejemplo de ello es la hipótesis desarrollada por Tomasello (2010), quien deriva el nacimiento del altruismo de la cooperación grupal, esto es, un proceso en donde todo el grupo se beneficiará de las conductas colaborativas si todos sus integrantes aportan su trabajo: el autor define a esta conducta como «mutualismo» (pág. 72).

Los antecedentes del paso de la actividad grupal a la colaborativa en los primeros humanos están asentados sobre tres procesos que Tomasello considera esenciales para el desarrollo del mutualismo tal y como lo conocemos hoy en día: (a) coordinación y comunicación basadas en habilidades cognitivas y sociales; (b) tolerancia y confianza entre los integrantes del grupo; y (c) las normas y las instituciones donde debieron darse las primeras pautas organizadas bajo un deber comunitario (Tomasello, Marengo, & Dweck, 2010, pág. 74).

La conducta altruista define un rasgo singular de nuestra especie, pues es casi imposible no encontrarla en cada rincón habitado de nuestro planeta, siendo una evidencia contrastable que nos ayudamos unos a otros con acciones cooperativas. Una explicación a este fenómeno de ayuda mutua tan extendido llamado «reciprocidad directa» la podemos hallar, según Nowak (2012), «en que la vida no se limita a una lucha por la supervivencia, sino que es también una colaboración por la supervivencia» (págs. 20-21).

Las investigaciones de F. de Waal (2014) sobre la cooperación en primates nos muestran algunas pistas sobre el funcionamiento de la cooperación en las sociedades humanas, las cuales remarcán tres aspectos fundamentales: «(a) que tal conducta no depende de vínculos familiares; (b) que la cooperación se

basa a menudo en la reciprocidad, pues los chimpancés recuerdan los favores recibidos; y (c) la cooperación puede estar motivada por la empatía» (págs. 56-57). Estos tres puntos sintetizan los beneficios de la cooperación y, posiblemente, sean una explicación a su extensión por la faz de la tierra.

Si bien la obra de Edward O. Wilson, *Sociobiología. La nueva síntesis*, ha estado marcada por largas controversias que llegan hasta nuestros días, lo cierto es que no podemos obviar (o por lo menos sospechar, según el grado de escépticismo de cada uno) la influencia que tiene la biología en el desarrollo social, sobre todo, su participación en los procesos adaptivos, favoreciendo e integrando las diversas creaciones culturales con el medio ambiente en un tenso equilibrio, por ejemplo, en la organización del trabajo, las formas de gobierno o nuestras relaciones cotidianas con amigos (Tomasello, Marengo, & Dweck, 2010, pág. 124). Romper el equilibrio nos puede abocar a sufrir diversas catástrofes, tales como pandemias o cambios severos en el clima.

Desde los primeros asentamientos agrarios hasta las redes sociales de hoy en día, la evolución nos ha mostrado, desde sus múltiples caras, que el camino de la supervivencia necesita de la cooperación. Si bien no contamos con evidencias concluyentes de cómo fueron los comienzos de la sociabilidad y la memoria autobiográfica, podemos atisbar, a partir de la reconstrucción de fragmentos incompletos, la historia de nuestra identidad como especie, la cual nuestros antepasados de todas las épocas han dejado plasmada en diversas manifestaciones sus experiencias de vida. Perdurar en el tiempo, de la forma que sea, es un anhelo que todos llevamos dentro, pues queremos ser recordados y para ello contamos con mecanismos cognitivos que la propia evolución nos ha dotado.

Para concluir, podemos decir sin temor a equivocarnos que la organización humana necesariamente es social. Nuestra estructura genética, bioquímica y sistemas cognitivos nos predisponen para el proceso de sociabilidad y no podríamos hacerlo de otra manera, ni siquiera evitarlo. Esta conspiración bio-social sería imposible sin la estructura cognitiva de la memoria, pues ella nos permite almacenar, recuperar y procesar información, en algunos casos tan esencial como el reconocimiento de los demás y el propio reconocimiento a través de la identidad. Algo fundamental que no debemos pasar por alto, también la memoria

nos permite proyectar hacia el futuro, de esta manera, subimos un peldaño en la escala de organización para ubicarnos en la cultura.

Hasta aquí el segundo eje del trabajo dedicado a la estructura social de nuestra naturaleza humana o sociabilidad. En el siguiente capítulo se abordará el tercer eje de la investigación donde se expondrán algunos de los debates que son centrales en la historia de la filosofía contemporánea como lo son, por ejemplo, la relación mente-cuerpo, el conductismo lógico y el materialismo. Considero que desde este punto podría plantearse la posibilidad de una disciplina neurofilosófica capaz de interpretar la relación memoria autobiográfica y sociabilidad.

Capítulo 7. Neurofilosofía

Con este capítulo iniciamos el recorrido del tercer eje de la investigación con el objetivo de delinear un marco teórico adecuado que nos permita construir una base donde asentar la investigación sobre la relación entre la memoria autobiográfica y la sociabilidad. En tal sentido, creo que la neurofilosofía, una joven disciplina en eclosión y con muchas áreas de investigación interconectadas, puede ofrecer un marco teórico adecuado a esta investigación, asimismo, preguntas y respuestas de cómo pueden articularse dos conceptos que solo en apariencia (y remarco esto último), parecen no estar comunicados: hablo de nuestros recuerdos y nuestras experiencias cotidianas con el medioambiente bio-psicosocial.

Con el nacimiento formal de la neurofilosofía, allá en la década de 1980, han aflorado diversas perspectivas que buscan explicar la relación entre la mente, el cerebro y la conducta humana (Diodado, Rodriguez Valls, & Arana, 2010). El alcance de estas investigaciones nos ha abierto un camino para investigar y preguntarnos sobre la configuración de la identidad humana y hasta dónde podemos llegar con las especulaciones filosóficas.

En una primera aproximación general, y apoyándonos en la obra *Neurophilosophy: toward a unified science of the mind-brain* de Patricia Churchland (1986), podemos definir a la neurofilosofía como una interfaz entre la filosofía, las neurociencias y la psicología. En esta relación, la filosofía aporta un elemento esencial para la revisión crítica de los conceptos básicos trabajados por las neurociencias y la psicología. La importancia de una revisión crítica viene suscitada por el desarrollo de las neurociencias que, en algunos casos, ha mostrado la incompatibilidad con algunos conceptos muy arraigados en la psicología. Asimismo, algunas posturas materialistas dentro de la neurofilosofía, que dan preeminencia al factor cerebral-nervioso, entran en controversia con otras de corte mentalista. Hoy en día, es abundante la bibliografía que aborda estas dos posiciones.

Si bien la neurofilosofía representa un revulsivo para el resto de disciplinas filosóficas, una idea central que defenderé a lo largo de todo nuestro itinerario, en muchos casos las estructuras burocráticas que sostienen el edificio de la investigación en nuestro medio socio-político representan un escollo infranqueable para el desarrollo de una metodología inter-transdisciplinar que combine campos teóricos y experimentales, tales como los hace nuestra disciplina en otras geografías. Este reclamo no surge de una obstinada visión particular, sino del éxito probado que tiene este modelo en entornos que apoyan sin restricciones una comunicación más fluida entre filósofos y científicos que, a la vista de los resultados, están reportando grandes beneficios —y no solo económicos— para la sociedad en general. No obstante, dejaré aparcada esta discusión para otro espacio más adecuado.

En la introducción del trabajo comentamos que abordar el estudio de la relación entre la memoria autobiográfica y la sociabilidad desde la neurofilosofía implica revisar conceptos, métodos de investigación y tecnología empleados por la filosofía, las neurociencias, las ciencias experimentales y las ciencias sociales. Por esta razón, considero que el camino más adecuado para emprender la investigación sobre los aspectos claves de la memoria autobiográfica que intervienen en la construcción de nuestro *ethos* social es el que señala la neurofilosofía, definida como una síntesis filosófica y neurocientífica.

En líneas generales, los objetivos que me he propuesto para este capítulo son: (1) exponer el marco teórico de la neurofilosofía desde la perspectiva materialista eliminativista y reduccionista asumida por Patricia S. Churchland, (2) analizar el alcance y la fortaleza de una posible «teoría unificada de la mente-cerebro», y (3) explicar por qué considero que esta perspectiva es la más adecuada para explorar el estudio de la relación memoria autobiográfica-sociabilidad en la filosofía contemporánea. De fondo estará presente el problema arrastrado desde los tiempos de Demócrito —y los atomistas— y Sócrates hasta nuestros días: la residencia del pensamiento, alma inmaterial o física de átomos, el cual está envuelto en un doble debate, filosófico y científico, aún en discusión.

7.1. Entre filosofía y neurociencia

«Los hombres deben saber que es del encéfalo, y solo del encéfalo, de donde surgen nuestros placeres, alegrías, risas, bromas, así como nuestras penas, dolores, tristezas y lágrimas. Concretamente, a través de él, pensamos, vemos, oímos y distinguimos lo malo de lo bueno, lo agradable de lo desagradable... Es lo mismo que nos vuelve locos o delirantes, nos inspira miedo o pavor, ya sea de día o de noche, nos produce insomnio, errores inoportunos, preocupaciones sin sentido, falta de motivación y actos contrarios a nuestros hábitos. Estas cosas que sufrimos proceden todas del encéfalo, cuando no está sano, pero se torna anormalmente caliente, frío, húmedo o seco, o sufre cualquier otra afección no natural a la que no está acostumbrado. La locura procede de su humedad. Cuando el encéfalo está anormalmente húmedo, necesariamente se mueve, y cuando se mueve ya no hay vista ni oído, pero vemos u oímos ahora una cosa y ahora otra, y la lengua habla según las cosas que se ven y se oyen en cada ocasión. Sin embargo, cuando el encéfalo está tranquilo, un hombre puede pensar correctamente»².

El texto precedente atribuido a Hipócrates nos revela que los antiguos griegos ya «eran conscientes de la relación entre las lesiones de la cabeza y la actividad convulsiva que afectaba al lado contrario del cuerpo» (Kandel E. R., 2001, pág. 911), más aún, llegaron a describir el «principal trastorno del estado de ánimo, la depresión unipolar» (Kandel E. R., 2001, pág. 1209).

Si bien la descripción de Hipócrates es en parte correcta, hemos de suponer que el objetivo de Kandel (2001) no va más allá del hecho de citar una referencia histórica que confirme que el interés por el encéfalo, como órgano central de la vida humana, no es una invención original y reciente de las neurociencias, tal y como afirma en el siguiente párrafo:

² Texto atribuido a Hipócrates de Cos, siglo V a. C., extraído de Principios de neurociencia de E. Kandel y citado en la bibliografía.

La neurociencia actual es neurociencia *cognitiva*, fusión de la neurofisiología, la anatomía, la biología del desarrollo, la biología celular y molecular, y la psicología cognitiva. Esta disciplina se fundamenta en la idea, establecida inicialmente por Hipócrates hace más de dos mil años, de que el estudio apropiado de la mente comienza en el cerebro. (pág. 1165)

También los médicos del Antiguo Egipto dejaron constancia de su interés por el cerebro. Un papiro traducido por James Breasted en 1930 y datado en el siglo XVII a. C. —posiblemente una copia de un original redactado hacia el año 3000—, describe «una lista de cuarenta y ocho casos de heridas en la cabeza y en el cuello [...] incluyendo en cada caso: título, examen, diagnóstico y tratamiento» (Changeux, 1985, pág. 15). Además, este papiro «contiene la primera mención del encéfalo de todos los documentos del ser humano» (Kandel E. R., 2001).

A principios de la década de 1980, el neurobiólogo francés Jean-Pierre Changeux reconoció en *El hombre neuronal* que, mientras el conjunto ciencias humanas avanzaban, las investigaciones sobre el sistema nervioso parecían entrar en un estado de estancamiento a causa de la complejidad del objeto y a cuestiones ideológicas, tanto políticas como religiosas, pues «toda investigación que, directa o indirectamente, se relacione con la inmaterialidad del alma pone la fe en peligro y está destinada a la hoguera» (Changeux, 1985, págs. 10-11). Hoy en día, es evidente que aquel «fervor ideológico» que envolvía en las tinieblas de la «superstición» a las ciencias del hombre hace unas décadas atrás, comienza a disiparse a fuerza de datos irrefutables (Churchland P. S., 2012, pág. 14). Sin embargo, a pesar de los esfuerzos de la ciencia, algunas tinieblas no han desaparecido en su totalidad.

Ahora bien, teniendo en cuenta el desarrollo de las neurociencias podemos preguntarnos ¿constituyen la mente y el cerebro dos entidades diferentes y por lo tanto deben ser estudiadas por separado?, ¿dónde está el límite entre ambas?, ¿son una y la misma cosa, pero con funciones diferentes en diferentes niveles? La complejidad de las posibles respuestas supone hacer frente a una problemática que la filosofía, entre otras muchas prestigiosas disciplinas, no ha podido —ni podrá— agotar por sí misma a través de los métodos introspectivos

o intuitivos de análisis, pues necesitará contar con la colaboración de las ciencias experimentales. Además, el problema para la neurociencia es su dependencia absoluta de la materia física y la ambigüedad para trabajar algunos de sus conceptos fundamentales, por ejemplo, «consciencia», «memoria» o «mente» que, por la dinámica de las investigaciones, se modifican con gran rapidez.

Por contra, en los orígenes de la filosofía, el problema de la mente se mezcló con el concepto de alma, arrastrando al pensamiento hacia una concepción de estructura mental dependiente de un espíritu inmaterial, lo cual hacía imposible atisbar cualquier mecanismo (físico) subyacente (Gazzaniga M. , 2019, pág. 23), una entidad a la que se puede acceder a través de la intuición.

Tomemos el camino de la filosofía, el de la psicología o el de la moderna neurociencia, cualquiera de ellos nos conducirá hacia un mismo punto de partida: René Descartes. Si bien es heredero de una larga tradición, para algunos, culpable —para otros no tanto— de haber separado la mente del cuerpo, el esfuerzo de Descartes por hallar un mundo de «ideas claras y distintas» es de valorar y no merece ser juzgado *a priori* de una forma intempestiva. Tampoco podemos caer en el juego de acusaciones retrospectivas, sino intentar entender el desarrollo de este difícil tema, si bien con profundos matices, aún en discusión.

Haciendo un poco de historia podemos ver que el comienzo del siglo XX fue el testigo privilegiado de nueva discusión entre empiristas y racionalistas sobre la naturaleza de la conciencia y los estados mentales. En 1907, el filósofo y psicólogo americano Williams James (1842-1910) indagó sobre la posibilidad de que un estado mental —por ejemplo, un deseo o una idea— pudiese afectar al cerebro, es decir, a la materia. (Gazzaniga M. , 2019, págs. 75-76). Posicionado desde el pragmatismo filosófico, James comenzó un movimiento contra el conductismo y el empirismo.

Los conductistas, entre los que se encontraban John B. Watson y B. F. Skinner, rechazaron cualquier explicación derivada de la investigación introspectiva que tuviese como objeto de estudio la experiencia mental y subjetiva, constituyendo como único camino de investigación la observación de la conducta a partir del método propuesto por las ciencias naturales.

El desacuerdo en las posiciones intelectuales se materializó en la década de 1950 con la irrupción en escena del cognitivismo, movimiento liderado por George A. Miller, Roger W. Sperry Allen Newell, Herbert Simon y Noam Chomsky (Gazzaniga M. , 2019, págs. 80-81). El propio Chomsky, a partir de las reglas gramaticales del lenguaje, fue el encargado de fulminar el concepto de «mente como *tabula rasa*» que sostenía las interpretaciones conductistas.

Según comenta Michael Gazzaniga (2019, pág. 85), las derivaciones psicológicas de las investigaciones de Chomsky sirvieron a G. Miller para explorar el funcionamiento del cerebro y la mente como un todo integrado. Este es el comienzo de una nueva etapa en el estudio de conciencia donde filósofos de la mente y neurocientíficos tendrán mucho que decir y, sobre todo, decirse.

Los extensos debates entre las diversas disciplinas científicas y filosóficas que han tenido como epicentro la relación mente-cerebro no han sido fáciles, lo cierto es que sus ecos aún suenan con fuerza. Está claro que llegar a acuerdos epistemológicos exentos de posturas —sean científicas o ideológicas— resulta una tarea imposible, sobre todo porque ni siquiera es posible llegar a un acuerdo dentro de una misma disciplina. Podemos suponer que esta situación no se debe solo a aspectos subjetivos de deliberación y entendimiento, sino a que los propios objetos de estudios —mente y cerebro— son inabarcables desde una mera conceptualización subjetiva.

Para cerrar este apartado me gustaría hacerlo con una cita del neurocientífico Francis H. Crick (1979) que he extraído del libro de M. Gazzaniga *El instinto de la conciencia* (2019), unas palabras que, en cierto sentido, arrojan un pequeño haz de luz sobre el estado de la cuestión en la relación filosofía-neurociencia:

No todo el mundo es consciente de que tiene un punto ciego, aunque es fácil de demostrar. Lo extraordinario es que no vemos un hueco en nuestro campo visual. La razón es, en parte, que no tenemos forma de detectar los bordes del hueco, y, en parte, que nuestro cerebro llena ese hueco con información visual que toma prestada de la zona inmediatamente circundante. Nuestra capacidad para engañarnos a nosotros mismos con respecto al funcionamiento de nuestro cerebro es casi ilimitada, sobre todo porque aquello de lo que podemos informar es solo una diminuta

parte de lo que ocurre en nuestra cabeza. De ahí que una gran parte de la filosofía haya sido estéril durante más de dos mil años, y es probable que siga siéndolo mientras los filósofos no aprendan a entender el lenguaje del procesamiento de la información.... [...] Eso no significa, sin embargo, que el estudio de nuestros procesos mentales mediante la introspección deba abandonarse por completo, como han pretendido hacer los conductistas. Hacer eso sería descartar uno de los atributos más significativos de lo que estamos tratando de estudiar. Pero el hecho es que la evidencia de la introspección nunca debe aceptarse al pie de la letra; debe explicarse en términos que no sean meramente los suyos propios... [...] El sistema nervioso superior parece ser una combinación extremadamente ingeniosa de cableado de precisión y redes asociativas... [...] La red se divide en muchas subredes de menor tamaño, algunas de ellas en paralelo, otras organizadas de forma más secuencial. Además, la parcelación de las subredes refleja tanto la estructura del mundo, externa e interna, como nuestra relación con él. (pág. 101)

El desafío que plantean tanto la neurociencia como las ciencias cognitivas está escribiendo un nuevo capítulo dentro de la historia de la filosofía contemporánea: estudiar la relación entre el funcionamiento y cómo piensa nuestro cerebro. Incorporar nuevos datos provenientes de la investigación experimental a los estudios teóricos y conceptuales nos ofrece la oportunidad de nuevas preguntas. Sin embargo, hallar una respuesta sólida capaz de acallar cualquier controversia no será fácil.

7.2 El problema del dualismo mente-cuerpo

Preguntarnos por la mente nos conduce de forma más o menos directa hacia un laberinto cognitivo por donde discurren la inteligencia, la conciencia, los sueños, los recuerdos y hasta la fantasía. Al igual que Sísifo debemos sobreponernos de la fatiga que supone, día tras día, la tarea de intentar clarificar el significado volátil de estos conceptos que, sumados al de «percepción», «memoria» y «razonamiento», hacen de la inteligencia una cualidad humana compleja e inasible en toda su magnitud (Damasio A. , 2021, pág. 45). No obstante,

las indagaciones filosóficas sobre la mente y sus cuestiones son tan antiguas como la propia filosofía (Pineda, 2012, pág. 11).

Para algunos investigadores centrados en diversos campos de las neurociencias, por ejemplo M. Gazzaniga (2019), uno de los grandes problemas de la filosofía ha sido la dependencia de un concepto inmaterial de la mente:

[...] Pero los antiguos no tenían un concepto explícito de conciencia. Les interesaba saber cómo funcionaba la mente, de dónde los pensamientos e incluso si ello implicaba un proceso puramente físico, pero la mayoría de los antiguos pensadores acabaron llegando a la conclusión de que la vida mental era el producto de un espíritu inmaterial. Y, cuando la conciencia se define como un espíritu inmaterial, se hace difícil empezar a pensar en los mecanismos subyacentes. (págs. 22-23)

Quizás, en parte, podemos estar de acuerdo con el neurocientífico y asumir el exceso de «inmaterialidad» con que nuestras disciplinas filosóficas han fundamentado las investigaciones sobre la mente. Sin embargo, sería faltar al sentido de justicia el realizar valoraciones retrospectivas y juzgar el desempeño teórico en las investigaciones sobre la mente que hemos realizado en la historia de la filosofía. Ahora bien, el desarrollo de las ciencias cognitivas y las neurociencias a lo largo del siglo XX, y todo lo que va del XXI, han permitido a la filosofía explorar nuevas áreas conceptuales que están reformulando el concepto «mente» y entender con mayor profundidad el funcionamiento del cerebro. Está claro que el interés por los resultados científicos en este campo ha sido motivado por diversos factores y ha adoptado múltiples perspectivas, desde opiniones favorables hasta las críticas más duras.

Para comenzar, el primer escollo que encontramos en el camino es el concepto de «mente». En la actualidad, las discusiones filosóficas no terminan de unificar criterios bajo una posición clara y considerar si el concepto «mente» designa una función, una consecuencia, un órgano, una invención de algo que no sabemos qué es ni cómo funciona, o bien, ninguno de los anteriores. Intentar desvelar este concepto es una tarea, por no decir la principal, de la filosofía de

mente, una disciplina vinculada, muy estrechamente, con la filosofía de la ciencia, la filosofía del lenguaje y las ciencias cognitivas. Tal es la complejidad de la cuestión que no solo cuesta poner de acuerdo a filósofos y científicos entre sí, sino a filósofos entre sí y científicos entre sí.

La filosofía de la mente es una heredera directa de la *philosophy of mind* nacida en la llamada tradición analítica o anglosajona, cercana en muchas áreas al desarrollo científico. Sin embargo, como bien hace notar Pineda (2012, pág. 13), entre el concepto *mind* y *mente* existen grandes diferencias, siendo su objetivo fundamental «ocuparse de los problemas filosóficos de la mente y que el empeño de resolver tales problemas es continuo cuando no coincidente con el empeño científico de comprender cómo funciona la mente humana» (Pineda, 2012, pág. 14). Podemos estar más o menos de acuerdo con esta cercanía entre las disciplinas científicas y filosóficas. Sin embargo, lo que no podemos es negar la necesidad de seguir de cerca el desarrollo científico y los resultados que tienen un impacto directo en este campo.

La relación entre la filosofía de la mente y las neurociencias se comprende con mayor claridad, tanto sus coincidencias como sus diferencias, cuando expomos los objetivos de estudio de la filosofía de la mente y la búsqueda de respuestas en saber qué son los estados mentales y su relación con los estados físicos, es decir, entre la mente y el sistema nervioso (Pineda, 2012, pág. 15).

La situación descrita se hace más evidente cuando analizamos, por ejemplo, «el concepto» y el «funcionamiento» (o fisiología). Aquí el trabajo de investigación entre la ciencia y la filosofía muestra una simbiosis, una necesidad de convivencia pacífica entre saberes que en algún punto serán complementarios. Pues, como bien comenta Pineda (2012):

Lo más importante para responder a la pregunta sobre la pertinencia de los estudios filosóficos acerca de la naturaleza de nuestros estados mentales es darse cuenta de que el análisis de los conceptos psicológicos es necesario para llevar a cabo con éxito la tarea de aclarar en qué consisten los estados mentales. La razón no puede ser más simple: si desconocemos cómo funcionan los conceptos difícilmente podemos averiguar la naturaleza de aquello denotado por ellos. (págs. 15-16)

Una de las cuestiones fundamentales que ha mantenido ocupados a filósofos de la mente, neurofilósofos, psicólogos y neurocientíficos ha sido determinar si existe una relación causal en entre los estados mentales y los estados cerebrales. A ello debemos sumar el problema de la «introspección o modo introspectivo», es decir, «el modo de conocimiento inmediato y no inferencial de nuestros estados mentales que es privado y nos confiere cierta autoridad epistemológica» (Pineda, 2012, pág. 18). Huelga decir que nos hallamos ante un problema ambiguo: si la respuesta es positiva (relación causal mente-cerebro), necesitamos un camino firme en cuanto a evidencias. En cambio, si es negativa, estamos frente a una paradoja (Pineda, 2012, pág. 54).

En tal sentido, es interesante el aporte descriptivo del concepto «mente» o «mental» realizado por Carlos Moya como «un conjunto bastante heterogéneo de propiedades y estados: sensaciones, creencias, deseos, sentimientos, emociones, intenciones, decisiones, rasgos de carácter, disposiciones y habilidades diversas» (2004, pág. 18). Si bien la descripción hecha por Moya es pertinente, necesitamos un soporte físico, pues como veremos, no somos pensamiento en extensión.

Volviendo a la cuestión del origen del problema, la mayor parte de los estudios nos conducen hasta René Descartes. Sin embargo, la cuestión del dualismo mente-cuerpo ya se encontraba presente en la filosofía de Platón y Aristóteles (Pineda, 2012, pág. 63). Dentro de la filosofía contemporánea los comienzos de la reflexión sobre lo mental suele atribuirse a G. Ryle y su obra *The Concept of Mind* publicada en 1949 (Rodríguez González, 2021, pág. 16).

Encontrar una línea causa-efecto entre los procesos mentales y los procesos corporales nos remite al problema central planteado por Descartes, como comenta Rodríguez González (2021):

El problema mente-cuerpo es el problema ontológico fundamental de la Filosofía de la Mente. Muchos especialistas subrayan que desde Descartes se formula sobre todo como el problema de cómo son posibles las relaciones causales entre los sucesos mentales y los corporales. Es decir,

¿cómo pueden los pensamientos y los actos voluntarios provocar movimientos corporales? O también, en el sentido inverso, ¿cómo es posible entender la estimulación puramente física de nuestros órganos sensoriales cause efectivamente sensaciones subjetivas y experiencias perceptivas? (pág. 25)

La doctrina cartesiana de la mente es concebida como un dualismo substancial que separa con claridad el cuerpo y la mente en dos entidades autónomas (Moya, 2004, pág. 27). Sin embargo, señala Moya, «el camino que conduce a Descartes al dualismo substancial es epistemológico: se trata de una reflexión sobre la verdad y la justificación de nuestras creencias» (2004, pág. 27). Las consecuencias de este dualismo radical es que el «yo pensante» diseñado por Descartes es una entidad distinta del cuerpo con existencia propia (Moya, 2004, pág. 29). Tal es esta distinción y tan profunda es la brecha entre la materia y el pensamiento que es imposible resolver de forma satisfactoria la interacción entre el cuerpo y la mente (Moya, 2004, pág. 37).

El dualismo cartesiano queda atrapado en una aporía que le hace incompatible con el fisicalismo, ello queda reflejado cuando intenta desarrollar la interacción entre particulares físicos y los datos de la conciencia que son particulares no físicos (Moya, 2004, pág. 41). Encontrar una perspectiva sólida a esta situación no es fácil, ya que su problema se viene arrastrando desde entonces, como bien lo explica José Hierro (2005):

El problema de la relación entre el cuerpo y la mente, que Descartes no supo solucionar, ha permanecido siempre en el trasfondo de la teorización posterior sobre la mente. En algunos casos, se ha planteado como la cuestión de la relación entre la mente y la materia, y más aún, como cuál es el significado de la mente en la naturaleza, o si se prefiere, cuál es el puesto del hombre en esta. (pág. 20)

Sin embargo, en el desarrollo contemporáneo que engloba los principales problemas de la relación mente-cuerpo podemos ver, como era de esperar, que no todas las opiniones han coincidido en una misma posición. Tal es el caso del filósofo de Oxford Stephen Priest para quien el problema del dualismo no ha sido

superado por el desarrollo de la ciencia moderna, en su opinión, ha sucedido todo lo contrario. Como ejemplo cita las investigaciones de John Eccles o Karl Popper, a quienes considera que se han tomado más en serio la oscura realidad de lo mental (Priest, 1994, pág. 18).

Al analizar en profundidad el desarrollo cartesiano de la relación mente-cuerpo, José Hierro-Pescador (2005) señala que el error de Descartes ha sido «imaginar la mente, por analogía con el cuerpo, como algo que obedece a las leyes físicas y funciona de la misma manera», este punto derivará, como se puede atisbar, en la imposibilidad de hallar ese «punto físico de contacto» entre ambas sustancias (pág. 23).

Según Pineda (2012), la mente para Descartes es una entidad no extensa y por lo tanto no puede tener partes ni ocupa espacio (pág. 80). Este es un punto fundamental de discusión en la neurofilosofía contemporánea, ya que mente y cerebro no pueden pensarse como entidades separadas. Será, pues, el dualismo cartesiano el «*planteamiento mismo del problema*, antes que una efectiva propuesta de solución» (Rodríguez González, 2021, pág. 26).

7.3 El conductismo lógico

Como hemos visto en el apartado anterior, el dualismo trajo consigo grandes desafíos a la filosofía, muchos de los cuales continúan si resolución. Si bien las controversias se diluyen en el tiempo, no desaparecen por completo.

Otra posición dentro la filosofía que planteó grandes desafíos fue el llamado conductismo lógico, «una tesis acerca de la naturaleza de los conceptos psicológicos» (Pineda, 2012, pág. 112) que, según Priest (1994), «hallarse en un estado mental es hallarse en un estado de conducta o de comportamiento» (pág. 55). Estos conceptos psicológicos nos permiten tanto la auto-atribución de estados mentales como la atribución a los demás, es decir, indagar en los más profundo de la vida mental de nosotros mismos y de los seres que nos rodean, y sobre todo, comprenderlos y comprendernos (Pineda, 2012, pág. 113). En este

punto es muy interesante la diferenciación que realiza José Hierro (2005) entre el conductismo lógico y el conductismo psicológico:

Llamaré «conductismo psicológico» a aquel enfoque de la Psicología que toma como objeto de ésta a la conducta de los organismos, prescindiendo de la mente y los estados mentales, o bien explica éstos como efecto de la influencia del medio ambiente. Es la posición que representan claramente Watson y Skinner. Llamaré «conductismo lógico» a aquella explicación del discurso mentalista (el discurso sobre la mente y los estados mentales) que recurre a manifestaciones de conducta o a disposiciones para ella. Sus representantes más conocidos son Wittgenstein en su segunda etapa, Ryle y Hempel. (pág. 48)

El conductismo psicológico desarrollado por Watson era una ciencia de la mente y de los estados mentales basada en el método introspectivo de investigación (Hierro-Pescador, 2005, pág. 50). Uno de los problemas fundamentales que encierra el método introspectivo es la complejidad epistemológica para encontrar un método válido de investigación.

Para Skinner, sin embargo, el conductismo es, básicamente, «una doctrina filosófica, no la ciencia de la conducta humana sino la filosofía de esa ciencia» (Hierro-Pescador, 2005, pág. 50), el cual distinguiría dos formas: el conductismo metodológico y el conductismo radical.

Si el conductismo metodológico excluía la introspección como un procedimiento válido de confirmación debido a su carácter privado y a la falta de intersubjetividad, el conductismo radical, en cambio, tal como Skinner lo entiende, acepta la autoobservación pero cuestiona la naturaleza de lo que es observado, limitando el contenido de la introspección al interior del cuerpo propio del observador y excluyendo de la observación una supuesta vida mental. (Hierro-Pescador, 2005, págs. 50-51).

En la obra *El concepto de lo mental* (1949), Gilbert Ryle expone la teoría del conductismo lógico y su crítica al dualismo cartesiano, una posición apartada de la concepción fisicista y del conductismo lógico de Hempel, «que consiste en

sustituir los enunciados mentalistas por enunciados sobre la conducta» (Hierro-Pescador, 2005, pág. 51).

El conductismo lógico es una tesis acerca de la naturaleza de los conceptos psicológicos... [...] los conceptos psicológicos son aquellos que utilizamos necesariamente en la auto-atribución de estados mentales a los demás. Son pues conceptos indispensables a la hora de describir nuestra vida mental y la de nuestros semejantes, y su manejo permite comprender al resto de los seres humanos con los que tenemos que convivir. Son el tipo de conceptos que utilizamos para clasificar los estados mentales en actitudes proposicionales, estados fenoménicos y emociones, esto es, conceptos como creencia, deseo, duda, temor, dolor, sensación cromática, amor, alegría, enojo, asco, etc. La tesis nuclear del conductismo lógico es que los conceptos psicológicos son una variedad de conceptos que denotan ciertas disposiciones a la conducta. (Pineda, 2012, págs. 112-113)

Según Pineda (2012), el conductismo lógico concibe al dualismo como «un error conceptual que proviene de una mala comprensión del funcionamiento de los conceptos psicológicos» (pág. 116). Para Ryle esto constituye un «error categorial», es decir, una ilusión producida por el mal uso del lenguaje ordinario (Priest, 1994, pág. 66).

Ludwig Wittgenstein fue otro de los filósofos que se enfrentó con el dualismo cartesiano, no obstante, sería un error ubicarlo dentro del conductismo lógico por la complejidad de su obra (Priest, 1994, pág. 57). Con respecto a la relación entre el conductismo lógico, Ryle y Wittgenstein, José Hierro (2005) comenta lo siguiente:

La influencia de Wittgenstein sobre Ryle no debe sugerir la idea de que las afirmaciones de Ryle sean reductibles a las de Wittgenstein, pues hay puntos en que no coinciden. Tal idea de reducción parece sugerida por el hecho de colocar a ambos autores bajo el nombre de «conductismo lógico», como estoy haciendo yo... [...] Me parece justificado considerar la posición de Ryle como una forma de conductismo lógico, aunque distinta de la de Hempel. Más cerca está de Wittgenstein, pues ambos rechazan el fisicismo de los neopositivistas como Hempel. Recordaré que entiendo

el conductismo lógico como doctrina que mantiene que para explicar expresiones sobre la mente y los estados mentales hay que recurrir a las manifestaciones de la conducta o a disposiciones para ella, y me importa recordarlo aquí porque quiero mencionar que algunos comentaristas no aceptan calificar de conductismo lógico la teoría de Ryle o la de Wittgenstein. Pienso que esto se debe a haber tomado el conductismo lógico en el sentido fisicista que tiene Hempel, y que ciertamente es ajeno a Ryle y a Wittgenstein. (pág. 55)

La obra *Investigaciones filosóficas* de Ludwig Wittgenstein (1988) tuvo una influencia muy fuerte sobre la problemática cuerpo-mente. Según comenta Hierro (2005), «el problema central para Wittgenstein es un problema filosófico-lingüístico: cómo explicar el significado de las expresiones que designan estados o procesos mentales» (pág. 55). La obra de Wittgenstein constituye una crítica a la concepción cartesiana de la mente, siendo sus propiedades puramente intrínsecas «en el sentido que no guarda ninguna relación interna o constitutiva con ningún proceso o entidad no mental como puede serlo el comportamiento o el mundo físico» (Moya, 2004, pág. 62).

Una de las posiciones más fuertes contra el dualismo cartesiano lo representa el argumento del último Wittgenstein contra la «posibilidad de un lenguaje privado» (Pineda, 2012, pág. 136), esto es, un lenguaje que no puede ser expreso vocalmente y que solo es conocido por uno mismo, siendo el caso de las vivencias internas (Priest, 1994, págs. 81-82).

Si bien el interés de Wittgenstein estuvo centrado en el lenguaje, el problema citado está focalizado en el problema mente-cuerpo. Según Pineda (2012), por lenguaje privado se entiende:

[...] un lenguaje epistémicamente privado, esto es, un lenguaje necesariamente no enseñable, o sea, un lenguaje necesariamente hablado por una sola persona, que no puede ser enseñado ni transmitido a otros. Un lenguaje de estas características sería un lenguaje cuyas palabras significasen entidades privadas solo accesibles al sujeto por introspección, esto es, el tipo de entidades con las que está comprometido el dualista cartesiano. La tesis de Wittgenstein puede formularse como la tesis de

que un lenguaje cuyos significados sean entidades privadas de este tipo es conceptualmente imposible, esto es, rechazable a priori, por razones conceptuales. (pág. 138)

Priest, en la misma línea, deja claro que la argumentación de Wittgenstein va contra toda posibilidad de lenguaje privado. Además, alerta sobre la importancia de dicha cuestión para la filosofía en general, pues «si tal lenguaje fuese posible, esas teorías tendrían que ser falsas» (1994, pág. 79). Sus alegaciones no pueden ser más claras en cuanto al alcance:

El argumento del lenguaje privado entraña un formidable poder de destrucción en filosofía. Si fuera válido, implicaría no solamente la refutación del dualismo, del idealismo (incluyendo al solipsismo), de la filosofía fenomenológica y del fenomenismo, sino que convertiría además en sinsentido el planteamiento de ciertas preguntas escépticas. Por ejemplo: la sugerencia de que no podemos saber que otras personas piensan ni que piensan y la sugerencia de que la experiencia que tú tienes puede ser absolutamente diferente de la mía parecen, igualmente, presuponer un lenguaje privado. Añadamos que, de acuerdo con más de una teoría influyente en filosofía del lenguaje, el significado de una palabra es una idea, es decir, algo «interior», privado y psicológico. Por otra parte, la tendencia filosófica denominada *empirismo*, globalmente considerada, es esencialmente la teoría de que todo el conocimiento que uno tiene, incluyendo el de su propio lenguaje, se deriva de la experiencia. Si la experiencia es «privada» entonces, de acuerdo con el empirismo, también lo sería el lenguaje. Síguese, pues, que si la empresa de Wittgenstein se viera coronada por el éxito, infligiría grave daño a las posturas de filósofos tan diversos como Descartes, Locke, Berkeley, Hume, Schopenhauer, Husserl, Russell y Ayer. (Priest, 1994, pág. 81)

La extensión del argumento sobre el lenguaje privado de Wittgenstein socava los cimientos que fundamentan la imagen cartesiana del mundo: la certidumbre y seguridad de la primera persona pensando, pues «la inteligibilidad de que con respecto a sí mismo según Descartes, gozaba su propia mente, vendría a gravitar toda ella sobre un vasto supuesto: un lenguaje público y el mundo de sentido común de los comunicadores humanos» (Priest, 1994, pág. 87).

Como podemos ver, las diferencias expuestas nos revelan la compleja barrera que debe superar cualquier teoría que aspire a unificar la relación mente-cerebro. En tal sentido, el conductismo lógico no ha hecho más que abrir una puerta más hacia la búsqueda, quizás, de un imposible. Para cerrar el apartado, creo conveniente esta cita sobre Wittgenstein de José Hierro (2005):

Wittgenstein está señalando el problema básico para la concepción conductista. La conducta solamente tiene valor psicológico si tiene algún significado, pero no lo tendrá al menos que trate de algo. ¿Y de qué trata la conducta? Aquello de lo trate la conducta, ese será objeto de la Psicología. ¿Cómo evitar que reaparezca la mente? (pág. 57)

Las palabras de Wittgenstein nos introducen en la encrucijada de hallar «la casa del significado» más allá de la conducta, o bien, otorgarle a la conducta algún sentido, el cual no puede estar más allá de la mente.

7.4 Materialismo eliminativo

La pertinencia de incluir un apartado dedicado al materialismo eliminativo está justificada por las líneas históricas que lo unen con una de las corrientes filosóficas contemporáneas más difundidas, tal es el caso de la neurofilosofía bajo el desarrollo de las investigaciones de Patricia S. Churchland y el equipo de neurocientíficos de la Universidad de California San Diego (UCSD).

Asimismo, cabe mencionar que de las fuentes de la neurofilosofía beben algunas de las disciplinas más relevantes dedicadas a las investigaciones en el campo de la filosofía de la mente, la neuroética, las ciencias cognitivas o la inteligencia artificial, por poner algunos ejemplos.

Como he dicho más arriba, en este trabajo la inclusión del materialismo eliminativo halla su justificación por tratarse de un hilo conductor natural para entender las discusiones teóricas y epistemológicas que envuelven a toda la estructura sobre la que se ha levantado el edificio de la neurofilosofía en el sentido

que la concibe P. S. Churchland. De más está decir que la discusión continúa abierta, si bien en un segundo plano.

Entrando ya en el tema, tanto el materialismo como su versión más radical, el materialismo eliminativo, ponen de manifiesto que algo existe por el mero hecho de «ser físico» (Priest, 1994, págs. 125-126). A simple vista lo que acabamos de decir contiene una verdad irrefutable en la que casi todos estaríamos de acuerdo. Y digo «casi todos» porque en los trabajos que tratan la relación entre la mente, el cerebro y sus procesos, la cuestión presenta importantes desacuerdos y discusiones (como hemos visto en los apartados anteriores), por ejemplo, cuando hablamos de pensamientos, recuerdos o fantasías. Ahora bien, siguiendo a Priest (1994):

El materialista no niega que pensamos. Lo que dice es que nuestros pensamientos son físicos, que la vida mental que tiene cada uno de nosotros es, en toda su integridad, una serie de acontecimientos físicos; o, como enseña la más moderna versión del materialismo, un conjunto de procesos electro-químicos del cerebro. La materia puede pensar. (págs. 125-126)

La eclosión del materialismo eliminativo podemos encontrarlo en el «fracaso de la teoría de la identidad en su versión inicial (básicamente, Place-Smart-Feigl)» (Rodríguez González, 2021, pág. 62), también es conocida esta teoría como *materialismo reduccionista* (Churchland P. M., 1992, pág. 50).

La crítica implacable (e impecable) realizada Feyerabend hizo diana en un punto neurálgico de la teoría de la identidad, ni más ni menos que en los fundamentos que describen que no solo «los procesos mentales tienen rasgos físicos, sino también que los procesos cerebrales poseen rasgos mentales». En tal sentido, si la teoría de la identidad es falsa, entonces debemos concluir que no existen procesos mentales (Rodríguez González, 2021, pág. 62).

La cuestión principal que ofrecía serias dudas era saber si se podía hallar una relación conceptual clara y biunívoca entre los conceptos de la psicología corriente (*folk*) y los conceptos de la neurociencia teórica, todo ello a partir de

una reducción interteórica (Churchland P. M., 1992, pág. 75), sobre todo, si tenemos en cuenta que nuestro conocimiento sobre el funcionamiento del cerebro no nos permite enlazar las identidades correspondientes (Churchland P. M., 1992, pág. 50).

Un aspecto clave, mejor dicho, crítico, se halla en la *reducción interteórica*. Podemos ver cómo la propia historia científica ha nutrido de ejemplos a los teóricos de la identidad para fortalecer su posición:

Sabemos que el relámpago es idéntico a una repentina descarga en gran escala de electrones entre las nubes, o entre la atmósfera y la tierra. Lo que ahora consideramos que son «estados mentales», argumenta el teórico de la identidad, son idénticos estados cerebrales exactamente de la misma manera. (Churchland P. M., 1992, pág. 51)

Los viejos principios que sostenían el marco de referencia de la teoría anterior son reemplazados por nuevos principios, por ejemplo, «calor», «está caliente» y «está frío» por las nociones de «energía cinética molecular», donde la explicación y la predicción aportan una descripción de la realidad mucho más acertada y correcta que en el modelo anterior: solo le queda ser desplazada por una teoría que explique mejor tales actividades (Churchland P. M., 1992, págs. 51-52).

El fracaso de la teoría de la identidad se debió, sobre todo, a que parecía improbable que la aparición de «una teoría materialista adecuada trajera consigo las correspondencias biunívocas exactas» (Churchland P. M., 1992, pág. 75), un falso optimismo en la reducción de la psicología *folk* por parte de un programa de la neurociencia. Por su parte, el eliminativismo declara que la perspectiva de la psicología *folk* «es una explicación radicalmente inadecuada de nuestras actividades internas, una explicación demasiado confusa y demasiado defectuosa como para ganarse la supervivencia a través de la reducción interteórica» (Rodríguez González, 2021, págs. 62-63).

El problema que aprecia Paul Churchland (1992) con respecto a la posibilidad de que una teoría eliminativista pueda conciliar una reducción de la psicología popular a una neurociencia teórica está en la falsedad del propio marco de referencia corriente del cual se parte:

A juicio del materialismo eliminativo, no podrán encontrarse las correspondencias biunívocas, y no se podrá efectuar una reducción interteórica del marco de referencia psicológico corriente, porque *el marco de referencia psicológico que utilizamos corrientemente es una concepción falsa y radicalmente engañosa sobre las causas de la conducta humana y la naturaleza de la actividad cognitiva*. Desde esta perspectiva, la psicología habitual no solamente constituye una representación incompleta de nuestra naturaleza interna, son que directamente constituye una *mala* representación de nuestros estados y actividades internos. En consecuencia, no es posible esperar que una explicación neurocientífica verdaderamente adecuada de nuestra vida interior proporcione las categorías teóricas que se corresponden escrupulosamente con las categorías de nuestro marco de referencia habitual. Consecuentemente, lo único que se debe esperar es que el antiguo marco simplemente sea eliminado y no que pueda reducirse por una neurociencia más desarrollada. (págs. 75-76).

Siguiendo las palabras del propio Paul Churchland (1995), la definición y el alcance del programa que emprende el materialismo eliminativo parece dejar muy claras cuáles serán sus intenciones:

El materialismo eliminativo es la tesis que sostiene que nuestra concepción de sentido común acerca de los fenómenos psicológicos constituye una teoría radicalmente falsa, una teoría tan esencialmente defectuosa que tanto sus principios como su ontología serán eventualmente desplazados, más que reducidos con fluidez, por una neurociencia completa [*complete neuroscience*]. Nuestra comprensión mutua y aun nuestra introspección podrán ser entonces reconstituidas dentro del marco conceptual de la neurociencia completa; una teoría que esperamos sea mucho más poderosa que la psicología de sentido común [*common-sense psychology*] a la que desplaza y, en general, mucho más integrada a la ciencia física... [...] Este enfoque representa un cambio en la suerte del materialismo. Hace veinte años se consideraba que las emociones, los

qualia y las «vivencias puras» [«raw feels»] eran los obstáculos principales para el programa materialista. Al disolverse estas barreras, el *locus* de la oposición ha cambiado. Ahora se sostiene que es el ámbito de lo intencional, el ámbito de la actitud proposicional, el que es irreductible a todo lo que pertenezca el marco materialista, e ineliminable respecto de él. (págs. 43-44)

La propia historia de la ciencia ha suministrado a los defensores del eliminativismo muchos ejemplos para mantener que la eliminación y superación teórica es solo cuestión de tiempo y de desarrollo científico, tal es el caso de la *teoría del calórico* sostenida en el siglo XVIII y XIX, donde se creía que el calor era un *fluído* sutil contenido en los cuerpos (Churchland P. M., 1992, pág. 76).

La posición del materialismo eliminativo, si bien es muy discutible y criticada, se ha construido sobre una sólida cimentación. ¿Por qué no puede efectuarse una reducción del macro conceptual de la psicología popular a la neurociencia? La respuesta aportada por el eliminativismo es que la psicología popular está anclada en una concepción «primitiva y profundamente confusa de las actividades internas» (Churchland P. M., 1992, pág. 79).

Según Churchland, nuestras opiniones negativas sobre el alcance de los postulados de la psicología corriente, a la vez son argumentos a favor del materialismo eliminativo, están ancladas en tres razones: (a) imposibilidad de la psicología popular para explicar, predecir y manipular; (b) la complejidad de los temas abordados por la psicología popular hace que cualquier intento de desplazamiento sea muy difícil en el corto tiempo; y (c) existe una ventaja del eliminativismo frente a la *teoría de la identidad* y el *funcionalismo*, ya que tendrá que crear un conjunto de requisitos lo suficientemente exigentes para crear una nueva teoría, y que esta «refleje fielmente la estructura conceptual específica que se va a reducir» (Churchland P. M., 1992, págs. 81-82).

Y lo cierto es que existen infinitamente más medios de constituir una neurociencia con potencia explicativa que *no* refleje la estructura de la psicología corriente, que de hacerlo al mismo tiempo que se *refleja* la propia

estructura específica de la psicología popular. Consecuentemente, la probabilidad a priori del materialismo eliminativo no es menor, sino sustancialmente *mayor* que la de cualquiera de sus dos rivales. Las intuiciones iniciales que tenemos aquí simplemente están equivocadas. (Churchland P. M., 1992, pág. 82)

Contra los postulados radicales del materialismo eliminativo, el propio Paul Churchland describe algunos argumentos. Por ejemplo, una de las críticas señala la falsedad del eliminativismo y defensa de la introspección que revela, por ejemplo, dolores, creencias, temores, etcétera. La respuesta eliminativista será que el sujeto comete el mismo error que pudo haber cometido un antiguo o medieval al observar la esfera celeste girando o la existencia de las brujas.

El hecho de que toda observación se produce dentro de algún sistema de conceptos, y los juicios de observación solo son tan válidos como lo es el marco de referencia conceptual dentro del que se expresan... [...] lo que se cuestiona es la validez de los marcos conceptuales básicos dentro de los que están expresados esos juicios de observación. (Churchland P. M., 1992, pág. 83)

Otra crítica, quizás más fuerte en argumentación, a los postulados del eliminativismo señala que «exagera los efectos de la psicología tradicional y menoscaba sus éxitos reales» (Churchland P. M., 1992, pág. 84). La misma señala que el desarrollo de una neurociencia más evolucionada necesitará que se realicen ciertos ajustes a determinados postulados de la psicología corriente, eso sí, alejados de una corrección total a gran escala. La situación nos ubica en «dos puntos extremos de un mismo espectro: reducción pura o eliminación pura». La propuesta de Paul Churchland pasa por una forma de «materialismo revisionista» antes que eliminativista (Churchland P. M., 1992, pág. 85).

Como acabamos de ver, la propuesta del materialismo eliminativo, si bien es fuerte en justificación, aún tiene que recorrer un largo camino si quiere ver los resultados sobre la psicología popular. Quizás, no en cuanto a su metodología, sino en el propio desarrollo de la neurociencia y que sus conceptos sean equiparables con el de la psicología popular.

El próximo apartado estará centrado en la neurofilosofía en cuanto proyecto interdisciplinar entre filosofía, psicología y neurociencia. Considero que su importancia es doble pues, por un lado, completa el tercer eje programático de este trabajo; por otro, condensa las ideas centrales desde las cuales pretendo avanzar hacia una línea de investigación centrada en la sociabilidad desde el programa neurofilosófico que expondré en el último capítulo a modo de propuesta personal.

7.5 Neurofilosofía y la interfaz filosofía, psicología y neurociencia

Para comenzar, me gustaría plantear una cuestión que estuvo presente desde los primeros esquemas que modelaron esta tesis, esta es si la filosofía del siglo XXI debería abordar la naturaleza de la mente y su relación con el cerebro al margen, por ejemplo, del desarrollo de las neurociencias encargadas de los estudios empíricos del cerebro y el sistema nervioso (Churchland P. M., 1992, págs. 21-22), y continuar con los métodos intuitivos de investigación.

Quizás, para ser más preciso, la formulación de la pregunta no tendría que estar centrada un «debería» que excluye y atenaza, sino en averiguar si las disciplinas filosóficas cuentan con las herramientas suficientes para abordar dicha relación en solitario y al margen de los resultados científicos. En todo caso, habrá que valorar si la independencia metodológica no expone a la investigación filosófica frente a un riesgo demasiado alto donde las conclusiones terminen siendo refutadas por los resultados empíricos.

La realidad es que numerosas disciplinas filosóficas contemporáneas están incorporando los aportes provenientes de la ciencia experimental. A fin de cuentas, los senderos que conducen a la verdad son múltiples y sinuosos. Por lo que concierne a esta investigación, la posición que asumo plantea la posibilidad, si bien lejana, de unificar mente y cerebro con prevalencia de la materia, una idea tomada del proyecto neurofilosófico diseñado por Patricia Churchland.

Desde mi punto de vista, la neurofilosofía abre una nueva perspectiva en la pregunta sobre nuestra naturaleza humana y su alcance antropológico-social, magistralmente explicada por la propia autora (Churchland P. S., 2013, pág. 11):

Mi cerebro y yo somos inseparables. Soy lo que soy porque mi cerebro es lo que es. Aun así, a menudo pienso en mi cerebro en términos diferentes de aquellos que uso cuando pienso en mí misma. Pienso en mi cerebro como «eso» y acerca de mi misma como «yo». Pienso en mi cerebro conteniendo neuronas, pero pienso en mí como teniendo recuerdos. Sin embargo, sé que mi memoria es todo acerca de las neuronas en mi cerebro. Últimamente, pienso en mi cerebro en términos más íntimos, como *yo*³.

Como hemos visto en los apartados anteriores, el problema mente-cuerpo, también conocido como el problema mente-cerebro, ha sido central en la construcción y desarrollo de la filosofía de la mente y las ciencias cognitivas. Aunque una buena parte de estas investigaciones, incluidas las neurocientíficas, remiten inicialmente a Descartes como uno de los precursores del tema, este problema ha estado presente desde la Antigüedad sin una posición «clara y distinta» que nos ayude a encontrar una solución definitiva, puesto que la complejidad de la cuestión va más allá de las posiciones asumidas en las investigaciones.

En *Materia y Conciencia*⁴, Paul Churchland describe cuatro problemas centrales en la filosofía de la mente que han reavivado las semipinternas discusiones sobre su naturaleza material o inmaterial: (a) el «problema ontológico», (b) el «problema semántico», (c) el «problema epistemológico» y (d) el «problema metodológico». De todos ellos, el que más interesa en esta parte de la investigación es el problema ontológico, es decir, la naturaleza de los estados y procesos

³ «My brain and I are inseparable. I am who I am because my brain is what it is. Even so, I often think about my brain in terms different from those I use when thinking about myself. I think about my brain as that and about myself as me. I think about my brain as having neurons, but I think of me as having a memory. Still, I know that my memory is all about the neurons in my brain. Lately, I think about my brain in more intimate terms —as *me*»

⁴ La edición original de *Matter and Consciousness* data del año 1984. En España, la primera edición es de 1992. Si bien una parte de los contenidos ha quedado expuesta al inexorable paso del tiempo, tomada en su totalidad, esta obra puede servirnos para analizar el proceso histórico de aspectos claves tratados por la filosofía de la mente desde la perspectiva del materialismo eliminativo, la cual ha ido cobrando mayor fortaleza con el progreso de las neurociencias.

mentales, cómo se llevan a cabo y cómo se relacionan con el mundo físico (Churchland P. M., 1992, págs. 16-17).

Los «estados» y «procesos mentales» son abordados desde dos posiciones antagónicas: el dualismo y el materialismo. Para el materialismo, que da prevalencia al mundo físico de la materia, estos serán parte de un sistema complejo, el cerebro. Para el dualismo, en cambio, estos procesos además de ser físicos constituyen un fenómeno de naturaleza esencialmente no física (Churchland P. M., 1992, pág. 17). Esta clasificación nos enfrenta a dos tipos de problemas: la mente como problema filosófico y el cerebro como problema científico. Las posiciones dualistas y materialistas incorporan una gran cantidad de versiones, eso sí, cargadas de matices particulares.

En tal sentido, Paul Churchland (1992, págs. 17-18) considera que las diferentes versiones del dualismo plantean serios inconvenientes para hacer frente al «problema semántico», es decir, cómo conocer el verdadero origen del significado de los términos que utilizamos para referirnos a esos estados mentales. Según este autor, esta situación conduce a un escepticismo, pues desconocemos el significado que dichos estados mentales tienen para los otros por la ausencia de una «experiencia directa», por ejemplo, del dolor.

La solución materialista para esquivar el escepticismo pasa por explicar los estados y procesos mentales a partir de la experiencia y del análisis de las «relaciones causales» las cuales se identifican, como no podría ser de otra manera, con el cerebro (Churchland P. M., 1992, págs. 18-19). Además, el problema mente-cuerpo necesita de una solución simultánea que aborde el problema semántico.

Ahora bien, con respecto a R. Descartes, sabemos que combatió al escepticismo desde el dualismo por medio de «evidencias de tipo matemática encontradas en el propio yo» (Sellés García & Solís Santos, 1994, pág. 122), un modelo geométrico garantizado por un Dios mecanicista que representaba «una escapatoria ordenada tanto de la anomia escéptica como de los heroicos furores radicales; adecuaba la idea de la naturaleza a las nuevas ciencias matemáticas

emergentes, y reconstruía la unidad del saber disgregado con la disrupción del aristotelismo» (Sellés García & Solís Santos, 1994, pág. 122).

Desde el punto de vista de J. P. Changeux (1985, pág. 23), este dualismo también rechaza la «tesis tripartita de Platón del alma única, inmaterial e inmortal». No obstante, más allá del paso de los siglos y de los «errores» cometidos, su interpretación sobre el cuerpo, el cerebro y la mente «siguen siendo influyentes en las ciencias y las humanidades de Occidente» (Damasio A. R., 1996, pág. 128).

En este punto es importante recordar que la filosofía de Descartes vio la luz en una época en donde aún la censura religiosa intervenía sobre todas aquellas ideas que pudiesen constituir una amenaza al orden creacionista establecido por Dios, sobre todo si se cuestionaba la inmaterialidad del alma (Changeux, 1985, pág. 23).

También, en esta misma época, el modelo vitalista aristotélico y iatrorquí-mico dominante en las ciencias biomédicas, defendido por filósofos y teólogos, comenzaba a perder espacio frente a la creciente perspectiva mecánico-corpuscular compartida por Gassendi, Descartes, Galileo o Boyle (Solís Santos & Sellés García, 2013, pág. 533).

Según Changeux (1985, pág. 23), en este contexto es difícil «distinguir la parte que responde al auténtico pensamiento del autor de lo que escribe para sobrevivir... [...] De este modo se comprende mejor el curioso sincretismo profesado por Descartes». Quizás, sea este el punto de vista más adecuado para acercarnos a su pensamiento, aunque más no sea a tientas.

Con respecto al mecanicismo cartesiano, este describe al hombre como una máquina automática, es decir, un cuerpo material «insensible como el de los animales» (*res extensa*) que es controlado por un alma espiritual inmaterial (*res cogitans*), cuya actividad principal reside en el pensamiento (Changeux, 1985, pág. 23), a diferencia del resto de los animales. Su noción dualista, «separó la mente del cerebro y el cuerpo» (Damasio A. R., 1996, pág. 128). Sin embargo,

«la base empírica de este enfoque no era distinta de la del paracelsismo o vitalismo aristotélico, pero se reinterpretaba, no si una buena dosis de imaginación en términos de mecanismos» (Solís Santos & Sellés García, 2013, pág. 533).

En el *Tratado del hombre*, Descartes confiere al cuerpo humano una entidad que podríamos denominar de segundo orden por detrás del alma y el espíritu. Tal descripción ha alimentado las suspicacias sobre su acercamiento a la doctrina cristiana, o bien, el querer evitar deliberadamente cualquier tipo de negación al incorporar por necesidad la figura de un dios creacionista:

Supongo que el cuerpo no es otra cosa que una estatua o máquina de tierra a la que Dios forma con el propósito de hacerla tan semejante a nosotros como sea posible, de modo que no solo confiere al exterior de la misma el color y la forma de todos nuestros miembros, sino que también dispone en su interior todas las piezas requeridas para lograr que se mueva, coma, respire y, en resumen, imite todas las funciones que nos son propias, así como cuantas podemos imaginar que tienen su origen en la materia y solo dependen de la disposición de los órganos... [...] Conocemos relojes, fuentes artificiales, molinos y otras máquinas similares que aun habiendo sido realizadas por el hombre, tienen capacidad para moverse de modos diversos en virtud de sus propios medios, y me parece que no acertaría a imaginar tantas clases de movimientos en esta máquina que supongo ha salido de la mano de Dios, ni tampoco alcanzaría a atribuirle tal artificio como que alguien no tuviera motivo para pensar que aún pudiera ser superior (Descartes, 1990, págs. 22-23).

De este modo los humanos, «a diferencia de los animales que no son más conscientes ni tienen más intencionalidad que los relojes» (Solís Santos & Sellés García, 2013, pág. 534), poseen voluntad, pensamientos y memoria. Además, un alma inmaterial ubicada estratégicamente en la glándula pineal —la glándula «H», según las explicaciones del propio autor y los grabados que ilustran *El Tratado del hombre*— que es consciente de todo lo que sucede en el interior y exterior por medio de las aferencias nerviosas, además de participar en la sensibilidad, la imaginación, la memoria y la causa de los movimientos corporales (Lokhorst, 2015).

Esta minúscula glándula ubicada en el tercer ventrículo cerebral posee una estructura de 6,5 milímetros, siendo principal función la elaboración de la hormona melatonina (Solís Santos & Sellés García, 2013, pág. 534). Para Changeux (1985, pág. 23), la solución de la glándula pineal propuesta por Descartes no resistió la crítica de científicos ni filósofos, entre ellos Spinoza.

Thomas Willis (1621-1675), «el principal iatroquímico británico» (Solís Santos & Sellés García, 2013, pág. 317) fue un pionero de las investigaciones neuroanatómicas. A diferencia de Descartes que había unido el alma y el cuerpo en la glándula pineal, Willis la ubica en el cuerpo caloso (puente de fibras nerviosas que comunica los hemisferios cerebrales), aunque persiste en la idea de un «alma razonable, propia del hombre e inmaterial» (Changeux, 1985, pág. 25).

Por esta misma época, Pierre Gassendi (1592-1655) reelaborará el atomismo griego y de Lucrecio que se había mantenido en la sombra durante la Edad Media (Solís Santos & Sellés García, 2013, pág. 111), como no podría ser de otro modo. Para Gassendi, los animales tienen un alma igual que el hombre, pero no está unida en un punto preciso. Esta proyección del «alma de las bestias» consigue, según Changeux, animalizar considerablemente al hombre (1985, pág. 25).

Este es un momento de transición importante para la filosofía y la antropología, pues comienza a cobrar peso la revolución biológica que conlleva un cambio en la metodología de la investigación, abandonando paulatinamente el modelo introspectivo en pos de un modelo experimental.

Derivados del enciclopedismo y la antropología física, a comienzo del siglo XIX surgieron algunos estudios sobre grupos humanos e individuos que dieron nacimiento a dos teorías importantes: el transformismo de Lamarck, y la frenología de Franz Joseph Gall quien, «constatando la identidad de las sustancias gris y blanca en el nivel central y periférico, consigue una “laicización” del cerebro considerablemente avanzada ya con La Mettrie y Cabanis» (Changeux, 1985, pág. 26).

Franz Gall (1758-1828) propuso una metodología empírica (Solís Santos & Sellés García, 2013, pág. 797) para analizar las funciones del cerebro y sus localizaciones, evitando de esa manera el método introspectivo (Changeux, 1985, pág. 26) tan extendido en los círculos filosóficos y científicos. Kandel (2001) comenta que existen tres ideas de Gall que han sido decisivas en el estudio de la fisiología cerebral relacionadas con la conducta:

Primero, defendió que toda la conducta emanaba del cerebro. Segundo, sostenía que determinadas regiones de la corteza cerebral controlaban funciones específicas. Gall afirmaba que la corteza cerebral no actuaba como un órgano único, sino que se dividía por lo menos en 35 órganos (posteriormente se añadieron otros), cada uno de los cuales correspondía a una facultad mental específica. Asignaba su zona en el cerebro incluso a las conductas humanas más abstractas, como la generosidad, la discreción y la religiosidad. En tercer lugar, Gall proponía que el centro de cada función mental crecía con el uso, de forma muy similar al aumento del volumen de un músculo con el ejercicio. (pág. 6)

Los estudios de Gall fueron puestos a prueba por medio de análisis experimentales llevados a cabo por el fisiólogo Pierre Flourens (1794-1867), quien «extirpó los centros funcionales de Gall de los cerebros de los animales de experimentación» (Kandel E. R., 2001, pág. 7). Sus resultados determinaron que las regiones descritas por Gall no eran las responsables de determinadas conductas, sino que «cualquier parte del hemisferio cerebral era capaz de realizar todas las funciones del hemisferio» (Kandel E. R., 2001, pág. 7).

La investigación de Flourens, denominada «teoría del campo global», más adelante sería cuestionada por los trabajos sobre epilepsia local del neurólogo británico J. Hughlings Jackson quién observó que la «lesión localizada de un sector cerebral particular no lleva nunca consigo la pérdida completa de la función. La localización de una lesión no puede identificarse con la de una función» (Changeux, 1985, pág. 34). Estos estudios constituyeron el fundamento que dio pie al llamado «conexionismo celular» desarrollado por K. Wernicke, C. Sherrington y Ramón y Cajal. Según Kandel (2001, pág. 7), «la mejor forma de ilustrar

las diferencias entre la teoría del campo global y el conexionismo celular es mediante el análisis de cómo el cerebro produce el lenguaje».

Cabe destacar que las investigaciones experimentales y la anatomía clínica motivaron un cambio de rumbo con respecto a la frenología iniciada por Gall. Paul Broca (1824-1880), Korbinian Brodmann (1868-1918) y Karl Wernicke (1848-1905) son los hacedores de esta «nueva frenología» que nace con el siglo, y que Changeux con acierto ha denominado como «la “*Belle Époque*” de las localizaciones cerebrales» (1985, pág. 32).

Broca, a diferencia de Gall, propuso que «el intento de localizar las funciones de la mente debe basarse en el estudio de los daños cerebrales producidos por lesiones clínicas, en lugar del examen de la distribución de los bultos en el exterior de la cabeza» (Kandel E. R., 2001, pág. 10). Brodmann, por su parte, «reactualiza la tentativa de Gall» cartografiando el cerebro en 52 áreas diferenciadas, «un mapa aún utilizado hoy en día» (Changeux, 1985, pág. 32).

Frente a la teoría del campo global, Wernicke mantuvo «que solo las funciones mentales más básicas, aquellas relacionadas con actividades perceptivas o motoras simples, están localizadas en áreas únicas de la corteza». Las funciones cognitivas complejas, en cambio, dependen de las interconexiones de varias zonas funcionales, proponiendo la idea del «procesamiento distribuido» (Kandel E. R., 2001, pág. 11).

Los trabajos de Broca —sobre la afasia—, Fritsch y Hitzig permitieron a Wernicke formular el primer modelo sobre la organización del lenguaje. Hoy en día, el «procesamiento del lenguaje es considerablemente más complejo que la concepción inicial de Wernicke» (Kandel E. R., 2001, pág. 13).

Sin embargo, la cartografía cerebral diseñada por los frenólogos no aportó pruebas suficientes sobre las áreas involucradas en las diferentes funciones cognitivas y motoras, pues «las regiones cerebrales no se ocupan de *facultades* de la mente, sino de operaciones elementales de procesamiento (Kandel E. R.,

2001, pág. 16) tales como la percepción, el movimiento, el lenguaje, el pensamiento y la memoria que funcionan en serie y en paralelo interconectando varias regiones del cerebro.

El estudio del lenguaje en el siglo XX abrió la puerta a las investigaciones sobre los objetos y procesos mentales, es decir, aquellos complejos biológicos de procesamiento de información en los que intervienen las operaciones mentales distribuidas en varias áreas cerebrales. En otras palabras, de este procesamiento dependen nuestras funciones cognitivas: cómo aprendemos, cómo almacenamos y cómo recuperamos la información, desde las más simples a las más complejas.

El desafío a que nos enfrentamos en el siglo XXI es saber si la biología y las neurociencias son capaces de explicarnos «la representación nerviosa de la conciencia y la conciencia de sí mismo» (Kandel E. R., 2001, pág. 16). No obstante, necesitaremos otra herramienta para comprender el aspecto semántico y conceptual que intervienen en la representación y si se ajustan los aspectos ontológicos y semánticos.

Hoy en día, sabemos que por medio de la experimentación con técnicas de imagen podemos medir los objetos y las imágenes mentales. Pero ¿cómo es posible acceder al estudio de un objeto mental si está situado en un nivel de organización al de la célula nerviosa? Según Changeux (1985):

La máquina cerebral es un conjunto de neuronas y nuestro problema consiste en investigar los *mecanismos celulares* que permiten pasar de un nivel a otro, hacer la disección y luego reconstruir los «objetos mentales» partiendo de actividades elementales de conjuntos definidos de neuronas... [...] Las imágenes mentales, los conceptos, son objetos de la memoria. Desde Pavlov, el behaviorismo y Skinner, la «reacción condicionada» se ha considerado como el mejor, incluso a veces el único modelo elemental de memoria (págs. 160-161).

Los objetos mentales (conceptos), como parte de la memoria, pueden medirse a partir de su evocación, es decir, a través de la huella material (electroquímica) que dejan en nuestro cerebro, lo cual nos acerca al estudio biológico

de los objetos mentales. En un tiempo se pensó que la «reacción condicionada» era el mejor modelo elemental de memoria, sin embargo, «al carecer de la generalidad esperada... [...] no pueden servir como punto de partida para construir objetos de memoria» (Changeux, 1985, pág. 162).

Como respuesta, Changeux (1985) propone, por su accesibilidad, recuperar la «hipótesis de parentesco entre imagen de memoria y percepto», es decir, la vinculación entre el percepto primario y la respuesta de los sentidos, y crear una «pasarela hipotética entre lo mental y lo biológico» (págs. 162-164).

La hipótesis de una «teoría biológica» (Changeux, 1985, pág. 160)—compartida por un amplio grupo de neurocientíficos en la décadas de 1970 y 1980—capaz de explicar los procesos mentales y su relación con la conducta nos ubica de lleno en las principales discusiones que han dado origen al surgimiento de la neurofilosofía y sus subdisciplinas. Sin embargo, como ya hemos comentado en otros párrafos, los roces y desacuerdos entre filósofos y neurocientíficos no han sido pocos (Bennett, Dennett, Hacker, & Searle, 2008).

Considero que lo comentado es un síntoma de la trascendencia objetiva de los temas tratados y de la posibilidad de movilizar una colaboración inter-transdisciplinar entre la teoría filosófica y experiencia científica, antes que un círculo vicioso de vacuas discusiones. No obstante, subyace aún la cuestión que magistralmente describe Kandel (2001):

[...] la principal razón por la que se ha tardado tanto tiempo en comprender qué actividades mentales están localizadas en cada región cerebral es que nos enfrentamos con el enigma más profundo de la biología: la representación nerviosa de la conciencia y de la conciencia de sí mismo (pág. 16).

7.6 Patricia Churchland y la neurofilosofía

¿Por qué una investigación sobre la memoria biológica y la sociabilidad con base en la neurofilosofía? En el contexto inter-transdisciplinar que planteo

esta investigación, parto del supuesto que la neurofilosofía nos aporta un marco propedéutico adecuado para describir aquellos aspectos clave de la memoria biológica que hacen posible la construcción de la sociabilidad humana, albergue del conocimiento y la moral.

Considero que la filosofía y la neurociencia cooperan en la creación de este nuevo campo de estudio a partir de dos acciones recíprocas: límite y expansión. Designo como «límite» a la acción de contención, de vigilancia conceptual y metodológica que ejerce una disciplina sobre la otra. La «expansión» hace referencia al avance de los contenidos en cada una de las disciplinas y que han sido contrastados por las evidencias.

En 1986, Patricia S. Churchland publicó *Neurophilosophy: Toward a unified science of the mind-brain*, un trabajo fundacional que describe «una concepción eliminativista de los procesos mentales que vendría a negar las diferencias entre filosofía y las neurociencias y en donde la primera acabaría siendo absorbida por las segundas» (Rodríguez Serón, 2002, pág. 150).

Considero que la interpretación de Rodríguez Serón (2002) se distancia de la perspectiva original de Churchland en cuanto a los objetivos de la neurofilosofía, pues en su opinión:

[...] el neologismo no tiene por qué remitir sistemáticamente al sentido definido por P. S. Churchland. La neurofilosofía puede ser concebida como un punto de encuentro fecundo entre filosofía y neurociencias donde las relaciones entre ambas no tienen que estar marcadas necesariamente por la eliminación o la absorción. (pág. 150)

Entiendo que las palabras de P. Churchland albergan, sobre todo, un deseo bien intencionado de equidad entre ambas disciplinas. No obstante, es evidente que hallar un equilibrio que satisfaga los intereses de todas las partes y que a la vez eluda la reducción del materialismo eliminativista constituye una meta muy difícil de alcanzar, por lo menos en el corto plazo.

Desde mi punto de vista, creo que la relación filosofía-neurociencia supondrá la búsqueda de un equilibrio entre dos fronteras, teoría y experiencia, en

constante estado de tensión, pues cada una alberga un conjunto de metodologías particulares. La lógica del planteamiento de Churchland no se entiende fuera del materialismo eliminativo, lo cual nos plantea la siguiente cuestión: ¿estamos ante una nueva forma de hacer filosofía?

Patricia Churchland (1986, págs. IX, Prefacio) afirma que la neurofilosofía es el inicio del camino hacia una «gran teoría unificada de la mente-cerebro» la cual recoge el desarrollo de las neurociencias, la neurotecnología y las ciencias cognitivas.

En un primer acercamiento, podemos ver que la síntesis propuesta por Patricia Churchland prioriza la necesidad metodológica de entender previamente el funcionamiento del cerebro para luego poder explicar aquello que sea «la mente» y su funcionamiento (Churchland & Di Francesco, 2007, pág. 179).

Ahora bien, a pesar de los ingentes esfuerzos realizados en pos de un programa unificado, la discusión continúa abierta a causa de la complejidad del objeto de estudio y la forma de acceder a dicha información, pues como bien dice la investigadora (Churchland P. S., 1986):

Nuestros cerebros son montículos enormes de células excitables que, de alguna manera, se las ingenian para contener una rica representación del mundo exterior, así como para permitir a los músculos lograr tales hazañas como atrapar una pelota, tocar el violín y hablar, además de, por supuesto, la alimentación básica, la huida, la lucha y la reproducción. Además, el cerebro humano, como el cerebro de otras especies, contiene información sobre sí mismo y sobre los otros cerebros, aunque para estar seguro, no aprehendemos de forma estándar la información bajo esa descripción⁵. (pág. 3)

⁵«Our own brains are massive mounds of excitable cells, which somehow contrive collectively to contain a rich representation of the outside world, as well as to enable the muscles to accomplish such feats as catching a ball, playing the violin, and talking, in addition of course to the fundamental feeding, fleeing, fighting, and reproducing. Additionally, the human brain, like the brains of other species, contains information about itself and about other brains, though to be sure, we do not standardly apprehend the information under that description».

Como podemos observar, tal complejidad material y fisiológica hace pendular las investigaciones entre dos entidades, al parecer diferentes: la mente y el cerebro. Esta es la razón por la que el programa neurofilosófico —materialista, reduccionista y dependiente en de las ciencias experimentales— realiza un ingente esfuerzo con el propósito de salvar esta división bajo la convicción de que una teórica unificada mente-cerebro pueda ser capaz de ofrecer datos más sólidos de los que contamos hasta el momento, eso sí, descartando cualquier metodología de investigación introspectiva (Churchland P. S., 1994, pág. 23). Finalmente, ¿terminará el cerebro absorbiendo cualquier entidad mental?

La interfaz entre neurociencia y filosofía que propone Churchland ha sido recibida con beneplácito en diversos ámbitos neurocientíficos. Sin embargo, hay que reconocer que su suerte ha sido dispar en los círculos filosóficos más tradicionales, sobre todo, después de que los neurofilósofos declarasen abiertamente que «la metodología *a priori* tipificada por los filósofos de la mente y el lenguaje en el siglo XX puede haber llegado a su fin»⁶ (Churchland & Di Francesco, 2007, pág. 179).

En cuanto al proyecto original, Patricia Churchland buscó introducir la filosofía a los neurocientíficos y la neurociencia a los filósofos. La autora ha remarcado con insistencia sobre determinados hechos empíricos acerca del funcionamiento del cerebro.

De la misma manera que lo hace Paul Churchland, su epistemología defiende una posición materialista eliminativista sobre cuestiones concernientes a la relación mente-cerebro, tales como: (1) el desarrollo de una alternativa a la teoría lógico-empirista de reducción interteórica, (2) responder a los argumentos de la propiedad dualista basada en la subjetividad y los qualia sensoriales, y (3) responder a los múltiples argumentos anti-reduccionistas.

⁶ «Whilst the *a priori* methodology typified by philosophers of mind and language in the twentieth century may have run its course, neurophilosophers now are playing an important role in synthesising and integrating data from many subfields».

Algo que debemos tener presente es que el programa neurofilosófico parte del análisis biológico-evolutivo que han experimentado los sistemas biológicos. Con respecto a las neurociencias, P. Churchland considera que las preguntas de carácter general se mueven en diversos niveles, altos y bajo, por ejemplo, cuando analizamos el funcionamiento de las retinas o el movimiento REM de los ojos (Churchland P. S., 1986, pág. 239).

Uno de los grandes problemas a que debe enfrentarse la autora es responder a la pregunta sobre la posibilidad de la integración entre neurociencia y psicología, donde la ciencia se verá obligada a buscar más allá de sus modelos y enfrentarse a las preguntas filosóficas.

Para explicar las relaciones entre ciencia, filosofía y el desarrollo de los diversos paradigmas, Churchland recurre a la comparación de la ciencia en la antigüedad, más ligada a la intuición que a la experimentación y el método, poniendo como ejemplo el caso de Aristarco y su concepción intuitiva (y por supuesto, equivocada) del sistema solar (1986, pág. 240).

En tal sentido, el ejemplo que utiliza la autora (Churchland P. S., 1986) para ilustrar las diferencias metodológicas, compara la naturaleza del movimiento de la Tierra con la naturaleza de la mente y su funcionamiento:

Supongamos que el objeto de la investigación no es la naturaleza del movimiento o de la Tierra, sino la naturaleza de la mente. ¿Es posible que la realidad de la naturaleza de la mente, de sus estructuras y procesos, también sea diferente de su apariencia? Si la respuesta es sí, significa que la neurociencia y la psicología experimental puedan dar descripciones de

estados y procesos internos que son profundos y diferentes —tal vez radicalmente diferentes— de cómo pensamos intuitivamente en los estados y procesos mentales sobre nosotros mismos⁷. (pág. 241)

En el artículo *The Impact of Neuroscience on Philosophy* (2008), Churchland explica la importancia para la filosofía del trabajo realizado por la neurociencia y, sobre todo, el impacto en la conducta humana. Desde ese punto de vista, los resultados experimentales de las neurociencias cognitivas nos pueden ayudar a entender, o por lo menos, elaborar nuevas preguntas sobre cómo funciona nuestra moralidad y los mecanismos neuronales que la gobiernan.

Como hemos analizado en los apartados anteriores, Descartes es una figura central en la discusión sobre el problema del cuerpo y la mente, su unidad y, sobre todo, si podría existir algún tipo de influencia recíproca entre estas dos entidades. En la actualidad, este problema ha cambiado de entidad, mejor dicho, de preguntas, pues ya no se dirigen al cuerpo, sino al cerebro y sus mecanismos. Para la neurociencia cognitiva la mente ya no es entendida ni estudiada como la mente clásica. Esto mismo ha sido trasladado a la filosofía, donde los análisis introspectivos de investigación han dado paso a una nueva metodología de investigación, por ejemplo, la llamada filosofía experimental, o bien, como en otros casos más teóricos, promoviendo la discusión sobre cuestiones conceptuales de la propia neurociencia, lo cual está alimentando el conocimiento sobre conceptos centrales tales como «libre albedrío», «conciencia», «valores», «empatía», «identidad», «recuerdos» o «moralidad» entre otros tantos.

Está claro que la influencia de la neurociencia no solo ha sido en la filosofía, sino en otros campos científicos cercanos como la psicología. Ello también ha puesto en vilo la llamada *psicología popular*.

⁷ Suppose the object of inquiry is not the nature of motion or of the Earth, but the nature of the mind. Is it possible that the reality of the mind's nature, of its structures and processes, is also different from the appearance? If the answer is yes, it means that neuroscience and experimental psychology may yield descriptions of inner states and processes that are deeper and different - perhaps radically different from how we intuitively think of mental states and processes in ourselves.

En esta segunda década del siglo XXI estamos viviendo la influencia de la inteligencia artificial, en sus múltiples formas y dispositivos, en todos los ámbitos del conocimiento y de la vida social. Esta situación nos enfrenta a nuevas cuestiones, quizás complejas, acerca de la relación mente-cerebro, ya que introduce dos nuevos elementos: la máquina y el software.

Estoy convencido de que estas cuestiones constituyen la fundamentación para seguir explorando los caminos iniciados por la neurofilosofía con el propósito de encontrar nuevas preguntas que ayuden a entender la complejidad de la realidad. En el próximo apartado se tratará la relación entre la neurofilosofía y la neuroética como filosofía práctica.

7.7 De la neurofilosofía a la neuroética

What happens when we accept that everything we feel and think stems not from an immaterial spirit but from electrical and chemical activity in our brains? (Churchland P. S., 2013)

Como si de un camino natural se tratase, el desarrollo de la neurofilosofía ha devenido en la neuroética. Ahora bien, ¿de qué hablamos cuando hablamos de neuroética y cuál es su importancia? Para comenzar a hablar de ello me gustaría ilustrarlo con el ejemplo de la pandemia que hemos atravesado y sus consecuencias menos visibles, pero persistentes en la memoria. La COVID 19 nos enfrentó a una nueva realidad y cambios como nunca antes, pues vivimos y experimentamos de forma directa situaciones límites de sufrimiento, muerte y, sobre todo, incertidumbre. Toda la situación viva impactó directamente en nuestra conducta, en nuestra forma de ver el mundo y la vida, en definitiva, en nuestra mente y nuestro cerebro: en nuestro cuerpo y la acción. El estrés soportado a causa de la incertidumbre social modificó nuestros hábitos más elementales, por ejemplo, como saludarnos o relacionarnos en el día a día, higienizar todo a nuestro paso y estar pendientes de la curva de incidencia que se informaba a todas horas. Una simple mascarilla modificó nuestros hábitos y relaciones sociales. Todo ello ha tenido un impacto directo en nuestra salud física y mental.

El ejemplo de la pandemia ilustra perfectamente el ámbito de actuación multidisciplinar de la neuroética, pues en él intervienen varios segmentos de la realidad que impactan en el sistema nervioso y modifica nuestra conducta, nuestros hábitos y, por supuesto, nuestros valores sociales, modificando nuestra sociabilidad. Como veremos a lo largo del apartado, la neuroética se encarga de estudiar, en parte, el origen de estos valores, ya sean genéticos o culturales, y su despliegue en vida social. Es importante tener presente que una de las principales características de nuestro cerebro es su naturaleza social (Gazzaniga M. S., 1993).

Los avances en genética, informática y ciencias de datos que marcaron las últimas tres décadas del siglo pasado prepararon el terreno en el que iban a eclosionar a inicios del siglo XXI las denominadas «ciencias neuro». Pronto, las investigaciones interdisciplinares entre genética, informática y neurociencias motivaron la curiosidad y el interés de diversos ámbitos científicos, quienes vislumbraban una base sólida en donde asentar sus edificios teóricos y prepararse para lo que vendría. Como era de esperar, las principales publicaciones de los resultados que se obtenían en los laboratorios no tardaron en contagiar su entusiasmo a muchas disciplinas. En filosofía, por ejemplo, las primeras repercusiones comenzaron a sentirse en los campos de las ciencias cognitivas, la ética y la estética. A partir de allí, paulatinamente, su alcance ha ido extendiéndose a otros ámbitos, por ejemplo, el derecho, la economía, la política y la psicología, por citar algunos.

En la década de los años ochenta del siglo pasado, la obra de Patricia Churchland dedicada a la neurofilosofía constituyó un punto de partida para las investigaciones sobre la neurofisiología y su impacto en la conducta.

¿Qué aporta a la filosofía el conocimiento de las redes neuronales y la neurofisiología? Acercarnos a la fisiología del cerebro nos puede brindar una perspectiva privilegiada de los laberintos mentales de nuestro funcionamiento interno.

Como he comentado en los apartados anteriores, las conclusiones obtenidas y la metodología de investigación empleada ya no tendrán su origen en el

marco habitual (y predominante) de la psicología *folk*, sino el abordaje se hará desde una neurociencia más desarrollada que «eliminará el antiguo marco de referencia» (Churchland P. M., 1992, págs. 76-76). A simple vista lo dicho puede resultar una cuestión menor, sin embargo, las ruinas de psicología popular han quedado sepultadas bajo la cimentación donde se asentará la construcción del futuro edificio de la neuroética en el siglo XXI.

Desde su nacimiento en 2002, la neuroética ha experimentado un avance extraordinario en investigaciones interdisciplinares centradas en las bases cerebrales que rigen la conducta moral y en el impacto de las neurociencias, incluida la neurotecnología, en los ámbitos jurídico, político y social.

Su modelo de actuación persigue un doble objetivo: por un lado, (a) permite a los investigadores compartir el alto coste que supone el mantenimiento de los laboratorios y la experimentación con tecnología de última generación, y (b) dada la complejidad del objeto de estudio, el funcionamiento del cerebro humano, la posibilidad de obtener resultados más completos.

Más allá de estos dos aspectos fundamentales, cabe señalar que la interdisciplinariedad encierra algunas dificultades, sobre todo a la hora de definir los lineamientos de una fundamentación y un programa metodológico autónomo. Sin lugar a dudas, este es uno de los retos principales que deberán afrontar los neuroeticistas con el propósito de unificar en mismo programa la ingente producción científica que se encuentra dispersa y que se identifica bajo el rótulo de «neuroética». Actualmente, ya podemos encontrar algunas investigaciones muy interesantes en esta dirección, por ejemplo, los trabajos de la investigadora sueca Kathinka Evers (2011) sobre la fundamentación para la neuroética desde una perspectiva del materialismo ilustrado, un marco propuesto, inicialmente, por el neurocientífico francés Jean Pierre Changeaux (1985).

Las fuentes que han permitido la eclosión de la neuroética debemos buscarlas en el desarrollo de los últimos treinta años de las neurociencias, el cual se ha focalizado en áreas muy sensibles tales como el sistema nervioso central (SNC) y el tratamiento de las patologías cerebrales. Como ejemplo podemos

mencionar las discusiones sobre el uso, quizás inadecuado o no, de las denominadas «smart drugs» o «nootrópicos»: una serie de fármacos destinados potenciar el rendimiento cognitivo en sistemas nerviosos sanos o «normales». Dichos estudios se realizaron en 1964 por el psicólogo y químico rumano, Cornelius Giurgea, quien sintetizó por vez primera el Piracetam, un nootrópico (potenciador cognitivo), siendo considerado el padre de las drogas inteligentes.

A finales del siglo XX, las primeras investigaciones que comenzaron a hablar de neuroética delinearon sus objetivos y alcances presentándola como una disciplina muy próxima a la bioética, centrada sobre todo en el examen de lo que es correcto o incorrecto, bueno o malo, acerca del tratamiento, perfeccionamiento, invasiones o manipulaciones del cerebro humano. Una definición adecuada, pero quizás incompleta, pues la neuroética es mucho más que una «bioética del cerebro» (Gazzaniga M. S., 1993, pág. 14). Haciendo un ejercicio de retrospección, vemos que las críticas sobre el alcance de la disciplina estaban en lo cierto.

Gazzaniga (1993) afirma que nuestra moralidad depende, en parte, de códigos (genéticos) de conducta anclados en el cerebro. Su investigación (coincidente con la opinión de otros neurocientíficos y neurofilósofos tales como Jesse Prinz, Patricia y Paul Churchland, Frans de Waal o Antonio Damasio entre otros) ha permitido ubicar a la neuroética más allá del territorio de la bioética, a la que define como «una filosofía de la vida con un fundamento cerebral» (pág. 15).

Para Neil Levy (2008, págs. 1-8), investigador de la universidad de Oxford, la posibilidad de la neuroética como una disciplina independiente encuentra su propia justificación más allá de cualquier reflexión ética (bioética) sobre las nuevas técnicas y tecnologías producidas por las neurociencias. Tal enfoque hace referencia a la forma en que los conocimientos nacidos de las neurociencias están iluminando a temas transversales de la historia de la filosofía, por ejemplo, la naturaleza y la fundamentación de la moralidad, la pérdida del autocontrol, el libre albedrío o la responsabilidad. Por lo tanto, nos hallamos ante dos ramas de la neuroética que interactúan entre sí produciendo un conocimiento nuevo, un

territorio común compartido por neurocientíficos, filósofos, psicólogos, sociólogos y juristas, entre otros: una ética de la neuroética y una neurociencia de la ética.

Esta bifurcación temática de la neuroética es conceptualizada como «neuroética aplicada o ética de la neurociencia» y «neuroética fundamental o neurociencia de la ética» (Evers, 2011, pág. 13). Mientras que la primera está relacionada con los problemas prácticos de la bioética y la biotecnología aplicada a las neurociencias (investiga, por ejemplo, los problemas éticos originados por el uso de las técnicas de neuroimágenes, la mejora cognitiva o el uso de la farmacología que modifican la conducta), la segunda, en cambio, se ocupa de cómo el conocimiento del funcionamiento del cerebro y su evolución pueden ayudarnos en la comprensión de nuestra identidad personal, la conciencia y la intencionalidad, asimismo, el desarrollo del pensamiento y la formación de juicios morales (Evers, 2011, pág. 13).

La neuroética fundamental tendrá que ser capaz de responder al modo en que las ciencias naturales pueden ayudarnos a la comprensión del pensamiento moral a la luz de los nuevos conocimientos aportados por las neurociencias cognitivas y su metodología de investigación.

Con respecto a esto último, en la primera parte de *Neuroética. Cuando la materia se despierta*, K. Evers pone de manifiesto la intromisión de las ciencias naturales en las investigaciones habituales de la ética y la moral, un fenómeno al que denomina de «irrupción inversa», pues esta vez no será la ética quien llame primero las puertas de los dominios científicos.

Las investigaciones realizadas por la neuroética fundamental son novedosas y pueden aportarnos conocimientos relevantes acerca de la conducta moral de los seres humanos. Ahora bien, al momento de hacer una evaluación de la situación actual de la neuroética como ética fundamentada en el cerebro, es importante tener en cuenta que una de las principales cualidades de las neurociencias, la velocidad del desarrollo de sus investigaciones, supone un aspecto positivo para los neurocientíficos. Sin embargo, esta misma situación representará

una cuestión problemática para la neurofilosofía, pues se debe procesar y sintetizar una ingente cantidad de información a una velocidad bastante menor, razón por la cual los neurocientíficos han tomado la delantera a la hora de presentar en sociedad las últimas novedades. No obstante, hay que reconocer que algunas de sus conclusiones no están exentas de contradicciones, sobre todo, al momento de adentrarse en los dominios de las teorías éticas y filosóficas que recorren toda la historia de la filosofía.

Es posible que la combinación, a partes iguales, de (a) una debilidad conceptual y (b) el exceso de confianza en las pruebas empírica sean parte de los motivos por el cual algunos sectores filosóficos han recibido a la neuroética con una gran dosis de escepticismo y reticencia en lo concerniente a que esta pueda ser considerada una disciplina capaz de explicar la fundamentación de nuestra moralidad basada en valores anclados en la genética de la fisiología cerebral. Sin embargo, los estudios experimentales desarrollados por la neurociencia confirman pequeñas evidencias que ponen en cuestión que nuestros valores morales provengan en exclusividad del campo de la cultura o, como en otros casos, de la religión o mandato divino.

En la intersección de las dos corrientes por los que transitan las investigaciones en neuroética —fundamental y aplicada— hallamos la influencia directa de las ciencias naturales y experimentales como soporte de sus investigaciones. Una vez más, el acercamiento de las Humanidades a las ciencias naturales parece reavivar el debate con las éticas normativas y religiosas, más afines a explicar la construcción de la moralidad a partir de una evolución socio-cultural, las cuales muestran ciertas reticencias para admitir una fundamentación cerebral de los valores éticos, más aún si tenemos en cuenta la fuerte dependencia de la genética y las teorías evolutivas que subyacen, incuestionables, en todo el desarrollo neuroética.

El estudio del cerebro —su constitución, fisiología y relación con el medioambiente— el protagonista de la escena científica en estas primeras décadas del siglo XXI, donde las Humanidades, por supuesto, no están exentas de su alcance y tratamiento. No obstante, considero que cualquier rechazo *a priori* del

desarrollo neurocientífico tendrá sus repercusiones directas en el propio desarrollo de las disciplinas humanísticas que, permeables a su influencia, asumen el desafío del análisis y de la síntesis devolviendo una investigación enriquecida de contenidos y puntos de vistas y, sobre todo, nuevas preguntas que ayudan a alimentar la investigación y desvelar aquellos resquicios claroscuros de verdad.

Antes de dar paso al próximo capítulo, me gustaría compartir unas palabras de Anil Seth (2023) que creo son muy pertinentes para nuestro tiempo:

En Praga, a finales del XVI, el rabino Judah Loew ben Bezalel extrajo arcilla de las orillas del río Moldava y, con ella, moldeó una figura de aspecto humano: un golem... [...] El golem del rabino Loew nos recuerda la soberbia en la que incurrimos cuando tratamos de fabricar criaturas inteligentes y sintientes a imagen y semejanza nuestras o de la mente divina. Rara vez sale bien. Desde el monstruo de *Frankenstein* de Mary Shelley hasta la Ava del filme *Ex_Machina* de Alex Garland, pasando por los robots epónimos de Karel Čapek, el *Terminator* de James Cameron, los replicantes de *Blade Runner* de Ridley Scott, o HAL de Stanley Kubrick, estas creaciones casi siempre se vuelven contra sus creadores y dejan a su paso regueros de destrucción, melancolía y confusión filosófica.

Durante la última década, más o menos, el rápido auge de la IA ha vuelto más urgente la necesidad de dar respuesta a ciertas preguntas que nos hacemos acerca de la conciencia de las máquinas. Actualmente, estamos rodeados de IA: está integrada en nuestros teléfonos, en nuestros frigoríficos y en nuestros automóviles, y está impulsada en muchos casos por algoritmos de red neuronal inspirados en la arquitectura del cerebro. Hacemos bien en preocuparnos por el impacto de esta nueva tecnología. ¿Nos quitará nuestros puestos de trabajo? ¿Deshebrará el tejido de nuestras sociedades? ¿Y terminará por destruirnos a todos, ya sea por un incipiente interés propio, o por una falta de previsión programadora que lleve a que se agoten los recursos de la Tierra, transformados en una inmensa montaña de chips sujetapapeles? Lo que subyace a muchas de esas preocupaciones, sobre todo a las de carácter más existencial y apocalíptico, es la suposición de que, en algún punto de su cada vez más acelerado desarrollo, la IA acabará por tener conciencia propia. Es el mito del golem hecho silicio.

¿Qué haría falta para que una máquina sea consciente? ¿Qué implicaciones tendría algo así? ¿Y cómo podríamos distinguir siquiera una máquina consciente de su equivalente zombi? (págs. 293-294)

Las preguntas que plantea Anil Seth son realmente inquietantes, a la vez que movilizadoras, pues nos desafían a encontrar respuestas a través de la acción colectiva de forma urgente, ya que está en juego nuestra civilización tal y como la hemos conocido, donde lo único que tenemos por seguro es la incertidumbre que reviste el mañana.

El siguiente apartado está dedicado a un sumario explicativo de los principales hitos que han marcado el desarrollo de la neuroética. Las explicaciones van acompañadas de diapositivas (D) que aparecen numeradas al final del trabajo en la sección de Diapositivas. Este breve resumen puede servir para tener una visión global de aquellos momentos relevantes que han marcado el desarrollo histórico de esta joven disciplina y la importancia que ha alcanzado en nuestros días.

7.8 Sumario: hitos de la neuroética

El objetivo de este apartado es mostrar de forma sumaria aquellos hitos y elementos relevantes de la joven historia de la neuroética, asimismo, sus relaciones con otros campos de la investigación. Al final del trabajo, en el apartado «Diapositivas», se hallan las ilustraciones en formato de diapositivas que complementan las descripciones aquí desarrolladas.

Ética: (*éthos* como «carácter» y *ethos* como «costumbre»). La ética es considerada una de las principales ramas de la filosofía que se ocupa del estudio racional de la moral, la virtud, el deber, la felicidad y el buen vivir. Requiere la reflexión y la argumentación. Sus indagaciones especulativas se remontan a los orígenes mismos de la filosofía en la Antigua Grecia, y su desarrollo histórico ha

sido amplio y variado. La ética estudia qué es lo moral, cómo se justifica racionalmente un sistema moral, y cómo se ha de aplicar posteriormente a nivel individual y a nivel social.

- Una doctrina ética elabora y verifica afirmaciones o juicios determinados.
- Una sentencia ética, juicio moral o declaración normativa es una afirmación que contendrá términos tales como «bueno», «malo», «correcto», «incorrecto», «obligatorio», «permitido», etcétera, referidos a una acción, una decisión o también las intenciones de quien actúa o decide algo.

Moral: moral o moralidad (del latín *mos*, *moris*, «costumbre») son las reglas o normas por las que se rige el comportamiento o la conducta de un ser humano en relación a la sociedad (normas sociales), a sí mismo y a todo lo que lo rodea. Otra perspectiva la define como el conocimiento de lo que el ser humano debe de hacer o evitar para conseguir un equilibrio social.

Neuroética: posee dos vertientes, ética de la neurociencia y neurociencia de la ética. Nace a principios del siglo XXI con el propósito de indagar cuáles las bases cerebrales que rigen nuestra conducta moral. En cuanto a metodología de investigación, uno de sus principales objetivos se encuentra en el desarrollo interdisciplinar de las temáticas que investiga. Asimismo, la revisión, crítica y recreación de conceptos que han recorrido toda la historia de la historia de la filosofía.

(D1) y (D2) El interés por el misterio de nuestro encéfalo se remonta a las antiguas dinastías egipcias

El conocido «papiro quirúrgico de Edwin Smith» se cree que fue escrito en siglo XVII a. C. en el antiguo Egipto y terminado de traducirse en 1930 por James Breasted. La palabra «encéfalo» aparece ocho veces en los documentos egipcios, seis de ellas están en estas páginas. Describen los síntomas, diagnóstico y el pronóstico de dos pacientes con fracturas craneales. El escrito completo está en la *Rare Book Room* de la *New York Academy of Medicine*.

(D3) Hipócrates

Hipócrates dedicó parte de sus indagaciones médicas al funcionamiento del cerebro y su alcance social como conductor de las emociones, del juicio moral y el juicio estético. Asimismo, resultan interesantes sus observaciones sobre la relación del encéfalo con los hábitos morales y la rectitud del pensamiento sereno. Aquí se aprecia con claridad la relación conceptual existente en la antiguedad entre la medicina y la filosofía.

(D4) Origen y nacimiento de la neuroética

El origen de la disciplina debemos buscarlo en el desarrollo de las neurociencias desde la última década del siglo XX hasta nuestros días. La neuroética comprende una red interdisciplinaria de investigación. Cuando hablamos de neurociencias y estudios neurocientíficos nos referimos a una compleja red de laboratorios, tecnología computacional, equipos legales, neurobiología, psicología experimental y psiquiatría entre otras disciplinas. Asimismo, la filosofía ocupa un lugar de relevancia, ya que da soporte a toda una estructura conceptual y al marco epistemológico de la investigación.

Las investigaciones sobre el sistema nervioso y las patologías cerebrales pronto entraron en contacto con las preocupaciones éticas y legales derivadas de dichos trabajos. Por ejemplo, se han analizado las consecuencias del uso de farmacología y dispositivos neurotecnológicos más allá del tratamiento de patologías, dicho de otro modo, los estudios sobre el sistema nervioso «normal» o «sano» con el propósito de mejorar su rendimiento y posible aplicación en otros ámbitos comerciales —deportivo, educativo o militar—.

Como vemos, esta red de saberes alcanza a cuestiones tan complejas y sensibles que abordarlas solo desde perspectivas unilaterales dejaría espacios inconexos en el camino de la investigación.

El desarrollo interdisciplinario de las neurociencias constituye el punto de partida desde donde comenzamos a preguntarnos por el nacimiento de la neu-

roética. Se toma como inicio de las deliberaciones el congreso celebrado en California en 2002, el cual fue patrocinado por la Fundación Dana, la Universidad de Stanford y la Universidad de California, San Francisco.

(D5) Definiciones

Del congreso de 2002 surgieron las primeras cuatro definiciones sobre neuroética:

- **William Safire:** para este periodista la neuroética incluye el examen de lo que es correcto o incorrecto, bueno o malo, acerca del tratamiento, perfeccionamiento, invasiones o manipulaciones del cerebro humano.
- **Neil Levy - Michael Gazzaniga:** estos investigadores abren la puerta a una nueva dimensión en el estudio de la disciplina, permitiendo de esa manera, una extensión de la neuroética más allá de la bioética. Para ellos la neuroética se define como «un intento de proponer una filosofía de la vida con un fundamento cerebral».
- **Kathinka Evers:** la filósofa sueca de la universidad de Uppsala propone dos ejes de investigación, «neuroética aplicada» y «neuroética fundamental», que corresponden a dos perspectivas, una ética de la neurociencia y una neurociencia de la ética.

(D6) Claves

- Campo interdisciplinar de investigación nacido en 2002.
- Dos ejes: neuroética aplicada y neuroética fundamental.
- La investigación tiene una dependencia de la neurociencia y neurotecnología.

(D7) Biología y moral. Antecedentes

Puede observarse en la tradición naturalista un eje histórico de autores. La neuroética en su desarrollo asume y renueva la tradición naturalista de Aristóteles, David Hume, Adam Smith y Charles Darwin, entre otros. La actualización de sus obras a partir de la biología evolutiva, la neurociencia y la genética nos permiten explorar con datos experimentales parte de las bases de lo que sean «la naturaleza social del hombre», «las emociones» o «la teoría de la evolución». Si bien la ciencia ha realizado grandes avances sobre la fisiología cerebral, la evolución o la genética, aún quedan escollos por resolver. No obstante, su aporte a la investigación sobre el origen de los valores morales ha sido determinante en

muchas áreas. Sin embargo, como reconocen algunos especialistas, no hay que caer en el error de creer que una excesiva presencia de cientismo en la neuroética comporta la explicación biológica de la moral. Quien así lo crea «exagera el alcance de la ciencia» (Churchland P. S., 2012).

(D8) Neurobiología y moral. Actualidad

Los cerebros operan en el mundo causal a partir de categorías de episodios de la misma manera que lo haría cualquier animal para sobrevivir, amoldándose a las condiciones ecológicas existentes que, en el mundo social, comprenden la conducta social de los individuos y las prácticas culturales (morales y legales).

La propuesta cimentada en la neurobiología y el naturalismo persigue una explicación de la moral desde un contexto evolutivo y experimental, una visión que por sus características metodológicas entra en fricción con aquellas teorías que son abordadas dese la ética normativa.

Los estudios sobre la oxitocina (OXT), una antigua cadena de péptidos, pueden explicar parte de la base que sostiene nuestras relaciones sociales. La importancia de este péptido, según P. S. Churchland, se basa en que constituye el centro de una complicada red de adaptaciones que poseen los mamíferos para el cuidado de los demás. Tomando como punto de partida la evolución del cerebro de los mamíferos, la oxitocina fue modificándose y adaptándose para atender las tareas del cuidado de la descendencia, y más tarde, ampliar el círculo de cuidado al grupo social.

Los avances en diversos campos de la neurotecnología, la IA y la ciencia de datos nos permiten visualizar, cada vez con mayor claridad, los lazos que unen la biología evolutiva con la moralidad y avanzar hacia una explicación sobre el origen nuestra conducta moral.

La neurobiología es el estudio de las células del sistema nervioso y la organización de estas células dentro de circuitos funcionales que procesan la información y median en el comportamiento. Es una **subdisciplina** tanto de la biología como de la neurociencia.

(D9), (D10) Neuroética aplicada

El psicólogo y químico rumano Corneliu Giurgea, en la década del 60' del siglo pasado, sintetizó los medicamentos llamados «nootrópicos». Tal cuestión nos plantea un dilema moral: ¿por qué esperar si podemos hacerlo ahora?, ¿aceptamos el uso de drogas inteligentes en ámbitos no competitivos, por ejemplo, la escuela?, ¿qué sucede con la meritocracia?

El ejemplo de Arquímedes de Agrigento nos muestra la relación entre el desarrollo científico y la industria militar, por ejemplo, armas de asedio: catapulta, rampa, tortuga, espejos para incendiar velas.

La ciencia en cuestión ahora no es la física sino la neurociencia y la cuestión que se plantea es si se puede controlar su militarización.

(D11) y (D12) P300

Ejemplos de aplicación de una prueba con neurotecnología (P300).

(D13) Investigación en neuroética aplicada

Investigación sobre temas de neuroética aplicada, dos representantes:

- Julian Savulescu
- Jonathan Moreno

(D14) Jonathan Moreno - DARPA

En su libro *Mind Wars: Brain Science and the Military in the 21st Century*, Jonathan Moreno (2012) denunció que la Agencia Gubernamental de Investigaciones Avanzadas (DARPA) financió y promovió la investigación en las siguientes áreas:

- Interface cerebro-máquina (prótesis neuronales). Posibilitaría a pilotos y soldados controlar armas tecnológicamente avanzadas con solo el pensamiento.
- «Robots vivos» cuyo movimiento puede ser controlado a través de implantes cerebrales. Esta tecnología parece que fue probada con éxito en «ratones-robots», los cuales podrían ser conducidos

- a través de mandos a distancia para detectar, por ejemplo, campos minados.
- Fabricación de cascos con capacidad para trazar mapas del estado cognitivo de los soldados con el objetivo de conocer su estado mental.
 - Tecnologías MRI (resonancias magnéticas cerebrales) que permitirían obtener «impresiones digitales cerebrales» en chequeo de sospechosos en aeropuertos o en interrogatorios. Este tipo de prácticas se enfrenta de lleno con 5^a enmienda de la Constitución americana que regula y prohíbe la «autoincriminación».
 - Fabricación de «armas de pulsos» u otros neurodisruptores que provocarían la confusión en los procesos de pensamiento de los soldados enemigos.
 - Fabricación de «armas neuronales» que utilizarían agentes biológicos para provocar la liberación de neurotoxinas, algo prohibido por la Convención de Armas Biológicas.
 - Nuevas drogas que posibilitarían reprimir la inhibición psicológica contra el matar, dejar de dormir durante días, borrar las memorias traumáticas, suprimir el miedo, etc.

El trabajo realizado por Jonathan Moreno es interesante desde el punto de vista que aporta una nueva arista al debate ético y político sobre el uso de la neurotecnología y sus fines. Además, cómo se financian estas investigaciones y si se utilizan fondos públicos.

(D15) Investigación en diversas áreas de la neuroética

- **Kathinka Evers:** neuroética fundamental y aplicada
- **Stephen Morse:** neurociencia y responsabilidad
- **Bernard Baertschi:** relación filosofía-medicina
- **James Giordano:** neurotecnología

(D16) The Human Brain Project – Brain Activity Map

Modelo virtual de cerebro. Científicos de más de 80 instituciones europeas trabajan en el Human Brain Project, una ambiciosa iniciativa que tiene por objetivo construir un modelo virtual de cerebro humano. Traerá grandes avances en neurociencia y en el tratamiento de enfermedades cerebrales.

Técnica renovada de imagen. El neurocientífico español Rafael Yuste, quien lidera la iniciativa del BAM, considera que es el momento oportuno para dar un gran salto cualitativo en la manera que tenemos de observar el cerebro, planteando un proyecto que permita mejorar la calidad de las imágenes, registrando mapas de 3D a gran resolución, desde el cerebro de una mosca Drosophila hasta los circuitos neuronales del cerebro de ratones y primates.

Pero queda todavía lejos del objetivo final del BAM: mapear simultáneamente millones de neuronas en lugar de centenares, hacerlo en 3 dimensiones en lugar de en un corte horizontal, diseñar técnicas que registren la actividad a tiempos más rápidos, incrementar la resolución para poder visualizar más partes de la neurona, construir métodos microscópicos que permitan medir estos patrones de actividad neuronal de manera no invasiva a través del cráneo, y eventualmente no estudiar sólo un ratón sino comparar los patrones de actividad neuronal de una persona sana con los de otra esquizofrénica, epiléptica, con principio de Alzheimer, en pleno proceso de aprendizaje, o durante cualquier otra función cerebral normal o patológica.

El verdadero objetivo del BAM es desarrollar nuevas técnicas para que, de la misma manera que ahora se puede secuenciar un genoma humano completo por mil dólares, dentro de 15, 20 o 30 años sea factible mapear la actividad de circuitos neuronales completos a un costo asequible.

(D17) Neurociencia de la ética

La neurociencia de la ética contribuye a plantear nuevas preguntas a la filosofía. Creo que la frase de Patricia Churchland es paradigmática, pues nos revela un panorama metodológico y de fundamentación diferente a cómo se venían planteando los problemas éticos. Actualmente, se habla sin pudor de una «filosofía experimental» y del trabajo del filósofo en el laboratorio científico.

(D18) El paradigmático caso de Phineas Gage y el estudio de las emociones

Phineas Gage trabajaba de capataz para el Ferrocarril Rutland & Burlington en Nueva Inglaterra. El 13 de septiembre de 1848, a la edad de veinticinco años, sufrió un terrible accidente laboral. Phineas «sobrevivió» del accidente y

su recuperación física fue asombrosa —según el relato de su médico personal, el doctor John Harlow—, aunque con el paso del tiempo comenzaron a manifestarse alteraciones importantes en su conducta. Se tornará un ser asocial y agresivo, un problema para la sociedad. Sin trabajo y abandonado por todos, muere a la edad de treinta y ocho años.

(D24) Damásio: «Gage ya no es Gage»

En *El error de Descartes*, Antonio Damasio (1996) expone este caso paradigmático de la historia de la neurología y las neurociencias. Para el investigador portugués, «Gage no es Gage» luego del terrible accidente.

Dos cuestiones fundamentales derivadas del análisis del caso Gage:

- su lesión mostraba por primera vez que había en el cerebro humano un sistema dedicado al razonamiento por sobre otra función y que esta afectaba específicamente a las dimensiones personales y sociales del razonamiento, es decir, a la conducta moral y ética del individuo.
- que una lesión cerebral semejante podía llegar borrar cualquier norma ética que el sujeto pudiese haber aprendido.

Gage presentaba una conducta inapropiada, había perdido todo el respeto a los valores y a las convenciones sociales, más proclive a los cuentos y fantasía que a los buenos modales.

Estos síntomas están asociados a la *desfrontalización*, una involución de los lóbulos temporales, y al síndrome orbitofrontal manifestado en el cambio de personalidad.

El caso Gage mostraba por primera vez una conexión directa entre una lesión cerebral en los lóbulos frontales y los trastornos de personalidad y comportamiento social. Para Damasio esta triste historia encierra un profundo significado y enseñanza para las neurociencias —y neuroética—, pues las convenciones sociales y las normas éticas podían perderse a causa de una lesión cerebral.

«Gage ya no es Gage». Gage es y no es Gage ¿por qué? Porque siguió manteniendo por muchos años intacta su capacidad de razonamiento. Tal y como hemos visto, esta situación es posible porque, más allá de la pérdida de masa cerebral del lóbulo frontal, de ningún modo se vio afectada su capacidad de razonamiento.

El propio Damasio considera que los pacientes que presentan un déficit en la estructura del lóbulo frontal —al igual que Gage— con deterioro de lo que denomina «emociones sociales» —simpatía, culpa y vergüenza—, aun sabiendo lo que es correcto o incorrecto desde el punto de vista social, fallan en tomar las decisiones adecuadas por el daño de las emociones sociales. Gracias a la lectura de imágenes cerebrales y la neurotecnología, podemos conocer la localización en el cerebro de estas áreas específicas que muestran la vinculación entre la capacidad de razonar, la de sentir y detectar los fallos emocionales que pueden generar una conducta antisocial.

(D20) y (D21) La mente no es algo intangible

Damasio (1996) afirma que los sentimientos de dolor o placer constituyen los cimientos de nuestra mente y, que más allá de cualquier idea de intangibilidad de los mismos, estos pueden ser conocidos en términos biológicos y neurobiológicos. La puerta de entrada a estos estudios fueron las lesiones cerebrales localizadas en determinadas áreas de pacientes neurológicos que exteriorizaban una sintomatología especial, tal como no experimentar vergüenza o sentir compasión, cuando correspondía sentir vergüenza o compasión. Lo asombroso de estos estudios fue que una lesión en un área determinada del cerebro no hacía desaparecer todos los sentimientos, concluyéndose la existencia de varios sistemas cerebrales que son capaces de controlar sentimientos diferentes. Asimismo, cuando un paciente perdía la capacidad para expresar una determinada emoción, perdía también la capacidad de experimentar el sentimiento correspondiente. Sin embargo, esto no ocurría al contrario, quienes perdían la capacidad de expresar determinados sentimientos todavía podían expresar las emociones correspondientes. Estos estudios de Damasio hicieron pensar que las emociones preceden a los sentimientos.

En *El gobierno de las emociones*, Victoria Camps (2011) reconoce el aporte significativo de la psicología cognitiva y las neurociencias en el tema de las emociones y los sentimientos, asimismo, la estrecha vinculación de estos en nuestras decisiones morales. Esta perspectiva —recogida en la Antigüedad por la magistral obra de Aristóteles— conecta a la ética con la educación más que con una lista de deberes y preceptos, propios de la Modernidad.

Si bien en la historia de la ética, a través de sus corrientes y autores, podemos encontrar una abundante bibliografía que hablan de la importancia de las pasiones, los afectos y los sentimientos en la conducta moral, los estudios acerca de las emociones fueron importados al acervo filosófico desde la psicología cognitiva y la neurología, quienes se han encargado de pulir los conceptos, identificando, a partir de los estudios neurológicos realizados con pacientes, una diferencia ordinal-secuencial entre las emociones y los sentimientos en la medida que se manifiestan: en esta secuencia primero aparecen las emociones las cuales producen o son síntomas de la existencia de determinados sentimientos. Desde la perspectiva filosófica, los estudios de las emociones interesan por la relación que puedan tener con la razón, considerándose actualmente una continuidad entre lo sensible y lo racional.

El ejemplo del caso Gage sirve para exponer cómo el desarrollo de las neurociencias —a partir de la neurotecnología— está actualizando el estudio de las emociones y su relación con la conducta moral. No es de extrañar que la ética contemporánea, la neuroética y las ciencias sociales se hagan eco de estos resultados e integren es sus líneas de investigaciones estos nuevos estudios.

(D27) Bioquímica y cooperación en la base de la moralidad

La vida social está fundamentada sobre la base de valores morales. En la raíz de nuestras costumbres o prácticas morales se encuentran los deseos sociales que pueden resumirse en tres grupos:

- apego a los miembros de nuestra familia
- atención a nuestras amistades

- la necesidad de pertenencia a un grupo

Estos valores son los que nos permiten mediar en los problemas que pueden amenazar nuestra supervivencia y bienestar. La gestión de los conflictos sociales —desde revueltas políticas hasta las reuniones de la comunidad de ocho propietarios— ha dado pie al nacimiento de las instituciones, prácticas culturales y convenciones sociales.

(D23) Los genes nuestra moralidad – Neuroendocrinología - OXT

Nadie cuestiona ya que los genes han modelado el cerebro y los han adaptado para la supervivencia, pero ¿de qué modo los genes han organizado el cerebro para valorar a los demás?, ¿pueden nuestros cerebros —y sobre todo las neuronas— preocuparse por algo? En los humanos, al igual que en el resto de los animales, el circuito neuronal persigue el bienestar propio y el cuidado de uno mismo con el objetivo esencial de preservar la propia vida. El objetivo de cada conducta es servir al bienestar del resto de los animales involucrados en ella, de lo contrario esta conducta se desecha porque implica un coste energético que puede poner en peligro la propia vida.

Aproximadamente, 350 millones de años evolución modelaron el carácter de sociabilidad y asociación voluntaria en los mamíferos, asimismo, los estilos de cooperación y cuidado de los demás. Estos cambios quedaron registrados en los cerebros de los mamíferos.

Una evidencia relevante de la manera en que la evolución modificó la organización neuronal de los mamíferos la podemos hallar en los estudios de neuroendocrinología. Estas investigaciones muestran cómo los humanos modificaron sus valores para generar otros nuevos y hacerlos extensivos al bienestar de terceros, esto es, un trabajo evolutivo que al principio solo incluía a la descendencia, y que luego fue extendiéndose paulatinamente a otros círculos —prole, amigos y desconocidos— dando nacimiento a lo que más adelante será la moralidad.

P. Churchland (2012) realiza una descripción de cómo la hormona oxitocina (OXT), una antigua cadena de péptidos, puede explicar parte de la base de las relaciones sociales.

La importancia de este péptido se basa en que constituye el centro de una complicada red de adaptaciones que poseen los mamíferos para el cuidado de los demás. Tomando como punto de partida la evolución del cerebro de los mamíferos, la oxitocina fue adaptándose para atender las tareas del cuidado de la descendencia, y más tarde, ampliar el círculo de cuidado al grupo social.

De esta manera y gracias al desarrollo de la neurotecnología contemporánea, los lazos entre la biología evolutiva y la moralidad se muestran cada vez más estrechos, sobre todo, de cara a la búsqueda de una explicación de la conducta moral humana.

Ahora bien, ¿cuáles son los mecanismos biológicos que hacen posible extender el cuidado más allá de la descendencia? Una posible respuesta a esta pregunta la encontraríamos, por ejemplo, en los mecanismos cerebrales y corporales que «maternalizan» el cerebro de la hembra y que están ligados a un conjunto de hormonas, entre las que se encuentran la oxitocina y la vasopresina arginina. La hipótesis que sostiene Churchland (2012) es que en un principio estos mecanismos no fueron seleccionados para servir a una serie amplia de finalidades sociales, sino todo lo contrario, buscaban asegurar a la hembra una cantidad pequeña de recursos necesarios —por ejemplo, amamantar y defender— que garantizasen el cuidado de las crías hasta que pudiesen valerse por sí mismas.

¿Cómo viajan las hormonas? La placenta del feto libera a través del torrente sanguíneo este conjunto de hormonas que actúa principalmente sobre las neuronas de las estructuras subcorticales, moldeando el carácter de la embarazada y «maternalizando» el cerebro de la madre. Este proceso hormonal está ligado a patrones comunes de conducta tales como: una mayor alimentación, la preparación del nido o el acondicionamiento de un lugar seguro para el parto. El nivel de producción de oxitocina se eleva, siendo fundamental en el período de lactancia. Una vez explicado el cuidado del círculo afectivo más cercano a partir

de la biología evolutiva y la química, es menester considerar su extensión, o bien, preguntarnos cómo es posible el cuidado de otros seres humanos más allá de este primer círculo.

Según Churchland (2012), esta extensión puede ser posible gracias a «pequeñas modificaciones adicionales» hasta alcanzar el cuidado de individuos ajenos a la propia descendencia, «pero cuyo bienestar es de alguna manera significativo para el bienestar propio y el de nuestra descendencia»

Es realmente interesante ver cómo desde la neuroquímica se puede hallar una primera explicación de lo que pueda ser el comienzo de la sociabilidad humana: no está en la palabra ni en la cultura, sino en la química. El cuidado de los demás más allá de la propia descendencia representa un cambio evolutivo en el cerebro de gran magnitud. Esto nos coloca en lo que podemos denominar «la antesala de la moralidad».

Estos cambios en el cerebro son apreciables a partir de dos modificaciones, (a) miedo y ansiedad ante la amenaza de separación de la cría y placer ante el reencuentro y (b) el aprendizaje grabado en la memoria relacionada con el placer y el dolor. Esta aptitud permite a los miembros de un grupo conocer las costumbres de los demás y anticiparse a los problemas. Estas modificaciones genéticas han permitido el desarrollo de la conciencia individual y grupal.

Otro indicio sobre las bases neuroquímicas de la moralidad lo podemos hallar en el término «apego». La neuroendocrinología lo utiliza para hacer una descripción de la disposición al cuidado de las demás personas y deseo de su compañía. Por el contrario, un sentimiento de dolor o tristeza por la separación.

La bioquímica hormonal influye sobre nuestro cuerpo modificándonos el comportamiento y la conducta para que podamos adaptarnos y sobrevivir en los diferentes círculos de nuestro desarrollo humano —ámbito familiar y ámbito social— por medio del «apego» y del «cuidado de otras personas». Esta disposición natural, según la neuroendocrinología y la psicobiología, está relacionada con otro de los pilares fundamentales que conforman la base de nuestra moralidad y nuestra cultura, nos referimos a la cooperación.

(D24) y (D25) Tomasello y la cooperación

Para Michael Tomasello (2010) la cultura humana posee un conjunto exclusivo de «habilidades cooperativas y motivaciones para colaborar» que se manifiestan en las instituciones sociales. Dichas formas de cooperación son posibles gracias a procesos biológicos denominados «intencionalidad compartida», esto es, una capacidad humana para generar junto a los demás intenciones, reglas y compromisos conjuntos para desarrollar las empresas cooperativas.

Existen diferencias con otras especies animales. La cooperación humana no solo se basa en el proceso de imitación y aprovechamiento, sino que fundamentalmente coopera (ejemplo de acción colectiva, la ciencia y la tecnología). La «inteligencia cultural» que desarrollamos desde niños comprende habilidades exclusivas de nuestra especie tendientes a la colaboración, comunicación y aprendizaje social que van constituyendo un proceso de pensar cooperativo.

Para el Tomasello, la base de la cooperación humana no estaría en el altruismo, que sería secundario, sino en el mutualismo ya «que abarca los actos en que todos nos beneficiamos con la cooperación, pero solo si trabajamos juntos, si colaboramos». Existe un beneficio colateral que obtenemos a partir acciones altruistas.

(D26) Biología y cultura - Mundos culturales

¿Cuál es el punto de confluencia entre la biología y la cultura? El éxito evolutivo de la humanidad estuvo ligado a la agricultura y a vida en las ciudades, desde donde comenzó a desplegarse y aumentar en población desde hace unos diez mil años. El surgimiento de las primeras organizaciones cooperativas dio paso a los problemas y la complejidad de las sociedades industriales contemporáneas.

¿Cómo fue posible llegar a las formas complejas de cooperación de las sociedades actuales? La respuesta estaría, «en la evolución de habilidades y motivaciones cooperativas —altruistas y de cooperación— a partir de interacciones de pequeños grupos humanos» (Tomasello, Marengo, & Dweck, 2010). La

colaboración es una distinción propia de la especie humana que es observada desde muy temprano en la ontogenia.

El éxito cultural. La biología ha preparado a los seres humanos para crecer y desarrollarse en un contexto cultural. Esto es evidente ya que un niño jamás podrá llegar a una edad adulta al margen del medio cultural que lo contenga, pues cada ser humano necesita de otros para aprender, muy diferente de la ontogénesis de los primates que pueden desarrollar habilidades cognitivas y sociales en diversos contextos sociales. Tanto la adaptación biológica como nuestras actividades de colaboración nos han permitido «crear mundos culturales a los cuales nos vamos adaptando continuamente».

(D27) Neuronas espejo y empatía

Los experimentos invasivos realizados con electrodos implantados en los cerebros de los monos *Rhesus* revelaron un hallazgo fascinante: las neuronas espejo. Dicho descubrimiento, fortuito, según comenta Marco Iacoboni (2011), alentó las expectativas de los científicos cognitivos con respecto al estudio de los mecanismos neuronales para comprender los estados mentales de los demás. Si bien las sospechas de la existencia de un mecanismo de funcionamiento similar no es algo novedoso, la historia de las neuronas espejo comenzó su camino en la Universidad de Parma, en el laboratorio de neurofisiología a cargo de Giacomo Rizzolatti.

Las neuronas espejo están localizadas específicamente en el área denominada F5. Los electrodos implantados en esta área fueron los encargados de registrar todos los cambios eléctricos —denominados por los neurofisiólogos «potenciales de acción»— que se producían, segundo a segundo, en la superficie de cada neurona cuando los monos, por ejemplo, realizaban ciertas tareas a cambio de pequeñas recompensas alimentarias. Estos cambios registrados en los potenciales de acción indicaban que una determinada neurona estaba activada en un momento dado. Cuando una neurona «se dispara», es decir, que muestra un estímulo, lo hace para codificar un «evento sensorial» (por ejemplo, al ver un objeto o una acción), un «acto motor» (coger una fruta) o un «proceso cognitivo» (recordar una acción determinada).

Los neurofisiólogos italianos interpretaron que los macacos eran capaces distinguir una *intención*. Si estaban en lo cierto y las neuronas espejo eran capaces de representar una finalidad o intención, dicho descubrimiento permitiría el salto a la comprensión de la neurobiología de la atribución mental (capacidad para atribuir pensamientos e intenciones).

Estos descubrimientos alentaron una hipótesis de carácter general que postulaba por primera vez que las neuronas espejo «pueden ser el correlato neuronal de los procesos de simulación necesarios para entender otras mentes», es decir, que poseen la capacidad para atribuir pensamientos e intenciones a otras personas. Lo que estaban diciendo Gallese y Goldman era que las neuronas espejo podrían explicar una «teoría de la mente».

La teoría de la simulación dice que, para que seamos capaces de sentir lo que sienten otras personas (por ejemplo, amor, odio o deseo) debemos simularlo nosotros mismos. Sin embargo, el problema que se plantea es saber de qué manera nuestro cerebro es capaz de realizar una simulación para que produzca tales resultados. La clave reside en que las neuronas de la corteza premotora simulan un movimiento cuando observamos que alguien lo realiza, es «como si quisieran hacerlo sin hacerlo en realidad».

Algo importante:

- si bien la hipótesis de la simulación podía explicar el mecanismo de atribución u objetivos, no obstante, no lograba resolver el mecanismo de funcionamiento de las creencias o las emociones.
- Las investigaciones realizadas a través de técnicas de imagen, como la tomografía por emisión de positrones (TEP) o la resonancia magnética (RMF) no pueden demostrar la existencia de neuronas espejo en humanos, solo pueden proporcionarnos «indicios indirectos de su existencia».

(D28) Empatía. Theodor Lipps

El psicólogo alemán Theodore Lipps describió, en una primera etapa, la relación entre la obra de arte y el espectador como *Einfühlung* (empatía). Luego, en una segunda etapa, hizo una ampliación de dicho concepto con el propósito

de incluir a las interacciones interpersonales. Para ilustrar la forma de cómo percibimos los movimientos de los demás y los imitamos en nuestro cerebro, tomó como ejemplo a un espectador que observa el trabajo de un volatinero en la cuerda. En el momento de la observación, todos los espectadores estaríamos dentro del acróbatas.

(D29) y (D30) Origen de la empatía en las neuronas espejo

Iacoboni (2011) remarca el papel fundamental que tiene la empatía en nuestra vida social, ya que nos permite compartir emociones, experiencias y necesidades comunes. Para el autor existen muchas pruebas empíricas llevadas a cabo por las neurociencias que nos pueden asegurar la existencia de un vínculo fuerte entre las neuronas espejo y la empatía. Sin embargo, desde otras posturas más prudentes, advierten que aún tenemos pocos datos que puedan confirmar que un sistema de neuronas espejo sea el substrato que permite atribuir intenciones a los demás. ¿Cuáles son estos indicios indirectos que ven los neurocientíficos para deducir la existencia de neuronas espejo en humanos? Los datos ofrecidos por las resonancias magnéticas muestran una mayor actividad en la zona cerebral que es homóloga al área del cerebro F5 de los monos. Sin embargo, estos estudios no pueden ofrecer una certidumbre en su totalidad.

Algunos neurocientíficos consideran que nuestras respuestas empáticas para entender los estados mentales de los demás se explican mejor a partir de *simulación*, es decir, que podemos entender, por ejemplo, el sufrimiento de los demás «porque nuestro cerebro simula una expresión facial de tristeza». Esta misma simulación puede trasladarse a otros estados, como el miedo, la ira, el desagrado, etcétera. La propuesta de Marco Iacoboni a partir de la simulación puede parecer, a primera vista, concluyente, sin embargo, es necesario aclarar que de momento no contamos evidencias suficientes que expliquen en su totalidad la correlación causa-efecto de su hipótesis.

Según Iacoboni, la imitación está estrechamente relacionada con la finalidad, el movimiento reflejo y un sistema de neuronas espejo. Todo este conjunto conforma el circuito central de imitación que se despliega en las acciones empáticas. Sin embargo, es muy importante estar atentos a los datos disponibles, ya

que estos no terminan de explicar cómo se produce la conducta imitativa. En tal sentido se han señalado dos objeciones que es oportuno tener en cuenta al momento de hacer una evaluación general de la *hipótesis de la imitación* basada en un sistema de neuronas espejo:

- Una primera objeción señala que el fenómeno espejo descrito en el experimento clásico con monos no es en realidad una imitación, puesto que el primate no imita lo que ve, ni sus músculos muestran algún movimiento relacionado.
- En segundo lugar, un metaanálisis ha demostrado que no existen datos fidedignos de que el área 44 del cerebro humano participe durante los procesos de imitación. Estos estudios no demuestran que el sistema de neuronas espejo no participa en la imitación.

(D31) Objetiones de Churchland

Los argumentos trabajados por P. Churchland (2012) buscan desmontar parte del optimismo puesto en la teoría de la simulación, además, se trata de ejemplos que podemos experimentar en forma personal y extraer nuestras propias conclusiones.

En la primera argumentación, la autora analiza, desde su punto de vista —textualmente, «desde su fenomenología»— qué es lo que siente cuando ve que alguien llora después que una avispa le ha picado en el pie, siendo el contexto de la situación muy variado, ya que puede depender si el que llora es su bebé o un intruso que se ha colado en su jardín. Llevemos al extremo la situación y supongamos que quien ha recibido la picadura es un familiar muy cercano, por ejemplo, nuestro hijo. En tal caso, nosotros no sentimos literalmente el dolor del agujón en nuestro pie. Lo que sí podemos sentir es un rechazo visceral al dolor y el impulso de asistir al infeliz con un antihistamínico.

El segundo argumento está relacionado con la observación, por ejemplo, de una persona enfadada que puede no generar ira en el observador, sino miedo, vergüenza o risa, todo esto dependiendo del contexto. Podemos reconocer el enfado, sin enfadarnos, y también el disgusto sin disgustarnos. Llevando el ejemplo al límite, también podemos sentir alivio o alegría si nuestro enemigo sufre

algún dolor. Sin embargo, estos y otros escollos en forma de estudios experimentales no han podido mitigar el entusiasmo por las neuronas espejo.

Hasta aquí el capítulo dedicado a la neurofilosofía en el cual he querido mostrar su nacimiento, carácter interdisciplinar y los diferentes caminos que pueden tomar sus investigaciones, tal y como hemos visto en el caso de la neuroética. Asimismo, me gustaría remarcar todo el aspecto epistemológico que falta por investigar, es decir, el trabajo en un modelo integral mente y cerebro, una discusión aún abierta.

Más allá de los hitos históricos marcados por la neurofilosofía, me gustaría remarcar su importancia de cara a los procesos de sociabilidad de las próximas décadas, donde la inteligencia artificial y la robótica tendrán un mayor protagonismo en todos los ámbitos humanos. En tal sentido, considero que nada de lo que ocurra en el medio psico-bio-social en el que nos movemos le será indiferente a nuestro cerebro y nuestra mente. Si algo nos han enseñado las neurociencias es sobre su plasticidad y conexión natural con otros cerebros, de nosotros dependerá su salud y el impacto de nuestra conducta en la sociedad.

En el próximo capítulo dejaré planteadas, a modo de propuesta personal, algunas ideas para construir una línea de investigación y poder continuar las iniciadas en esta tesis. Está claro que el desafío de los frentes abiertos es grande y la realidad compleja, pero más grande es la curiosidad que nos mueve a seguir buscando aquellos fundamentos que hacen que nuestra especie sea por naturaleza social.

Capítulo 8. Propuesta: hacia una neurofilosofía de la sociabilidad

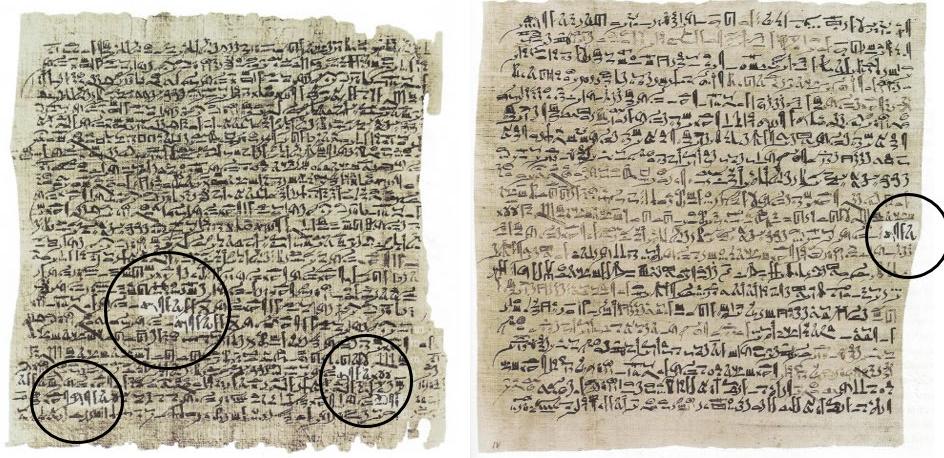


Figura 5. Papiro quirúrgico de Edwin Smith, del siglo XVII a.C.

Las figuras son conocidas como «el papiro quirúrgico de Edwin Smith» y datan, según fechas aproximadas, de la Dinastía XVIII de Egipto (siglo XVII a.C.). El documento está redactado en escritura hierática y describe algunas observaciones anatómicas y quirúrgicas. Más allá de la importancia histórica del hallazgo, aparecen remarcados con círculos unos símbolos que se repiten y que representan la palabra «encéfalo», la cual podría pronunciarse como «ys» (Kandel, Schwartz, & Jessell, 2008). El papiro contiene «las primeras descripciones de suturas craneales, de la meninge, la superficie externa del cerebro, del líquido cerebroespinal y de las pulsaciones intracraneanas» (Wilkins, 2020).

Como vemos, el interés humano por el encéfalo ha quedado registrado en la historia desde hace miles de años. Ya sea con símbolos o palabras, hemos manifestado nuestro deseo por conocer, aunque sea de forma intuitiva, cómo era aquel espacio que elaboraba nuestras ideas, recuerdos, temores, creaciones, el mundo de las sombras y de los sueños.

A través del recorrido histórico de la filosofía, el arte y la ciencia, podemos comprobar que esa curiosidad por saber qué esconde nuestro encéfalo no ha variado. Claro está, hoy podemos jactarnos del desarrollo científico y tecnológico, y renovar un gran número de preguntas, otras en cambio, las damos por respondidas. Sin embargo, nuestro conocimiento no es lineal, sino que vamos dando saltos de avance y retroceso, una situación que enriquece el saber y nutre nuestra curiosidad siempre insatisfecha.

Esa misma curiosidad es la que ha movilizado este trabajo. En el recorrido de la tesis ha habido avances, retrocesos, preguntas, respuestas y dudas que han colaborado a lo largo de su realización. Sin embargo, desde el primer momento del proyecto tuve la certidumbre que este trabajo podía convertirse en la cimentación para una eventual línea de investigación que pudiese completarlo.

En tal sentido, he querido dedicar este capítulo a trazar las primeras líneas de esa continuación a partir de un proyecto que integre el estudio de la sociabilidad humana y la neurofilosofía como interfaz filosofía, neurociencia y psicología.

La neurotecnología (relación humano-máquina), la inteligencia artificial (realidad extendida, metaverso, etcétera), el desarrollo genético y la robótica están planteando serios desafíos a nuestras sociedades contemporáneas, sobre todo, en el campo de las relaciones humanas, las patologías sociales y mentales, la intromisión ideológica, nuestra relación con el medio ambiente y una importante influencia sobre nuestra identidad. Todas ellas tienen una incidencia en nuestra salud mental y cerebral y, directamente, en la construcción de la realidad.

Como comenta P. Churchland (2012), tenemos capacidades innatas para identificarnos con los demás, para empatizar y saber lo que le pasa a uno le puede a todos los demás:

La mayoría de los seres humanos tienen la habilidad para identificarse con los demás, y, en términos más generales, para «leerles la mente», es decir, para saber lo que los demás sienten, intentan, quieren, etcétera. Cuando esa capacidad empieza a fallar, por ejemplo, como resultado de una demencia frontotemporal, que implica degeneración del tejido neural

de las cortezas frontal y temporal, el efecto es verdaderamente catastrófico. Este dato nos recuerda la enorme importancia que tienen las habilidades que acostumbramos a ejercer sin esfuerzo, de un modo fluido y rutinario. (pág. 133)

Está demostrado que nuestro cerebro y mente son permeables, no sólo a la influencia externa (sociedad, medioambiente y cultura), sino también interna, como lo hemos visto en la relación entre la memoria autobiográfica y la sociabilidad.

El proyecto neurofilosófico que nace del desarrollo interdisciplinar es consecuente con una metodología experimental de investigación (Churchland P. S., 2007, pág. 187). En tal sentido, una neurofilosofía de la sociabilidad como la línea de investigación que se propone tendría que estar alineada dentro de esta metodología.

La propuesta interdisciplinar de investigación iniciada por neurofilosofía desde sus diversas perspectivas ha demostrado que es capaz de realizar interesantes propuestas y debates sobre conceptos clave, tanto para la filosofía como para el campo de las ciencias, tales como: «conciencia», «experiencia», «voluntad», «mente», «libre albedrío», «justicia», «verdad», «inteligencia» o «memoria autobiográfica».

El avance de la neurofilosofía ha permitido allanar el camino a otras disciplinas y saberes que beben de sus fuentes. Como hemos visto en los apartados anteriores, tal es el caso de la neuroética que focaliza sus investigaciones en el origen de la moralidad, nuestro carácter social o la naturaleza de los valores morales, entre otros temas, partiendo muchas veces de los estudios neurológicos sobre las lesiones cerebrales y neuropatologías: otra puerta de acceso a la compleja fisiología cerebral por medio de la neurotecnología (Churchland P. S., 2007, pág. 185).

Estas cuestiones plantean nuevos desafíos al desarrollo de la inteligencia artificial, sobre todo a la hora de procesar una ingente cantidad de datos, además, claro está, la captación legal de los mismos. Como podemos intuir, subyace

en el recorrido de la propuesta una pregunta para las próximas décadas y es si la filosofía podrá trabajar al margen de las neurociencias y de la ciencia de datos.

Considero que otro punto a destacar a la hora de iniciar una línea de investigación sobre la neurofilosofía de la sociabilidad lo constituye la *neurobiología de las habilidades sociales*, pues ella nos ayudará en la indagación de las bases biológicas que rigen determinados hábitos que modelan nuestra moralidad. Su cooperación con las neurociencias cognitivas y la neurotecnología están abriendo nuevas preguntas para hallar una fundamentación de la moralidad basada en los recientes descubrimientos de las funciones cerebrales.

El avance de las neurociencias y la neurotecnología nos plantea la posibilidad de que dichas disciplinas puedan integrarse dentro del modelo planteado de neurofilosofía de la sociabilidad y trabajar de forma cooperativa con el objetivo de revisar críticamente los trabajos de la tradición filosófica y ofrecer un abanico de preguntas actualizadas a las sempiternas cuestiones, por ejemplo, la naturaleza humana, cómo nos constituimos en seres sociales o por qué nuestra especie es capaz de construir un *ethos* moral.

Para P. Churchland (2012, pág. 179) la neurobiología de las habilidades sociales es una disciplina «joven pero vigorosa» capaz de sintetizar los últimos descubrimientos entre la psicología, la biología y las neurociencias.

El libro *El error de Descartes* (Damasio A. R., 1996) es un claro ejemplo de cómo la neurobiología puede realizar grandes aportes al ámbito de la moral y la construcción de la sociabilidad humana, un terreno cultivado históricamente por las ciencias sociales y la filosofía⁸. En esta obra, A. Damasio rescató de los anaqueles de la historia de la neurología el paradigmático caso de Phineas P. Gage, un trabajador ferroviario del siglo XIX que pasó a los anales de la medicina a partir de un accidente cráneo-encefálico con pérdida de masa cerebral que sufrió mientras manipulaba explosivos. Documentada al detalle por el médico

⁸ Aunque la síntesis propuesta por Damasio es innovadora, su análisis acerca de la filosofía cartesiana ha sido muy criticado por tratarse de una evaluación retroactiva e intempestiva del pensamiento de Descartes. Véase Juan Arana (Diodado, Rodríguez Valls, & Arana, 2010, págs. 203-215).

John Harlow, la historia de Gage se transformó en un caso paradigmático para la investigación neurológica, psiquiátrica y filosófica. Tras el accidente, vivió trece años, pero «Gage ya no era Gage», se había transformado en otra persona y, por su conducta, ya no se le reconocía como tal (Damasio A. R., 1996, pág. 27).

El caso Gage mostró la existencia en el cerebro humano de un sistema de razonamiento por sobre otra función, y que esta afectaba específicamente a las dimensiones personales y sociales del razonamiento, es decir, a la conducta moral del individuo. Por otro lado, se vio que una lesión cerebral semejante a la sufrida por Gage podía llegar a borrar cualquier norma ética que el sujeto hubiera aprendido en el transcurso de su vida. Asimismo, el caso Gage demostraba una conexión directa entre una lesión cerebral en los lóbulos frontales, los trastornos de personalidad y el comportamiento social. Según Damasio (1996), esta historia confirma que «[...] la práctica de las convenciones sociales y normas éticas adquiridas previamente (al accidente) podían perderse como resultado de una lesión cerebral, aun cuando ni el intelecto básico ni el lenguaje parecían hallarse comprometidos» (pág. 29).

Los pacientes que presentan un déficit en la estructura del lóbulo frontal—similar al de Gage— con deterioro en las «emociones sociales» (simpatía, culpa y vergüenza), «aun sabiendo lo que es correcto o incorrecto desde el punto de vista social, fallan en tomar las decisiones adecuadas a causa del daño en el centro que regula las emociones sociales» (Ruiz Rey, 2009). En la actualidad, el desarrollo de la neurotecnología de imágenes nos puede informar con precisión la localización de las áreas cerebrales que muestran la vinculación entre la capacidad de razonar, sentir y detectar los fallos emocionales que pueden generar una conducta antisocial.

A lo largo del trabajo comenté otro punto que considero fundamental para la construcción de una neurofilosofía de la sociabilidad: la hipótesis de la simulación y las neuronas espejo.

Las investigaciones con neuronas espejo han hallado una correlación entre la imitación y la empatía, dos capacidades necesarias para un adecuado

desarrollo de las habilidades sociales y morales (incluidas desde la más primigenia relación entre la madre y su cría, hasta la complejidad sociopolítica de las sociedades contemporáneas). Además, sin ellas no podríamos comprender, entre otras cosas, los pensamientos, las creencias y los deseos de las demás personas. Es decir, la vida en sociedad sería inviable. Estos estudios sugieren que es el mecanismo fisiológico del sistema de neuronas espejo quien posibilita dichas capacidades. No obstante, las evidencias directas de la existencia de las neuronas espejo están limitadas a los experimentos con electrodos implantados en macacos, por tal motivo debemos suponer (remarco «suponer») que en el cerebro humano subyace un mecanismo espectral de análogo funcionamiento al del mono, capaz de atribuir intenciones y objetivos a otras personas.

Patricia Churchland aporta una síntesis bastante completa sobre las bases neurobiológicas en que se asienta la construcción de la sociabilidad humana: unas investigaciones focalizadas en desvelar el proceso de cómo el hombre se constituye en un ser social, aquello que Aristóteles denominó la «naturaleza política del hombre» (Aristóteles, 2007a, pág. 58).

Para esta investigadora, la sociabilidad humana está fundamentada en los deseos sociales (valores sociales) que sostienen nuestras costumbres, y que ha resumido en tres grupos: (a) apego a los miembros de nuestra familia, (b) atención a nuestras amistades, y (c) la necesidad de pertenencia a un grupo. Estos valores nos permiten mediar en los problemas que pueden amenazar nuestro bienestar, gestionar conflictos sociales y crear convenciones sociales y culturales (Churchland P. S., 2012, pág. 23), es decir, alcanzar un nivel de sociabilidad que garantice la supervivencia de la especie por sobre el individuo.

En el capítulo 6 hablamos de un factor fundamental que ha favorecido el desarrollo de la sociabilidad de los mamíferos: la neuroquímica. La evolución modeló el carácter de sociabilidad en los mamíferos, la asociación voluntaria, los estilos de cooperación y el cuidado de los demás.

Según indica Churchland (2012, pág. 25), el registro evolutivo que modificó la organización neuronal puede verse con claridad, por ejemplo, en la aso-

ciación para la reproducción o en el cuidado de la descendencia. La neuroendocrinología nos ofrece pistas de cómo los cerebros humanos modificaron sus valores para generar otros nuevos y hacerlos extensivos al bienestar de terceros, es decir, un desarrollo evolutivo que al principio solo incluía a la descendencia directa, y que luego se extendió a otros círculos menos próximos como familiares de segunda línea, amigos, vecinos y desconocidos, configurando de esa manera las primeras formas de lo que hoy conocemos por moralidad.

Otro aspecto fundamental que relaciona la neuroquímica y los procesos de sociabilidad lo constituyen dos hormonas, oxitocina (OXT) y vasopresina arginina (VPA), que son reguladoras de los mecanismos cerebrales que actúan en el cerebro de la hembra maternizándolo durante el embarazo. La placenta del feto libera el grupo de hormonas a través del torrente sanguíneo que actúa, principalmente, sobre las neuronas de las estructuras subcorticales, moldeando el carácter y maternalizando el cerebro de la madre.

Este proceso hormonal, en plena ebullición durante los meses de embarazo, está ligado a patrones comunes de conducta tales como una mayor alimentación, la preparación del nido o el acondicionamiento de un lugar seguro para el parto.

El nivel de producción de oxitocina durante el embarazo siempre se regula al alza, siendo fundamental durante la lactancia. Una vez cubierto el cuidado hacia círculo afectivo más próximo, su extensión será posible gracias a «pequeñas modificaciones adicionales» hasta alcanzar el cuidado de individuos ajenos a la propia descendencia, «pero cuyo bienestar es significativo para el bienestar propio y el de nuestra descendencia» (Churchland P. S., 2012, pág. 44).

Junto a las modificaciones realizadas por la oxitocina (OXT) en nuestro cerebro, hemos visto en los capítulos anteriores otros dos cambios que han sido fundamentales para el desarrollo del proceso de sociabilidad, me refiero al apego/desapego y a la sensación de placer/dolor para estimular la capacidad de aprendizaje y la memoria que permite a los individuos del grupo conocer las costumbres, hábitos y realizar previsiones de futuro.

Toda esta bioquímica estará enraizada con otro de los fundamentos donde se apoya la sociabilidad: la cooperación (Tomasello, Marengo, & Dweck, 2010, págs. 12-17). La biología nos ha preparado para crecer y desarrollarnos en un contexto cultural. Esto es evidente ya que un niño jamás podrá llegar a una edad adulta al margen de un nicho socio-cultural que lo contenga, pues cada ser humano necesita de otros para aprender, una característica muy diferente, por ejemplo, de la ontogénesis de los primates quienes pueden desarrollar habilidades cognitivas y sociales en diversos contextos sociales. Tanto la adaptación biológica como nuestras actividades de colaboración nos han permitido «crear mundos culturales a los cuales nos vamos adaptando continuamente» (Tomasello, Marengo, & Dweck, 2010, pág. 124).

El avance de las neurociencias en esta primera parte del siglo XXI está teniendo efectos colaterales positivos en un gran abanico de disciplinas científicas y humanísticas. La filosofía no solo ha sido capaz de sintetizar gran parte de sus resultados, sino que a su vez está colaborando en devolver una neurociencia renovada y revisada, tanto en conceptos como en diversidad temática.

En la tesis he querido mostrar y dejar constancia que de dicha renovación y del trabajo interdisciplinar han surgido dos disciplinas que son complementarias, la neurofilosofía y la neuroética, cuyos objetivos, según mi punto de vista, son plantear nuevas preguntas, quizás más completas y complejas.

Desentrañar el funcionamiento de la mente ha sido una empresa que ha estado presente en el pensamiento filosófico de todos los tiempos. A diferencia de otras épocas, hoy contamos con el apoyo de la neurotecnología para nuestras investigaciones, lo cual no significa que tengamos resueltas todas las dudas ni mucho menos, sino que el camino hacia las evidencias se nos presenta algo más despejado.

Creo que la propuesta nacida en la neurofilosofía es una alternativa válida y completa de la que podemos rescatar la premisa de que para conocer nuestra mente es necesario conocer el funcionamiento de nuestro cerebro.

En la actualidad, contamos con ingentes conocimientos en genética, neurobiología y neurofisiología que, sumados a las investigaciones en ciencias sociales, informática y humanidades, nos permitirán escrutar las principales teorías sociales y filosóficas. Todos ellos, junto al desafío que supone la IA, constituyen una fuente de información muy completa para escrutar los caminos que una neurofilosofía de la sociabilidad nos puede ofrecer.

Capítulo 9. Conclusiones

Este trabajo partía de la hipótesis de lo que he considerado como la existencia de una influencia recíproca entre la memoria autobiográfica y la sociabilidad la cual permite a nuestra especie su desarrollo bio-sociocultural. Teniendo en cuenta esta hipótesis, se proponían dos objetivos principales: (a) revisar aquellos datos que aporten evidencias que la relación memoria autobiográfica-sociabilidad constituye lo que podemos considerar como la «antesala de la moralidad»; y (b) proponer una línea de investigación que he denominado *neurofilosofía de la sociabilidad* con el propósito de continuar las indagaciones aquí iniciadas.

Más allá de la importancia que este trabajo pueda tener en el plano ético o político, considero que también podrían tenerse en cuenta otros fenómenos derivados como pueden ser los mecanismos psico-neurológicos de resistencia ante la mentira, la desinformación o la falta de confianza. Sobre todo, si se tienen en cuenta la expansión social de la IA y de la realidad extendida.

El modelo de neurofilosofía planteado por Patricia Churchland ha servido de guía a esta investigación en muchos de los aspectos que enmarcan los temas dentro de la relación filosofía-neurociencias. Sin embargo, bajo este modelo subyace un problema epistemológico derivado del materialismo eliminativista que excede los límites de esta tesis. Este problema de fondo afecta a todo el marco epistemológico de la neurofilosofía. Por ello, como se ha remarcado en varios puntos del trabajo, la discusión con el eliminativismo aún no está cerrada. En tal sentido, la propuesta de seguir avanzando en el análisis del eliminativismo como soporte del modelo neurofilosófico queda abierta para un proyecto de investigación futuro en el que se trate de forma exclusiva.

A continuación, las quince conclusiones que extraigo de este trabajo. La última conclusión está centrada en la propuesta personal de investigación, integrando un nuevo elemento que será central: la identidad.

1. La relación existente entre la memoria autográfica y la sociabilidad, es decir, entre un proceso cognitivo interno e individual y un proceso externo y social es una vía de comunicación bidireccional. Dicho flujo de comunicación forma parte del proceso evolutivo de nuestra especie, tal como lo ha probado la biología y la neuroendocrinología encargada de los estudios hormonales. Asimismo, la psicología experimental y la neurociencia cognitiva en relación con los estudios de la memoria.
2. Para la investigación de la relación entre memoria autobiográfica y sociabilidad se destaca el estudio de determinadas patologías mentales que deterioran el proceso cognitivo, por ejemplo, demencias seniles o Alzheimer. En fases avanzadas se observa el impacto de dichas patologías en los procesos de sociabilidad y reconocimiento del medioambiente bio-socio-cultural.
3. Se determina el concepto de «investigación inversa» para estudiar la relación memoria autobiográfica-sociabilidad. Dicho concepto implica la observación y el análisis del impacto de las patologías mentales, tanto en los procesos cognitivos que intervienen en los mecanismos y formación de la memoria autobiográfica como en los procesos sociales (por ejemplo, el desapego o la falta de empatía social, soledad no deseada). Asimismo, este proceso de investigación inversa puede aplicarse al estudio y observación de las patologías sociales y su impacto en la estructura cognitiva y emocional de los sujetos (por ejemplo, el impacto de la pandemia y su influencia sobre las bases estructurales de sociabilidad).
4. En una primera fase de estudio, se determina que la estructura de la memoria autobiográfica comprende tres conceptos relacionados entre sí que tienen connotaciones personales y sociales: *narrativa*, *experiencia* y *recuerdos*. La investigación de estos conceptos puede abordarse desde la filosofía, la psicología experimental o la neurofilosofía de la sociabilidad. Lo óptimo es que dicha investigación pueda realizarse en un marco interdisciplinar.
5. La narración de nuestras experiencias vitales nos permite al acceso a nuestra memoria autobiográfica. En tal sentido, es una experiencia subjetiva y social,

donde la influencia transita en dos sentidos y permite la particular visión de la realidad.

6. La memoria es un sistema cognitivo que no solo permite el almacenamiento y la recuperación de información, además, posibilita la creación y re-creación hacia el futuro de nuestras experiencias.

7. El sistema cognitivo de la memoria autobiográfica es permeable a la influencia de las emociones. Esta influencia puede generar mecanismos psíquicos de defensa por medio de recreaciones y fantasías que incorpora el sujeto a su vida como realidades.

8. En la actualidad, el aumento desproporcionado de las modalidades de sociabilidad virtual a través de programas de realidad extendida (por ejemplo, el metaverso, los videojuegos o aplicaciones móviles) plantean un serio desafío educativo, sanitario y político a todas las sociedades. Está claro que el sistema cognitivo de la memoria autobiográfica, tal como está definido, exige una revisión que incorpore estas nuevas formas de experiencias y narrativas.

9. La importancia para la filosofía del estudio de la relación memoria autobiográfica-sociabilidad reside en que esta nos permite ubicarnos en la base de la construcción de la moralidad, sobre todo, porque se analizan aquellos elementos genéticos y cognitivos que dan nacimiento y forma a nuestras habitudes y costumbres, de las más elementales a las más complejas. A partir de aquí, las construcciones del discurso filosófico podrán terminar de modelar las diversas teorías éticas, sociales y políticas. La relación nos ubica en «la antesala de la moralidad».

10. Las problemáticas y relaciones planteadas en este estudio constituyen un interés más que relevante para el ámbito de las empresas tecnológicas que desarrollan diversos programas informáticos como los de IA, los cuales pueden tener una influencia ideológica directa en la construcción de la identidad personal y colectiva.

11. Dos de los principales descubrimientos de la neurociencia del siglo XX han sido, por un lado, la determinación del cerebro humano como un «cerebro social», esto implica un cambio de paradigma en los estudios que se venían realizando desde una perspectiva individualista e inconexa. Por otro lado, la plasticidad cerebral, es decir, la capacidad de nuestro cerebro de amoldarse y modelarse en el medioambiente bio-psico-social. Estos descubrimientos pueden integrarse perfectamente dentro de la perspectiva de investigación de una neurofilosofía de la sociabilidad.

12. En un sentido de aplicación pragmática y social, la relación memoria autobiográfica-sociabilidad, al estudiar la estructura de los valores morales y sociales, puede constituir una herramienta propedéutica y educativa de gran aporte, tanto para empresas del sector privado como para instituciones, sobre todo si se tienen en cuenta los ODS que se deben cumplir y las secciones dedicadas a la RSC. Profundizar en los estudios de la naturaleza humana, de nuestros hábitos y valores puede hacer que ciertos ámbitos modifiquen sus estructuras en pos de una mayor y más saludable sociabilidad.

13. En cuanto al ámbito político, resultará de interés conocer aquellas estructuras que intervienen en la construcción de la sociabilidad, es decir, todo el proceso biológico, cultural, psicológico y emocional que se pone en marcha. Asimismo, el fomento de actividades que favorezcan un equilibrio relaciones y en la salud de la población.

14. La propuesta realizada de una línea de investigación interdisciplinar de neurofilosofía de la sociabilidad resulta de interés a todo el ámbito educativo y pedagógico. Como se ha comentado, el desarrollo a pasos acelerados de la realidad extendida y el metaverso desafían a toda la estructura social. Ofrecer una respuesta conceptual y, a la vez, pragmática queda contemplada desde esta línea de investigación. El desafío es hacer frente a la vulnerabilidad ideológica que tiene una influencia directa en la construcción de la identidad.

15. Por último, el estudio aquí presentado permitirá profundizar en la investigación, como he comentado en la propuesta *Hacia una neurofilosofía de la sociabilidad*.

bilidad, sobre la cimentación de la sociabilidad humana desde el marco epistemológico de la neurofilosofía. Las aportaciones derivadas de dicha línea de investigación plantean nuevas cuestiones filosóficas, cognitivas y políticas. En tal sentido, desde la perspectiva de la neurofilosofía de la sociabilidad, he construido un primer esquema que asume tres conjuntos conceptuales que están relacionados entre sí y que pueden ayudarnos a elaborar nuevas preguntas sobre la naturaleza social de nuestra especie: (a) *sociabilidad*; (b) *memoria autobiográfica*; y (c) *identidad*. El concepto de sociabilidad integra a su estructura de formación a las *relaciones ideológicas*, las cuales poseen una estructura biológica, cognitiva y cultural (valores y disvalores). La formación de la memoria autobiográfica también incorpora las mismas tres estructuras de la sociabilidad (biológica, cognitiva y cultural), además, de la suya propia (experiencia, narrativa y recuerdos). El tercer conjunto conceptual corresponde a la conexión de los otros dos conjuntos. Estos conjuntos abrirán paso a la formación de la *identidad*, tanto individual como social, donde su compleja estructura recibe la influencia de los elementos de los otros dos conjuntos. Considero que relación sociabilidad-memoria autobiográfica-identidad constituye un aporte novedoso para plantear nuevas preguntas y hallar evidencias sobre nuestra naturaleza humana como *zoon politikón* (*ζῶον πολιτικόν*).

Bibliografía

- Amaral, D. G. (2001). Organización anatómica del sistema nervioso central. En E. R. Kandel, J. H. Schwartz, & T. M. Jessell, *Principios de neurociencia* (4^a ed., págs. 317-336). Madrid: McGraw-Hill Interamericana.
- Arana Carnedo-Argüelles, J. (2015). *La conciencia inexplicada. Ensayo sobre los límites de la comprensión naturalista de la mente*. Madrid: Biblioteca Nueva.
- Aristóteles. (2007a). *Ética a Nicómaco* (Vol. BT 8239). Madrid: Alianza Editorial.
- Aristóteles. (2007b). *Política* (1231 ed., Vol. BT 8206). Madrid: Alianza Editorial.
- Ayan, S. J. (Julio/Septiembre de 2004). Neuronas especulares. *Revista Mente y Cerebro*(8), 79-81. Recuperado el 01 de Abril de 2020
- Baddeley, A. D. (1999). *Memoria humana: teoría y práctica*. (G. Evangelista, Trad.) Madrid: McGraw-Hill.
- Bear, M. F., Connors, B. W., & Paradiso, M. A. (2008). *Neurociencia: la exploración del cerebro* (3^a ed.). Barcelona: Wolters Kluwer Health España.
- Beltrán-Jaimes, J. O., Moreno-López, N. M., Polo-Díaz, J., Zapata-Zabala, M. E., & Acosta-Barreto, M. R. (2012). Memoria autobiográfica: un sistema funcionalmente definido. (J. M. Cuartas, Ed.) *International Journal of Psychological Research*, 5(2), 108-123.
- Bennett, M., Dennett, D., Hacker, P., & Searle, J. (2008). *La naturaleza de la conciencia: cerebro, mente y lenguaje* (Vol. 72). Barcelona etc.: Paidós.
- Bergson, H. (1911). *Matter and Memory*. Londres: Allen y Unwin.
- Berntsen, D., & Rubin, D. C. (2012). *Understanding Autobiographical Memory. Theories and Approaches*. Cambridge: Cambridge University Press.
- Bickle, J. (2009). *The Oxford Handbook of Philosophy and Neuroscience*. Oxford: Oxford University Press.
- Boyano Moreno, J., & Mora Mérida, J. (06 de 2012). Algunos antecedentes históricos del concepto de memoria autobiográfica. *Revista de Historia de la Psicología*, 33(2), 47-60.
- Brewer, W. F. (1986). What is autobiographical memory. En D. C. Rubin, *Autobiographical memory*. Nueva York: Cambridge University Press.

- Butcher López, E. A., Pachecho Railey, C. C., & Tirado Medina, H. (2006). Plasticidad cerebral. En A. Téllez López, *Atención, aprendizaje y memoria. Aspectos psicobiológicos* (1^a ed., págs. 11-40). Alcalá de Guadaíra, Sevilla: Trillas Eduforma.
- Camps, V. (2011). *El gobierno de las emociones*. Barcelona: Herder.
- Cañas, J. J., & Bajo, M. T. (1998). Memoria autobiográfica. En J. M. Ruiz-Vargas, *Psicología de la memoria* (págs. 369-382). Madrid: Alianza Psicología.
- Casebeer, W. D., & Churchland, P. S. (enero de 2003). The Neural Mechanisms of Moral Cognition: A Multiple-Aspect Approach to Moral Judgment and Decision-Making. *Biology & Philosophy*, 18, 169-194. Obtenido de <https://doi.org/10.1023/A:1023380907603>
- Cela Conde, C. (1985). *De genes, dioses y tiranos. La determinación biológica de la moral*. Madrid: Alianza.
- Cela Conde, C., & Ayala, F. (2001). *Senderos de la evolución humana*. Madrid: Alianza.
- Changeux, J. P. (1985). *El hombre neuronal*. Madrid: Espasa Calpe.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of personality and social psychology*, 76(6), 893-910. Obtenido de <http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=1999-05479-002>
- Churchland, P. (1995). El materialismo eliminativo y las actitudes proposicionales. En E. Rabossi, *Filosofía de la mente y ciencia cognitiva* (págs. 43-68). Barcelona: Paidós.
- Churchland, P. (2019). *Conscience. The Origins of Moral Intuition*. New York: W. W. Norton & Company.
- Churchland, P. M. (1981). Eliminative Materialism and the Propositional Attitudes. *The Journal of Philosophy*, 78(2), 67-90.
- Churchland, P. M. (1992). *Materia y Conciencia. Introducción contemporánea a la filosofía de la mente* (1^a ed.). Barcelona: Gedisa.
- Churchland, P. M. (1998). Toward a cognitive neurobiology of the moral virtues. *Topoi*(17), 1-14.
- Churchland, P. S. (1986). *Neurophilosophy: toward a unified science of the mind-brain*. Cambridge, Massachusetts: MIT Press.

- Churchland, P. S. (1994). Can neurobiology teach us anything about consciousness? *Proceedings and Addresses of the American Philosophical Association*, 67(4), 23-40. Obtenido de <http://www.jstor.org/stable/3130741>
- Churchland, P. S. (18 de noviembre de 2006). *The big questions: do we have free will?* Recuperado el 01 de noviembre de 2011, de NewScientist.com: <https://www.newscientist.com/issue/2578/>
- Churchland, P. S. (01 de Octubre de 2007). Neurophilosophy: the early years and new directions. *Functional Neurology*, 22(4), 185-195. Obtenido de <http://www.functionalneurology.com/common/php/portiere.php?ID=56a5e947f832405161b7bb4e45cc85ec>
- Churchland, P. S. (06 de noviembre de 2008). The impact of neuroscience on philosophy. *Neuron*, 60(3), 409-411. Obtenido de <http://linkinghub.elsevier.com/retrieve/pii/S0896627308008969>
- Churchland, P. S. (2012). *El cerebro moral. Lo que la neurociencia nos cuenta sobre la moralidad.* Barcelona: Paidós.
- Churchland, P. S. (2013). *Touching a nerve. The self as brain.* New York: W.W. Norton.
- Churchland, P. S., & Di Francesco, M. (octubre de 2007). Neurophilosophy: an introduction and overview. *Functional Neurology*, 22(4), 179-180. Obtenido de http://www.functionalneurology.com/index.php?PAGE=articolo_dett&ID_ISSUE=284&id_article=2618
- Coeckelbergh, M. (2023). *La Filosofía Política de la Inteligencia Artificial. Una Introducción.* Madrid: Cátedra.
- Colombo, M., & Piccinini, G. (2023). *The Computational Theory of Mind.* Cambridge: Cambridge University Press.
- Cortina, A. (2010). Neuroética: ¿Las bases cerebrales de una ética universal con relevancia política? *Isegoría. Revista de Filosofía Moral y Política*(42), 129-148.
- Cortina, A. (2011). *Neuroética y neuropolítica. Sugerencias para la educación moral.* Madrid: Tecnos.
- Crick, F. (1979). Thinking About the Brain. *Scientific American*(241), 219-232.
- Damasio, A. (2021). *Sentir y saber. El camino de la conciencia.* Barcelona: Ediciones Destino.
- Damasio, A. R. (1996). *El error de Descartes: La emoción, la razón y el cerebro humano.* Barcelona: Crítica.

- de Waal, F. (2014). Raíces del espíritu cooperativo. *Investigación y Ciencia*(458), 55-57.
- Decety, J., Bartal, I. B.-A., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of prosocial: highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions B*, 371(1686), 20150077. doi:<https://doi.org/10.1098/rstb.2015.0077>
- Dejours, C. (2009). *Trabajo y sufimiento. Cuando la injusticia se hace banal*. Madrid: Modus Laborandi.
- Dennett, D. (2013). Eo yo como centro de gravedad narrativa. *Logos. Anales del Seminario de Metafísica*, 46, 11.25. doi:http://dx.doi.org/10.5209/rev_ASEM.2013.v46.42862
- Descartes, R. (1990). *El tratado del hombre* (Vol. 663). (G. Qintás, Ed.) Madrid: Alianza.
- Diodado, C., Rodriguez Valls, F., & Arana, J. (2010). *Neurofilosofía. Perspectivas contemporáneas*. Madrid: Thémata Plaza y Valdés.
- Dudai, Y. (2002). *Memory from A to Z: keywords, concepts, and beyond*. Oxford University Press.
- Eceiza, M., Arrieta, M., & Goñi, A. (2008). Habilidades sociales y contextos de la conducta social. *Revista de Psicodidáctica*, 13(1), 11-26.
- Eichenbaum, H. (2003). *Neurociencia cognitiva de la memoria: una introducción* (1^a ed.). Barcelona: Ariel.
- Elliot, K. (2022). *Values in Science*. Cambridge: Cambridge University Press.
- European Commission. (2013). *Human Brain Project*. Recuperado el 19 de febrero de 2013, de Ethics Programme: <https://www.humanbrainproject.eu/ethics-and-society;jsessionid=1x3el6o2oimtm1nsapot0b2xe1>
- Eustache, F., & Desgranges, B. (2010). Hacia un modelo unificado de la memoria. *Mente y Cerebro*(43).
- Evers, K. (2011). *Neuroética. Cuando la materia se despierta* (Vol. 3071). Buenos Aires; Madrid: Katz.
- Fernández , N. (2021). Posthumanismo cyborg: la propuesta de Andy Clark y Donna Haraway. *Instantes y azares. Escrituras nietzscheanas*, 26, 83-96.
- Fernández Ruiz, J., & López García, J. (1998). La neuropsicología de la memoria. *Ciencias*(49), 18-25.
- Feyerabend, P. (1963). Comment: "mental Events and the brain". *Journal of Philosophy*, 60(11), 295-296.

- Fivush, R. (2011). The development of autobiographical memory. *Annual Review of Psychology*, 25.
- Fivush, R., & Reese, E. (1992). The social construction of autobiographical memory. En M. Conway, D. Rubin, H. Spinnler, & W. Wagenaar, *Theoretical Perspectives on Autobiographical Memory* (pág. 518). Dordrecht: Springer Science+Business Media Dordrecht.
- Fivush, R., Habermas, T., Waters, T. E., & Zaman, W. (2011). The Making of Autobiographical Memory: Intersections of Culture, Narratives and Identity. *International Journal of Psychology*(46), 321-345.
doi:10.1080/00207594.2011.596541
- Gabriel, M. (2016). *Yo no soy mi cerebro. Filosofía de la mente para el diglo XXI*. Barcelona: Pasado y Presente.
- Gallese, V., & Goldman, A. (1998). Mirror Neurons and the Simulation Theory of Mind-Reading. *Trends in cognitive sciences*, 2(12), 493-501.
doi:[https://doi.org/10.1016/S1364-6613\(98\)01262-5](https://doi.org/10.1016/S1364-6613(98)01262-5)
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 593-609.
- García García, E. (2008). Neuropsicología y educación. De las neuronas espejo a la teoría de la mente. *Revista de Psicología y Educación*, 3(1), 69-89.
- García García, E. (2018). *Somos nuestra memoria. Recordar y olvidar*. Madrid: Bonalletra Alcompas.
- García-Bajos, E., & Migueles, M. (2016). Recuperación de experiencias autobiográficas positivas y negativas del pasado y futuro. *Ciencia Cognitiva*, 10(3), 89-91.
- García-Molina, A. (2012). Phineas Gage y el enigma del córtex prefrontal. *Neurología*, 27(6), 370-375. doi:[dx.doi.org/10.1016/j.nrl.2010.07.015](https://doi.org/10.1016/j.nrl.2010.07.015)
- Gazzaniga, M. (2019). *El instinto de la conciencia. Cómo el cerebro crea la mente*. Barcelona: Paidós.
- Gazzaniga, M. S. (1993). *El cerebro social* (Vol. 1646). Madrid: Alianza Editorial.
- Giordano, J. (2012). *Nerotechnology. Premises, Potential, and Problems*. New York: CRC Press.
- Graf, P., & Schacter , D. (1985). Implicit and explicit memory for new associations in normal and amnesic subjet. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 11, 501-518.
- Haidt, J. (2001). The Emotional Dog and its Rational Tail: a Social Intuitionist Approach to Moral Judgment. *Psychological Review*, 814-834.

- Hebb, D. O. (1985). *Organización de la conducta*. (T. del Amo Martín, Trad.) Madrid: Debate.
- Hernando, A. (15 de noviembre de 2013). *Europa se da diez años para construir un modelo virtual del cerebro*. Recuperado el 18 de noviembre de 2013, de <http://www.agenciasinc.es/Reportajes/Europa-se-da-diez-anos-para-construir-un-modelo-virtual-del-cerebro>
- Hierro-Pescador, J. (2005). *Filosofía de la mente y de la Ciencia cognitiva*. Madrid: Akal.
- Iacoboni, M. (2011). *Las neuronas espejo. Empatía, neuropolítica, autismo, imitación o de cómo entendemos a los otros* (1^a ed., 2^a ed.). Buenos Aires: (Arg): Katz.
- Kandel, E. R. (2001). De las células nerviosas a la cognición: la representación celular interna necesaria para la percepción y la acción. En E. R. Kandel, J. H. Schwartz, & T. M. Jessell, *Principios de neurociencia* (4^a ed., págs. 381-403). Madrid: McGraw-Hill Interamericana.
- Kandel, E. R. (2001). Neuronas y conducta. En E. R. Kandel, J. H. Schwartz, & T. M. Jessell, *Principios de neurociencia* (4^a ed., págs. 19-35). Madrid: McGraw-Hill Interamericana.
- Kandel, E. R. (2001). Principios de neurociencia. En E. R. Kandel, J. H. Schwartz, & T. M. Jessell. Madrid: McGraw-Hill Interamericana.
- Kandel, E. R. (2007). *En busca de la memoria: el nacimiento de una nueva ciencia de la mente*. Katz.
- Kandel, E. R. (2019). *La nueva biología de la mente. Qué nos dicen los trastornos cerebrales sobre nosotros mismos*. Barcelona: Paidós.
- Kandel, E. R., Kupfermann, I., & Iversen, S. (2001). Aprendizaje y memoria. En E. R. Kandel, J. H. Schwartz, & T. M. Jessell, *Principios de neurociencia* (4^a ed., págs. 1227-1246). Madrid: McGraw-Hill Interamericana.
- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2008). *Neurociencia y Conducta* (Reimp ed.). Madrid etc.: Prentice Hall.
- Kauffman, S. (2021). *Más allá de las leyes físicas. Un largo camino desde la materia hasta la vida*. Barcelona: Tusquets Editores.
- Kim, J. (2002). El problema mente-cuerpo tras cincuenta años. *Azafea. Revista de Filosofía*, 4, 45-63.
- Kinsbourne, M. (1989). The boundaries of episodic remembering: Comments on the second section. En H. Roediger, & F. Craik, *Varieties of memory and consciousness. Essay in honour of Endel Tulving*. Hillsdale N. J.: Erlbaum.

- Köber, C., & Habermas, T. (2016). El peso de la memoria autobiográfica. *Mente y Cerebro*(78), 11-15.
- Koch, C. (2006). *La conciencia: una aproximación neurobiológica* (1^a ed.). Barcelona: Ariel.
- Konorski, J. (1948). *Conditioned reflex and neuron organization*. Cambridge: Cambridge Univ. Press.
- Lakin, J., Jefferis, V., Cheng, C., & Chartrand, T. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior*, 27(3), 145-162. Obtenido de <http://dx.doi.org/10.1023/A:1025389814290>
- Lara, F., & Savulescu, J. (2021). *Más que humanos. Biotecnología, inteligencia artificial y ética de la mejora*. Madrid: Tecnos.
- Levy, N. (14 de 02 de 2008). Introducing Neuroethics. *Neuroethics*(1), 1-8. doi:<https://doi.org/10.1007/s12152-008-9007-7>
- Levy, N. (2014). *Consciousness and Moral Responsibility*. Oxford: Oxford University Press.
- Lizón, Á. (2016). *La sociabilidad humana. Estudios de evolución y cognición social*. Madrid: Síntesis.
- Lokhorst, G.-J. (2015). Descartes and the pineal gland. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Otoño 2015 ed.). <http://plato.stanford.edu/archives/fall2015/entries/pineal-gland/>.
- López Corredoira, M. (2010). Algunas respuestas a las críticas al materialismo en el problema mente-cerebro. En C. Gómez Diosdado, F. Rodríguez Valls, & J. Arana (Edits.), *Neurofilosofía: Perspectivas contemporáneas* (Vol. 7, págs. 129-141). Sevilla: Thémata.
- López Muñoz, F., & Álamo, C. (2000). El tratado del hombre: Interpretación cartesiana de la neurofisiología del dolor. *Asclepio*, 52(1), 239-267.
- Manzanero, A. (08 de 10 de 2017). *Psicología de la memoria*. Obtenido de <http://psicologiadelamemoria.blogspot.com/p/memoria-corto-plazo.html>
- Meadows, D. (2022). *Pensar en sistemas*. Madrid: Capitán Swing Libros.
- Mele, A. R. (2009). *Effective intentions. The power of conscious will*. New York, USA: Oxford University Press.
- Mendoza García, J. (2004). Las formas del recuerdo. La memoria narrativa. *Athenea Digital* , 1-16.
- Mora Teruel, F. (2007). *Neurocultura. Una cultura basada en el cerebro*. Madrid: Alianza.

- Moreno, J. (2012). *Mind Wars: Brain Science and the Military in the 21st Century*. New York: Bellevue Literary Press.
- Morse, S. J. (2011). Neuroscience and the future of personhood and responsibility. En B. W. Jeffrey Rosen (Ed.). Washington D.C.: Brookings Institution Press. Obtenido de <http://ssrn.com/abstract=2033748>
- Moya, C. (2004). *Filosofía de la mente*. Valencia: Universitat de Vàlencia.
- Moya, C. (2012). El pronblema de lo mental. *Themata. Revista de Filosofía*, 43-55.
- Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.
- Nowak, M. (Octubre de 2012). ¿Por qué cooperamos? *Investigación y Ciencia*(433), 18-23.
- Pineda, D. (2012). *La mente humana. Introducción a la filosofía de la psicología*. Madrid: Ediciones Cátedra.
- Prades, J. (2006). Filosofía de la mente: el estado de la cuestión. *Theoria*, 57, 315-332.
- Priest, S. (1994). *Teorías y Filosofías de la Mente*. Madrid: Ediciones Cátedra.
- Reber, A. (1992). The cognitive unconscious: An evolutionary perspective. *Consciousness and Cognition*(2), 93-133.
- Redolar Ripoll, D. (2014). *Neurociencia cognitiva*. Madrid: Editorial Médica Panamericana.
- Rizzolatti, G., & Sinigaglia, C. (2006). *Las neuronas espejo. {Los} mecanismos de la empatÃA emocional* (1Âª ed., Vol. 63). Barcelona: PaidÃ³s.
- Robinson, J. (1986). Autobiographical memory: a historical prologue. En D. Rubin, *Autobiographical memory* (págs. 19-24). New York: Cambridge University Press.
- Rodríguez González, M. (2021). *Filosofía de la mente*. Madrid: Ediciones Complutenses.
- Rodríguez Serón, A. (2002). La neurofilosofía como punto de encuentro entre filosofía y neurociencias. *Contrastes. Revista Interdisciplinar de Filosofía*, VII, 149-166.
- Roediger, H. L., Yadin, D., & Susan M., F. (2007). *Science of memory: concepts*. Oxford University Press, USA. Obtenido de <https://books.google.es/books?id=c6R0-I6C7M8C>
- Rozin, P. (1976). The evollution of intelligence and access to the cognitive unconscious. *Progress in Psychobiological and Physiological Psychology*(6), 245-280.

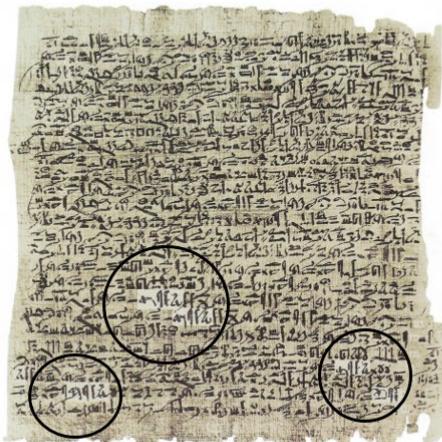
- Rubin, D. C. (1986). *Autobiographical Memory*. New York: Cambridge University Press.
- Ruggieri, V. (2013). Empatía, cognición social y trastornos del espectro autista. *Revista de Neurología*, 56(1), 13-21.
- Ruiz Martín, H. (2022). *Los secretos de la memoria. Las historias humanas que relaron qué es y cómo funciona la memoria*. Barcelona: Penguin Random House.
- Ruiz Rey, F. (2009). Libre albedrío y neurociencias. Tercera Parte. Neuroética: neurociencia de la ética, acrecentamiento de habilidades, acción voluntaria y responsabilidad. *Psiquiatría.com*, 13(3). Recuperado el 13 de abril de 2013, de <http://www.psiquiatria.com/revistas/index.php/psiquiatriacom/article/view/437/>
- Ruiz-Vargas, J. (1998). Bases conceptuales y metodológicas. En J. Ruiz-Vargas, *Psicología de la memoria* (págs. 27-84). Madrid: Alianza.
- Ruiz-Vargas, J. (2004). Claves de la memoria autobiográfica. En C. Fernández, & M. Hermosilla, *Autobiografía en España: Un balance* (págs. 183-220). Madrid: Visor.
- Ruiz-Vargas, J. M. (1994). *La memoria humana: función y estructura* (Vol. 10). Madrid: Alianza.
- Ruiz-Vargas, J. M. (1998). *Psicología de la memoria* (1a , 5a reimp ed., Vol. 35). Madrid: Alianza.
- Ryle, G. (1949). *The concept of mind*. San Francisco, California: Hutchinson.
- Sacks, O. (2019). *El río de la conciencia*. Barcelona: Anagrama.
- Santamaría, A., & Montoya, E.-M. (23 de 06 de 2014). La memoria autobiográfica: el encuentro entre la memoria, el yo y el lenguaje. *Estudios de Psicología*, 29(3), 333-350.
- Schacter, D. (1984). Toward the multidisciplinary study of memory: Ontogeny, phylogeny, and pathology of memory systems. En L. Squire, & N. Butters, *Neuropsychology of memory*. Nueva York: Guildford.
- Schacter, D. (1990). Perceptual representation system and implicit memory. Toward a resolution of the multiple memory systems debate. *Annals of the New York Academy of Sciences*, 543-571.
- Sellés García, M., & Solís Santos, C. (1994). *La revolución científica* (Vol. 15). Madrid: Síntesis.
- Seth, A. (2023). *La creación del yo. Una nueva ciencia de la conciencia*. Madrid: Sexto Piso.

- Sherry, D., & Schacter, D. (1987). The evolution of multiple memory systems. *Psychology Review*(94), 439-454.
- Soler, M. (2009). *Adaptación del comportamiento: comprendiendo al animal humano*. Madrid: Síntesis.
- Solís Santos, C., & Sellés García, M. (2013). *Historia de la ciencia*. Madrid: Espasa.
- Squire, L. (1986). Mechanisms of memory. *Science*(5), 1612-1619.
- Sutton, J. (2007). *Philosophy and memory traces. Descartes to connectionism*. Cambridge: Cambridge University Press.
- Téllez López, A. (2006). *Atención, aprendizaje y memoria: aspectos psicobiológicos* (1^a ed.). Alcalá de Guadaíra, Sevilla: Trillas Eduforma.
- Téllez López, A. (2006). La memoria. En A. Téllez López, *Atención, aprendizaje y memoria: aspectos psicobiológicos* (1^a ed., págs. 105-178). Alcalá de Guadaíra, Sevilla: Trillas Eduforma.
- Tirapu-Ustároz, J., Pérez-Sayes, G., Erekatxo-Bilbao, M., & Pelegrín-Valero, C. (2007). ¿Qué es la teoría de la mente? *Revista de Neurología*, 479-489.
- Tomasello, M., Marengo, E., & Dweck, C. S. (2010). *¿Por qué cooperamos?* (1^a ed., Vol. 2030). Buenos Aires; Madrid: Katz.
- Tulving, E. (1972). Episodic and semantic memory. En E. Tulving, & W. Donaldson, *Organization of memory*. Nueva York: Academic Press.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford: Oxford University Press.
- Tulving, E. (1985). How many memory systems are there? *American Psychologist*(40), 385-390.
- Tulving, E. (1989a). Remembering and knowing the past. *American Scientist*(77), 361-367.
- Tulving, E. (1989b). Memory: Performance, knowledge, and experience. *The European Journal of Cognitive Psychology*(1), 3-26.
- Tulving, E. (1991). Concepts of human memory. En L. Squire, N. Weinberger, G. Lynch, & J. McGaugh, *Memory: Organization and locus of change*. Nueva York: Oxford University Press.
- Tulving, E. (1993). Varieties of consciousness and levels of awareness in memory. En A. Baddeley, & L. Weiskrantz, *Attention: Selection, awareness and control. A tribute to Donald Broadbent*. Oxford: Oxford University Press.

- Tulving, E., & Schacter, D. (1990). Priming and human memory systems. *Science*(247), 301-306.
- Tulving, E., Schacter, D., McLachlan, D., & Moscovitch, M. (1988). Priming of semantic autobiographical memory: A case study of retrograde amnesia. *Brain and cognition*(8), 3-20.
- Valdés, M. (1980). La objeción de Kripke a la teoría de la identidad mente-cuerpo. *Diánoia*, 26(26), 195-213.
- Viard, A. (2010). La memoria autobiográfica. *Mente y Cerebro*, 57-61.
- Vilarroya, Ó. (2019). *Somos lo que nos contamos. Cómo los relatos construyen el mundo en que vivimos*. Barcelona: Ariel.
- Wilkins, R. (20 de 05 de 2020). *Neurosurgical Classic-XVII Edwin Smith Surgical Papyrus*. Obtenido de Cyber Museum of Neurosurgery: <http://www.neurosurgery.org/cybermuseum/pre20th/epapyrus.html>
- Wittgenstein, L. (1988). *Investigaciones filosóficas*. (A. García Suárez, & N. Moulines, Trads.) Barcelona: Céltica.

Diapositivas

Papiro quirúrgico Edwin Smith - S. XVII a.C.

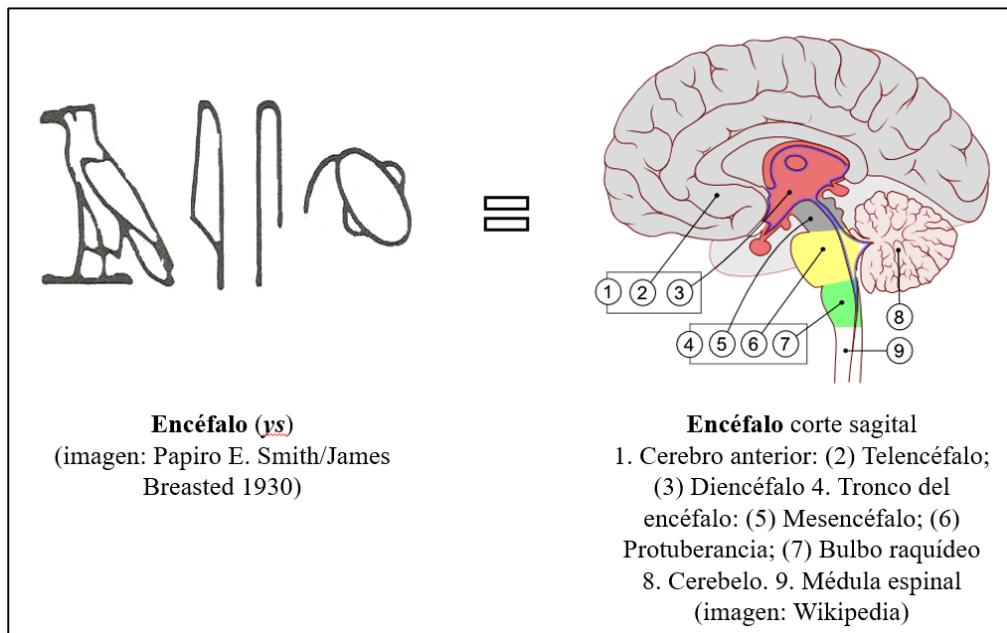


Columna II

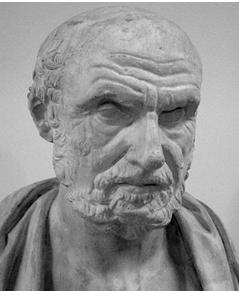


Columna IV

(D1)



(D2)



Hipócrates

«Los hombres deben saber que es del encéfalo, y solo del encéfalo, de donde surgen nuestros placeres, alegrías, risas, bromas, así como nuestras penas, dolores, tristezas y lágrimas. Concretamente, a través de él, pensamos, vemos, oímos y distinguimos lo malo de lo bueno, lo agradable de lo desagradable... Es lo mismo que nos vuelve locos o delirantes, nos inspira miedo o pavor, ya sea de día o de noche, nos produce insomnio, errores inoportunos, preocupaciones sin sentido, falta de motivación y actos contrarios a nuestros hábitos. Estas cosas que sufrimos proceden todas del encéfalo... Sin embargo, cuando el encéfalo está tranquilo, un hombre puede pensar correctamente»

(Texto atribuido a Hipócrates, siglo V a.C. Extraído de Eric R. Kandel et al., Principios de Neurociencia)

(D3)

Origen y nacimiento de la neuroética

- Desarrollo de las neurociencias (4 décadas aprox.)
- Congreso California 2002
- Cuestiones éticas y legales

Impacto/Interés en: \Rightarrow

- Informática
- Neurobiología
- Psiquiatría
- Psicología
- Filosofía
- Política
- Economía
- Defensa
- Derecho, etc.



(D4)

Definiciones

- **William Safire** (exdirector de *The New York Times*): «Bioética del cerebro»
- **Michael Gazzaniga** (neurocientífico): «Un intento de proponer una filosofía de la vida con un fundamento cerebral»
- **Neil Levy** (neurofilósofo): «Algo más que bioética». Dos ramas que producen un **conocimiento nuevo**
- **Kathinka Evers** (filósofa): Dos caminos, (1) «**neuroética aplicada**» o «**ética de la neurociencia**»; y (2) «**neuroética fundamental**» o «**neurociencia de la ética**»

(D5)

Claves

- **Campo interdisciplinar de investigación nacido en 2002**
- «**Ética de la Neurociencia**» y «**Neurociencia de la Ética**»
- **Dependencia de la neurociencia y la neurotecnología**

(D6)

Biología y moral. Antecedentes

Desde la Filosofía ⇒ Tradición Naturalista



Aristóteles

virtudes
éticas y
políticas

D. Hume

naturaleza
humana
sentimientos

A. Smith

sentimientos
economía

Ch. Darwin

evolución
instintos
emociones



**Ética o Filosofía Moral, Filosofía Política, Antropología
Economía, Filosofía de la Mente, Filosofía de la Ciencia, etc.**

(D7)

Neurobiología y moral. Actualidad

- Patricia y Paul Churchland
 - Michael Gazzaniga
 - António Damásio
 - Francis Crick(†)
 - Judy Illes
 - Martha Farah
 - Julian Savulescu
- } filósofos y neurocientíficos
- ↓

«La explicación ofrecida por la neurobiología quizás no sea la mejor, pero sí la más adecuada» (P S. Churchland)

(D8)

Neuroética aplicada

«El hombre no tiene que esperar pasivamente durante millones de años para que la evolución le ofrezca un cerebro mejor... Desarrollar una farmacología de acción del cerebro, en un sentido nootropico, me parece que tiene su lugar en el objetivo de la investigación humana»

(Corneliu E. Giurgea)

(D9)

Algunos debates



⇒ Tratamiento vs mejora con *smart drugs*



¿Es éticamente aceptable en un espacio competitivo el uso de drogas inteligentes?

¿Acaso la mejora no constituye, en la mayor parte de los casos, una trampa, un abandono de la meritocracia como criterio de distribución de oportunidades?

(D10)

SUÉSOS Antonio Losilla está acusado de matar a Pilar Cebrán, pero su cuerpo no ha aparecido

Un juez autoriza aplicar una técnica pionera sobre el presunto autor de la muerte de su esposa

- Basada en la neurofisiología, analiza los cambios eléctricos de la actividad cerebral
- Se pretende orientar en la búsqueda de los restos de la mujer, aún sin hallar
- El marido fue detenido en octubre de 2012, seis meses después de la desaparición

«No es una ‘**prueba de verdad o mentira**’, ha aclarado Jodrá, quien ha añadido que no se requieren tampoco una declaración verbal del individuo, sino la evidencia de una actividad cerebral...»

«Se llevará a cabo para **esclarecer un delito** de “extrema gravedad”...»

«Esta técnica, conocida como **P300 ODD-BALL**, se aplicará al presunto autor de la muerte...»

(D11)

Caso Marta del Castillo

La Policía pide que Carcaño se someta a la prueba que se realizó al presunto descuartizador de Ricla

Europa Press. Madrid | Actualizada 26/01/2014 a las 13:45

8 Comentarios | A- A+

El presunto asesino de Ricla se sometió a la prueba P300, pionera en España, en diciembre.

Twittear

Recomendar

Tuenti

Menéame

Share

Compartir

Enviar

Imprimir



Equipo que realiza la prueba P300... Guillermo Mestre

«La Brigada de Policía Judicial de Zaragoza **valoró positivamente** esta experiencia si bien el cuerpo aún no ha aparecido tampoco...»

«La Policía Nacional **defiende que esta prueba no supone la violación de ningún derecho personal de Carcaño**, sino que busca comprobar que el sujeto que se somete a ella tiene **recuerdos sobre las cuestiones** que se le plantean...»

(D12)

Investigación

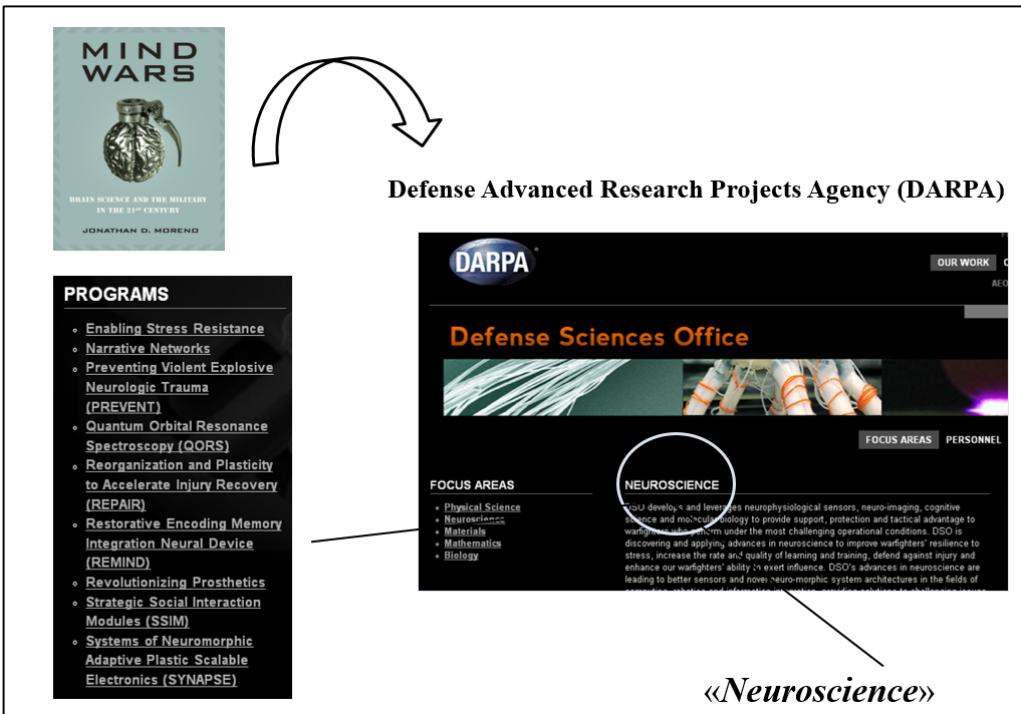
Julian Savulescu (The Oxford Centre for Neuroethics)

- ética de la genética, eugenésia, «*enhancement*» o mejora genética y clonación
- mejora cognitiva
- frontera entre la conciencia y el deterioro neurológico grave
- libre albedrío-responsabilidad-adicción
- neurociencia de la moralidad y la decisión-acción



Jonathan Moreno (U. Pennsylvania)
discusión de los dilemas éticos de la aplicación
de tecnología neurocientífica al desarrollo
militar

(D13)



(D14)

Kathinka Evers
(U. Uppsala)
neuroética
aplicada y
fundamental
*Human Brain
Project*



Stephen Morse
(U. Pennsylvania)
neurociencia y
responsabilidad



Bernard Baertschi
(U. Genève)
relación filosofía-
medicina-
neurociencias

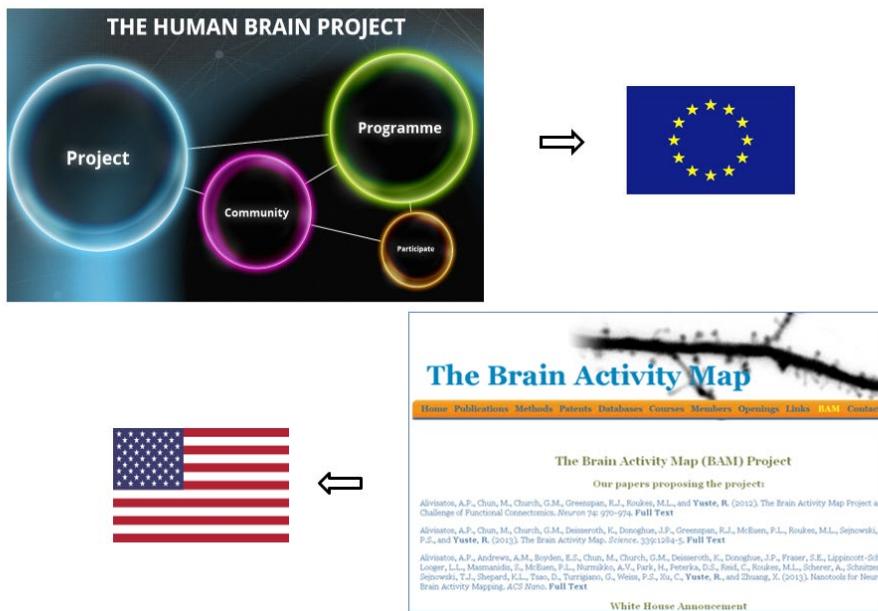


James Giordano
(U. Stanford)
neurotecnología
neuroética del dolor



(D15)

Retos para la próxima década (2020-2030)



(D16)

«What happens when we accept that everything we feel and think stems not from an immaterial spirit but from electrical and chemical activity in our brains?»

(Patricia S. Churchland)

(D17)

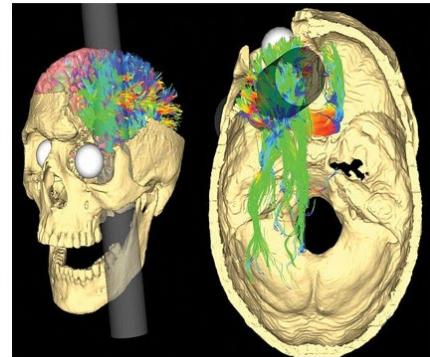
Phineas Gage y el estudio de las emociones



Phineas P. Gage



Cráneo de P. Gage y barra



Representación de la zona afectada

(D18)



António Damásio: «Gage no es Gage»



«ser obsceno e inmoral para la sociedad»

lesión en el lóbulo frontal

**No afecta la capacidad de razonamiento Sí afecta sistema de emociones sociales
(simpatía-culpa-vergüenza)**

fallo para tomar decisiones adecuadas

(D19)

sentimientos de «dolor y placer» son los cimientos de nuestra mente



puerta de acceso biología-neurobiología-lesiones cerebrales



la mente no es un «algo intangible»

(D20)

**Neurociencias
Psicología cognitiva**



**actualizan el estudio de las
emociones**



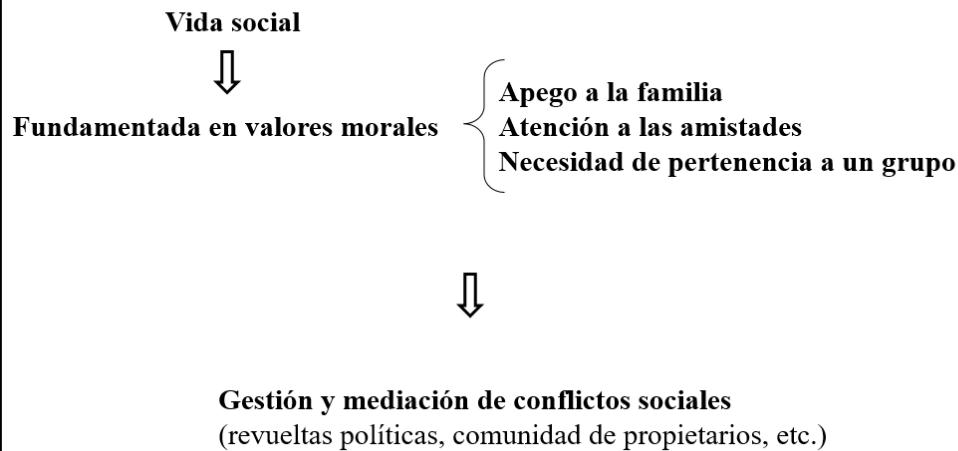
decisiones morales



**racionales
emocionales**

(D21)

Bioquímica y cooperación en la base de la moralidad

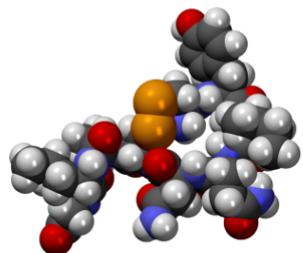


(D22)

¿De qué modo los genes han modelado nuestro cerebro para adaptarlo a la supervivencia?
¿De qué modo valoramos y cuidamos a los demás?
¿Pueden nuestros cerebros preocuparse por algo?



Neuroendocrinología



Oxitocina

⇒ «antesala de la moralidad»



- **sociabilidad de los mamíferos**
- **apego/tristeza por separación**
- **cuidado de los demás**

(D23)

moralidad

```
graph TD; Moralidad[moralidad] --> Oxitocina[oxitocina]; Moralidad --> Cooperacion[cooperación]
```



Michael Tomasello



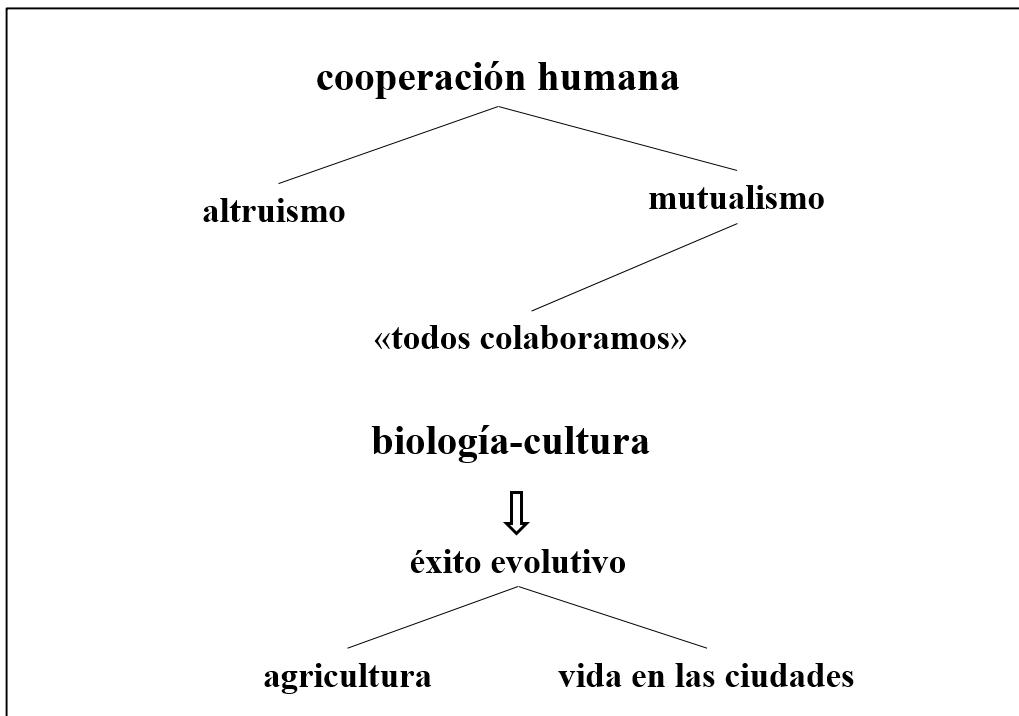
Patricia S. Churchland

cultura humana

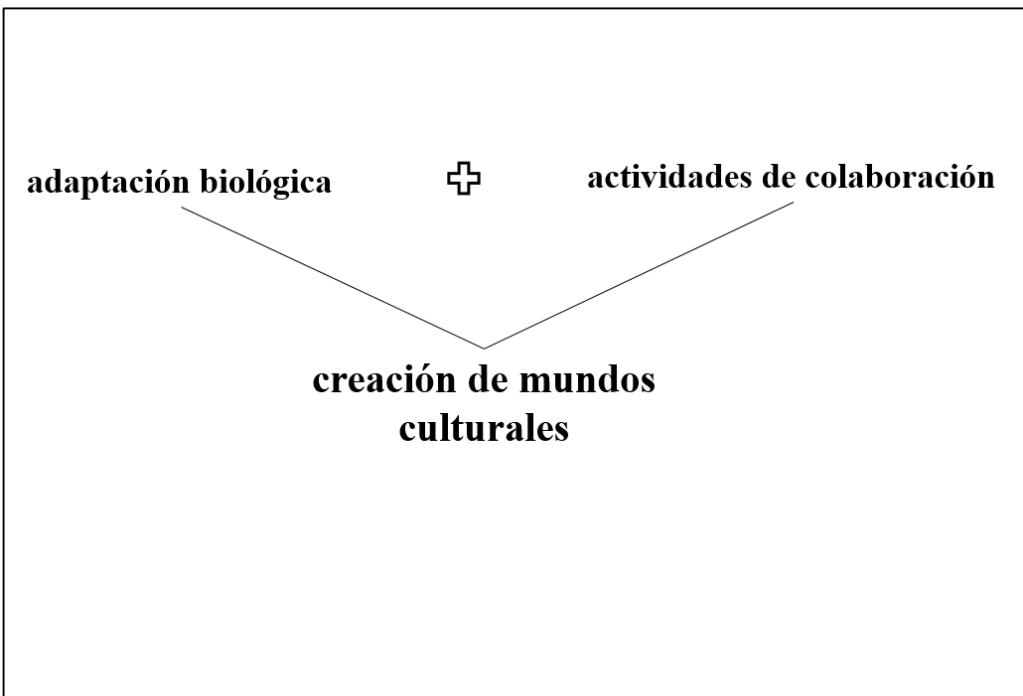
↓

«habilidades y motivaciones para colaborar»

(D24)



(D25)



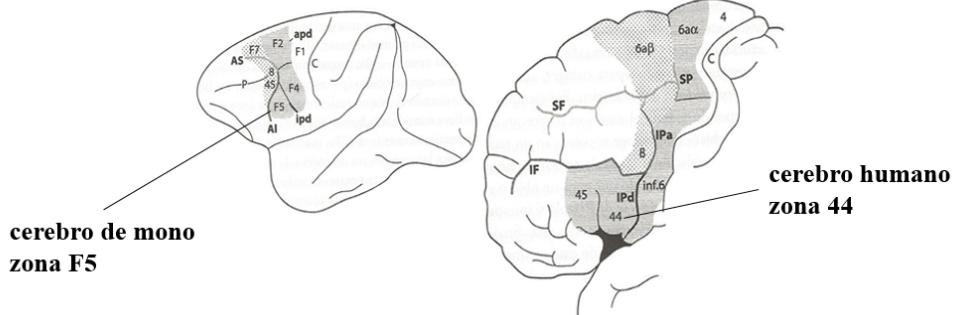
(D26)

Neuronas espejo y empatía

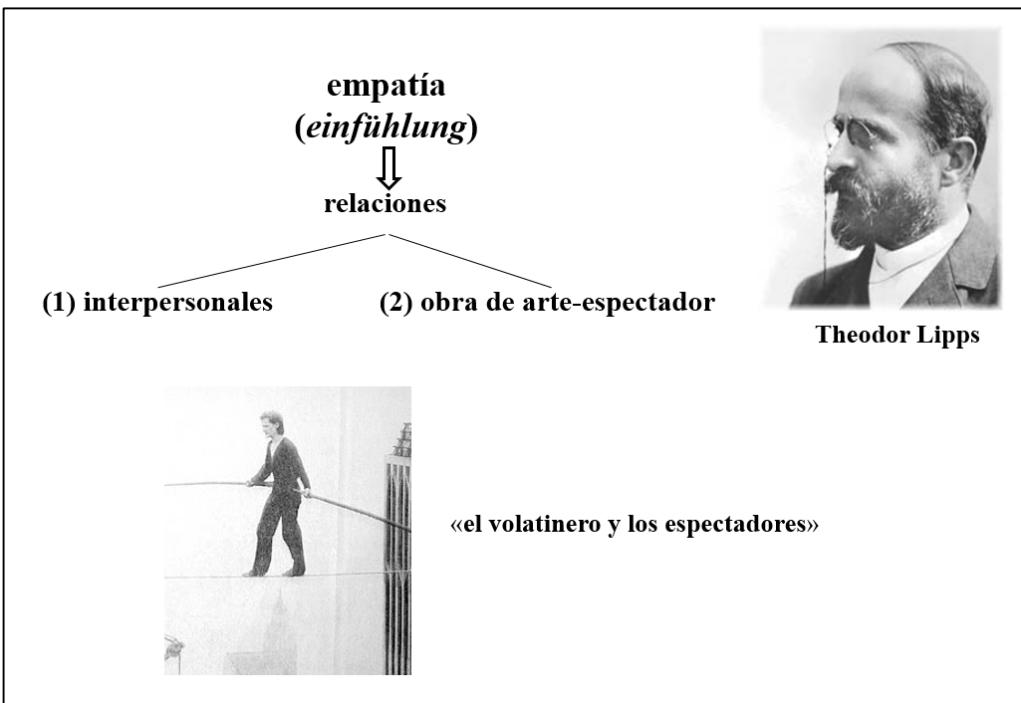
1996 Laboratorio de Parma de Rizzolatti-Fogassi-Gallese



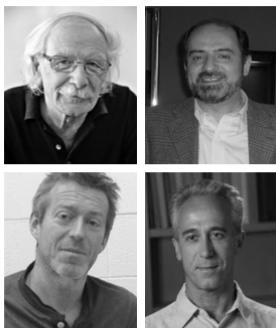
Experimentos con electrodos en macacos *Rhesus*



(D27)



(D28)



G. Rizzolatti

V. Gallese

C. Sinigaglia

M. Iacoboni

origen de la empatía



mecanismo asentado en las neuronas espejo

«circuito de imitación o simulación central»



**«comprendemos la tristeza o alegría ajena porque
nuestro cerebro simula una expresión facial»**

(D29)



(D30)

Patricia Churchland



dos argumentos contra la teoría de la simulación

(D31)

Anexos

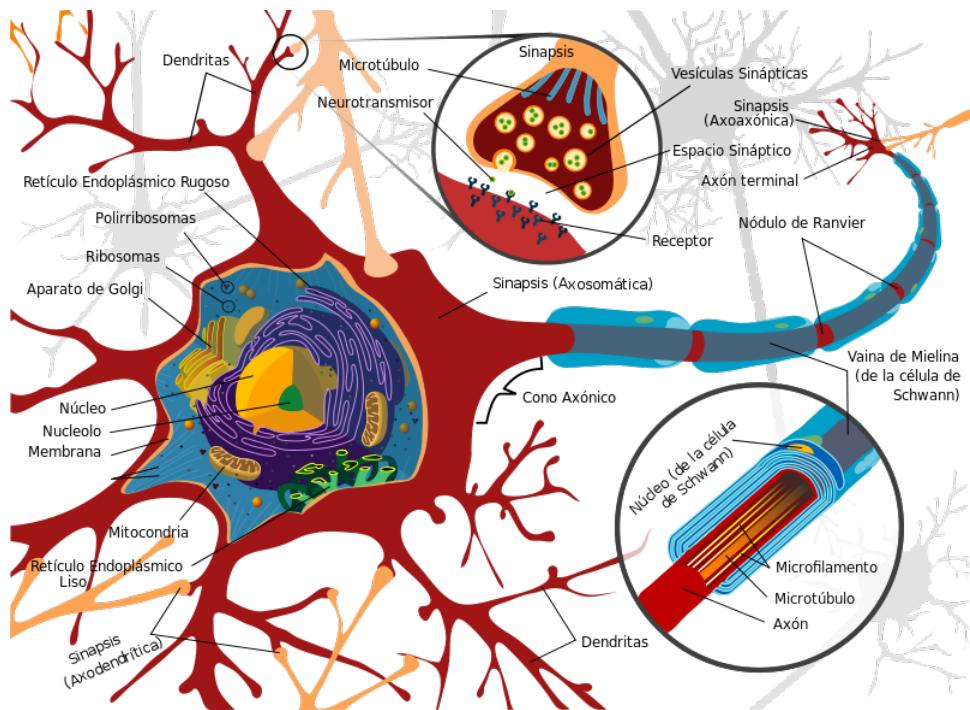


Figura 1. Estructura básica de una neurona

Fuente: Wikipedia <https://commons.wikimedia.org/wiki/File:Complete_neuron_cell_diagram_es.svg>

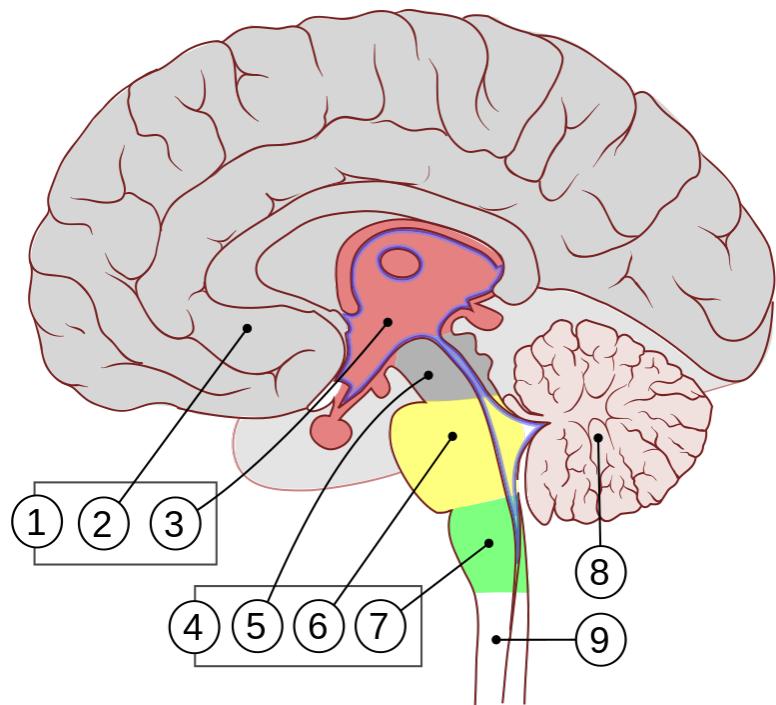


Figura 2. Corte sagital del encéfalo humano

Fuente: Wikipedia <<https://commons.wikimedia.org/w/index.php?curid=8405926>>

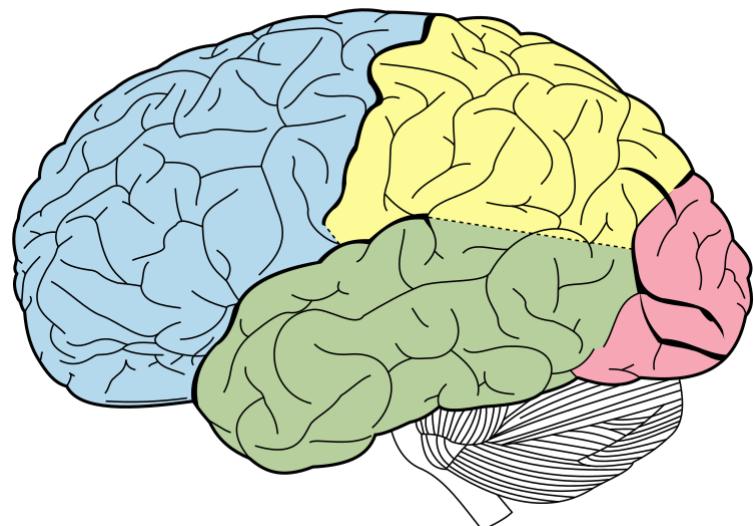


Figura 3. Lóbulos de la corteza cerebral

Fuente: Wikipedia <<https://commons.wikimedia.org/w/index.php?curid=1676555>>

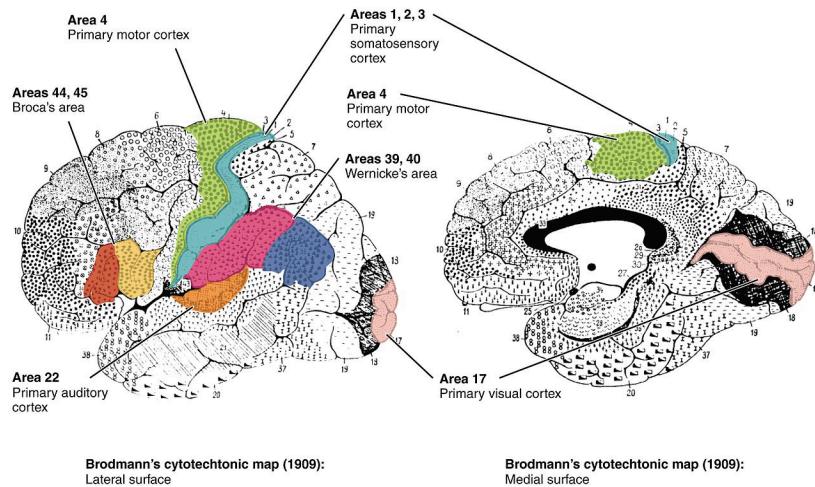


Figura 4. Áreas de Brodmann

Fuente: Wikipedia <<https://commons.wikimedia.org/w/index.php?curid=30147951>>