

## *Introducción*

DESDE SIEMPRE LOS SERES HUMANOS hemos distinguido entre nuestros procesos corporales ordinarios, como comer, beber, caminar o correr, reproducirnos o dar patadas a un balón, y procesos que no nos parecen tan ordinarios, sino incluso sorprendentes, como razonar, tener creencias, sentir emociones, recordar eventos pasados o experimentar deseos. Los procesos del segundo tipo los llamamos procesos mentales y, ya al menos desde Aristóteles, se entiende que son objeto de la ciencia psicológica, es decir, de una ciencia que se ocupa de la psique, entendida como un principio vital especial.

De manera paralela, la reflexión sobre algunos problemas profundos planteados por los procesos mentales, como su particular naturaleza o su relación con los procesos corporales ordinarios, también es antigua, pudiendo remontarse, al menos, hasta Platón y Aristóteles. Tal reflexión es de índole filosófica con lo que cabe hablar de una filosofía de la mente.

Ahora bien, la filosofía de la mente (al igual que la filosofía del lenguaje) no se constituye como disciplina académica sino recientemente. Por supuesto que René Descartes, al distinguir de manera radical entre cuerpo y mente, puso en marcha lo más básico y central de la temática en filosofía de la mente, pero habría que esperar hasta fines del siglo XIX, con la constitución de la psicología científica, para que la filosofía de la mente fuese una disciplina filosófica con datos científicos sobre los que discutir.

Con todo, el carácter académico e institucional de la filosofía de la mente es aún más reciente. Entiendo que cabe fecharlo en torno a 1950, con la difusión de las ideas de Gilbert Ryle y de Ludwig Wittgenstein. Con estos autores aparece la filosofía de la mente clásica, entendida fundamentalmente como filosofía de la psicología.

En torno a 1977, año en que se crea la revista *Cognitive Science*, aparece lo que he denominado en otro lugar «la nueva filosofía de la

mente», caracterizada por ser no sólo una filosofía de la psicología sino también una filosofía de la inteligencia artificial. En substancia, la nueva filosofía de la mente es una filosofía de las ciencias cognitivas, constituyendo éstas un campo de investigación interdisciplinar cuyo tema aglutinador es el estudio del fenómeno de la cognición, tanto en seres humanos, como en animales o en máquinas. Por ello podemos hablar de mentes humanas, mentes animales y mentes mecánicas.

En la actualidad el campo de la filosofía de la mente se ha extendido llegando hasta los límites de otras disciplinas filosóficas. Cabe pensar que tal extensión era de esperar, ya que lo mental es, en principio, lo característico de cualquier agente cognitivo. En todo caso, la filosofía de la mente actual, tal como puede apreciarse en el presente volumen, es filosofía de las ciencias cognitivas (psicología cognitiva, inteligencia artificial, lógica para la inteligencia artificial, neurociencia cognitiva), pero también permanece unida desde hace tiempo a partes de la filosofía del lenguaje, a la metafísica del sujeto, a la teoría del conocimiento y a la filosofía de la ciencia.

El presente libro contiene trece trabajos, escritos por profesores universitarios, siete de Universidades españolas y seis de Universidades extranjeras.

A fin de facilitar la lectura de los mismos, se han organizado en seis apartados: I) Posiciones en filosofía de la mente, II) Intersubjetividad, creencias e intenciones conjuntas, III) Metafísica y filosofía de la mente, IV) Teorías de las emociones, V) Neurobiología y filosofía, y VI) Cognición, lógica y ciencia.

En el primer apartado, sobre posiciones en filosofía de la mente, se incluyen sendos estudios de los profesores Andy Clark y Herbert A. Simon.

Andy Clark, de la Universidad de Sussex (Reino Unido), en su trabajo «Incorporización y la filosofía de la mente», pone de relieve cómo, en las ciencias cognitivas, el énfasis ha ido desde formas aislacionistas de solución de problemas a las interacciones agente-ambiente. Comentando la posición anti-representacionista, Clark señala que la cuestión no es si hay representadores, sino si existen estados o procesos internos distintos e identificables cuyo papel funcional o sistemático consiste en sustituir características o estados de cosas. Además, la noción de que el caso interactivo es biológicamente básico es compatible con la tesis de que la razón ambientalmente desacoplada no es la mera punta del iceberg. Clark concluye que, en vez de oponer representacionismo y dinámica interactiva, deberíamos abrazar una visión del ámbito representacional interno más amplia y buscar las continuidades cruciales entre

estrategias conductuales fuertemente acopladas y el espacio más «cartesiano» de la razón ambientalmente desacoplada.

A su vez, Herbert A. Simon, quien falleció el pasado 9 de febrero de 2001, profesor de la Universidad Carnegie Mellon (USA), Premio Nóbel de Economía en 1978, en su trabajo «Teorías computacionales de la cognición», defiende la posición ortodoxa, que él contribuyó a formar, en ciencias cognitivas. Simon concibe el pensamiento humano como procesamiento de información, es decir, como manipulación de símbolos, aunque reconoce que aún hoy la psicología del punto de vista del procesamiento de la información no ha calado en todos los psicólogos. Simon acepta, ante los conexionistas, que hay un amplio lugar para el procesamiento tanto serial como paralelo en el pensamiento humano. Por otra parte, sostiene que la simulación por computador puede capturar en general el pensamiento humano, analizando los casos más diversos. En resumen, en este trabajo Simon sigue proponiendo la hipótesis del sistema de símbolos físico considerándola una adecuada teoría computacional de la cognición.

El segundo apartado de este libro se titula «Intersubjetividad, creencias e intenciones conjuntas», y comprende estudios de Antoni Gomila Benejam, Josefa Toribio Mateas y Raimo Tuomela.

Antoni Gomila, de la Universidad de las Islas Baleares, en su trabajo «La perspectiva de segunda persona: mecanismos mentales de la intersubjetividad», plantea un modo de reconocer y atribuir estados mentales en los demás que no responde ni al modelo empático-proyectivo de la primera persona, ni al teórico-predictivo de la tercera. Según él, la perspectiva de la segunda persona es la actitud que adoptamos espontáneamente en las situaciones comunicativas de interacción cara a cara, constituyendo el mecanismo cognitivo de la intersubjetividad. Tras considerar las observaciones de diferentes autores sobre la naturaleza de la intersubjetividad, Gomila concluye que quienes han insistido en planteamientos intersubjetivos han tomado el concepto como primitivo, o como garantizado por el lenguaje, pero sin atacar el problema de su mera posibilidad psicológica. Para Gomila, una perspectiva de segunda persona explica además la socialidad, no mediante un mecanismo de cálculo autointeresado, sino mediante un mecanismo de comprensión social en términos mentalistas, que permite interacciones intersubjetivas, creación de significados compartidos y coordinación práctica.

Por su parte, Josefa Toribio, de la Universidad de Sussex (Reino Unido), presenta un estudio titulado «Responsabilidad semántica y naturalismo». En él distingue entre creencias no-responsables, que son simples codificaciones de información, y creencias responsables, que

son estados mentales susceptibles de constituirse en objetos de reflexión crítica. Asimismo sostiene la tesis de que la capacidad de reflexión crítica es central no sólo en la justificación sino también en la determinación de la naturaleza del contenido de los estados mentales. Las creencias responsables son las que de manera más genuina pertenecen al ámbito semántico, ya que al sustentar este tipo de creencias nos revelamos como participantes en una variedad de prácticas de carácter normativo. Toribio concluye argumentando que la naturalización de las creencias no-responsables es relativamente simple, pero que naturalizar las creencias responsables resulta ser no sólo un proyecto mucho más complejo sino quizás imposible.

Raimo Tuomela, de la Universidad de Helsinki y de la Academia de Finlandia, en «Intención conjunta y colectiva», analiza de manera pormenorizada las características de estas intenciones. Tuomela sostiene que, dependiendo del contexto, son necesarios diferentes tipos de intenciones colectivas para explicar respectivamente los diferentes tipos de actividades colectivas. También argumenta que la actuación conjunta colectiva y social debe depender de intenciones conjuntas en el «modo-nosotros». Para Tuomela, una auténtica intención conjunta, que está basada en un plan, es el tipo más fuerte de intención colectiva. Asimismo se ocupa del tema de tener la intención desde un colectivo y de cómo esto se diferencia de las acciones privadas agregadas de tener la intención. Finalmente Tuomela discute con cierto detalle los puntos de vista de otros autores sobre la intención conjunta, lo cual le permite aclarar su propia posición así como relacionarla con escritos suyos anteriores.

El tercer apartado de este libro lleva por título «Metafísica y filosofía de la mente» e incluye trabajos de Manuel Liz y Pascual F. Martínez-Freire.

En «El problema del mundo externo y el problema de la experiencia subjetiva», Manuel Liz, de la Universidad de La Laguna, se plantea las aparentes imposibilidades de que, por un lado, nuestra experiencia subjetiva acceda a un mundo externo y, por otro lado, de que el mundo externo acoja experiencias subjetivas. Liz analiza detalladamente ambas dificultades, que se muestran como el problema epistemológico del mundo externo y el problema ontológico de la experiencia subjetiva. Argumenta que sólo tiene sentido plantearse el problema ontológico de las relaciones que vinculan nuestra experiencia subjetiva con un mundo «externo», si damos por supuesta alguna solución al problema epistemológico acerca de las relaciones que vinculan los contenidos de nuestra experiencia subjetiva con tal mundo «externo». Asimismo Liz

defiende que hay ingredientes muy valiosos en el realismo directo como estilo de respuesta al primer problema, y que estos ingredientes pueden ser perfectamente exportables al terreno en que se sitúa el segundo problema.

A su vez, Pascual Martínez-Freire, de la Universidad de Málaga, en su estudio «La realidad desde la mente», traza el recorrido filosófico y científico que puede llevar desde nuestros procesos mentales hasta la realidad ajena a nuestras percepciones. Para ello, arranca del sujeto solipsista, como el punto de vista más radical, que aparece como absurdo e inútil. A continuación se coloca en el sujeto natural, como punto de partida más inmediato, que se revela como un sujeto de habla. Este sujeto se muestra como sujeto de representación, en el sentido de las ciencias cognitivas, y asimismo como un sujeto incorporizado, en términos filosóficos y científicos. Más aún, sobre tal noción de sujeto de representación cabe montar la categoría de sujeto cognitivo, resolviendo también en términos filosóficos y científicos, el problema clásico de las otras mentes. Finalmente, Martínez-Freire distingue entre la realidad ajena a las percepciones y los diferentes mundos percibidos, con diferencias no sólo entre los individuos humanos sino también entre los humanos y otras especies animales. Se concluye que, gracias a que poseemos básicamente un mismo sistema de representación así como una misma competencia lingüística, es posible el conocimiento objetivo del mundo percibido por los humanos y, a través de él, de la realidad.

El título del cuarto apartado de esta obra es «Teorías de las emociones» y comprende sendos trabajos de Fernando Broncano y Carlos J. Moya.

En «Las emociones: territorios intermedios en la mente», Fernando Broncano, de la Universidad de Salamanca, destaca y analiza la heterogeneidad de las emociones que las convierte en un territorio intermedio en el espectro de los estados mentales. El sistema emotivo contiene una ligadura esencial con sistemas motores y fisiológicos, contiene elementos modulares inalcanzables desde los sistemas cognitivos superiores y al mismo tiempo sus producciones se configuran culturalmente. Broncano considera las actitudes eliminacionistas respecto de las emociones y defiende una explicación funcional pura en la que podrían haber sido los memes los encargados del sostenimiento del sistema emotivo, al tiempo que éste ancla al sistema cognitivo en un espacio de evaluaciones. Para Broncano, las emociones conforman un sistema de señales o mensajes que porta y procesa contenido, aunque quizás este contenido es no-conceptual. La teoría funcional defendida es una teoría explicativa que, sin embargo, permanece neutra ante la discusión sobre el peso de la preadaptación evolucionaria de tales o cuales emociones.

Por su parte, Carlos J. Moya, de la Universidad de Valencia, en su estudio «Emociones, racionalidad y responsabilidad», se ocupa de abordar algunos problemas que plantean las emociones a la posibilidad de la responsabilidad moral. Moya comienza con la consideración de que, en la medida en que hacemos a un sujeto responsable de sus emociones, presuponemos que posee algún grado de control voluntario en relación con ellas. Analiza con detalle la concepción cognitiva de las emociones, pero sostiene que hay importantes razones para el escepticismo con respecto a tal concepción, así como para moderar el optimismo asociado a ella sobre la posibilidad del control racional de las emociones. En particular señala que, desde el punto de vista evolutivo, es plausible pensar que el funcionamiento de los sistemas emocionales primarios, relacionados con la supervivencia y la reproducción, no podía esperar al desarrollo de los sistemas cognitivos superiores. Moya concluye que el conflicto potencial entre racionalidad y emociones puede ser en último término ineliminable, y que la educación emocional es un arte, no una ciencia, un proceso sujeto a contingencias azarosas y en el que la suerte no deja de desempeñar un papel.

El quinto apartado del libro versa sobre «Neurobiología y filosofía», e incluye sendos estudios de los esposos Patricia S. Churchland y Paul M. Churchland.

En «¿Puede la neurobiología enseñarnos algo sobre la conciencia?», Patricia S. Churchland, de la Universidad de California en San Diego, sustenta la postura de que la neurobiología puede ayudarnos a comprender la naturaleza de la conciencia. Considera las distintas opciones contrarias a la investigación neurobiológica de la conciencia, como la que señala que el objetivo es absurdo, que es inconsistente con la realizabilidad múltiple, la teoría de Searle o la teoría de Dennett sobre la conciencia, o la tesis del misterio de la conciencia, encontrando que ninguna de ellas es sólida. Propone, a su vez, tomarse en serio las aportaciones de Crick sobre la conciencia visual, las de Llinás sobre los estados dormir/sueño/desperto, y las teorías de Damasio sobre la representación del cuerpo y autorepresentación. Patricia Churchland sostiene que estas tres aproximaciones ampliamente neurocientíficas son estrategias complementarias para encarar diferentes pero superpuestos segmentos del problema amplio y desconcertante de la conciencia.

Por otro lado, Paul M. Churchland, de la Universidad de California en San Diego, en su trabajo «Hacia una neurobiología cognitiva de las virtudes morales», propone una reconstrucción de los fenómenos cognitivos morales en términos neurobiológicos cognitivos. Paul M. Churchland contempla una explicación sistemática y unificada, esboza-

da en términos de redes neuronales, de los diferentes fenómenos morales: conocimiento, aprendizaje y percepción morales, ambigüedad y conflicto morales, el razonamiento moral, las virtudes y el carácter morales, patología y corrección morales, diversidad y progreso morales, realismo y unificación morales. Todo ello se apoya en la capacidad de los modelos recientes de redes neuronales (de la actividad cerebral de micronivel) para reconstruir en forma explicativa los rasgos principales de la actividad cognitiva de nivel global. Por ejemplo, desde la perspectiva neurocomputacional adoptada por Paul Churchland, el progreso moral sólo se muestra diferente en su enfoque ontológico del llamado progreso científico, pero en un caso como en el otro tenemos adquisiciones de un conjunto de habilidades, explicables mediante los mismos mecanismos neuronales.

Finalmente el sexto apartado del libro se titula «Cognición, lógica y ciencia», comprendiendo los estudios de Alfredo Burrieza Muñoz y Anna Estany.

En «Conciencia, lógica e inteligencia artificial», Alfredo Burrieza, de la Universidad de Málaga, examina el tratamiento lógico de la noción de conciencia dentro del campo de la inteligencia artificial. Basándose en el trabajo de Fagin y Halpern, se estudia la conciencia como un operador lógico añadido a lenguajes formales de conocimiento y de creencia, cuyo significado viene dado, en el contexto de una semántica de mundos posibles, por el conjunto de fórmulas de las que un agente es consciente en cada mundo. Por otra parte, Burrieza analiza cómo la conciencia desempeña el papel de un filtro, de tal manera que el conocimiento o creencia explícita es el resultado de conocimientos o creencias implícitas que deja pasar dicho filtro. Esta visión sintáctica de la conciencia es además una herramienta adecuada para tratar el problema clásico de la omnisciencia lógica. Asimismo Burrieza examina algunas otras interpretaciones de la conciencia en diversos campos del razonamiento sobre creencia y conocimiento en inteligencia artificial.

Anna Estany, de la Universidad Autónoma de Barcelona, en su estudio «Ventajas epistémicas de la cognición socialmente distribuida», aplica a la filosofía de la ciencia el modelo de cognición socialmente distribuida desarrollado recientemente en las ciencias cognitivas. El problema planteado en filosofía de la ciencia consiste en cómo aunar la falta de un algoritmo para la elección de teorías científicas con la objetividad de la ciencia. Estany considera los criterios racionales para la elección de teorías en varios autores destacados. Su nuevo paso consiste en recurrir, en primera instancia, a la idea de actividad situada en ciencias cognitivas, y, en segunda instancia, al modelo de Edwin Hutchins de la cognición

socialmente distribuida. Hay muchas actividades humanas regidas por sistemas con múltiples actores que buscan una interpretación coherente de una serie de fenómenos; el equipo de navegación y el de la cabina de un avión son dos ejemplos analizados por Hutchins, pero también cualquier comunidad de científicos es un sistema de este tipo. Estany concluye que todo lo referente a la naturaleza y estructura de la cognición del grupo es relevante para la epistemología.

Para terminar quiero manifestar mi más profundo agradecimiento al profesor Manuel Toscano Méndez, Secretario de Redacción de la revista *Contrastes*, por su inestimable colaboración en la edición del presente libro.

*Pascual F. Martínez-Freire*  
Universidad de Málaga