

Teoría de la mente y Neuronas Espejo

**Ponencia presentada en II Jornadas de Filosofía de la Mente. Facultad de
Filosofía. Universidad Complutense. Madrid.**

Emilio García García

**Dpto. Psicología Experimental, Procesos Cognitivos y Logopedia
Universidad Complutense. Madrid**

Introducción

En nuestro tiempo sigue pendiente un problema en las Ciencias Cognitivas, y particularmente en la Neurociencia: explicar cómo surgen los procesos mentales a partir de la materia física del cerebro. El desafío para esta década, y no sabemos cuántas más, está en conocer el código por el cual los aproximadamente cien mil millones de neuronas y sus complejÍsimas conexiones pueden generar la conciencia personal, la propia identidad, los cualia mentales. Comentamos dos aproximaciones a la mente: la Teoría de la Mente y las Neuronas Espejo.

Teoría de la Mente

La psicología popular asume que las personas tienen mente, y que el comportamiento de las personas se debe a lo que tienen en su mente. La mente, entendida como un sistema de conocimientos e inferencias que permite interpretar y predecir la conducta de los demás, merece el calificativo de “teoría”, puesto que no es directamente observable y sirve para comprender, predecir y modificar el comportamiento. En cierta medida se puede comparar con los conceptos y teorías que emplean los científicos para explicar, predecir y modificar el campo de realidad que estudian. Las teorías de los científicos tampoco son observables, pero permiten explicar una determinada realidad.

Teoría de la mente hace referencia a la capacidad de “leer la mente”, de “mentalizar”, de suponer lo que está pensando y sintiendo la otra persona, de comprender y prever su comportamiento. Cuando leemos la mente de otro estamos suponiendo sus estados mentales, sus pensamientos y sentimientos, y somos capaces de

anticipar lo que va a hacer. Captamos lo que pasa por la mente de otros valiéndonos de conocimientos y razonamientos para entender y predecir su conducta. Tenemos un acceso reflexivo, consciente y lingüístico a la mente. Suponemos las intenciones, creencias, deseos y sentimientos del otro porque nuestro cerebro nos permite construir teorías acerca de las otras personas. Esta capacidad de predecir y explicar la conducta tiene éxito y es condición necesaria para que las relaciones interpersonales y el mundo sociocultural sean posibles y tengan sentido. Las teorías modulares de la mente han propuesto módulos cerebrales innatos y específicos, para explicar los procesos de mentalización.

Tradicionalmente se ha hecho hincapié en determinados factores, como posición bípeda, conformación de la mano, fabricación de herramientas y útiles, como las fuerzas impulsoras en el proceso de hominización. Se ha destacado la importancia de lo que podemos llamar inteligencia técnica o capacidad para fabricar instrumentos y medios con objeto de satisfacer necesidades de supervivencia, caza, defensa, cobijo, etc. Pero más importante y decisivo en la filogénesis de la mente probablemente fue la capacidad para resolver los problemas de orden social, la inteligencia social, como motor del proceso de hominización y desarrollo de la mente. La inteligencia social sería la fuerza determinante en la conquista de superiores niveles de inteligencia y desarrollo de la mente.

La vida en grupo de nuestros antepasados planteaba problemas muy complejos que requerían alto grado de cooperación, colaboración y organización para la caza, defensa y ataque frente a otros, la división del trabajo, jerarquía y orden social, relaciones sexuales y pautas de crianza. Enfrentarse con estas demandas sociales requería capacidades mentales complejas. La mente habría evolucionado ante la presión, precisamente, de la exigencia de la vida en grupo. Esas conquistas mentales podrían aplicarse, después, a otras situaciones problemáticas del mundo físico.

La hipótesis de la existencia de dos grandes tipos o dominios de mente, la física y la social, ha recibido confirmación desde diferentes ámbitos. Premack y Woodruff, en 1978, publicaron un trabajo clave sobre la teoría de la mente de los chimpancés. La cuestión se planteaba así: ¿tiene el chimpancé una teoría de la mente? Los antropoides superiores no hablan sobre la mente, ni elaboran teorías sobre la mente, ni expresan verbalmente sus pensamientos, deseos o sentimientos; pero puede que sí atribuyan

mente a otros individuos de su misma especie o próximos, como el hombre. Los seres humanos a determinadas edades tempranas tampoco hablan sobre la mente, ni son conscientes de que tienen mente, y sin embargo sí atribuyen mente a los demás. Para estudiar las atribuciones o inferencia de estados mentales que los chimpancés pueden hacer, Premack y Woodruff realizaron ingeniosos experimentos, que se han seguido desarrollando en investigaciones posteriores con primates.

En el estudio sobre teoría de la mente el engaño ocupó, desde sus inicios, un lugar privilegiado. Cuando un individuo engaña está suponiendo pensamientos en la mente de otro, e intenta alterar tales pensamientos para que haga algo que le interesa. Por tanto, el engaño es un indicador fiable de la teoría de la mente. Woodruff y Premack realizaron otra investigación para ver si los chimpancés engañaban. Cuatro chimpancés disponían de información sobre la localización de comida, que estaba fuera de su alcance. En unos casos una persona "cooperadora" entraba a la sala donde se encontraban los chimpancés y les acercaba la comida; en otros casos otra persona "competidora" cogía la comida y se quedaba con ella. Los chimpancés aprendían a discriminar entre situaciones en las que resultaba adaptativo "informar correctamente" cuando la persona era cooperadora, o engañar y ocultar información cuando era competidora, como dirigir la mirada a lugares muy diferentes a donde se encontraba la comida.

En el engaño, un individuo sabe que otro tiene una representación errónea de la realidad, o llega a provocar en el otro un conocimiento o representación equivocada, para aprovecharse y sacar partido en propio interés y beneficio, al predecir el comportamiento del otro a partir del conocimiento erróneo que éste tiene. Las conductas de engaño son muy frecuentes en temas de alimento, sexo y poder.

La vida en grupo plantea a los primates -y a los seres humanos- muchos problemas que exigen unas capacidades mentales específicas, una inteligencia social para poder afrontarlas con eficacia. Continuamente los individuos tienen que sopesar los pros y los contras de cooperar o competir. Se compete por sexo, alimento, poder, territorio, compañero de juego, de acicalamiento; y también se puede cooperar según circunstancias. Resolver este tipo de problemas sociales parece más difícil que la capacidad para utilizar instrumentos, como ramas de árbol para sacar y comer hormigas, o la capacidad para reconocer y orientarse en el territorio. El engaño es la otra cara de la cooperación y colaboración. Engañar y colaborar son indicadores de la inteligencia

social. La misma mente social puede ser maquiavélica o solidaria, pues lo que la caracteriza es disponer de un sistema de conocimientos e inferencias sobre los estados mentales de otros individuos, que posibilita saber quiénes son amigos o enemigos, y qué cabe esperar de ellos.

No se da unanimidad entre los investigadores respecto a si los chimpancés poseen una teoría de la mente en el sentido más pleno, de modo que incluya intenciones, sentimientos y deseos. Se considera esencial en la teoría de la mente la presencia de representaciones que puedan resultar verdaderas o falsas. Se ha planteado que las habilidades mentalistas de los chimpancés puedan no ser tan mentalistas, y sus comportamientos se puedan explicar desde supuestos más básicos, como procesos de aprendizaje operante o por imitación.

Es mayor el consenso en el estudio de la teoría de la mente en seres humanos. Las personas continuamente hablan y expresan sus estados mentales. La investigación ha sido muy variada e importante en diferentes áreas, como la filosofía de la mente, la psicología evolucionista, la psicología evolutiva, la psicolingüística, la psicopatología y la psiquiatría. Parece muy razonable aceptar que el estudio de la mente en las personas es más prometedor y fácil que en los grandes simios, ya que los humanos podemos expresar, gracias a la comunicación lingüística, las intenciones, creencias, sentimientos y deseos, y atribuir a los estados mentales la causa del propio comportamiento.

En la investigación de la mente infantil se ha utilizado, y con gran éxito, el paradigma de la falsa creencia. En un estudio clásico de H. Wimmer y J. Perner, un niño contempla una situación en la que el experimentador y otro niño, Juan, están juntos en una habitación. El experimentador esconde un trozo de chocolate bajo una caja que se encuentra delante de Juan. Entonces Juan sale un momento de la habitación y, mientras está ausente, el experimentador cambia el chocolate a otro escondite. Se le pregunta al niño dónde está realmente el chocolate, y dónde lo buscará Juan cuando entre en la habitación. El niño tiene que distinguir entre lo que sabe que es cierto, o sea dónde está realmente ahora el chocolate, y lo que sabe del estado mental de Juan, de lo que piensa Juan. Además, tiene que inferir que el comportamiento de búsqueda del chocolate por parte de Juan dependerá de las representaciones mentales de Juan y no de la realidad.

A la edad de tres años los niños no resuelven correctamente el problema y responden en función de la situación real que ellos conocen. No comprenden que el

protagonista se comportará según su creencia falsa. Entre los cuatro y cinco años los niños ya no tienen dificultad para resolver la tarea. En el primer ejemplo el niño tiene un conocimiento verdadero de dónde está escondido realmente el chocolate, pero el otro niño, Juan, tiene una creencia falsa, Juan actuará en función de su creencia equivocada y buscará en la caja donde pensaba que estaba el chocolate, cuando realmente el niño sabe que no está ahí. Para responder correctamente a las preguntas de dónde buscará Juan, el niño debe saber que los demás tienen pensamientos y deseos, y que se comportan a partir de ellos, y que esos pensamientos pueden ser verdaderos o falsos, y que la gente se comporta según sus pensamientos y creencias. Además, el niño es capaz de separar sus propias creencias de las creencias que tiene Juan, que está equivocado. Diferencia entre contenido proposicional: "el chocolate está realmente en ..." de la actitud proposicional: "Juan cree que el chocolate está en ...".

La investigación en teoría de la mente ha tenido especial relevancia en psicología clínica y psicopatología, particularmente en el autismo. Una de las teorías para comprender el Trastorno de Espectro Autista sería el déficit en la capacidad de mentalizar, de leer la mente, que explicaría las dificultades en la interacción interpersonal y comunicación social. Los niños con autismo presentan graves dificultades para comprender que los demás puedan tener representaciones mentales diferentes a las propias, y además puedan no corresponder con la realidad. Se planteó entonces si el autismo mostraría un déficit en unas capacidades mentales concretas y propias de un módulo específico, que podría dejar preservadas otras capacidades mentales, correspondientes a otros módulos. El autismo podría entonces explicarse como consecuencia de un déficit específico del módulo de teoría de la mente. Este módulo estaría codificado genéticamente y se desplegaría en un momento dado del desarrollo cerebral, y funcionaría de modo específico, encapsulado, automático, independiente de otras capacidades mentales del individuo, de modo similar a como ocurre con el lenguaje en la teoría chomskiana.

Neuronas Espejo

Además de la teoría de la mente hemos mencionado una segunda aproximación a la mente en la Teoría de la simulación. Para explicar los estados mentales y los comportamientos de los otros recurrimos a la simulación. Comprendemos lo que los

otros hacen o van a hacer en una situación concreta, imaginando que nosotros estamos en esa situación. Proyectamos los propios estados mentales, motivacionales, cognitivos y emocionales, mediante un proceso empático. Llegamos a la mente de los demás sirviéndonos de nuestra propia mente como modelo. En lugar de disponer de un sistema conceptual o una teoría de la mente, la persona imagina qué estados mentales tendría si estuviera en la situación del otro, y proyecta en el otro sus propios estados mentales.

Lo que vemos en el cuerpo, y particularmente en el rostro y la mirada del otro, son sus estados mentales, sus emociones e intenciones, su misma mente. La mente no está oculta sino abierta y patente a nuestro alcance en la misma expresión gestual. No la suponemos, atribuimos o inferimos, sino que la percibimos, reconocemos e identificamos en la expresión, y respondemos a la expresión con otras expresiones en una interacción comunicativa y cooperativa. En esta tradición fenomenológica de la filosofía de la mente situamos la investigación sobre Neuronas Espejo.

Las Neuronas Espejo son un tipo particular de neuronas que se activan cuando un individuo realiza una acción, pero también cuando él observa una acción realizada por otro individuo. No necesitamos realizar razonamientos lógicos para comprender la mente de los otros; las neuronas espejo nos permiten comprender las intenciones, sentir las emociones de otras personas con sólo observar sus gestos y comportamientos. Las investigaciones sobre neuronas espejo proporcionan explicaciones sobre las bases neuronales de la teoría de la mente.

En los comienzos de la década de 1990, un equipo de neurobiólogos italianos, dirigido por G. Rizzolatti, con L. Fadiga, L. Fogassi y V. Gallese, en la universidad de Parma, se encontró con unos datos inesperados en el transcurso de la investigación con monos macacos, que tenían microelectrodos implantados en la corteza premotora del cerebro, para registrar la actividad eléctrica de las neuronas, cuando los monos mostraban ciertos comportamientos, como agarrar un palo o comer una uva. En el córtex promotor es sabido que se planean los movimientos. En determinadas ocasiones sucedía algo desconcertante, al activarse el aparato de registro sin que el mono realizase ninguna actividad. Las neuronas se activaban sin que el mono hiciese un movimiento, bastaba con que viera que otro realizaba una acción. Los científicos italianos habían identificado un tipo de neuronas desconocidas hasta ese momento, las denominaron

neuronas espejo. El artículo se publicó en 1992. Esta primera publicación inició un camino de múltiples y espectaculares investigaciones hasta la actualidad.

La investigación comenzó en monos macacos y proporcionó claves para explicar los procesos cerebrales que estarían a la base de los comportamientos interindividuales en un contexto determinado. A lo largo de la década de 1990 se realizaron investigaciones que constataron repetidamente cómo al observar un comportamiento se activaban en el cerebro del mono unos patrones similares a los que se originaban cuando el mono llevaba a cabo esas conductas. Se analizaron comportamientos diversos, como coger alimentos, llevarlos a la boca, comerlos, manipular objetos, hacer gestos con manos o boca, etc. Se comprobó que las neuronas espejo se activaban con distinta intensidad según se tratase de conductas en diversos contextos, y en la mayoría de los casos se activaban cuando el mono observaba la conducta ajena de modo similar a cuando ejecutaba la acción.

Pronto se planteó la existencia de neuronas espejo en seres humanos, pero las técnicas para investigar las neuronas espejo en los macacos no eran aplicables. No se pueden implantar electrodos intracraneales en personas para objetivos experimentales. Se aplicaron entonces otras tecnologías como la estimulación magnética transcraneal, la tomografía por emisión de positrones, la resonancia magnética funcional, e incluso el registro de neuronas en operaciones de neurocirugía. En el ser humano se han identificado sistemas de neuronas espejo en la corteza motora primaria, principalmente el área de Broca, el área parietal inferior, la zona superior de la primera circunvolución temporal, el lóbulo de la ínsula, el hipocampo y sistema límbico.

Los sistemas de neuronas espejo posibilitan el aprendizaje de gestos por imitación: sonreír, caminar, bailar, jugar al fútbol, etc., pero también sentir que nos caemos cuando vemos por el suelo a otra persona, la pena que sentimos cuando alguien llora, la alegría compartida. El intercambio complejo de ideas y prácticas que llamamos cultura, los trastornos psicopatológicos, los trastornos de lenguaje, el trastorno de espectro autista, pueden encontrar en las neuronas espejo claves de explicación.

Las neuronas espejo forman parte de un sistema de redes neuronales que posibilita la percepción-ejecución-intención-emoción. La simple observación de movimientos de la mano, pie o boca, activa las mismas regiones específicas de la corteza motora, como si el observador estuviera realizando esos mismos movimientos. Pero el proceso va más

allá de que el movimiento, al ser observado, genere un movimiento similar latente en el observador.

Las investigaciones han constatado que la mera observación de las acciones de los demás activa en el observador las mismas áreas cerebrales, como si fuera él mismo quien ejecutara las acciones. Los descubrimientos se han aplicado en programas de rehabilitación de pacientes con lesiones cerebrales en las áreas motoras, por ictus cerebral. Las partes del cuerpo paralizadas pueden recuperar funciones mediante ejercicios de rehabilitación, ya que las áreas cerebrales próximas pueden asumir las funciones de las lesionadas. Ello exige a los pacientes intensa y continuada práctica de ejercicios.

En determinadas circunstancias la activación de las neuronas espejo depende de lo familiarizados que estemos con las imágenes vistas. Por ejemplo, en el aprendizaje de un nuevo deporte, quien no haya jugado nunca al tenis o no haya practicado la natación, no es probable que pueda imitar en su mente los movimientos precisos. La observación de acciones especializadas, por ejemplo, de bailarines profesionales, activan las neuronas espejo de forma muy diferente dependiendo de si quien lo observa es también un bailarín profesional, familiarizado con cada uno de los movimientos, o si por el contrario es una persona ajena al mundo de la danza. En este último caso la comprensión de lo que hace el bailarín no es inmediata ni empática, está más intelectualizada y pertenece a un dominio más semántico y consciente.

Cuando una persona realiza acciones en contextos significativos, tales acciones van acompañadas de la captación de las propias intenciones y emociones al realizarlas. Se conforman sistemas neuronales que articulan la propia acción asociada a la intención o propósito que la activa y a la emoción. Formadas estas asambleas neuronales de acción-ejecución-intención-emoción en un sujeto, cuando ve a otro realizar una acción, se provoca en el cerebro del observador la acción equivalente, evocando a su vez la intención y emoción con ella asociada. El sujeto, así, puede atribuir a otro la intención y emoción que tendría tal acción si la realizase él mismo. Se entiende que la lectura que alguien hace de los estados mentales del otro es, en gran medida, atribución desde las experiencias propias. Si alguien entiende la acción de otra persona es porque tiene en su cerebro un modelo para esa acción, basado en sus propias experiencias. La observación de alguien emocionado provoca una reacción de sistemas neurales especulares, de modo que el observador vivencia en su cerebro similar emoción, el observador siente y experimenta directamente el mismo estado emocional.

La alegría, la tristeza, el dolor, el miedo, la ira, el asco, la sorpresa son emociones susceptibles de ser compartidas por quien las observa. Cuando vemos a una persona sonriente, inmediatamente sintonizamos con su estado emocional, y parece que nos contagiamos de su alegría. Cuando vemos a otra persona en apuros, parece que inconscientemente simulamos tales apuros en nuestra mente, como si sintiéramos las sensaciones negativas de la otra persona, y ello nos llevara actuar para aliviar su situación. Las investigaciones demuestran que respondemos a las emociones, alegría, tristeza, dolor de los demás con análogos patrones fisiológicos de activación, como si nos ocurriera a nosotros. Literalmente sentimos los estados emocionales de los demás como si fueran propios. La capacidad de simular lo observado tiene una especial relevancia para la interacción social, creando un espacio de acción compartido, necesario para las relaciones interindividuales

Otra emoción muy estudiada es el asco y rechazo. La visión de expresiones faciales de asco provoca en el observador la activación de la región anterior de la ínsula, por lo que la activación de esta área cerebral es crítica, no sólo para desencadenar sensaciones y reacciones de asco, sino también para percibir un estado emotivo semejante en la cara de otras personas. Los daños en la ínsula provocan en los que lo padecen incapacidad de sentir asco, pero también de reconocer expresiones tanto visuales como sonoras de asco en los demás.

Los sistemas de neuronas espejo posibilitan comprender los estados mentales de otros, sus intenciones, pensamientos, emociones y también sus recuerdos. Cabe pensar que las redes de neuronas espejo están muy interconectadas y distribuidas por todo el cerebro, desde las áreas prefrontales a las temporales. Las redes de neuronas espejo en las áreas temporales, responsables de los aprendizajes y memorias, se activarían también en los recuerdos. Al realizar una acción en un contexto determinado, se activarían redes neuronales que codifican el recuerdo o memoria episódica.

Conclusiones

La investigación en teoría de la mente y neuronas espejo se ha desarrollado desde marcos teóricos y con metodologías muy distintas. La teoría de la mente ha sido tema de investigación en la psicología cognitiva, evolutiva y psicopatología, mientras que las neuronas espejo proporcionan explicaciones desde la neurociencia. La teoría de la mente, en la década de 1980 y 1990, aceptaba los postulados modulares y entendía el

cerebro organizado en regiones especializadas y localizadas. Las neuronas espejo, desde los años 1990, propone circuitos y redes neuronales muy interconectados y distribuidos por todo el cerebro, para explicar los procesos mentales.

Los procesos cognitivos y emocionales humanos son muy complejos, desde los que se sitúan a nivel más automático e inconsciente, a los más conscientes y controlados. Podemos distinguir procesos *bottom-up*, que son respuestas automáticas de abajo arriba, y procesos *top-down* de arriba abajo, que implican consciencia y mayor control emocional. Cuando vemos una cara con determinados gestos e intuimos rápidamente su estado mental de tristeza, lo hacemos de modo automático e inconsciente. Cuando recordamos una experiencia triste que hemos vivido, están presentes procesos más deliberados, conscientes y lingüísticos. En la primera situación dominan los procesos *bottom-up* de carácter más afectivo; en la segunda están más presentes procesos *top-down* y más cognitivos.

La explicación de los procesos mentales, la consciencia, el yo, sigue pendiente hoy. No disponemos en la actualidad del código por el cual el cerebro humano traduce un lenguaje de impulsos eléctricos y químicos en los estados mentales. Explicar cómo las tormentas de impulsos nerviosos que recorren el sistema nervioso se traducen en percepciones, cogniciones y emociones sigue siendo el desafío de la neurociencia.

Para los autores místéricos como Chomsky, Fodor, Pinker, McGinn es objetivo imposible porque nuestro cerebro no evolucionó para responder tales cuestiones, sino que es resultado de la evolución para sobrevivir y dejar descendencia. Pensar que los *homo sapiens* poseen capacidades cognitivas ilimitadas es ir en contra del principio darwiniano de selección natural.

Para los místéricos, la mente humana presenta un “cierre cognitivo” para determinados problemas. Así como perros y gatos no pueden entender los números primos el cerebro humano está cerrado a algunas cuestiones. Responder a cuestiones básicas de la neurociencia como la forma en que los procesos físicos del cerebro dan lugar a los procesos mentales, la consciencia, resulta imposible para la mente humana, es un misterio.

Para estos autores, en el caso más extremo, podemos llegar a formular alguna teoría, pero no llegaremos a comprenderla. Algo así como ocurre en la física cuántica y la teoría de la relatividad. Para la mecánica cuántica las partículas pueden estar en dos

lugares a la vez o surgir del espacio vacío, aunque resulta difícil de comprender los procesos de la física cuántica permite hacer predicciones científicas.

Los autores que hemos calificado de problemáticos, que son la mayoría de los neurocientíficos actuales, como Damasio Gazzaniga, Edelman, Ramachandran consideran atrevido y arrogante calificar de misterios determinados problemas que plantean las ciencias, particularmente la neurociencia. Nuestro cerebro no evolucionó para captar los rayos X, los ultravioleta, las ondas gravitacionales, y necesitamos determinadas prótesis tecnológicas para observarlos como ocurre con otros fenómenos del macrocosmos y microcosmos.

Estamos cerrados perceptivamente a la luz ultravioleta, pero con determinadas tecnologías podemos tener acceso. Gracias a la cultura y conocimientos acumulados, la interacción de humanos llega a resultados asombrosos. La mente extendida es una gran conquista. Frente al pesimismo sobre las capacidades cognitivas y las limitaciones biológicas que defienden los místicos, están los logros alcanzados por la cultura científica. Resulta dogmático considerar determinadas cuestiones inabordables para la mente humana y parece más prudente y razonable considerar los diversos interrogantes que tenemos como problemas a resolver. Entre ellos está el problema mente-cerebro.

Las técnicas de neuroimagen están comenzando a mostrar las complejas, interactivas y distribuidas redes neuronales que están a la base de las capacidades humanas. Cabe esperar que en los próximos años estén disponibles nuevas tecnologías en el marco de los programas que se están desarrollando, como son el *Proyecto Brain*, el *Proyecto Human Brain Project* y el *Connectome Project*, que mostrarán las redes y subredes neuronales en la conectividad funcional del cerebro humano.