

COMPETENCIAS MENTALISTAS Y BASES NEURALES: SOBRE UNA PROPUESTA MIXTA

Fernanda Velázquez

Introducción

Como es sabido, los estados mentales no se pueden observar directamente y tampoco existe una correlación simple entre estados mentales y conducta. Sin embargo, cotidianamente atribuimos estados mentales a otras personas, y explicamos y predecimos su conducta en base a éstos. La mayoría de estas inferencias se producen automáticamente, sin la mediación del pensamiento o de la deliberación.

La competencia de percibir y comunicar estados mentales es considerada una habilidad básica de la cognición social, que subyace al aprendizaje social y a las interacciones sociales exitosas, factores necesarios para el progreso de nuestra especie y de la tecnología. El estudio de la habilidad de atribuir estados mentales asume que las personas atribuyen estados mentales, y que *leer* la mente de los otros es importante porque son sus estados mentales los que determinan sus acciones. La idea de que las acciones son causadas por estados mentales se ha denominado *actitud intencional* (Dennett, 1987) o *poseer una teoría de la mente* (Premack & Woodruff, 1978).

Desde Premack y Woodruff (1978), la expresión *Teoría de la Mente* ha designado un dominio de investigación específico, cuyo propósito es proveer una explicación de la habilidad de comprender, explicar y predecir las acciones propias y ajenas. Llamativamente, este estudio seminal se preguntaba si los chimpancés tenían una teoría acerca de la mente de otras criaturas o no. Esta pregunta ha resultado difícil de contestar y continúa siendo una cuestión abierta si los primates no humanos poseen una Teoría de la Mente (Povinelli & Bering, 2002; Tomasello, Call, & Hare, 2003).

El impacto de este artículo se extendió del dominio de la primatología a la psicología y la filosofía. En sus comentarios a Premack y Woodruff (1978), Dennett (1978), Bennet (1978) y Harman (1978) señalaron cuáles son las condiciones que debe reunir una conducta observable para poder inferir, a partir de ésta, que la criatura posee una concepción de la mente

del otro. Por un lado, un organismo que posee una Teoría de la Mente tiene que ser capaz de tener creencias acerca de las creencias de los otros y de distinguirlas de las propias; por el otro, tiene que ser capaz de hacer o predecir algo en relación a esas creencias atribuidas y que distingue de las propias. Por esto, un test adecuado se debe basar en la generación de una predicción sobre la conducta de otro agente, a partir de la atribución de una creencia equivocada (falsa). En general, cuando una conducta está relacionada al estado de cosas en el mundo, puede ser predicha en base al conocimiento de las regularidades que conectan la situación y la conducta. La predicción de que alguien va a ir a recoger frutillas donde las hay, no permite distinguir entre mis creencias y las creencias del otro. En cambio, si uno predice que alguien va en busca de frutillas al lugar en que uno sabe que no las hay, entonces tal predicción está basada en la representación equivocada que ese alguien tiene sobre la situación.

Wimmer y Perner (1983) adaptaron la sugerencia y, sirviéndose del lenguaje, elaboraron un paradigma experimental claro para utilizar con niños: la tarea de falsa creencia (de primer orden). La tarea original (existen variaciones como la tarea de 'Sally-Anne' de Baron-Cohen, Leslie & Frith, 1985) implica un personaje (Maxi) que coloca una golosina en un lugar particular de la habitación y se retira. Durante su ausencia, la golosina es movida de lugar, y entonces se le pregunta al niño dónde buscará Maxi la golosina cuando regrese. Para llevar a cabo exitosamente la tarea, el niño debe entender que Maxi piensa que la golosina está aún donde la colocó, es decir, el niño debe entender que Maxi tiene una creencia errónea (una falsa creencia). Esta tarea se adecua a los requisitos señalados anteriormente porque, para hacer una predicción correcta, el niño debe ser capaz de contrastar la propia percepción de la situación (el estado del mundo) con la creencia que tiene el personaje de la historia (la creencia de Maxi). En otras palabras, para resolver la tarea de falsa creencia (*TFC*) es preciso que el niño se represente no sólo el estado de los hechos, sino también la capacidad de personaje de representarse la situación en función del acceso informativo que ha tenido a ella.

El hallazgo de Wimmer y Perner (1983) fue que los niños pasan con éxito la tarea entre los 4 y 6 años de edad, por lo cual sugirieron que en este período del desarrollo parece emerger una nueva habilidad cognitiva: la de representarse creencias erróneas y de entender cómo una persona reaccionará en base a estas creencias (Wimmer & Perner, 1983). Esta

capacidad se denomina *metarepresentación* (Pylyshyn, 1978) y es considerada el supuesto básico de la Teoría de la Mente (*TdM*).

Desde entonces, los psicólogos se han encargado de replicar la tarea, redefinirla y, a su vez, han tratado de hacer descender la edad crítica de adquisición. Este hallazgo se ha repetido en otras tareas como en el *paradigma de la esponja* (Flavell, Flavell, & Green, 1987), que consiste en mostrar al niño una esponja pintada de modo tal que parece una piedra y preguntarle: “¿qué es?”. El niño debe contestar que “es una piedra”; a continuación se le pide que la toque y determine qué es. Entonces el niño contestará que “es una esponja”. Finalmente, se le pregunta “¿qué parece?”, y para tener éxito en la tarea, el niño debe contestar que “parece una piedra”. Sin embargo, los niños menores de 4 años contestan que parece una esponja. Los investigadores, al explicar estos resultados, sugieren que el niño aún no tiene una concepción de la capacidad representacional de la mente en este período del desarrollo y, por esto, no capta cómo algo puede ser de un modo y aparecer de otro.

Otra tarea que apoya esta observación es la del *tubo de smarties* (Perner, Leekman, & Wimmer, 1987), en la cual se muestra al niño un tubo de caramelos y se le pregunta por el contenido del mismo. El niño debe responder “caramelos”. A continuación se le muestra que se ha cambiado el contenido del tubo y que en lugar de caramelos, hay lápices. Se hace entrar a otro niño y se le pregunta al primero “¿qué dirá el nuevo niño cuando se le pregunte por el contenido del tubo?”, a lo cual debe responder “caramelos”. Sin embargo, los niños menores de 4 años contestan “lápices”.

En general, las tareas para testear Teoría de Mente se sirven del lenguaje, ya que por medio de éste los seres humanos atribuimos explícitamente deseos y creencias, o predecimos manifiestamente conductas. Sin embargo, es preciso señalar que estudiar la *TdM* sirviéndose del lenguaje puede llevar a minusvalorar las capacidades mentalistas de los niños pequeños y de los primates no humanos.

Los dos enfoques

Desde siempre, los filósofos han tenido interés en el conocimiento (adulto) de la propia mente y de la mente de los otros, pero no se han interesado en la adquisición de este conocimiento, asumiendo sin más, que estaba ligada a algún proceso de enculturación. Sin embargo, desde los años '80 ha cambiado esta situación y los filósofos han volcado su interés en las

cuestiones de la adquisición. Una de las razones de este cambio fue el desafío propuesto por la alternativa simulacionista (Gordon, 1986; Heal, 1986).

Tal alternativa teórica planteó nuevos argumentos y reformuló el debate acerca de cuáles son las habilidades y recursos que subyacen a la habilidad de mentalizar. Hasta el momento, los teóricos no han llegado a un acuerdo en este punto. En principio, se puede decir que existen dos enfoques dominantes, aunque se presentan en distintas formulaciones con diferencias importantes.

Las dos corrientes teóricas principales en filosofía de la mente son la Teoría de la Teoría (*TT*) y el simulacionismo o Teoría de la Simulación (*TS*). Algunos autores señalan la existencia de una tercera corriente teórica de raigambre racionalista o hermenéutica (Goldman, 2006). En este trabajo se tratarán los dos primeros enfoques que han tenido extensiones en el estudio de las bases neurales de la habilidad de atribuir mente o mentalizar.

El enfoque *TT* implica la idea básica de que la atribución de estados mentales se lleva a cabo recurriendo a principios o reglas generales del comportamiento (Fodor, 1987). Éstas pueden tener la siguiente forma: “generalmente, la gente cree en lo que ve” o “si alguien quiere algo, y cree que hay algo que puede hacer para obtenerlo, entonces, generalmente, llevará esto a cabo, *ceteris paribus*”. Tales principios y reglas están representados internamente, y conforman una teoría psicológica tácita o *Folk Psychology*, a partir de la cual, por medio de algún tipo de inferencia, se subsumen los casos particulares a los principios generales para sacar conclusiones acerca del comportamiento de los demás y del propio (Stich & Nichols, 1992; Botterill & Carruthers, 2003).

A diferencia de la *TT*, el enfoque de la simulación mental considera que atribuimos estados mentales y explicamos y predecimos la conducta de otros, sin recurrir a una teoría compleja. Los defensores de la Teoría de la Simulación consideran que la dificultad principal de la Teoría de la Teoría consiste en que las reglas o principios de la teoría postulada no se han podido reconstruir, ni pueden pensarse como adquiridas por un niño pequeño en el momento en que ya es capaz de hacer atribuciones de este tipo. Según esta perspectiva, lo que subyace a las habilidades mentalistas ya maduras no es algún tipo de teoría, sino la habilidad de proyectarnos imaginativamente a nosotros mismos en la perspectiva del otro, simulando la actividad mental del otro con la propia actividad mental (Gordon, 1992;

Goldman, 1989, 1992, 1993). La idea es que podemos entender los estados mentales de los otros en base a nuestros propios estados mentales (Gallese & Goldman, 1998), mediante un proceso de simulación en el que usamos los mecanismos mentales propios para calcular y predecir lo que ocurre en la mente de las otras personas (Goldman, 1989, 1992). De este modo, la capacidad de mentalizar es conducida por un proceso y no por conocimiento como en el caso de la *TT* (Goldman, 1992).

Existen diferentes versiones de *TS* que, a grandes rasgos, se pueden distinguir según la rutina de simulación se lleve a cabo explícita y concientemente (Gordon & Baker, 1994) o de modo implícito, inconciente, automático y subpersonal, basado en un procesamiento neural automático (Gallese & Goldman, 1998).

En la versión *off-line* (Goldman, 1992) el proceso de simulación se explica de la siguiente manera: A partir de lo que podemos observar en el comportamiento de las otras personas, consideramos cuál será la situación en la que se encuentra y generamos deseos y creencias hipotéticos. Sucesivamente, estos últimos alimentan (como *input*) nuestro sistema de toma de decisiones, que es utilizado como modelo de la persona en cuestión (Goldman, 1989, 1992; Stich & Nichols, 1992). El resultado de este proceso (*output*) se proyecta en la otra persona como explicación de su comportamiento, o como un modo de predecir su comportamiento. Así, el proceso de toma de decisiones no culmina como en el caso de ser utilizado para el comportamiento propio; la decisión no se ejecuta sino que se proyecta en la otra persona. En este sentido, en el sistema se utiliza *off the line*.

Apoyo empírico

En el ámbito de la psicología, los investigadores se han adherido al enfoque de *TT*, pero discrepan acerca del origen de tal teoría *Folk*. Se han propuesto al menos tres opciones respecto al mecanismo de adquisición. La teoría *Folk* puede ser resultado de: 1) la maduración de un mecanismo o módulo innato, donde las experiencias juegan el rol de meras disparadoras del proceso de maduración (Baron-Cohen, Leslie, & Frith, 1985; 1986; Baron-Cohen, 1995); 2) un aprendizaje por teorización (el enfoque del niño científico) (Perner, 1991; Gopnik & Meltzoff, 1997), ó 3) un aprendizaje por enseñanza y enculturación (Astington, 1996). En las últimas dos opciones, se asume la existencia de alguna base innata inicial que da el

puntapié al proceso de adquisición, pero éste es concebido propiamente como aprendizaje.

El enfoque *TT* ha encontrado apoyo empírico, principalmente, en los estudios de psicología del desarrollo. En general, se ha considerado como indicador de la adquisición de la competencia de atribuir estados mentales al momento en que los niños pueden atribuir *falsas creencias*, esto es, cuando son capaces de advertir que los demás pueden tener creencias erróneas sobre los estados del mundo y/o creencias que difieren de las que el propio niño posee. Como señalamos anteriormente, esta capacidad se adquiere alrededor de los cuatro años (Wimmer & Perner, 1983).

En relación al estudio de la *TdM* en el ámbito del desarrollo, es preciso mencionar la hipótesis ampliamente aceptada. El autismo, en alguna medida, implica un déficit de *TdM* y se define como un déficit en la socialización, la imaginación y la comunicación. Según el estudio de Baron-Cohen, Leslie y Frith (1985), tal déficit se debe a una falla en la atribución de estados mentales que se presenta disociada del retraso mental. Estos autores observaron que los niños con autismo tuvieron una peor *performance* en la tarea de falsa creencia, en contraste con el desempeño de los niños con desarrollo típico y con síndrome de *Down*, a pesar de contar con una mayor edad mental. Este estudio ha sugerido que el autismo podría constituir el déficit cognitivo específico de la *TdM*. De este modo, brinda apoyo a la hipótesis de la modularidad de la *TdM*, puesto que se espera que exista un déficit cognitivo específico para que un sistema pueda considerarse un sistema modular de dominio específico.

A su vez, el enfoque de la simulación también puede dar cuenta de algunos hallazgos en el desarrollo. Principalmente, es plausible que la capacidad de enrolarse en el juego de ficción esté ligada a la capacidad de poder alimentar nuestro sistema de toma de decisiones con creencias o deseos ficticios (Gordon, 1986; Goldman, 1989). Sin embargo, el enfoque de *TT* puede explicar este (Leslie, 1987) y otros fenómenos con plausibilidad. Por ello, se considera que mientras la *TS* es compatible con los datos de los estudios de desarrollo, la *TT* parece mejor fundada en éstos (para un desarrollo extenso de esta discusión, ver Stich & Nichols, 1992 y Botterill & Carruthers, 2003).

Con el desarrollo de las técnicas de neuroimagen en los '90, los estudios del cerebro originaron una nueva fuente de evidencia para la *TS*. El impulso de esta investigación se debió, principalmente, al hallazgo de las neuronas

espejo. Los datos sugieren que en los humanos existe un mecanismo neural similar al de las neuronas espejo descubiertas en los macacos (Fadiga et al., 1995). Este sistema, que equipara la observación y la ejecución de la acción, puede constituir la base de la capacidad de simular (Gallese & Goldman, 1998).

Por su parte, los representantes del enfoque modularista de la *TdM* han sugerido la posibilidad de que exista un mecanismo cerebral especializado en la atribución de estados mentales (Frith & Frith, 1999). En este caso, puesto que la *TdM* es una habilidad compleja, el mecanismo neural en que se apoye deberá consistir en una red o circuito. A su vez, se espera que tal base neural presente variaciones si las inferencias son llevadas a cabo por un niño o por un adulto y también, si la atribución se refiere a uno mismo o a otro agente.

Bases neurales de la Teoría de la Mente

La neurociencia social es el subcampo de la neurociencia cognitiva que investiga la fisiología que subyace a los procesos cognitivos que intervienen durante las interacciones sociales. Habilidades sociales básicas tales como la capacidad de atender, de reconocer y de recordar estímulos socialmente relevantes, constituyen el foco de estas investigaciones. De hecho, se ha definido la agenda de la neurociencia social como el esfuerzo por tratar de entender los fenómenos en términos de interacciones a partir de tres niveles diferentes de análisis: el nivel social, que se encarga de establecer los factores motivacionales y sociales que influyen en el comportamiento y la experiencia, el nivel cognitivo, que se ocupa de los mecanismos de procesamiento de información que dan lugar a los fenómenos del nivel social y el nivel neural, que se ocupa de analizar los mecanismos cerebrales que instancian los procesos de nivel cognitivo (Ochsner & Lieberman, 2001).

El interés de la neurociencia en la habilidad de mentalizar es reciente (Saxe, 2006) y las investigaciones y los datos son dispares. Se han aplicado una gran variedad de métodos para estudiar cómo los humanos hacen inferencias acerca de los estados mentales de los otros y, entre éstos, han cobrado especial relevancia los estudios de neuroimagen por resonancia magnética funcional (*IRMf*). También son frecuentes los estudios de neuroimagen utilizando la tomografía por emisión de positrones (*TEP*) y la tomografía por emisión de fotón único (*SPECT*). A su vez, se han

llevado a cabo estudios de estimulación magnética transcraneana (*EMT*) que permiten provocar efectos, temporales y reversibles, que emulan lesiones cerebrales. Con el método de electroencefalograma (*EEG*) se ha medido la actividad superficial del cerebro mientras se realizan tareas de mentalizar. Además, se han llevado a cabo estudios de lesión a pacientes con lesiones cerebrales acotadas a regiones vinculadas con la cognición social, y en pacientes con desórdenes neurodegenerativos específicos, como la demencia frontotemporal en su variante frontal (*fvFTD*), donde se ve afectada la corteza prefrontal. Es preciso mencionar que los estudios con animales también han sido importantes, especialmente aquéllos que registran la actividad de neuronas únicas.

Algunos investigadores han cuestionado el valor de los estudios de neuroimagen para proporcionar nuevas ideas acerca de la naturaleza de los procesos cognitivos en la cognición social (Coltheart, 2006; Willingham & Dunn, 2003). Sin embargo, otros se han encargado de destacar que los estudios de neuroimagen han re-direccionado el campo de estudio (Mitchell, 2008). En esta línea se sugiere que tales estudios no sólo proveen nuevos datos, sino que han encaminado la investigación hacia un conjunto de cuestiones sobre la organización de la cognición social que la neuroimagen funcional está especialmente equipada para abordar. Especialmente, este método empírico puede contribuir en la disquisición acerca de si hay procesos cognitivos dedicados especialmente al pensamiento social o no.

En principio, la evidencia ha sugerido que existen mecanismos especializados en la cognición social, en base a la observación de un patrón específico de activación durante tareas que requieren atribuir estados mentales. Fletcher et al. (1995) observaron una activación diferencial de la corteza prefrontal (*CPF*) en tareas de leer historias que requerían atribuir estados mentales, frente a aquellas historias que implicaban causalidad física. Blakemore, Winston y Frith (2004) observaron una variación en la activación prefrontal, según el estímulo fuera una atribución mentalista explícita, en contraposición a una actividad de lectura pasiva. Sin embargo, todavía se debe precisar cuáles son las partes constituyentes del sistema de la cognición social, y echar luz sobre ciertas cuestiones que han sugerido los datos. A saber, cómo es posible que la cognición social difiera y se solape a la vez, con formas de cognición no social (Mitchell, 2008).

Como hemos señalado, los métodos de neuroimagen permiten ir más allá de las correlaciones observadas en el comportamiento y testear, por

ejemplo, si el razonamiento acerca de las otras mentes depende de bases neurales especializadas o no. Sin embargo, es preciso señalar que los métodos de neuroimagen funcional proporcionan evidencia limitada, ya que sólo permiten detectar el nivel funcional de las regiones. Esto es, sólo describen qué funciones están asociadas a determinadas tareas psicológicas y no permiten determinar qué relaciones son necesarias.

Las investigaciones de *IRMf* y *TEP* han proporcionado observaciones consistentes acerca de las áreas que están asociadas a los procesos que ocurren durante la atribución de estados mentales. Estas áreas son la corteza prefrontal medial (*CPFm*), la juntura temporoparietal (*JTP*), el precúneo o la corteza paracingulada anterior (*CPA*), la amígdala, el surco temporal superior (*STS*) y los polos temporales. Como señalamos anteriormente, existe una discusión teórica de larga data acerca de los procesos que se ponen en marcha durante la atribución de estados mentales. Se han propuesto dos operaciones diferentes que hemos llamado más arriba, los enfoques de *TT* y de *TS*.

Un gran número de estudios de neuroimagen han hallado, consistentemente, que las mismas áreas cerebrales se activan cuando se experimenta un estado afectivo (dolor, miedo, asco) y cuando se observa a otra persona experimentarlo. Se ha observado que la corteza cingulada anterior se activa durante la experiencia subjetiva de dolor y al mirar videos de personas que sufren daño corporal, por ejemplo, lastimarse un dedo (Singer et al., 2004). Algunas regiones de la ínsula anterior muestran activación al oler un olor desagradable y al ver a alguien expresar asco inducido por un olor (Wicker et al., 2003). Estas observaciones se han extendido más allá de la experiencia afectiva. Se ha observado el solapamiento neural de áreas implicadas en llevar a cabo acciones dirigidas a una meta y al observar la misma acción en otros (Rizzolatti & Craighero, 2004). También se ha observado que al reportar las propias preferencias y opiniones se activan las mismas áreas que al inferir preferencias y opiniones en otros (Mitchell, McRae, & Banaji, 2006). Las observaciones sobre el solapamiento neural han llevado a postular la existencia de áreas cerebrales compartidas para el procesamiento de representaciones del yo y del otro. El descubrimiento del sistema espejo en el cerebro, para las emociones y acciones, ha posibilitado el estudio de la base fisiológica de procesos sociales fundamentales como la imitación, la resonancia emocional y la empatía, proveyendo datos experimentales confiables. La observación de

que las mismas áreas son activadas cuando tenemos ciertas experiencias y cuando observamos experiencias similares en otros, es compatible con la teoría de la simulación, que sostiene que los estados mentales de otras personas pueden entenderse sobre la base de los estados mentales propios (Gallese & Goldman, 1998).

Ahora bien, una cuestión distinta es sostener que la cognición social procede en su totalidad por vías de la simulación. Algunos investigadores consideran que los procesos a los que da lugar el sistema espejo sólo ponen en marcha el proceso de mentalización, y que no son suficientes para que la atribución de estados mentales se lleve a cabo. C. Frith y U. Frith (2006) sostienen que experimentar la misma emoción que el otro o imitar sus movimientos, implican sólo un primer paso en el proceso de mentalizar. Básicamente, esto no es suficiente porque no permite inferir la causa de las emociones, ni los objetivos o intenciones que subyacen a la acción.

En esta línea de investigación, se han estudiado los correlatos neurales de la *TdM* aplicando neuroimagen a sujetos voluntarios normales, y utilizando una amplia gama de paradigmas como leer historias, observar figuras moviéndose, ver películas de dibujos animados o jugar juegos interactivos. A pesar de sus diferencias, todas estas tareas tienen como requerimiento común tener que pensar acerca de los estados mentales de otras personas.

La evidencia, aunque no es concluyente, señala tres áreas cerebrales como intervinientes en la habilidad de mentalizar (Frith & Frith, 2003): la porción posterior del surco temporal superior (*pSTS*) junto al área adyacente denominada juntura temporoparietal (*JTP*), los polos temporales y la corteza prefrontal medial (*CPFm*). Si bien se sabe que existen conexiones anatómicas entre estas tres áreas, aún no se conocen los vínculos causales entre ellas. A su vez, los roles precisos de esas regiones cerebrales y los factores comunes que existen entre tales roles, recién comienzan a ser esbozados. A continuación se presentan algunas de las hipótesis acerca de la posible contribución de estas áreas a la *TdM*, basadas en lo que ya se conoce sobre sus roles funcionales. Cabe señalar que en estos estudios quedan excluidas ciertas áreas señaladas más arriba como base neural de la cognición social. Se trata de las que intervienen en los procesos de empatía afectiva (la amígdala y la corteza paracingulada anterior).

La evidencia sugiere que la *pSTS / JTP* parece estar implicada en el procesamiento de los movimientos del ojo y provee información acerca del

lugar al cual se está mirando (Pelphrey, Viola, & McCarthy, 2004). A su vez, esta región parece estar involucrada en la capacidad de representarse el mundo desde diferentes perspectivas (Aichhorn et al., 2006). C. Frith y U. Frith (2006) señalan que es probable que esta función pueda contribuir a la capacidad de mentalizar, ya que saber hacia dónde se dirige la mirada y qué es lo que puede estar dentro del campo de visión, permite identificar, por ejemplo, la causa del miedo que observamos en otra persona. En relación a esto, esta área puede ser la encargada de posibilitar que veamos el mundo como lo ven los otros, y esto nos permite advertir que nuestros congéneres pueden tener un conocimiento diferente al nuestro y que pueden poseer, incluso, creencias erróneas (o falsas creencias) acerca de los estados del mundo (Frith & Frith, 2006). A su vez, cierta evidencia sugiere que la *JTP* parece estar especialmente implicada en el procesamiento de las falsas creencias (Apperly et al., 2004).

Se ha reportado que el daño en los polos temporales se correlaciona con un déficit en la habilidad de utilizar el conocimiento que se adquiere por la experiencia; por esto se ha sugerido que esta región interviene en el procesamiento de tal habilidad (Fullnell, 2001). Básicamente, el conocimiento obtenido por experiencia se refiere a personas específicas (cómo se llaman, dónde viven) y a los cambios que, momento a momento se producen en las situaciones en que se encuentran las personas. Es decir, esta área se relaciona con cambios en el conocimiento, los sentimientos y las disposiciones, que afectan el comportamiento de las personas. A su vez, esta observación es consistente con el hallazgo de que los polos temporales son zonas de convergencia donde características simples de diferentes modalidades se reúnen para definir, por su conjunción, situaciones e individuos únicos (Damasio et al., 2004). Tal convergencia se relaciona con las modificaciones contextuales de los objetos. C. Frith y U. Frith (2006) sugieren que la contribución de los procesos instanciados en los polos temporales a la *TdM* consiste en la aplicación del conocimiento general y en la posibilidad de usar la información acerca de una persona particular, que se actualiza momento a momento (por ejemplo, conocer cuáles van siendo sus sentimientos en el momento de interacción).

A diferencia de los casos anteriores, el rol preciso de la *CPFm* permanece controversial. Existe evidencia de que el daño en la *CPFm* se correlaciona con fallas en tareas de *TdM* (Stuss, Gallup, & Alexander, 2001), aunque no se ha podido localizar críticamente el daño que implica tal déficit.

Si bien se ha reportado un caso de daño *CPFm* que no implica fallas en tareas de *TdM* (Bird et al., 2004), C. Frith y U. Frith (2006) sugieren que esta contradicción en los datos puede deberse a el hecho de que la *CPFm* intervenga en los procesos de *TdM* cuando pensamos acerca de los estados mentales, pero puede que no sea necesario su reclutamiento en las tareas de falsa creencia.

Se ha sugerido que la *CPFm* está implicada en la planificación del futuro y en la representación de los estados anticipados (Shallice, 1988). De este modo, es posible que la contribución de la *CPFm* a la habilidad de atribuir estados mentales esté relacionada con la anticipación de qué sentirá o pensará una persona, y con la predicción de su acción. Sin embargo, aún no es claro cómo se llevan a cabo estas predicciones.

Como señalamos anteriormente, se ha observado que la *CPFm* interviene al pensar acerca de los estados mentales de otros y de uno mismo (Frith & Frith, 2003), y tal observación es consistente con el simulacionismo. Sin embargo, parece que esto sólo tiene lugar cuando la persona cuyo comportamiento queremos predecir es similar a nosotros mismos y, en relación a esto, se ha reportado la segregación de funciones de las áreas de la corteza prefrontal en relación a diferentes aspectos del mentalizar. Existe evidencia de activación ventral cuando se hacen atribuciones mentalistas a agentes similares a uno mismo, mientras que se activa la porción dorsal de la *CPF* cuando se los juzga como diferentes (Mitchell, McRae, & Banaji, 2006). Amodio y Frith (2006) han sugerido que, a su vez, existe otro tipo de segregación funcional en la *CPF*. La corteza prefrontal ventral (u orbital) monitorea las emociones del yo y del otro, mientras que la región dorsal (postrostral), que incluye la sección cognitiva de la corteza cingulada anterior, monitorea las acciones del yo y del otro. En este caso, hay disgregación entre las atribuciones emocionales y aquéllas acerca de acciones. Una tercera porción (la anterior rostral), ubicada entre estas dos áreas, se encarga de procesar estímulos de los agentes similares y de la intencionalidad comunicativa. El estudio de Hynes, Baird y Grafton (2006) reportan activación prefrontal orbital en tareas que exigen una toma de perspectiva emocional, mientras que se observó una activación dorsal en la toma de perspectiva en general. Estos dos últimos estudios coinciden en la segregación dorsal en el aspecto emocional de la atribución.

Hasta aquí hemos mencionado las hipótesis acerca de la contribución funcional a la *TdM* de las áreas cerebrales implicada según los estudios

de neuroimagen. Ahora bien, para algunos investigadores la búsqueda del sustrato neural de la habilidad de atribuir estados mentales implica que la *TdM* es una habilidad modular. Si esto es así, también será preciso llevar a cabo otros estudios íntimamente ligados con esta hipótesis, por ejemplo, estudios de genética molecular que intenten determinar cuáles son los genes que codifican tal habilidad. En esta línea, otros estudios que cobran importancia son los trabajos sobre la contribución de los andrógenos en la *TdM* (Baron-Cohen, 2003; Mazzola et al., 2007), puesto que en las áreas cerebrales que parecen ser reclutadas para la *TdM* hay abundantes receptores de andrógenos. Esta línea de investigación es conocida como la hipótesis del cerebro masculino.

Problemas en relación a la búsqueda de bases neurales de la TdM

En el apartado anterior se hizo referencia al problema general de la búsqueda del sustrato neural de las funciones cognitivas. Se trata de uno de los métodos más utilizados que sirve para determinar el perfil funcional de determinadas regiones cerebrales, funciones que, en última instancia, se correlacionan con funciones psicológicas. Sin embargo, es preciso señalar que este método sólo describe qué funciones están asociadas con tareas psicológicas pero no describe si las relaciones pueden ser necesarias o no. Para determinar si una función cognitiva es un componente necesario de otra función, los investigadores deben recurrir a otras metodologías.

A su vez, se plantean problemas específicos en la búsqueda de las bases neurales de *TdM*. Podemos señalar tres dificultades. En primer lugar, bajo el término *TdM* se agrupa un amplio abanico de habilidades que, por ejemplo, puede implicar capacidades tales como la toma de perspectiva emocional y cognitiva (Hynes, Baird, & Grafton, 2006). En segundo lugar, es preciso averiguar cuáles son las funciones cerebrales de cada zona. Para ello, las expectativas se han puesto en dos técnicas: en los estudios que registran la actividad de neuronas individuales (estudios de neurona única), que sólo se pueden realizar en animales (Frith & Frith, 1999), y recientemente se ha recurrido a la combinación de neuroimagen con pruebas neuropsicológicas (Frith & Frith, 2003). En tercer lugar, en la búsqueda del sustrato neural de la atribución de estados mentales se han hallado conjuntos diferentes de áreas cerebrales, según se busquen las bases neurales de la *TdM* o de la simulación mental.

En consecuencia, el campo de estudio de las bases neurales de la cognición social se ha polarizado. Por un lado, algunas investigaciones hacen foco en el sistema de las neuronas espejo de la corteza premotora (Gallese & Goldman, 1998), así como en áreas parietales y temporales para la atribución de intención. Mecanismos similares a éstos se han postulado para la atribución emocional (Gallese et al., 1996; Gallese, Keysers, & Rizzolatti, 2004). En línea con esta investigación surge la hipótesis de los circuitos compartidos (*CC*) de la que nos ocuparemos más abajo. Por otro lado, el enfoque modularista de la *TdM* ha encontrado evidencia de la activación de estructuras medias del cerebro (corteza prefrontal medial) en la atribución de estados mentales a otros (Gusnard et al., 2001; Frith & Frith, 2003).

Recientemente, ha aparecido un intento de conciliar tal contradicción en la literatura. Keysers y Gazzola (2007) sostienen que, ante la polarización del campo, hay que intentar establecer por qué los investigadores encuentran dos conjuntos diferentes de áreas cerebrales como las más prominentes y de qué modo los mismos interactúan. Para estos autores, la contradicción se debe a la utilización de diferentes ejemplos de cognición social en los estudios llevados adelante por cada una de las corrientes. Estos investigadores señalan que los fenómenos estudiados en cada campo difieren dramáticamente en la cantidad de pensamientos explícitos que acompañan a las tareas utilizadas. En los estudios característicos de *TdM*, como la Tarea Falsa Creencia (Baron-Cohen, Leslie, & Frith, 1985), se exige a los sujetos que infieran concientemente los estados mentales de otras personas o personajes. En cambio, en los estudios de neuroimagen para los circuitos compartidos (*CC*), los sujetos simplemente observan videos de acciones, emociones o sensaciones sin que se les pida que reflexionen sobre el significado del estímulo, o sobre las creencias y pensamientos de los actores.

De este modo, podría ser que el hecho de que se hayan activado conjuntos diferentes de áreas cerebrales se deba al estudio de fenómenos distintos respecto a la cantidad de pensamientos explícitos requeridos por la tarea (Keysers & Gazzola, 2006, 2007). Para estos investigadores, el planteo de una dicotomía entre procesos de simulación y de teoría, que implica que uno u otro deben explicar todas las condiciones de las cogniciones sociales (Gallese & Goldman, 1998; Saxe, 2005), ha obviado esta cuestión.

Keysers y Gazzola (2006, 2007) presentan un modelo teórico del modo en que pueden interactuar ambos conjuntos de áreas cerebrales. En él se propone la existencia de circuitos neurales compartidos (*CC*) para las acciones, sensaciones y emociones, a partir de evidencia que sugiere que las mismas áreas cerebrales están implicadas en la perspectiva de primera persona (yo siento, yo hago) y en la perspectiva de tercera persona (ella siente, ella hace). En los humanos, el *CC* para la acción está compuesto por regiones de los lóbulos temporal, parietal y frontal. En el lóbulo temporal se reclutan, específicamente, el *STSp* y el giro medio temporal (*GMT*). El lóbulo parietal inferior (*LPI*) y ciertas regiones del lóbulo frontal, a saber, las Áreas de Brodman (*AB*) 44/6. Las cortezas somatosensoriales primaria y secundaria constituyen los *CC* para el tacto. Los *CC* para las emociones están compuestos, en el caso del asco, por la ínsula anterior y es probable que la amígdala constituya una de las regiones que forme parte de un *CC* para el miedo (Keysers & Gazzola, 2006).

En este modelo, los *CC* tienen como función transformar lo que vemos que otras personas hacen o sienten en algo muy conocido para nosotros: lo que nosotros sentimos o hacemos. Como resultado de esto, los *CC* pueden proveernos un entendimiento simple o una captación intuitiva de la vida interna de los otros. Además, tienen vital importancia en la imitación y, probablemente, en el aprendizaje del lenguaje hablado (Keysers & Gazzola, 2006).

Estos investigadores retoman cierta evidencia proveniente de la investigación de las bases neurales de la *TdM*, como fundamento para asociar a los *CC* con tales bases neurales. Específicamente, se refieren a los reportes de estudios de *RMf* y *TEP* para tareas de *TdM* (Frith & Frith, 2003) de los que se siguen las siguientes conclusiones: (a) la corteza prefrontal medial (*CPFm*) tiene la capacidad de asociar estados de nuestro cuerpo y de nuestra mente, puesto que la evidencia sugiere que ésta se activa durante la introspección y reporte de las emociones propias (Gusnard et al., 2001); (b) la *CPFm* se activa en tareas que requieren más mentalización; (c) por su estructura anatómica, la *CPFm* recibe *inputs* indirectos de nuestro propio cuerpo (representaciones motoras, somatosensoriales y viscerales) que pueden permitir una segunda representación de un estado corporal (Frith & Frith, 2003). Específicamente, se puede diferenciar un primer nivel de representación (lo que nos ocurre: asco al vomitar) y un segundo nivel de representación (el entendimiento conciente de lo que nos ocurre: reflexión y respuesta en palabras de cómo nos sentimos al vomitar).

En este punto, la simulación se liga con los conceptos de *TdM* (Keysers & Gazzola, 2006). La *CPFm* genera representaciones secundarias (concientes, palabras) de los estados del cuerpo, a partir de representaciones primarias (representaciones somatosensoriales, motoras y viscerales que llegan indirectamente). La idea de Keysers y Gazzola es que, a través de los circuitos compartidos (*CC*), la acción, la sensación y la emoción del otro se transforman en el lenguaje corporal de nuestras acciones, emociones y sensaciones. Esto da lugar a una representación primaria, que puede generar un compartir implícito y un entendimiento de los estados de los otros. Si hay que pensar explícitamente qué ocurre en la mente del otro, uno tendrá que generar una representación secundaria más conciente, explícita y cognitiva. Puesto que este estado ya ha sido traducido al lenguaje de los propios estados por los *CC*, Keysers y Gazzola sugieren que tal tarea no se diferenciará de la de reflexionar sobre los propios estados y, por esto, se activan las mismas áreas que al pensar acerca de uno mismo.

En resumen, para Keysers y Gazzola (2006), los circuitos compartidos (*CC*) actúan como traductores de los estados de otros en representaciones primarias de nuestros estados. Estos autores consideran que, generalmente, el procesamiento de la información social termina en este punto. Compartimos la tristeza de nuestros congéneres sin pensar más en ello, y estos son los casos en los que han pensado los proponentes de la simulación como mecanismo subyacente a las atribuciones mentalistas. En otras ocasiones, reflexionamos sobre los estados mentales de los otros, utilizando el mismo mecanismo por medio del cual reflexionamos acerca de estados propios (representación secundaria). Tales reflexiones proveen un entendimiento más cognitivo, elaborado y diferenciado de los otros individuos. De estos casos se han ocupado los investigadores de la *TdM*.

Discusión y conclusiones

Un punto central en el análisis de Keysers y Gazzola (2006, 2007) es que los estudios de bases neurales encuentran conjuntos de áreas cerebrales diferentes, porque los investigadores están estudiando diferentes ejemplos de cognición social. Así, unas tareas de cognición social implican más pensamientos explícitos que las otras.

Aceptado este planteo, podemos introducir una división de tipos de estados psicológicos, tradicional en filosofía, que nos permite notar otra diferencia entre los ejemplos planteados. Esta clasificación distingue

entre sensaciones y actitudes proposicionales. Las sensaciones refieren a percepciones, es decir, a la información que proviene del mundo externo, y a propiocepciones, es decir, a información que proviene del propio cuerpo. Las actitudes proposicionales son aquellos estados mentales que tienen una estructura del tipo *S cree / desea que P*, donde *P* puede ser verdadero o falso, y creer o desear constituye una actitud hacia esa proposición.

Ahora bien, en el modelo de los *CC*, las áreas neurales propuestas como base neural de la simulación se encargan de aportar las representaciones primarias somatosensoriales, motoras y viscerales. Éstas, según la clasificación introducida, pueden entenderse como sensaciones. En cambio, las áreas superiores (*CPFm*) que procesan las representaciones secundarias y llevan a cabo los procesos inferenciales que suponen alguna teoría (o reflexionar, como señalan los autores), se encargarían de las actitudes proposicionales (intenciones, deseos, creencias). De esta manera, podríamos decir que existe una división de tareas entre los sistemas cerebrales. Las áreas cerebrales que subyacen a la simulación procesan sensaciones (en el sentido introducido), mientras que las áreas cerebrales superiores que subyacen a la *TdM* procesan las actitudes proposicionales.

Como señalamos anteriormente, en filosofía de la mente se denominan actitudes proposicionales a cierto tipo de estados psicológicos, que se caracterizan por estar conformados por (a) una actitud (una creencia, un deseo, un presentimiento, o cualquier otra cosa) y por (b) una proposición a la que se refiere tal actitud. Así, si decimos que Juan cree que llueve, la actitud es ‘creer’ y la proposición es ‘que llueve’. Las proposiciones pueden ser verdaderas o falsas, en virtud de su relación con el mundo. Si en un día de lluvia Juan cree que llueve, entonces su creencia es verdadera. Sin embargo, el rasgo destacado de las actitudes proposicionales es que es posible para uno poseer tales actitudes proposicionales respecto de cosas que no son el caso. De este modo, uno puede creer que llueve, cuando de hecho no llueve (Para un tratamiento extenso de la cuestión ver Fodor, 1987, capítulo 1). En resumen, las actitudes proposicionales prototípicas son las creencias, los deseos y las intenciones. Éstas se encuentran en la base de una psicología del sentido común, que describe, explica y predice la conducta propia y ajena en base a reglas del tipo: “si alguien quiere algo, y cree que hay algo que puede hacer para obtenerlo, entonces, generalmente, llevará esto a cabo, *ceteris paribus*”.

En base a la clasificación mencionada anteriormente, podemos decir que la interacción entre las áreas propuestas para la simulación y las áreas propuestas para *TdM* implica una división de tareas respecto a los estados psicológicos que cada conjunto de áreas procesa. Unas se ocupan de sensaciones mientras que las otras se encargan de las actitudes proposicionales. Las segundas implican alguna teoría, las primeras no.

Este trabajo no pretende resolver la cuestión de si este intento de conciliar los hallazgos de las bases neurales resuelve el problema de Teoría de Teoría (*TT*) vs. Teoría de Simulación (*TS*). Sin embargo, cabe señalar una dificultad de esta propuesta. Los teóricos de la *TS* consideran que el mecanismo de simulación también puede dar cuenta de las actitudes proposicionales. En este caso, de la interacción propuesta parece desprenderse una división de tareas entre simulación y teoría, que implica que la simulación cede la capacidad para dar cuenta de las actitudes proposicionales a la teoría.

Más allá de esta y de otras dificultades que puedan surgir del modelo presentado, resulta interesante la posibilidad de un enfoque mixto de *TT* y *ST*. Si bien es claro que existe un entendimiento implícito entre las personas, pensar que tal procesamiento está en la base de toda cognición social es, al menos, problemático. En primer lugar, no es cierto que el proceso implícito sea un proceso tan simple e inmediato (como pueden sugerir los circuitos compartidos) porque es preciso distinguir si esta representación, que en principio es compartida, se refiere al yo o al otro. Esto implica poder explicar cuáles son los mecanismos que llevan a cabo esta distinción en cualquier procesamiento empático (Lieberman, 2007) y, particularmente, en la propuesta de la cognición social basada en los *CC* (Keysers & Gazzola, 2006).

En segundo lugar, en ciertos casos, el éxito de las interacciones sociales precisa algo más que una cognición social basada en el entendimiento intuitivo del otro. Los *insights* que provee este tipo de entendimiento implican inherentemente asumir que somos iguales, pero la cognición social puede ser más sofisticada e ir más allá de este caso. Por ejemplo, puede ser que cierta acción que a uno le haga feliz pueda entristecer a otro. Keysers y Gazzola (2007) advierten esta dificultad y sugieren que los procesos cognitivos superiores pueden tener un rol crítico en la tarea de atribuir estados mentales sofisticados.

La discusión acerca de cuáles son los mecanismos y recursos que subyacen a la competencia de atribuir estados mentales ha conducido a

propuestas teóricas que intentan hacer interactuar los mecanismos de simulación y de teoría. En este trabajo se ha comentado una propuesta sustentada en evidencia proveniente de la búsqueda de las bases neurales de la cognición social. Se podría decir que este enfoque concibe a la teoría como un enriquecimiento de la operación de simulación, cuando la situación social requiere que el entendimiento no sea meramente intuitivo (Keysers & Gazzola, 2006, 2007).

Para tratar de entender el modo en que estos mecanismos, simulación y teoría, interactúan, en la propuesta se introdujo una clasificación de estados psicológicos que distingue entre sensaciones (percepciones y propiocepciones) y actitudes proposicionales. A la luz de ésta, se puede advertir una división de tareas en la propuesta de los CC. En tal modelo teórico, por medio de la simulación se procesarían las sensaciones, mientras que por medio de la teoría se procesarían las actitudes proposicionales. En relación a esto, se ha señalado que la propuesta de los CC no se compromete con la capacidad de la simulación para dar cuenta de las actitudes proposicionales.

■ Referencias

- Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladumer, G. (2006). Do visual perspective need theory of mind? *Neuroimage*, 30, 1059-1068.
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268-277.
- Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to Theory of Mind: Neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience*, 16, 1773-1784.
- Astington, J. (1996). What is theoretical about the child's theory of mind? A Vigotskian view of its development. En P. Carruthers & P.K. Smith (Eds.) *Theories of theories of mind* (pp. 184-199). Cambridge: Cambridge University Press.
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and TOM*. Cambridge MA: MIT Press.
- Baron-Cohen, S. (2003). *The essential difference: Male and female brains*

- and the truth about autism. Cambridge MA: Basic Books.
- Baron-Cohen, S., Leslie, A., & Frith, U. (1984). Mechanical, behavioural and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology*, 4, 113-125.
- Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a theory of mind? *Cognition*, 21, 37-46.
- Bennet, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences*, 1, 557-560.
- Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on 'Theory of Mind' and cognition. *Brain*, 127, 914-928.
- Blakemore, S. J., Winston, J., & Frith, U. (2004). Social cognitive neuroscience: where are we heading? *Trends in Cognitive Sciences*, 8, 216-222.
- Botterill, G., & Carruthers, P. (2003). *The philosophy of psychology*. Cambridge: Cambridge University Press.
- Coltheart, M. (2006). What has functional neuroimaging told us about the mind (so far)? *Cortex*, 42, 323-331.
- Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition*, 92, 179-229.
- Dennett, D. (1978). Belief about beliefs. *Behavioral and Brain Sciences*, 1, 568-570.
- Dennett, D. (1987). *The intentional Stance*. Cambridge MA: MIT Press.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, 73, 2608-2611.
- Flavell, J. H., Flavell, E. R., & Green, F. L. (1987). Young children's knowledge about the apparent-real and pretend-real distinctions. *Developmental Psychology*, 23, 99-103.
- Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., & Frith, C. D. (1995). Other minds in the brain: a functional imaging study of 'theory of mind' in story comprehension. *Cognition*, 57, 109-128.
- Fodor, J. (1987). *Psychosemantics*. Cambridge: MIT Press.
- Frith, C. D., & Frith, U. (1999). Interacting minds-A biological basis. *Science*, 286, 1692-1695.

- Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50, 531-534.
- Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. En C. D. Frith & D. Wolpert (Eds.), *The neuroscience of social interaction* (pp.45-75). Oxford: Oxford University Press.
- Fullnell, E. (2001). Evidence for scripts in semantic dementia: Implications for theories of semantic memory. *Cognitive Neuropsychology*, 18, 323-341.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in premotor cortex. *Brain*, 119, 593-609.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12), 493-501.
- Gallese, V., Keysers, C., & Rizzolatti, G. (2004). An unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8(9), 396-403.
- Goldman, A. (1989). Interpretation psychologized. *Mind & Language*, 4, 161-185.
- Goldman, A. (1992). In defense of the simulation theory. *Mind & Language*, 7, 104-119.
- Goldman, A. (1993). The psychology of Fol. *Psychology. Behavioural and Brain Science*, 16, 15-28.
- Goldman, A. (2007). *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*. NY: Oxford University Press.
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.
- Gordon, R. (1986). Folk Psychology as Simulation. *Mind & Language*, 1, 158-171
- Gordon, R. (1992). The simulation theory: Objections and misconceptions. *Mind & Language*, 7, 11-34.
- Gordon, R. M., & Baker, J. A. (1994). Autism and the theory of mind debate. En G. Graham & G. L. Stephens (Eds.), *Philosophical psychopathology* (pp. 163-181). Cambridge, MA: MIT Press.
- Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *PNAS*, 98, 4259-4264.
- Harman, G. (1978). Studying the chimpanzee's theory of mind. *Behavioral*

- and Brain Sciences, 1, 515-526.
- Heal, J. (1986). Replication and functionalism. En J. Butterfeld (Ed.), *Language, mind, and logic* (pp. 135-50). Cambridge: Cambridge University Press.
- Hynes, C., Baird, A. A., & Grafton, S. T. (2006). Differential role of the orbitofrontal lobe in emotional versus cognitive perspective taking. *Neuropsychologia*, 44, 374-383.
- Keysers, C., & Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Progress in Brain Research*, 156, 379-401.
- Keysers, C., & Gazzola, V. (2007). Integrating simulation and theory of mind: from self to social cognition. *Trends in Cognitive Sciences*, 11(5), 194-196.
- Leslie, A. (1987). Pretence and representation: the origins of 'theory of mind'. *Psychological Review*, 94, 412-426.
- Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annual Review of Psychology*, 58, 259-289.
- Mazzola, F., Siegal, A., MacAskill, A., Corden, B. Lawrwnce, K., & Skuse, D. H. (2007). Eye tracking and fear recognition deficits in Turners syndrome. En R. Saxe & S. Baron-Cohen (Eds.), *A special Issue of Social Neuroscience: Theory of Mind* (pp.259-269). NY: Psychology Press.
- Mitchell, J. P. (2008). Contributions of functional neuroimaging to the study of social cognition. *Current Directions in Psychological Science*, 17(2), 142-146.
- Mitchell, J. P., McRae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, 50, 655-663.
- Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience, *American Psychologist*, 56, 717-734.
- Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When Strangers Pass: Processing of Mutual and Averted Social Gaze in the Superior Temporal Sulcus. *Psychological Science*, 15, 598-603.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge MA: MIT Press/Bradford Books.
- Perner, J., Leekman, S. R., & Wimmer, H. (1987). Three year olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5, 125-137.

- Povinelli, D. J., & Bering, J. M. (2002). The mentality of apes revisited. *Current Directions in Psychological Science*, 11, 115-119.
- Pylyshyn, Z. W. (1978). When is attribution of beliefs justified? *Behavioral and Brain Sciences*, 1, 592-593.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1, 515-516.
- Rizzolatti, G., & Craighero, L. (2004). The Mirror-neuron System. *Annual Review Neuroscience*, 27, 169-92.
- Saxe, R., & Baron-Cohen, S. (2007). *The neuroscience of theory of mind*. NY: Psychology Press.
- Saxe, R. (2005). Against simulation: the argument of error. *Trends in Cognitive Sciences*, 9(4), 174-179.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, 16, 235-239.
- Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C.D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303, 1157-1162.
- Stich, S., & Nichols, S. (1992). Folk psychology: Simulation or Tacit theory. *Mind & Language*, 7(1-2), 35-71.
- Stuss, D. T., Gallup, G. G., & Alexander, M. P. (2001). The frontal lobes are necessary for 'Theory of Mind'. *Brain*, 124, 279-286.
- Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states- the question is which ones and to what extent. *Trends in Cognitive Sciences*, 7, 153-156.
- Wicker, B., Keysers, C., Plilly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, 40, 655-664.
- Willingham, D. T., & Dunn, E. W. (2003). What neuroimaging and brain localization can do, cannot do and should not do for social psychology. *Journal of Personality & Social Psychology*, 85, 662-671.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs. *Cognition*, 13, 103-128.