

EL TÚ Y EL YO EN LA NUEVA CIENCIA COGNITIVA

FERNANDO SOLER TOSCANO

Profesor de Filosofía, Universidad de Sevilla



¿Cómo podríamos conocer mejor los fenómenos mentales que nos caracterizan como humanos? Esta es una de las preguntas que se hace la ciencia cognitiva (E. García, 2007). En una primera etapa, desde los años 40 a los 80, más o menos, la ciencia cognitiva estuvo muy marcada por los programas de inteligencia artificial. Esta etapa fue importante, porque se cuestionaron los supuestos teóricos del conductismo, que no permitían que la psicología científica hablara de contenidos y procesos mentales. Además, hubo un importante trabajo interdisciplinar entre neurocientíficos, informáticos, psicólogos, lingüistas, antropólogos y filósofos. El paradigma durante esta primera etapa del cognitivismo se conoce como *funcionalismo* y su tesis principal es la metáfora del computador: los procesos mentales son estados internos que funcionan como el software que rige el comportamiento del cerebro. Se trata de encontrar el programa mental para implementarlo en una máquina. Este software mental se creía independiente de su soporte físico. Esto cambia en la segunda

etapa del cognitivismo, a partir de los años 80. Ahora el soporte físico (cerebro) pasa a verse como fundamental. La explicación naturalista se hace dominante, tomando argumentos de la física, la química y la biología. Uno de los autores que abren esta segunda etapa es Searle. Es conocido por su crítica al funcionalismo y proponer que el estudio de la mente no puede ignorar los problemas fundamentales, siendo el principal de ellos la conciencia. Propone el *Principio de Suficiencia Neurofisiológica*, que sostiene que: «para cualquier fenómeno mental hay condiciones causalmente suficientes en el cerebro». Así pues, estudiemos el cerebro y conoceremos mejor nuestra naturaleza. Pero aún no conocemos lo suficiente de nuestro cerebro para responder a las preguntas más interesantes sobre nuestra naturaleza. Así que, mientras llega ese día, si es que llega, tenemos que aprovechar otras vías de acceso a nuestra vida mental.

Algo filosóficamente sorprendente es el éxito de la llamada *psicología popular*. Se trata del tipo de discurso que solemos utilizar para explicar o predecir

la conducta de otras personas atribuyéndoles contenidos mentales, como pueden ser creencias o deseos. Por ejemplo, puedo predecir que Fulano no irá a la fiesta porque está enfadado con Mengano. Utilizamos este tipo de atribuciones y es tan sorprendente que funcionen (luego Fulano no va a la fiesta), como que todos solemos hacer el mismo tipo de atribuciones siguiendo, aparentemente, las mismas reglas. De hecho, podemos comprender textos literarios escritos hace cientos o miles de años porque la psicología popular no ha cambiado. La psicología científica sí ha cambiado, así como otros conocimientos populares, pero la psicología popular no.

En 1978, dos primatólogos, Premack y Woodruff se preguntan si los chimpancés tienen una capacidad semejante a nuestra psicología popular. Por entonces estaba muy de moda el estudio de las capacidades mentales de los chimpancés, en lo referente al manejo de herramientas, pero sobre todo a la interacción social. La idea es que el mayor desafío cognitivo de un chimpancé no es de tipo técnico sino relativo a las relaciones sociales dentro de su manada. Premack y Woodruff se preguntan si los chimpancés tienen una «teoría de la mente». Acuñan este nombre para referirse a la capacidad que tenemos las personas de distinguir contenidos mentales diferentes a los propios que en cada momento tenemos. Yo tengo una teoría de la mente si puedo distinguir entre mis contenidos actuales y lo que estaba pensando hace un rato, o lo que otra persona está pensando ahora. Es una capacidad muy demandante porque supone mantener como dos «líneas mentales» separadas: mis contenidos actuales y los otros que estoy considerando.

A lo largo de los años 80 se diseñan varios experimentos para averiguar si los chimpancés tienen una teoría de la mente, pero lo interesante es cuando estos mismos experimentos se adaptan para aplicarse a humanos. En las últimas décadas se han ido descubriendo diversas capacidades, hasta donde se sabe exclusivamente humanas, que pueden explicarse por la teoría de la mente. Lo primero que se investigó es la detección de la falsa creencia. Por la importancia que ha tenido vamos a comentar los experimentos más usados:

■ **Experimento del juguete desplazado.** Intervienen dos personajes. El primero de ellos guarda una pelota en una caja y luego se marcha a dar un paseo. Entonces entra en juego el segundo personaje que cambia de sitio la pelota. Cuando el primero vuelve de dar su paseo, ¿dónde buscará la pelota? Antes de los 4 años, los niños suelen responder que la buscará donde la puso el segundo personaje, porque es ahí donde está. No distinguen entre

la realidad y la creencia del personaje que estuvo ausente.

■ **Experimento de la caja engañosa.** Se enseña al niño un bote de caramelos con el que está familiarizado y se le pregunta qué hay dentro. Responde que hay caramelos. Entonces lo abre y descubre que hay lápices de colores. Después de un minuto de jugar con los lápices se le pregunta qué es lo que dijo que había en la caja antes de abrirla. El niño, antes de los 4 años responde que dijo que había lápices. Incluso cuando se le pregunta qué es lo que respondería a la pregunta un amigo suyo también dice que lápices.

Sorprende que hasta los 4 años un niño no sea capaz de superar estas pruebas. Si a los 3 años tiene ampliamente desarrollado el lenguaje, y es capaz de comprender preguntas sobre el pasado, ¿cómo se olvida de la creencia que tuvo hace un minuto? ¿Por qué no reconoce la falsa creencia ajena? Si bien hasta los 4 años no es capaz el niño de superar estas pruebas, se han ido descubriendo algunas capacidades que requieren la teoría de la mente y aparecen mucho antes. Uno de ellos es el *gesto de apuntar*, que aparece alrededor de los 11 meses. Se requiere la doble línea mental para situarnos ante una persona con contenidos mentales distintos de los nuestros, cuya atención queremos orientar hacia algo. También merece la pena destacar el *juego simbólico*, alrededor de los 18 meses. Es también la doble línea mental que permite distinguir entre el cepillo que tiene el niño en la mano y el coche con que simula estar jugando. Parece ser algo exclusivamente humano: atender a un contenido ficticio que se sabe que es ficticio.

¿Qué nos enseña la teoría de la mente de nosotros mismos? La doble línea mental que nos ofrece la teoría de la mente nos da la posibilidad de ponernos en el pellejo de otra persona. Esto nos permite situarnos en lo que se conoce como la *perspectiva de segunda persona*. Vamos a tratar de explicarlo siguiendo las ideas de A. Gomila (2002). El modo tradicional de acercarnos al estudio de la mente, desde el «pienso, luego existo» de Descartes hasta muchas corrientes en psicología cognitiva, es la perspectiva de primera persona: la conciencia, la autoconciencia. La perspectiva que sólo acepta el ámbito íntimo de la subjetividad como el de la certeza. Esto nos lleva a muchos problemas, entre ellos la dificultad de justificar la existencia de otras mentes y el solipsismo. Por otra parte, muchos acercamientos contemporáneos a la mente se sitúan en lo que se conoce como perspectiva de tercera persona. Enfoques tan diversos como el del conductismo, la neurociencia o el funcionalismo se sitúan ante la men-

te desde fuera, como el científico que se coloca ante su objeto de estudio. Desde esa posición, aparecen muchos problemas para dar cuentas de la subjetividad, por ejemplo.

Dice Antoni Gomila (2002) que «la idea básica de la perspectiva de segunda persona consiste en que en ciertas situaciones de interacción personal la atribución mental no sólo es recíproca –esto es, que ambos participantes en la interacción interpretan al otro y son interpretados por el otro–, sino que además se dan cuenta de esa mutua atribución, lo que condiciona, cuando no determina, el contenido de la atribución». Nos dice, además, que es la perspectiva que aparece primero en el desarrollo ontogenético. Antes de que el niño pueda situarse en la perspectiva de primera persona y descubrir su yo autoconsciente, es capaz de interactuar con otras personas. *La captación de la interioridad ajena y la atribución de contenidos mentales sería una capacidad exclusivamente humana.*

Durante los años 90 hubo varias teorías que trataron de explicar por qué nuestra mente (y nuestro cerebro) es capaz de atribuir contenidos mentales ajenos (E. García, 2007). Sin entrar en el debate, sólo indicaremos que algo decisivo fue el descubrimiento en 1996 por parte de Rizzolatti y otros neurobiólogos italianos de lo que se ha conocido como *neuronas espejo*. Son unas neuronas especiales que se activan cuando un individuo realiza una acción pero también cuando percibe esa misma acción realizada por parte de un congénere. Dichas neuronas nos permitirían asociar la acción observada en otra persona con la intención que suele producir en nosotros la misma acción. Hacen posible imitar movimientos como abrir la boca, desde las pocas horas del nacimiento. El descubrimiento de las neuronas espejo produjo una gran euforia e inicialmente se las hizo responsables incluso de cosas como el altruismo o la empatía. Pero para explicar la teoría de la mente hacen falta muchas más cosas que las neuronas espejo, porque éstas también están en primates que no tienen nuestras capacidades de atribución de contenidos mentales. De hecho, es notable que los chimpancés no hayan desarrollado el gesto de apuntar. Disponen de otros gestos como el de mendigar. Son capaces de llamar la atención de un congénere y de conducirlo de la mano a algún lugar. Pueden atribuir percepciones ajenas, pero su interpretación del cuerpo ajeno (el que está apuntando) no les lleva a la doble línea característica de nuestra teoría de la mente (T. Bejarano, 2003).

Estamos debatiendo qué nos distingue de los animales y de las máquinas. Una de las facultades humanas que tradicionalmente se han opuesto al determinismo mecánico o biológico es la libertad.

¿Puede la teoría de la mente decirnos algo al respecto? El intento más notable que conocemos es el realizado por T. Bejarano (2005), que vamos a presentar a continuación. Como ella, nos centramos en la libertad moral. En sentido fuerte, elegir entre el bien y el mal. Hoy día, quienes más defienden el naturalismo pueden aceptar diversos tipos de altruismo que no implican libertad moral. Por ejemplo el altruismo genético, el esfuerzo de los animales por sacar adelante a sus crías aun a costa de su propia vida. Pero esto no requiere para nada libertad moral, es un mecanismo tan poco libre como la búsqueda de comida. Tampoco la requiere el altruismo recíproco, «hoy por ti y mañana por mí», que es una conducta determinada por los propios intereses. ¿Cómo podría darse un altruismo auténtico, una auténtica libertad sin recurrir al dualismo? Por supuesto que la respuesta no la tenemos, pero sí podemos darle vueltas al problema y ver si lo que hemos presentado de la teoría de la mente podría abrir alguna vía interesante.

Hasta donde sabemos hoy, la captación de los estados mentales ajenos es una capacidad exclusivamente humana. El chimpancé tiene una capacidad que es un precedente remoto de nuestra teoría de la mente. Un chimpancé es capaz de darse cuenta de lo que un congénere, a partir de su ubicación y postura, puede o no puede ver. Es decir, puede atribuir percepciones. Pero esa interioridad que atribuye al congénere no se la puede atribuir cuando éste le mira a él. La percepción de los estados mentales ajenos de una persona que se dirige a mí, me mira y puede querer algo de mí, parece ser una característica específicamente humana. ¿Cómo eso podría producir una conducta de ayuda realmente libre? Pensemos en una ayuda que sea costosa y que no contribuya a nuestros propios intereses, ni a la mejora de nuestra propia imagen. Un altruismo realmente libre que no pueda explicarse por nuestro interés biológico.

Volvamos a la percepción de una mente ajena. El rostro de una persona en quien descubrimos una necesidad. Esa necesidad no tiene por qué movernos a la acción, porque la segunda línea mental, por la que la percibimos, no está conectada a nuestras fuerzas biológicas tanto como lo que yo vivo y me afecta. Pero la percepción nos llega como algo objetivo: es una persona como yo que tiene un problema, y puedo imaginar qué significa tener ese problema. Esto no basta por sí solo para producir la conducta de ayuda. Si bastase, tampoco garantizaría nuestra libertad. Pero tenemos un recurso importantísimo que puede jugar un papel clave: el lenguaje.

El lenguaje nos permite otra capacidad humana que tiene mucho que ver con la teoría de la mente: la au-

torregulación de las emociones y la conducta. El lenguaje interior, hablarnos a nosotros mismos, puede conseguir movernos a la acción hacia algo a lo que no nos orientan las fuerzas biológicas primarias. Esto ocurre con otros tipos de conducta. A diferencia del deseo animal, que es fundamentalmente deseo de un objeto presente, nosotros somos capaces de sacrificar nuestros deseos primarios para lograr un fin que consideremos más elevado y que no tenemos presente. Muchas personas son capaces de sacrificar numerosos deseos primarios para ahorrar un dinero que les permita adquirir algún producto que aumente su prestigio social. O para conseguir un aspecto físico que produzca admiración. O para aprobar un examen para el que hay que estudiar largo tiempo. Estos sacrificios de las propias fuerzas biológicas no están motivados por el altruismo, pero es importante subrayar que se llevan a cabo ayudados por el lenguaje interior: la persona se tiene que convencer a sí misma de que no debe hacer algo que ahora mismo le apetece muchísimo, por una meta que quiere lograr a largo plazo. Es interesante, como apuntamos antes, que el lenguaje interior aparece sobre los siete años, que es cuando tradicionalmente se hablaba de la «edad de la razón». Es también la edad en que los niños son capaces de autorregular su conducta (en el experimento clásico, los niños son capaces de aguantar cierto tiempo sin comerse un trozo de chocolate que tienen delante, por la promesa de recibir uno mayor). También es importante que esa meta que nos planteamos a largo plazo no es algo que vemos, sino que evocamos, capacidad también según parece exclusivamente humana. La evocación, junto al lenguaje interior, nos sirve para fortalecer esa segunda línea mental que es mucho más débil que la primera, al no conectar a nuestros intereses primarios. Del mismo modo, el mecanismo del lenguaje interior podría servirnos, ante la captación de la necesidad ajena, para fortalecer el

impulso hacia una conducta auténticamente altruista. De este modo, la percepción de un tú, la perspectiva de segunda persona, estaría a la base de la moral (A. Gomila, 2008). *Para la experiencia moral, lo primero que aparece sería la obligación ante la segunda persona, antes que el derecho de la primera persona.* No estamos muy lejos del «me dueles, luego existo» de Carlos Díaz. El «me dueles» es una percepción objetiva, como hemos visto, exclusivamente humana. A

REFERENCIAS

- T. Bejarano Fernández, «El gesto de apuntar: una curiosa exclusividad humana». *Thémata*, vol. 30, 2003.
- T. Bejarano Fernández, «Libertad moral y evolución biológica: un interrogante que conviene ya ir planteando». *Thémata*, vol. 34, 2005.
- E. García García, «Teoría de la mente y ciencias cognitivas». En García García, E., *et al.*, *Nuevas perspectivas científicas y filosóficas sobre el ser humano*. Universidad Pontificia de Comillas, 2007.
- A. Gomila, «La perspectiva de segunda persona de la atribución mental». *Azafea. Rev. Filos.*, vol. 4, 2002.
- A. Gomila, «La relevancia moral de la perspectiva de segunda persona». En D. Pérez y L. Fernández Moreno (editores), *Cuestiones filosóficas: ensayos en honor de Eduardo Rabossi*, Catálogos, 2008.

ejercicio de la libertad. La alusión a Kant tiene que ver con eso del establecimiento de las «condiciones de posibilidad». En otras palabras, trato de sugerir que lo característico del hombre, aquello en que todos nos parecemos, consiste precisamente en lo que hace posible que cada uno de nosotros sea diferente. Esto puede parecer un retruécano, pero creo que ayuda a explicar las facetas tan contradictorias que ofrecemos cuando se nos somete a examen. Veamos si soy capaz de explicarme con un mínimo de claridad. En primer lugar hay que advertir que, si el término con el que debe ser confrontada la naturaleza del hombre es la libertad, pierden valor e importancia otras contraposiciones, como las ya mentadas naturaleza/artificio, naturaleza/cultura o también naturaleza/educación. Si lo artificial, lo cultural y lo aprendido se consideran ajenos a lo natural es porque se entiende que son más bien productos de la libertad, pero sucede que la libertad no puede ser *producida*, al menos por el hombre, sino tan sólo *ejercida*. Eso significa que los productos de la libertad revierten a la naturaleza, se *naturalizan*, de la misma manera que la reiteración de una decisión consciente se convierte en costumbre y hasta en automatismo si no somos capaces de reafirmarla en su primitivo carácter. Esto enseña algo acerca del carácter *histórico* del hombre: nuestra naturaleza no es invariable, sino que recapitula de algún modo la biografía de nuestra libertad. Cuentan que el sociólogo Max Weber fue uno de los más decididos opositores de la fiebre belicista que condujo a la primera guerra mundial. Cuando se le preguntaba por qué se había resignado con tanta paz a la catástrofe que se había producido por no seguir sus consejos, respondía: «Lo que antes fue producto de la insensatez, hoy es ya destino inexorable». Según esto, es legítimo decir que la naturaleza del hombre es producto de su libertad; no obstante, en cierto modo también es cierto lo contrario, es decir, que la libertad es el rasgo más genuino posibilitado por nuestra naturaleza. Así que estamos ante un problema parecido al de qué fue primero, el huevo o la gallina. Creo que, de todas formas, algo se avanza si se advierte que la libertad misma no puede ser identificada con la naturaleza, precisamente porque no es una *cosa*, no es algo que se pueda decir con exactitud *en qué consiste*. En cambio, sobre los artefactos,

la cultura, la educación sí se puede teorizar: ahí están la tecnología, las enciclopedias, los planes de estudios, etc. ¿Dónde están en cambio los manuales de instrucciones para el uso de la libertad? Nadie se hace bueno por saber mucha ética. De hecho, los que se dedican a la ética corren un riesgo bastante considerable de volverse cínicos con respecto a la libertad, por lo menos así lo parece últimamente. En definitiva, no es posible decir en qué consiste la libertad, de manera que, si creemos que podemos saber



en qué consiste el hombre, por fuerza habremos de negar que sea libre. Hay un texto de Karl Popper que me parece muy luminoso en este sentido:

Puedo mencionar aquí que creo que, en principio, pueden construirse máquinas capaces de realizar *cualquier tarea especificada que puedan realizar los seres humanos*. El énfasis está aquí en las palabras 'cualquier... especificada': porque puede usarse una especificación, en principio, para construir una máquina. Por esta razón, no estoy preparado para aceptar el desafío del mecanicista, '¡Especifique con *precisión* una prueba que sólo pueda pasarla un ser humano y que sea imposible pasarla, en principio, para una máquina!'. Si *especificamos con precisión* las pruebas que hay que realizar para descubrir si estamos frente a un hombre o a una máquina, tenemos que hacer frente a la posibilidad de que se construya una máquina para *esas especificaciones* y pase así la prueba especificada. Esto, sin embargo, no significa que, si nos enfrentásemos a una máquina parecida al hombre, tuviésemos la menor dificultad de encontrar infinidad de pruebas que la máquina no fuera capaz de pasar, especialmente si conociéramos su especificación (o incluso si procediésemos simplemente por el método de ensayo y error), aunque pocos

seres humanos fallarían dichas pruebas. (Karl R. Popper, *El universo abierto. Un argumento en favor del indeterminismo*, Madrid, Tecnos, 1986, p. 92.)

Es decir, todo lo que se especifica, todo lo que se determina, es natural, cerrado, reproducible de modo mecánico. La pregunta por lo natural en el hombre es entonces la pregunta por lo que en él es mecánico, repetitivo, sometido al reino de la necesidad. Y la respuesta a esa pregunta no puede ser más que ésta: *todo, absolutamente todo* es natural en el hombre una vez que hemos conseguido cuadrarlo, preverlo, analizarlo, deconstruirlo, hermeneutizarlo o someterlo a la acción de cualquier verbo similar. Somos como el rey Midas, con la diferencia que todo lo que tocamos lo convertimos en naturaleza, en necesidad, en máquina. Por lo tanto, más valiera que no insistiésemos demasiado en toquetearnos a nosotros mismos, no sea que nos convirtamos en una máquina más, como con todo gozo pretende hacer el inefable gurú de la inteligencia artificial, Minsky (M. Minsky: «¿Serán los robots quienes hereden la Tierra?», en *Investigación y Ciencia. Vida en el Universo*, Diciembre 1994, pp. 86-92). Por dar una respuesta más concreta en la medida de lo posible, diría para cerrar mi intervención que todo es natural en el hombre, excepto la irreprimible tendencia a no conformarse con su propia naturaleza. 