

La explicación de la conciencia: ¿Qué se puede hacer hoy?

Teresa BEJARANO
(Universidad de Sevilla)

Es detectable a mi entender un cierto cansancio ante el tema de la conciencia. La ya prolongada polémica entre los eliminativistas y los fenomenólogos no tiene visos de dar fruto. Los defensores de 'la primera persona' describen, pero no explican; los eliminativistas de uno u otro tipo explican algo, pero no lo experiencial prelingüístico, y entonces, al reducir la conciencia no lingüística a lo meramente informacional, colocan a los animales muy por debajo de lo que un punto de vista evolutivo aconseja. El conseguir una explicación no reduccionista, como promete Flanagan, sería, por supuesto, maravilloso. Pero ¿es factible en nuestros días? Esta sensación de incomodidad teórica ha llevado a Block, 1995, a denunciar como híbrido y confuso el concepto de conciencia. Pero la distinción -- conciencia fenomenal vs. conciencia de acceso -- con la que él pretende poner remedio, ¿es la que de verdad resultaría conveniente hacer?

Propondré que en el momento actual sólo es oportuno ocuparse de los estados mentales de segundo orden, y en concreto, de la metacrecencia. Afortunadamente para nuestra generación, esa parcela, aunque pequeña, es la más interesante para una curiosidad antropológica.

1. La excesiva impaciencia de Flanagan

Flanagan, 1992¹, parte de un rechazo hacia los autores —él los llama ‘nuevos místéricos’— que, arguyendo, o bien el abismo entre los puntos de vista de primera y tercera persona (Nagel, 1974 y 1986), o bien la imposibilidad de que un sistema, la conciencia, se capte a sí mismo (McGinn, 1991), desconfían de que la conciencia pueda alguna vez ser objeto de explicación científica. Pero, a pesar de su rechazo del ‘misterianismo’, Flanagan no comparte tampoco la postura de los autores eliminativistas. Estos, es decir, Dennett o los Churchland, buscan, sí, la explicación científica, pero en esa tarea reducen la conciencia a sólo dos extremos opuestos, pero igualmente observables y medibles: por un lado, la mera sensibilidad informacional, que se da en los vegetales y hasta en las máquinas, y, por el otro, la expresión lingüística. Y frente a esa reducción, Flanagan reivindica “como también detectables” (p. 73) los *qualia* subjetivos y, en suma, la “sensibilidad experiencial” (p. 55), para la que el giro cognitivista, la teoría de la mente triunfante desde 1970, no supuso —Flanagan (pp. 3-8) explica pormenorizadamente varias causas— el menor retorno. Y cuando muestra los *qualia* como detectables, pretende (comentaremos en seguida esto) estar mostrando que ya podemos empezar a explicarlos científicamente. “Gracias al trabajo reciente en neuropsicología, ciencia cognitiva y fenomenología, tenemos los comienzos de una teoría creíble de la conciencia. La cuestión es que nos atengamos al *método natural* de construir esa teoría, es decir, que tratemos esas tres líneas diferentes de análisis con igual respeto por cada una” (p. 11).

Antes que nada, quiero dejar claro que participo de la esperanza de Flanagan de que la ciencia llegará a explicar cómo la organización de la materia en el cerebro llega a producir “los estados mentales como discernidos desde el punto de vista de la primera persona”. Pero creo que el optimismo de Flanagan es demasiado impaciente. Ninguno de sus intentos de mostrar cómo se puede trabajar ya ahora con su ‘método natural’ me convence lo más mínimo.

Es verdad que “el extraordinario número de experiencias sensibles diferentes en la modalidad de la vista, del gusto..., en cualquiera de las modalidades sensoriales, es explicado por el número extraordinario, pero no infinito, de diferentes valores que pueden ser tomados por cada célula receptora, y

¹ Flanagan, O., *Consciousness Reconsidered*, MIT, 1992.

posibilidades combinatorias entre esas pocas células” (p. 52). Pero lo que así queda explicado es que sean diferentes las experiencias subjetivas, no que sean experiencias subjetivas.

¿Y cuando pretende que quienes no se dan por satisfechos con esos tipos de explicación estarían exigiendo demasiado, exigiendo “*standards* de inteligibilidad de una altura imposible” (p. 115)? También en este punto me parece Flanagan desacertado. Es verdad que ninguna explicación científica de lo real capta la realidad toda, ni pasa, por tanto, de lo superficial y parcial. Pero cuando una explicación científica ha tenido lugar, hay algo que siempre se consigue: si alguna vez en adelante se observa falta de correlación entre los fenómenos cuya conexión se explicó, esa falta resultará, no sólo chocante por su irregularidad, sino realmente más incomprensible (o aun más incomprensible —le concedemos el matiz a Flanagan—) que los casos regulares. Y a mi entender, las supuestas explicaciones contempladas por Flanagan no tienen en absoluto ese efecto: el que una determinada pauta de activación cerebral no tuviera su esperada consecuencia subjetiva, sería, claro está, chocante, pero no sería ni siquiera un punto más incomprensible de como es la correlación normal.

Un argumento particularmente atractivo que el libro recoge es la prolongación del famoso ‘experimento pensado’ del espectro de colores invertido. Un individuo vería verde lo que todos verían rojo, y al revés. Pero, como desde que nació es así, aprendió a llamar “verde” lo que él percibe como rojo. Los eliminativistas consideran que no tiene sentido atender a esa subjetividad individual indetectable. Pero se puede invocar (pp. 71-72) el efecto de excitación del rojo, o de relajación del verde para defender que el espectro invertido puede ser revelado en la conducta o en el cerebro, o en la conducta lingüística (Flanagan menciona los trabajos de Lakoff, G. y Johnson, M. sobre las metáforas). Desde luego, para que el argumento pueda siquiera empezar a andar, hay que añadir una puntualización. Ha de convenirse —yo personalmente convendría de buena gana— que el vínculo entre el rojo de sangre y de fuego y la excitación se habría forjado filogenéticamente. En efecto, si fueran las asociaciones experimentadas por el individuo las responsables, la percepción verde del individuo anómalo estaría vinculada a las amenazas de la sangre y el fuego. Pero, una vez incorporada esa puntualización, ¿qué pasa con el argumento? Creo que logra superar la postura eliminativista acerca del espectro invertido: Está claro que hay vías para detectar la inversión, unas vías más sutiles que el uso del término “rojo” o “verde”. Pero respecto a lo que interesa, es decir respecto a cómo se explica

la sensibilidad experiencial, no se ha adelantado nada; la única novedad es que la gama de conductas observables se ha revelado más amplia de lo que clásicamente se admitía.

Y ese dictamen hay que repetirlo ante otros puntos del libro. Una y otra vez Flanagan defiende convincentemente que los estados subjetivos pueden ser detectables desde el exterior, pero eso no equivale ni mucho menos al éxito en lo que se está debatiendo, es decir, en la cuestión de si la sensibilidad experiencial puede ya empezar a ser objeto de explicación científica. Precisamente ese mismo resultado frustrante es el que aporta el capítulo, precioso para mi gusto, "The missing shade of you" (El título hace referencia al párrafo de Hume sobre cómo una persona imagina un matiz de azul nunca por ella percibido que falta en una carta ordenada de azules). "Yo puedo tomar tu experiencia, según yo la comprendo, dentro de mi mente, y así tratar de comprender como qué son las cosas para tí. Esa comprensión no es nada misteriosa, sino que se consigue gracias a una sensitiva detección de lo que los otros dicen, hacen y pretenden. Los seres humanos somos buenos "detectores mentales", es decir, buenos detectores de los invisibles estados mentales ajenos" (p. 102-103). Flanagan sólo quiere aquí argüir contra la impenetrabilidad de lo subjetivo ajeno, o, diciéndolo con el consabido término, a favor de su detectabilidad. Por eso, es en plan de concesión marginal como hace justicia a la diferencia entre comprender la subjetividad ajena y comprender la propia: "Nosotros tenemos recursos para intentar imaginar el matiz especial que es tu subjetividad. ¿Quiere eso decir que, cuando yo lo consigo, te estoy captando a tí? Claro que no. Eso sería sobrepasar mis límites, y el sistema nervioso lo impide. Pero tenemos inteligencia e imaginación, y algo al menos, aunque sea poco, podemos lograr respecto a asomarnos a otras mentes," (p. 89, con dobles en 106 y 107). Y, sin embargo, esa diferencia que frente a la sensibilidad experiencial primaria es propia de la imaginación o comprensión de la subjetividad ajena, puede ser leída en clave positiva, no como algo que hay que reconocer a regañadientes, sino como la mejor vía hacia aquella explicación científica que sí es ahora emprendible. Pero no adelantemos las cosas: esa línea la desarrollaremos en un apartado posterior.

Nuestra conclusión es que Flanagan no ha logrado convencernos de que nuestros recursos científicos sean adecuados ante el tema de la sensibilidad experiencial, y por eso no ha logrado superar ni a la postura misteriosa ni a la eliminativista, que, tan opuestas una de otra como son, participan, sin embargo —lo tendremos ya claro— de una misma idea de base, la de la descon-

fianza en esos recursos científicos. Frente a esa desconfianza, no veo yo más solución —ya lo adelanté— que remitirse al futuro y al largo plazo. Si Newton en su XVII hubiese centrado su curiosidad en cuál es la diferencia entre lo vivo y lo inerte, habría desperdiciado su talento; y lo mismo Saussure, si hubiera enfocado la *parole* cuando todavía nadie había sistematizado la *langue*. Pero tras esta recomendación tan conocida² quiero comentar dos puntos.

2. Lo que el optimismo de plazo largo permite decir sobre la sensibilidad experimental

¿Podemos imaginar cuándo será el tiempo oportuno para tratar la ahora inabordable cuestión? ¿Vislumbramos cuál será la reformulación bajo la cual la pregunta básica empezará a ser contestada? Yo creo que la situación estará madura cuando seamos capaces de comprender cómo se diferencia del tropismo vegetal, la pauta consumatoria innata de los animales. La pauta consumatoria, he dicho: eso supone que el nivel más primario de la sensibilidad experiencial o conciencia animal hay que buscarlo en el deseo, y no en la percepción³. Y si en el deseo, eso supone también, por tanto, que lo que estamos buscando es el deseo anterior a toda percepción, como el de la paloma novata de (Craig) / Lorenz, es decir, la pauta que pide ser rellenada, aunque no dé pistas sobre cómo será exteroceptivamente el estímulo apropiado. ¿Se me pregunta con qué fundamento atribuyo sensibilidad experiencial a organismos animales no humanos? Yo replico (ahora, porque para criticar a Flanagan no era necesaria, claro, tal réplica) que la especie humana, por muy revolucionarios que hayan sido los retoques evolutivos que la constituyeron, y por inacabable que sea la catarata de consecuencias de tales retoques, tuvo sin embargo que heredar de sus ancestros primates la mayor parte de su constitución. Un ‘Buen Truco’⁴, hay, desde luego, necesidad de postularlo al nivel

² Esa recomendación de paciencia, aparte de ser seguramente la ocurrencia más popular sobre el tema a nivel de calle, constituye el punto esencial de Cottrell, A., 1995, una recensión del libro de Flanagan significativamente titulada con un interrogante escéptico. “Tertium datur?”. *Philosophical Psychology*, vol. 8, nº1, pp. 85-103.

³ Precisamente el deseo, el movimiento, y la separación entre vegetal y animal son los temas que se echan de menos en Humphrey, 1992. *A history of the Mind*.

⁴ El ‘Good Trick’ y su relación con el lenguaje es, a mi entender, la idea que ante todo debemos recoger de Dennett, 1991, aunque yo reduzca a los que considero límites sensatos

del surgimiento de lo humano, pero para que actúe, no desde un grado cero de conciencia, sino desde la sensibilidad experiencial largamente elaborada por toda la evolución animal. En definitiva, la versión que yo considero la más radical del interrogante por la conciencia sería la siguiente. El animal busca el alimento, y el vegetal busca también la luz: ¿por qué, y cómo —eso es lo que hay que explicar— el animal se siente hambriento, tiene deseo y experiencia consciente?

En este punto creo que la cuestión del epifenomenalismo⁵ o no de la conciencia se pondrá en el centro del tapete. El epifenomenalismo lo enfoca en varias ocasiones Flanagan (por supuesto, siempre como una postura implausible⁶, aunque reconociendo —acertadamente, creo yo— que “la sospecha epifenomenalista puede ser sana hasta un cierto punto, ya que tendemos a sobrevalorar tanto el tamaño de aquella parte de nuestra vida mental que es consciente, como su papel causal”, pp. 7-8). Pero cuando de verdad llegará a ser útil la pregunta por la finalidad de la conciencia será —apuesto por ello— en ese seguramente lejano futuro en que se pueda abordar la diferencia entre esos dos tipos de hechos —fototropismo y hambre, p.e. Para aclarar un poco

—esto ya lo he dicho— los poderes del Buen Truco, y aunque más adelante concretaré la influencia del lenguaje de un modo distinto (La postura de Dennett respecto a esto, inspirada seguramente en la historieta prehistórica de Sellars, “El empirismo y la filosofía de lo mental”, *Ciencia, percepción y realidad*, Tecnos, 1971, sobre todo, pp. 199-201, y quizá también en Jaynes, J. *The origins of consciousness in the breakdown of the bicameral mind*, Boston, Houghton Mifflin, 1976, puede verse en Dennett, 1991, pp. 194-196, donde apela a una modularidad entendida de un modo fortísimo, de un modo inverosímilmente fuerte, a mi entender, y también en su ilustración 7.3, un rostro con una flecha desde boca a oreja —que se parece a la que, aludiendo a una propuesta mucho más sólida, ha sido tomada como emblema por Karmiloff-Smith, A., *Beyond Modularity*, MIT, 1992, y “Précis of *Beyond Modularity*”, *Behavioral and Brain Sciences*, 1994, 17, pp. 693-707, véase p. 706-).

⁵ Epifenomal en el sentido de ‘sin función’, no en el de ‘no detectable’ (Concretamos así en recuerdo de la distinción de dos acepciones en Dennett, *Consciousness Explained*, Little Brown and Company, 1991, p. 402)

⁶ Un argumento un poco dudoso que esgrime contra el epifenomenalismo es el de la ‘vista ciega’, esa curiosa patología en la que el individuo se siente ciego, pero sin embargo en tareas visuales es capaz de acertar muy por encima del azar. Flanagan arguye que el déficit funcional de esos enfermos muestra que sin sensibilidad experiencial, la capacidad de respuesta se resiente. Ahora bien, para que ese argumento pueda empezar a ser siquiera evaluado, habría que contar con el dato, que aún no se tiene, de si el grado de déficit funcional y el de déficit experiencial van siempre parejos. Pero ante todo, a mi entender, en toda esa discusión se olvida que esos enfermos, aunque carentes de los *qualia* de la modalidad visual, poseen aquello que es lo verdaderamente asombroso de la hazaña perceptiva, a saber, la concepción de un mundo exterior al organismo.

mejor lo que estoy queriendo decir, voy a recurrir a una cita recogida por Dennett, 1991, p. 451, aunque la llevaré a un molino que no es ya adonde él la llevaba. Se trata de un párrafo de Dawkins, M.S. (no el Dawkins del gen egoísta, sino la estudiosa del sufrimiento animal y sus implicaciones morales): “Si se les daba a elegir, las gallinas elegían siempre la jaula con el suelo cubierto de paja. ¿Sufrían cuando no la tenían así? Entonces introdujimos una complicación. Las aves tenían que ‘trabajar’ para llegar desde el comedero hasta la jaula con paja. Y lo que sucedió fue que las gallinas no se molestaban en llegar hasta la jaula de sus preferencias”. A esto le podríamos añadir los ya clásicos datos⁷ acerca de las ratas hiperfágicas, o ratas a las que artificialmente se les ha anulado el ‘avisador de saciedad’: Esas ratas siguen y siguen comiendo, pero sólo si no han de hacer esfuerzo para conseguirla. Tenemos, pues, en uno y otro caso unas tendencias que no son capaces de poner en acción a los animales, de lanzarlos a la búsqueda. Esos casos no son los deseos animales normales. Y no es la causa de la anomalía —un bien no necesario, un organismo alterado— lo que nos interesa, sino la anomalía misma, porque en ella vemos la perfecta contrafigura del deseo animal⁸. Si el vegetal no necesita conciencia, ni siquiera en la vertiente —la del deseo— más primaria, es porque no actúa, no busca en el exterior (“No habría ningún

⁷ Manning, *Introducción a la conducta animal*, Alianza, 1977, (orig. ing. 1972), p. 104, y 121.

⁸ ¿Sufrén los animales? Entonces, la selección natural, que pivota sobre el fracaso de los más débiles, estaría pivotando sobre el sufrimiento. Así que, si en los animales sólo hubiera conducta ligada al estímulo, si su interioridad fuera vacía al estilo zombie, nos ahorraríamos abordar una cuestión bien dura. Pero ¿y si, como hemos nosotros hecho, atribuimos sensibilidad experiencial, y no sólo informacional, a los animales superiores? El atribuir sensibilidad experiencial a animales no es defender la equivalencia entre sufrimiento animal y sufrimiento humano. El dolor humano físico (concretemos así para no rehuir la comparación) tendría siempre un plus añadido, pues la persona contempla su situación no como meramente lo que en el presente hay, sino como algo que podría haber sido de otra manera, como una posibilidad — la posibilidad que se realizó— entre otras, que, al no haberse realizado, sólo son ya para el sujeto expectativas o creencias cuya falsedad él tiene constatada. En cambio, el animal que sufre alberga, sí, (y esto no sólo lo admito, sino que insisto en ello) la expectativa de una situación mejor para el momento siguiente, pero no atiende a ninguna alternativa del momento presente mismo. El animal estaría así polarizado en la expectativa, en el momento siguiente (después de todo, atender a eso, que es lo aún modificable, es la única atención útil en primera instancia biológica), y no estaría nunca centrado en el presente, ni, tampoco, por tanto, en su dolor cuando su presente sea malo. Las líneas que acabo de escribir han sido, ya se habrá adivinado, mero tantear a ciegas y sin base. Conclúyase de ellas, eso sí, que el asunto me parece apasionante e importantísimo.

‘punto’ evolutivo en que un árbol al que se le están cortando las ramas tuviera la capacidad de sufrir en silencio” es la espléndida formulación de M. S. Dawkins). El deseo animal va con, y sirve a, la capacidad de salir a buscar fuera. Para la conducta de las gallinas respecto al lecho de paja, o de las ratas hiperfágicas respecto a la comida, podría haber bastado un mecanismo análogo al de los tropismos vegetales.

Y una vez admitido que lanzar al animal a la búsqueda en el exterior es la función y la clave del deseo, se nos aparece ya claro cuál ha de ser nuestro asidero en el tema de la vertiente perceptiva de la conciencia. En efecto, el que la percepción la hayamos considerado derivada y al servicio del deseo, nos encajará ahora perfectamente con un rasgo importantísimo de toda percepción. Para ello, permitamos aquí que nos cale el asombro: se llega a percibir que las cosas están fuera del cuerpo, y, sin embargo, el control sensorial es sólo de la piel para dentro. La hazaña característica de la percepción es, pues, que coloca en el exterior, en ese exterior al que ha de salir a luchar el animal, las cosas percibidas. Este rasgo de la sensibilidad experiencial es subrayado por Revonsuo, 1995⁹, pp. 48-49: Nos asombramos —viene éste a decir— de que, siendo los sueños internos al cerebro, estando el caballo de mi pesadilla sólo en mi interior, me parezca, sin embargo, mientras la padezco, que el caballo está enfrente de mí; pero eso, aunque, en efecto, es asombroso, no lo es más que la percepción normal y la existencia toda experimentada fuera del cerebro. “La conciencia, no menos que los sueños, es una ‘realidad virtual’ construida”¹⁰. Yo, eso sí, a diferencia de Revonsuo, insisto en que el deseo sería lo primario.

¿Cómo se originaron esas capacidades del deseo y la percepción cuando la capacidad para el movimiento las hizo necesarias? ¿Cómo el deseo se diferenció del tropismo cuando eso fue requerido por los nuevos organismos capaces de moverse? Esos interrogantes, interrogantes por el cómo¹¹,

⁹ Revonsuo, A. “Consciousness, dreams and virtual realities”, *Philosophical Psychology*, 1995, vol. 8, nº 1, pp. 35-58. En el mismo sentido también podríamos citar a Alroy, D., 1995, “Inner light”, *Synthese*, 104, pp. 147-160.

¹⁰ Aunque no por ello —apostilla interesantemente Revonsuo — haya que distinguir, ni en los sueños ni en la percepción, las ‘presentaciones’, por un lado, y las ‘audiencias’, por otro, como sería la acusación de Dennett. “Muy al contrario del ‘Teatro Cartesiano’ que atrae las iras de Dennett, el mundo experimentado es fundamentalmente uno, y no hay lugar para diferenciar el sujeto y el mundo externo” (Revonsuo, 1995, p. 48).

¹¹ Coincidimos con Dennett, 1991, en que el cómo es la clave. “No es que no puedas imaginar un robot consciente. Es que no puedes imaginar cómo un robot podría ser consciente”

son hoy inaccesibles. Pero la posibilidad de darles respuesta será —repito lo que dije párrafos atrás— lo que seguramente señale que el tiempo oportuno habrá llegado, el tiempo ese anhelado en el que las correlaciones que hoy la investigación empieza a establecer dejarán de ser mero andar por las ramas y servirán por fin para construir la verdadera explicación científica de la conciencia.

3. La cuestión que es factible estudiar ya, ahora

Añadir esta parte resultaba obligado. En efecto, he manifestado mi opinión de que hoy, en nuestro final del XX, no es todavía una tarea planteada la de trabajar por una explicación científica propiamente tal de la conciencia, y, aunque a largo plazo soy, sí, optimista, ese optimismo no alivia el amargor. Debo, pues, atender a la pregunta de si no habrá acaso un premio de consolación. Saussure, después de todo, si bien encontró vetado el tema de la *parole*, tenía la *langue* para en ella investigar con fruto. ¿No tendremos nosotros una parcela en la que el asunto de la conciencia esté ya maduro, y susceptible a una verdadera indagación explicativa?

A mi entender, lo hay, ese premio de consolación que se quiere, y precisamente tal que llamarlo con ese término resignado y grisáceo es de lo más inadecuado. Si nuestra curiosidad se centra en lo antropológico, si queremos enfocar precisamente la mente humana, entonces para nosotros el primer premio, el premio por excelencia, lo constituye justo una parcela que, aunque no haya en ella todavía apenas construcción, está ya, ahora mismo ya, empezando a ser accesible. En efecto, si (como por razones evolutivas parece - ya lo he dicho— sensato y casi necesario hacer) se les concede a los animales la sensibilidad experiencial, entonces el problema mente / cuerpo, tan tradicional y venerable en la psicología humana y en la antropología, quedará, sin embargo, alejado del centro de la diana de esas disciplinas. Y la pregunta que de verdad nos apasiona a quienes querríamos explicar cómo se constituye lo específicamente humano, viene, pues, a enfocar esa parcela que hemos anunciado como accesible.

(p. 432). Pero mientras que él considera que en su libro se ha hecho: “¡Lo hemos hecho! Con el nivel de descripción centrado en el ‘software’ hemos suministrado el puente explicativo. Hemos imaginado cómo un cerebro podría producir experiencia consciente” (p. 434), mi opinión al respecto es mucho más pesimista. Mi optimismo, ya se sabe, concierne sólo al futuro y no al corto plazo.

Ahora bien, la diferenciación que nos va a servir¹², la que se traza entre la conciencia irreflexiva (o de primer orden) y la reflexiva o propiamente humana, está delineada desde hace mucho tiempo (digamos sesenta años, para no rebuscar más). Y cuando la conciencia de segundo orden ha recibido el nuevo tratamiento de Dennett, 1991 (pp. 303-309, sobre todo, donde se basa explícitamente en Rosenthal), la principal aportación ha sido la de esgrimir el segundo orden como compensatorio de la pérdida aceptada por ese autor, es decir, de la renuncia, en él definitiva, a tratar de la sensibilidad experiencial (o sea, la que no es meramente informacional, ni necesariamente lingüística)¹³. Así pues, si el acotamiento de la conciencia humana es una noticia ya antiquísima, y si sus retoques novedosos van en una dirección, la del reduccionismo, que no nos gusta, ¿cuál es entonces —se me preguntará— ese maravilloso premio que le ha tocado a nuestra generación, y cuyas mieles podríamos ya empezar a saborear?

La feliz novedad a la que me estoy refiriendo es la de que ahora, y no antes, contamos, respecto a la investigación de la conciencia exclusivamente humana, con un programa de preguntas, de preguntas concretas y ante las que

¹² Block, N. "On a confusion about a function of consciousness", *Behavioral and Brain Sciences*, 1995, 18, pp. 227-287, pretende (en la introducción, ya lo mencionamos) que la distinción salvadora es la que separa la conciencia fenoménica y la conciencia de acceso, y ello a pesar de que también reconoce otros tipos, y entre ellos el que a nosotros nos interesa (p. 235, "el pensamiento de orden más alto o metacognitivo, tal como lo ha definido Rosenthal"). A mi entender, Block ha elegido mal su bifurcación básica. Más que a una distinción que esté realmente fundamentada, estaría más bien aludiendo a la duplicidad de corrientes que, ya separadas, ya confundidas, se aprecian en la bibliografía, donde bajo el único rótulo de conciencia tan pronto se nos habla de "la experiencia y lo fenoménico" como de "modelos de espacio de trabajo global, o de procesamiento de información". Pero el que haya dos metodologías explicativas diferentes no autoriza sin más a hacer una distinción pareja en la realidad. En cambio, un distinto momento de aparición evolutiva sí que es capaz de autorizar un corte de la realidad, y a eso nos hemos atenido nosotros. Curiosa y reveladoramente Block en p. 236, invoca la distancia entre surgimientos evolutivos como criterio contra Johnson-Laird, es decir, contra una confusión que no es la que él está denunciando a todo lo largo de su artículo, sino justo la confusión entre pensamiento de primer orden y pensamiento de segundo orden.

¹³ En realidad, las diatribas (Dennett, 1991, *passim*) contra el Teatro Cartesiano no son un buen medio de explicitar cuál es exactamente el blanco contra el que se dispara. Se cumple en ello la conocida regla de que la claridad de una postura no queda bien servida cuando para exponerla se escoge un enemigo demasiado cargado, demasiado perverso: ¿es sólo el extremo de esa perversidad lo que es rechazado, o, por el contrario, es ya contra el núcleo mismo sin aditamentos agravantes contra el que se dirigen las críticas? El lector tiene que recurrir a otro tipo de párrafos, o a la entrevista de 1995 (*Journal of Cognitive Neuroscience*, 7, 3, pp. 408-414), para encontrar perfilada con más claridad la postura de Dennett.

se puede al menos diseñar una vía por donde buscar la respuesta. En 1978¹⁴, y en un comentario al trabajo de Premack y Woodruff, 1978, en el que éstos se preguntaban si el chimpancé tendría una ‘teoría de la mente’, Dennett dejó abocetado un marco para estudiar el problema de la atribución a un organismo de una ‘teoría de la mente propia y ajena’. Si el sujeto tiene —propuso Dennett— un estado mental (creencia o deseo) acerca de otro estado mental suyo o ajeno, el primero de esos estados mentales será un estado mental de segundo orden, y el sujeto que lo tiene poseerá una ‘teoría de la mente’¹⁵. Como se ve, en esa formulación latían varios dilemas. Creencia o deseo: ¿valen lo mismo para el caso, o, por el contrario, uno de esos estados mentales es más importante que el otro? Y también, dado que la creencia puede ser verdadera o falsa: ¿tiene una de ellas un papel más primario? Y todavía hay que seguir. Puesto que el estado mental que es tomado como objeto puede ser propio del sujeto o ajeno, ¿cuál es el acceso originario al segundo orden, el acceso interpersonal o el intrapersonal? Pues bien, esos dilemas e interrogantes que latían en la formulación de Dennett son los que hizo suyos la psicología evolutiva. Y el estudio de cómo el niño va adquiriendo la ‘teoría de la mente’ ha puesto ya en claro bastantes puntos.

En primer lugar, aprendimos que lo difícil, la adquisición tardía, es el estado mental acerca de una creencia¹⁶. Los estados mentales acerca de un deseo han resultado ser de un rango muy inferior. Al principio pudo ello parecer sorprendente, pero ahora lo vemos lógico. En efecto, es sólo la creencia captada como tal lo que constituye un duplicado mental de lo que para nosotros es la realidad. El deseo, por el contrario, es una realidad más del mundo, otra fuerza física, que hay que tener en cuenta (en efecto, el deseo no es sino

¹⁴ Dennett, D., “Beliefs about beliefs”, *Behavioral and Brain Sciences*, 1, pp. 568-570.

¹⁵ En Dennett, 1978, no se habla de conciencia, sino de intencionalidad — de intencionalidad en sus distintos grados —. En cambio, en el pasaje, antes citado, de *Consciousness explained* donde se recoge lo de 1978, ya no aparece el término ‘intencionalidad’.

¹⁶ Algunos autores —Leslie, A. M. (“Some implications of pretense for mechanisms underlying the child’s theory of mind”, Astington, J., Harris, P., y Olson, D., eds. *Developing Theories of Mind*. Cambridge Univ. P., 1988) como cabeza de serie — han pretendido relacionar la ‘creencia acerca de creencia’ con el juego simbólico, que es una capacidad muchísimo más temprana. Pero, dejando aparte cuán dudoso sea ese vínculo (en mi trabajo “Las emociones ante la ficción”, *Thémata*, 1995, defendiendo otra explicación para el juego simbólico), ni siquiera dentro de la propuesta de Leslie pierden las metacreencias o ‘creencias acerca de creencia’ su alto rango en la jerarquía de dificultad: En efecto, Leslie ha de descomponer en módulos diferentes, y escalonados en su adquisición por el niño, la capacidad unitaria y englobante que él pretende.

intención de actuar tan pronto como la actuación en cuestión sea posible) y, por tanto, captar un deseo no requiere concebir un dominio en contraposición a la realidad.

Y respecto a cuál creencia es más fácil que sea captada como objeto por un estado mental de segundo orden, si la creencia verdadera o la creencia falsa, ¿qué ha aportado la investigación con el niño? Al principio, el enfocar la creencia falsa fue una mera necesidad metodológica: en efecto, sólo de aquella creencia que el sujeto considere falsa, es de la que no podemos dudar que está siendo pensada por el sujeto como creencia, y no como la realidad misma. Pero más adelante, empezó a plantearse la cuestión de la necesidad, no ya a nivel de la investigación, sino, según el enfoque funcional, a nivel del organismo. ¿Para qué, en efecto, en el caso de la creencia que el sujeto juzga verdadera, iba a ser útil en primera instancia el desdoble y diferenciación entre lo creído y la creencia? En un primer nivel (en niveles superiores ya sería distinto¹⁷, pero no es eso ahora nuestro tema), en el origen, repito, el desdoble en el caso de la creencia verdadera sería epifenomenal, superfluo, un superferolítico, digámoslo así, ejercicio cartesiano. Más allá del niño, también encontramos algunas observaciones que pueden argüirse en pro de esta nuestra respuesta. P. e., leemos en Edwards, D. & Potter, J., *Discursive psychology*, Sage, 1992, p. 47: “La gente tiende a atribuir cualquier desviación respecto de lo que *ellos perciben* que es la verdad a factores distorsionantes tales como la personalidad, la carencia de competencia, y demás factores psicológicos o sociológicos del narrador. En cambio, consideran que no hay nada que explicar, y, en consecuencia, no invocan ninguno de esos factores en el caso de narraciones o testimonios que ellos consideran ‘factuales’”.

Ahora bien, las creencias que un sujeto está en el momento presente considerando falsas, pueden ser, o bien propias, y entonces, necesariamente, ya pasadas, o bien ajenas, en cuyo caso pueden seguir en ese momento vigentes. Los estudios con niños han mostrado que, tanto el recordar la propia creencia falsa pasada¹⁸, como el inferir la creencia falsa ajena, son capacidades que

¹⁷ En efecto, dejada atrás su inutilidad en la primera instancia, el captar como mera creencia, y no como la realidad misma, la creencia que para uno mismo está vigente, podría en niveles superiores llegar a ser un factor previo indispensable a la resolución creativa de problemas, o sea, a la inteligencia propiamente humana (En “La advertencia sentido/referencia y la resolución creativa de problemas”, *Revista de Filosofía*, 1996, 1, ya traté este punto, y en una recensión —*Thémata*, 16, 1996, pp. 331-335 — de Karmiloff-Smith, 1992/1994, hago alguna sugerencia que va un poco más allá).

¹⁸ La propuesta de Dennett, 1991, de las Versiones Múltiples que acerca de un mismo

sólo se adquieren hacia los cuatro años. Y con esto ya hemos llegado al tercer dilema. ¿Es más primario el acceso intrapersonal o, por el contrario, el interpersonal?

Lo importante que hay aquí que considerar es el tipo de cuestiones que ahora están reformulándose de un modo nuevo. ¿Es intrapersonal o interpersonal el origen de la conciencia exclusivamente humana? He ahí una pregunta antropológica clave y clásica, que puede ser, desde luego, abordada por muchas vías. Así —para poner un ejemplo de tales vías, uno que no tiene nada de especial y que escojo simplemente porque me es familiar—, si el acto de habla que se clasifica a sí mismo como tal acto de habla (tal es una buena definición de los performativos) se caracteriza por una peculiarmente intensa atención del hablante al punto de vista de la audiencia (véase mi “Nota sobre ‘permite que’”, *Theoria*, 1992, pp. 941-951, en concreto p. 944), ello puede prestar cierto apoyo, por mínimo e indirecto que sea, al origen interpersonal de la capacidad por la cual el sujeto contempla y describe sus propias acciones. Y seguramente otras muchas vías parecidas a ésta podrían también invocarse —vuelvo a reconocerlo. Pero la mencionada pregunta sobre el origen de la conciencia es ahora susceptible de reformularse bajo la forma de

momento pueden sucederse en la conciencia, todas con el mismo derecho, podría estar usufructuando precisamente este fenómeno del olvido de la propia creencia no vigente. Si, según el antiguo experimento al que recurre Dennett, un punto de color rojo se ilumina para apagarse en seguida, y muy poco después, se ilumina otro punto, esta vez verde, muy cerca del primero, y, como resultado, a los sujetos les parece que un solo punto se ha movido y ha cambiado de color a la mitad del camino, “¿cómo somos capaces de colocar el cambio de color antes de que llegue el segundo punto?”. Este es el experimento real, que después Dennett, al reformularlo para que resulte más pedagógico o más susceptible de ilustración gráfica —ilustraciones 5.9 y 5.10— (mujer sin gafas, que en la segunda versión de la conciencia se nos aparece como con gafas), altera en un punto clave, justo en el punto por el que el experimento real conecta con el fenómeno del olvido de la propia creencia pasada. En efecto, mientras que la ausencia de gafas nunca es desmentida por los hechos, hay, en cambio, en el experimento real una creencia o expectativa —la que durante algunos milisegundos el sujeto alberga de que el punto no se moverá, y de que todo acabó con el apagarse del rojo— que resulta desmentida por el suceso inmediato. Y es a causa de ser desmentida, por lo que esa creencia pierde, a despecho de su simultaneidad con el momento del apagón, su título de informadora de tal momento (lo pierde en beneficio, claro, de la creencia surgida tras el reencendido). Desde luego, el olvido de la propia creencia pasada es propio de niños, y el experimento en cuestión tiene a adultos como sujetos, pero, admitiendo que el olvido infantil se da porque el niño no es capaz de simbolización del episodio en el lenguaje interior, cabe admitir también que esa misma incapacidad afectaría a los adultos en el caso de un lapso de tiempo excesivamente corto entre el episodio y su desmentido.

cuál es el acceso primario a la creencia de que una creencia es falsa. Y esa reformulación, no sólo es que incida en lo que, a mi entender, constituye el núcleo originario de lo metacognitivo, sino que también —y es esto lo que queríamos subrayar— es susceptible por fin de un abordaje operativo, de un abordaje para empezar el cual contamos hoy con algunos medios.

Una vez que hemos subrayado la importancia de lo metacognitivo, incluido lo metacognitivo interpersonal, estará ya claro aquel comentario que hicimos acerca del capítulo “The missing shade of you” de Flanagan. La captación de los estados mentales ajenos, aunque es mucho más insegura, más parcial e indirecta que la sensibilidad experiencial propia de uno, supone un proceso psíquico muy superior, y exclusivamente humano, y aún más, quizá constituya el proceso en el que justo se origina la exclusividad humana. En su comparación, la inenarrable riqueza de matices de los *qualia* subjetivos queda muy por debajo, y, lo que es más decisivo, como algo bastante periférico respecto a la antropología.

4. Los interrogantes sobre lo que hoy se está trabajando

Hasta aquí, hemos hecho dos propuestas acerca de la explicación de la conciencia: primero, que hoy las posibilidades se centran en la génesis de la conciencia exclusivamente humana, y segundo, que esta cuestión puede reformularse bajo la forma de cuál sería el acceso primario a la captación de la falsa creencia. Ahora, veremos una panorámica del debate acerca de tal acceso. ¿Cuál es la falsa creencia que antes se capta?, ¿es antes la propia pasada o, por el contrario, la presente ajena? (estas alternativas constituyen la primera interpretación, o interpretación débil, que se les viene dando a los rótulos intrapersonal e interpersonal). Y si nos centramos en la falsa creencia ajena, ¿hasta qué punto es interpersonal o intrapersonal el proceso de tal captación? (he aquí la segunda, y más fuerte, interpretación de los rótulos). Y otra pregunta más: ¿Cuál es el acceso primario?: ¿es por vía lingüística o no?.

Pero antes que nada, un comentario previo. Los experimentos que se manejan contemplan siempre la captación mediante inferencia no lingüística. Ante la mirada del niño aparece un personaje —personaje *A*— que esconde una canica en un cajón, y se va. Después aparece un personaje —personaje *B*— que, sacándola del cajón, la esconde en un florero, y se va. Cuando vuelve *A*, se le pregunta al niño “¿Dónde buscará *A* la canica?”. Y la respuesta correcta la alcanzan hacia la edad de cuatro años. Pero yo quiero insistir en

que las primeras predicaciones funcionalmente tales¹⁹ del niño ya nos muestran que el niño ha captado la falsa creencia ajena. Supongamos que la madre ha dicho: “El kiosco está ahora cerrado”, y que el niño ha alcanzado a ver luz y movimiento en el kiosco. Entonces, el niño dirá “Cerrado, no”. Ahí el niño ha oído una creencia de la madre que él ha constatado en seguida que era falsa. Pues bien, este tipo de captación de creencia falsa se da a una edad bastante anterior a los cuatro años. Según eso, el acceso por vía de audición lingüística a la falsa creencia ajena sería la forma más primitiva de captación de creencias falsas. Esta consideración yo la tengo muy presente, y quizá me empuje —lo reconozco— a favorecer, primero, que la captación de la creencia falsa ajena sería la primaria, y, segundo, que el proceso de tal captación sería interpersonal lingüístico²⁰. Pero, una vez confesado esto, me concentraré en el debate tal como lo encuentro.

El trabajo de Mitchell, P. y Lacohee —Mitchell, P. “Realidad y representación del mundo: ¿Una distinción innata?”, *Mundo Científico*, 1993, 131, 13, pp. 82-83— se presenta con la pretensión de constituir un apoyo a la tesis proinnatista y prointrapersonal, pero, a mi entender, favorecería más bien todo lo contrario. En efecto, allí se muestra que si los niños introducían en un buzón un dibujo de su primera creencia mientras todavía no había sido ésta desmentida, se lograba rebajar la edad para la tarea de la falsa creencia propia pasada. Ahora bien, el buzón natural por excelencia, ¿no es acaso la cabeza de aquél a quien el sujeto, cuando aún confiaba en la creencia falsa, la contó?. Pero, una vez que se ha desmontado la supuesta prueba proinnatista, se nos plantea el problema de cuál de los dos accesos es más primario.

Según los datos de los experimentos, la creencia falsa ajena, el niño pare-

¹⁹ Entiendo que la predicación es el tipo de acto de habla que se originó por y para la función de corregir creencias que el hablante supone que el oyente tiene (corregir para mejorarlas, o, en el caso de la mentira, para empeararlas). Por supuesto, muchas oraciones de forma predicativa no responderán a tal función. Por eso en el texto he precisado “predicaciones funcionalmente tales”.

²⁰ El acceso no lingüístico a la falsa creencia ajena sería, según eso, bastante más difícil. Y así ese comentario nuestro converge curiosamente con la crítica que Millikan hace a lo que ella considera el mito nuclear de la psicología intencional, a saber, el mito de que las predicciones de la conducta ajena se basarían en la (realmente, dice Millikan, difícilísima) atribución de creencias y deseos a los otros. Ella propone que el instrumento normal para la predicción lo constituyen “las observaciones de regularidades conductuales pasadas del grupo o del individuo” (Millikan, “*White Queen Psychology and Other Essays for Alice*”, MIT, 1993, p. 185), pero yo prefiero —ya se sabe— insistir en que, al lado de la, desde luego, difícil atribución, hay un acceso muy fácil a las creencias de otro, a saber, el oírlas.

ce empezar a *inferirla* (ya se sabe el por qué de mi subrayado) por el mismo tiempo que conserva su propia falsa creencia pasada. Así pues, a lo único a que podríamos, parece, recurrir aquí para decantarnos por un origen, es al enfoque funcional: ¿qué es más útil en primera instancia? ¿lo es que el sujeto recuerde su anterior creencia falsa, o, por el contrario, que capte la creencia falsa de otro? Y desde ese punto de vista, como han señalado Cassidy, 1993, y Tomasello, 1993²¹, el origen interpersonal parece más verosímil. En efecto, seguir teniendo presente una creencia propia que se le haya revelado ya como falsa, ¿qué sería en primera instancia²² para el sujeto sino una posible fuente de confusiones, y una carga inútil para la atención?

Uno de los autores que se han opuesto a la primacía de la captación de la falsa creencia ajena es Goldman (en el debate Gopnik/Goldman, 1993²³). Pero éste, que se sitúa claramente en la línea de Searle —reivindicación de la evidencia fenomenológica de primera persona— malinterpreta en cierto modo el debate. En efecto, si Gopnik ha puntualizado que la comprensión de la cual está tratando incluye la comprensión de que las creencias puedan ser falsas para el sujeto —p. 3, “misleading (...), misrepresentation (...) false”—, está claro que es la creencia de segundo orden lo que se está debatiendo. Sin embargo, en Goldman, que como ejemplos invoca ante todo sensaciones de dolor, la separación entre el primer y el segundo orden parece estar bastante difuminada. Y así esta crítica sólo nos obliga a volver a insistir en la ya conocida diferencia entre los dos órdenes (es decir, en lo que venimos defendiendo desde el principio del trabajo).

Pero la consideración del proceso de captación de la creencia falsa ajena ha provocado un fuerte debate. Y, para más complejidad, ese debate se presenta entrecruzado con otro. Y así, tenemos a quienes optan por un proceso de captación interpersonal (interpersonal aquí, claro, en la interpretación fuerte) no simulacionista, y enfrente, a los partidarios de una captación intra-

²¹ Cassidy, K. W., 1993, “There’s more to mental states than meets the inner ‘I’”, *Behavioral and Brain Sciences*, 16, p. 34-5, y Tomasello, M. “Where’s the person?”, *Behavioral and Brain Sciences*, 16, 84-85.

²² Repito aquí también lo de ‘en primera instancia’, porque más allá, es decir, en combinación con alguna otra capacidad exclusivamente humana, su utilidad sí sería clave. p.e., para la capacidad numérica (El clásico experimento de Buytendijk: Descripción relevante de la realidad en cada caso = la fracasada predicción anterior + 1).

²³ Gopnik, A. “How we know our minds: The illusion of first-person knowledge of intentionality”, *Behavioral and Brain Sciences*, 1993, 16, pp. 1-14. Goldman, A. I. “The psychology of folk psychology”, *ibidem*, pp. 15-28.

personal simulacionista. Por eso, antes de preguntarnos cuál de esos dos bandos nos convence más, habrá que dirimir la cuestión de si aceptamos o no el empaquetamiento con que se nos presentan los rasgos. El decantarse a favor de que el proceso (el proceso de captación de la falsa creencia ajena, ya se sabe) sería interpersonal, ¿obliga necesariamente, o no, a rechazar la simulación?

Lo primero que ahora nos conviene es ver con qué carácter se presenta en Gopnik la tesis del proceso interpersonal. Gopnik está en la órbita de Dennett. En efecto, no separa explícitamente “los dos componentes de la comprensión de la creencia, a saber el primario e indispensable comprender lo que uno cree, y el sofisticado comprender que la creencia es una creencia” (como en su comentario denuncian justamente Zelazo & Frye, *ibidem*, 1993). Pero, a diferencia de lo que veíamos en Goldman, esa falta de separación no se da en Gopnik porque se centre en los estados mentales de primer orden, sino justo por lo contrario. Así Gopnik (p. 6) intenta usar el olvido infantil de la falsa creencia pasada, como un argumento a favor de que los niños no tendrían experiencia fenomenológica de primera persona, aunque para ello esté movilizándolo ilegítimamente como dato relevante el lapso de la introspección adulta. Asimismo, el escamoteo del primer nivel y la consecuente confusión entre los dos niveles, llevan a nuestra autora a no saber valorar la verdadera dificultad de la metacreencia, pues invoca la condición de expertos de los adultos para explicar el supuesto dato de “que las creencias acerca de los estados psicológicos propios, parezcan ser perceptualmente inmediatas” (Gopnik, 1993, p. 11). Pero ahí, a mi entender, sobran las palabras “las creencias acerca de”. Pues esa percepción de inmediatez es la de los estados mentales de primer nivel. En efecto, decir que la metacreencia acerca de la propia vigente creencia es percibida de modo inmediato sería tan equivocado como decir que unas gafas bien graduadas y limpias son inmediatamente perceptibles para el que las lleva puestas. La consecuencia de toda esta postura de Gopnik la vemos en la forma como se concreta su propuesta. En efecto, en esta autora, el término ‘teoría’ que figura en el rótulo ‘teoría de la mente’ es tomado completamente en serio, como construcción teórica. Y ello provoca muchas hostilidades (Gustafson, D. “Eighteen months on the planet and already a psychological theorist”, *Philosophical Psychology*, 1995, 8, 2, pp. 125-137, o Horst, S. “Eliminativism and the ambiguity of ‘belief’”, *Synthese*, 1995, 104, pp. 123-145) que revierten contra la primacía de la captación de la creencia falsa ajena (Y tal revertir es injusto, porque, a mi entender, este acceso sería originariamente —ya lo dije antes — una cuestión, no de inferir

la creencia ajena, sino de oír: Los niños de 18 meses que, en palabras de Gopnik, 1993, p. 3, “usan palabras como ‘no’ para comentar la no existencia de objetos hipotéticos”, estarían precisamente respondiendo a la falsa creencia ajena que acaban de oír, respondiendo, p.e., después de comprobar que no hay caramelos, a un ‘Quedan caramelos’ dicho por el adulto, o a un lingüístico o protolingüístico ‘Más’ (= ‘Dame más caramelos’).

Pues bien, si queremos desligar y separar de la línea eliminativista o instrumentalista, la postura a favor de que la captación de la falsa creencia ajena sería la primaria (postura que, como antes dije, me parece preferible en virtud del enfoque funcional), puede ser conveniente mirar hacia la simulación o *role-taking* que autores como Harris²⁴ han esgrimido a favor de un proceso intrapersonal de captación de la falsa creencia ajena. Creo, en efecto, frente a las conclusiones últimas de Harris, por una parte, y frente a la propuesta al estilo de Gopnik a favor de la interpersonalidad, por otra, que la simulación o imitación es el factor al que hay que atender para que el enfoque interpersonal (entiéndase siempre en adelante ‘interpersonal’ en el sentido fuerte, o sea, referido al proceso) adquiera por fin concreción y generalidad a la vez.

Antes que nada, reconoceré mi fortísimo sesgo a favor de todo lo que signifique atender a la imitación. Siempre, en efecto, me ha fascinado el significativo conjunto de hechos que aureolan la capacidad de imitación. Recordemos brevemente tales hechos. Uno, el reconocimiento de que el esquema corporal propio y el ajeno son homólogos, que es el paso previo para cualquier imitación verdaderamente motora (es decir, no mera emulación de los resultados externos), empieza en la escala evolutiva justo en el chimpancé, como lo demuestra la capacidad que este y sólo este animal tiene de autorreconocerse en el espejo. Dos, absolutamente todos los animales se muestran incapaces de imitar secuencias motoras nuevas, por más que, como vimos, el chimpancé, y sólo él, pueda imitar posturas o movimientos simples, no secuenciales. Tres, la cultura humana, incluido el lenguaje, es aprendida, y, por tanto, imitada. Cuatro, la particular especialización hemisférica que, añadiéndose a la ampliamente compartida especialización por contralateralidad, es exclusiva del ser humano consiste en que los movimientos aprendi-

²⁴ Esta postura estaba ya en Johnson, C.N., “Structure of conscious experience”, en Astington, J.; Harris, P. y Olson, D., eds., *Developing theories of mind*, Cambridge, U.P. 1988. P. 58: “La simulación evita la necesidad de una teoría de la mente gracias a que aprovecha la estructura de la mente misma (...), concretamente la capacidad de colocar el yo de uno en los zapatos de otro.”

dos por imitación se ubican todos en el hemisferio llamado dominante por más que se hagan con la mano ipsilateral o con las cuerdas vocales. Pero, una vez que he intentado contagiar de mis prejuicios al lector, pasemos ya a ver cómo emplea Harris ese factor de la capacidad imitativa o simulación.

En el niño pequeño, nos dice, la capacidad de simulación o imitación de lo ajeno es todavía demasiado poco flexible. “Mientras mayor sea el número de escenarios *by default* (es decir, de los escenarios que se escogen en ausencia de oposición o alternativa) que requieran suspensión temporal, más difícil será la simulación” (Harris, 1991²⁵, p. 284). Hasta aquí no tengo nada que comentar.

Pero pasemos a leer su propuesta de captación simulatoria intrapersonal de la creencia falsa ajena. Empieza Harris por auto-objetarse que, dado que no hay ninguna cualidad introspectiva asociada con la creencia falsa, el éxito en la tarea de la falsa creencia parece inexplicable desde su propuesta de la proyección imaginativa o simulación. Pero a continuación replica que “una combinación de imaginación e introspección puede ser la respuesta. Consideremos la tarea de la falsa creencia con más detalle. El niño observa a un personaje depositando chocolate en una caja y marchándose. Un segundo personaje llega, coloca el chocolate en un sitio nuevo y se va. El primer personaje vuelve y se dispone a buscar el chocolate. Entonces se le hace al niño la pregunta clave: ¿dónde buscará el personaje? Los niños mayores de 4 años, que son los únicos que predicen correctamente que el personaje buscará donde el chocolate fue depositado originariamente, ¿cómo lo logran? Esos niños simplemente se imaginan lo que ellos harían si estuvieran en los zapatos del primer personaje. Específicamente, se imaginan a ellos mismos fuera de la habitación mientras el chocolate se cambia de sitio, e incapaces por tanto de ver tal cambio de sitio. Ellos se imaginan a sí mismos volviendo a la habitación y habiendo antes colocado el chocolate en su sitio original” (Harris, 1990²⁶ p. 224). ¿Qué nos parece esa explicación de Harris?

Lo primero que haremos será preguntarnos si acaso la capacidad de imaginarse en el emplazamiento y la postura ajenos, basta para proporcionar al sujeto la creencia falsa del primer personaje. Esta pregunta resulta tanto más interesante cuanto que el chimpancé es capaz de calibrar lo que un congénere en un determinado emplazamiento y una determinada postura verá o no verá. ¿Bastará entonces con aplicar dos veces esa simple capacidad para obte-

²⁵ Harris, P.L. “The work of the imagination”, en Whiten, A., ed. *Natural Theories of Mind*. Blackwell, Oxford, 1991.

ner la captación de la creencia falsa? En mi opinión, no. ¿Cómo ha de ser realmente ese imaginarse, ese imitar al primer personaje, si ha de dar como resultado la captación de la creencia falsa?. Se necesita que el niño haya mantenido la imitación del primer personaje durante toda la secuencia. La imitación ha tenido que ser, pues, imitación compleja y no de movimientos o posturas simples.

Pero, una vez puntualizado esto, centrémonos en el punto que venimos persiguiendo. Cuando el sujeto imita la creencia falsa captándola como falsa, ¿es acaso todavía ese estado mental un estado mental propio suyo? A mi entender, está claro que justo en ese momento —en el momento, repito, de captar la creencia como falsa, y de acceder, pues, a la metacreencia— el sujeto ha de tener ya desenganchada de sí mismo la creencia falsa. Sería entonces precisamente cuando la conciencia del tú surgiría, y también, a la vez, la conciencia propia exclusivamente humana. Pensemos que, en caso contrario, el sujeto se encontraría con dos creencias opuestas sobre la misma realidad, lo cual lo inhabilitaría para habérselas con su entorno. Así pues, podemos decir que en la captación de la creencia falsa ajena, la desembocadura, al menos, del proceso tiene que salir fuera de la introspección ‘de la primera persona’.

¿Y para las etapas anteriores del proceso? ¿Se puede mantener la introspección para explicar lo anterior a la desembocadura? ¿O, por el contrario, el mencionado proceso de imitación compleja tendría que apoyarse todo él sobre una ruptura previa de la interioridad, es decir, sobre un anterior desgajamiento que hubiera dado lugar a la conciencia del yo y a la conciencia del otro? La clave del dilema estriba, creo, en lo siguiente: Una imitación latente secuencial paralela al flujo sensorio-motor propio, ¿es acaso capaz por sí sola de hacer surgir una aprehensión que choque con la aprehensión que uno tiene sobre justo la misma realidad? Si se contestara que no —que la imitación latente secuencial no tendría capacidad de suscitar la distinción entre la conciencia del tú y la conciencia del yo—, entonces habríamos de proponer que esa distinción tendría que ser anterior y derivada de la comunicación lingüística. Pues bien, mientras que respecto al origen en el niño, creo que hay razones —ya lo mencioné— para resolver este dilema y dar primacía al acceso por vía de audición lingüística, respecto al origen histórico, en cambio, la

²⁶ Harris, P. L., “The child’s theory of mind and its cultural context”, en Butterworth, G. & P. Bryant, eds. *Causes of Development: Interdisciplinary Perspectives*, 1990, Harvester Wheatsheaf.

cuestión sigue abierta²⁷ Lo único que ahí se puede hacer es puntualizar que, en todo caso, tal imitación latente secuencial sería distinta, y, en concreto, mucho más demandante, que aquella otra clase de imitación latente secuencial por la que el sujeto se limita a imaginarse en otra situación y otro ámbito de actividad diferentes a los que en ese momento están vigentes para él. En efecto, mientras que la imitación secuencial o pauta imaginada discorra en un ámbito distinto al propio, no hay posibilidad de choque, ni, por tanto, de captación de una creencia falsa como falsa.

Pero, sea esto como sea, la condición interpersonal de la desembocadura del proceso está clara. Y así podemos proponer una síntesis entre el enfoque simulacionista y la consideración interpersonal del proceso. Así lo ha sugerido Tomasello, 1993²⁸, p. 502, quien acertadamente la llama 'imitación más vigotskiana' (no, por supuesto, porque Vigotski la describiera, sino porque en una tal imitación se cumpliría el Principio de Vigotski de que los procesos superiores se originan en lo interpersonal y después se intrapersonalizan). En efecto, el imitador accedería a un nivel metacognitivo sólo al imitar la conciencia ajena como ajena, es decir, como contradictoria e inasimilable con la suya propia, y esa capacidad de meta creencia se aplicaría a continuación a las propias creencias del sujeto pasadas (y más tarde, y en ocasiones excepcionales, incluso a las vigentes).

Como se ve, en torno al asunto de la conciencia de segundo orden, conforme se van contestando algunas preguntas, surgen otras muchas nuevas, y éstas son cada vez más concretas. ¿Se admite que eso es un buen indicio de su inminente madurez como objeto científico? Entonces las páginas anteriores habrían prestado cierto apoyo a lo que he propuesto es hoy la estrategia más rentable dentro de los estudios sobre la conciencia. Resumo en qué con-

²⁷ Y, si se aceptara también aquí la opción lingüística, ¿cuál sería el mínimo exigible de lenguaje para que pudiera originarse históricamente la captación de la creencia falsa ajena? Sería suficiente atribuir al 'creyente' la capacidad de producir mensajes protolingüísticos, es decir, mensajes que sean todavía holofrásticos, o sin sintaxis, y cuya función, expresada meramente por lo entonatorio, sea sólo de petición o llamada, pero que ya tengan semántica, una semántica verdadera capaz de discriminar si se trata de una llamada (llamada a madre, p.c.) o una petición (de cuidados maternos). La suficiencia de este mínimo es clave para que la opción lingüística siga abierta en el nivel histórico. En efecto, dado que la función comunicativa propia de la predicación es la de corregir la creencia que sobre el tema albergue (a juicio del hablante) el oyente, no podríamos invocar la audición de una predicativa para explicar el origen histórico de la captación de una creencia falsa ajena.

²⁸ Tomasello, M., Kruger, A.C., and Ratner, H.H., "Cultural learning", *Behavioral and Brain Sciences*, 1993, 16, 495-552.

siste tal estrategia. Primero, mantener bien separados los dos órdenes de estados mentales, reconociendo la sensibilidad experiencial de la mera percepción, pero sin atribuirle a tal sensibilidad una introspección reflexiva. Y segundo, y esto sería lo importante, concentrarse en la metacreencia, y en general, en los estados mentales de segundo orden, para cuyo estudio los de primer orden, o, dicho de otro modo, los no exclusivamente humanos, que pasarán en el futuro algún día a ser *explanandum*, ocuparían ahora, en cambio, sólo el papel de dato.