

La mente extendida

Andy CLARCK y David J. CHALMERS*

Traducción de Eva ALADRO

Propuesto: 12 de diciembre de 2010
Evaluado: 25 de diciembre de 2010
Aceptado: 12 de enero de 2011
(Abstracts y palabras clave al final del texto)

INTRODUCCIÓN

¿Dónde acaba la mente y dónde empieza el resto del mundo? La pregunta invita a dos respuestas típicas. Algunas personas aceptan las demarcaciones de la piel, y el cráneo, y afirman que lo que se halla fuera del cuerpo está también fuera de la mente. Otros sin embargo se inclinan por el argumento de que los significados de las palabras “simplemente no están en la cabeza” y sostienen este externalismo del significado, extendiéndolo a toda la mente. Nosotros queremos proponer una tercera posición. Abogamos por un tipo diferente de externalismo: un externalismo activo, basado en el papel activo que el entorno tiene en la consecución de los procesos cognitivos.

COGNICIÓN EXTENDIDA

Consideremos estos tres problemas humanos:

1. Una persona está sentada frente a una pantalla de ordenador que exhibe imágenes de diversas formas bidimensionales y se le pide que responda a preguntas sobre si encajan esas formas en una configuración que aparece dibujada. Para poder establecer ese encaje, la persona debe rotar mentalmente las formas y alinearlas con sus conexiones.
2. Una persona está sentada delante de una pantalla de ordenador similar, pero esta vez puede escoger entre rotar físicamente la imagen en la pantalla, presionando un botón de rotación, o hacerlo mentalmente como antes. Podemos

* Este artículo fue publicado en 1998 en la revista *Analysis* (58: 10-23). David J. Chalmers es investigador del Departamento de Filosofía de la Universidad de Washington y Andy Clarck es profesor de la Universidad de Arizona.

también suponer, con bastante verosimilitud, que habrá cierta ventaja de velocidad en la operación de rotación física.

3. En algún momento de un futuro *ciberpunk*, una persona está sentada frente a la pantalla del ordenador. Esta persona, sin embargo, disfruta de un implante neuronal que puede llevar a cabo la operación de rotación tan rápidamente como el ordenador del ejemplo anterior. Este agente debe de todos modos escoger qué recurso interno usar (el implante o la vieja y eficaz rotación mental), pues cada uno de estos recursos demandan diferente atención y otras actividades cerebrales concurrentes.

¿Cuán presente está la *cognición* en estos casos? Sugerimos nosotros que los tres casos son similares. El caso 3 del implante neurológico parece alinearse con el caso 1. Y el caso 2, del botón de rotación en el ordenador, parece mostrar el mismo tipo de tecnología computacional del caso 3, aunque la distribución del dispositivo entre el sujeto y el ordenador, es diferente cuando está internalizado dentro del sujeto. Si la rotación en el caso 3 es cognitiva, ¿cómo podemos considerar el caso 2 como diferente a éste? No podemos simplemente señalar la piel o el cráneo como límite cognitivo para justificarnos, pues la legitimidad de ese límite está precisamente en cuestión. Pero en todo lo demás parecen casos similares.

El tipo de caso que acabamos de describir no es en absoluto exótico como pudiera parecer. No es la presencia de los recursos avanzados digitales lo que ha suscitado este debate, sino más bien la tendencia de los seres humanos razonadores a depender profundamente de soportes ambientales. Así, consideremos el uso de un bolígrafo y de un papel para desarrollar una larga multiplicación (McClelland *et al.* 1986, Clark 1989), el uso de reconfiguraciones de piezas físicas con letras para activar la memoria léxica en el juego Scrabble (Kirsh 1995), el uso de instrumentos como la brújula náutica (Hutchins 1995) y la parafernalia general del lenguaje, los libros, los diagramas, y la cultura. En todos estos casos el cerebro individual lleva a cabo diferentes operaciones, algunas de las cuales se delegan a manipulaciones de medios externos. Si nuestros cerebros fueran diferentes, la distribución de tareas sin duda sería diversa.

De hecho, incluso los ejemplos de rotación mental descritos en los escenarios 1 y 2 son reales. Ambos reflejan opciones disponibles para los jugadores del juego de ordenador *Tetris*. En el *Tetris*, el descenso de formas geométricas debe dirigirse con rapidez hacia el lugar apropiado de una estructura de huecos que va emergiendo. Un botón de rotación se usa a este efecto. David Kirsh y Paul Maglio (1994), calculan que la rotación física de la forma en 90 grados se realiza en 100 milisegundos, sumados a los 200 milisegundos que precisa la selección del botón. Para conseguir el mismo resultado en una rotación mental tardaríamos unos 1000 milisegundos. Kirsh y Maglio continúan, presentando una evidencia arrolladora de cómo la rotación física no se usa simplemente para colocar la pieza en el lugar idóneo que rellene un hueco, sino para ayudar a *determinar* si la forma y el hueco son compatibles. Este uso constituye un caso de lo que Kirsh y Maglio denominan una “acción epistémica”. Las *acciones epistémicas* alteran o actúan en el mundo para ayudar y aumentar procesos cognitivos como el reconocimiento y la búsqueda. Las acciones meramen-

te *pragmáticas*, en cambio, alteran el mundo porque obedecen a la necesidad o deseo de un cambio físico estrictamente determinado (por ejemplo, rellenar un agujero de cemento).

La acción epistémica, sugerimos nosotros, implica la expansión del *crédito epistémico*. Si, cuando abordamos una tarea, una parte del mundo funciona como un proceso que, *si se hiciera dentro de nuestra cabeza*, no dudaríamos que es parte del proceso cognitivo, entonces esa parte del mundo (defendemos nosotros) es parte del proceso cognitivo. ¡Los procesos cognitivos no están (todos) dentro de la cabeza!.

EL EXTERNALISMO ACTIVO

En estos casos, el organismo humano se ve unido a una entidad externa en una interacción en dos direcciones, que crea un *sistema ensamblado* que podemos considerar un sistema cognitivo por derecho propio. Todos los componentes del sistema juegan un papel causal activo, y gobiernan conjuntamente la conducta, del mismo modo que lo hace la cognición usualmente. Si eliminamos el componente externo del sistema de competencia conductual, es como si elimináramos una parte del cerebro mismo. Nuestra tesis es que este tipo de proceso ensamblado es igualmente un proceso cognitivo, ocurra completamente dentro o no, de la cabeza.

Este externalismo difiere grandemente de la variedad típica que estudiaron Putnam (1975) y Burge (1979). Cuando yo creo que el agua está mojada y mi gemelo cree que el agua gemela está también mojada, los rasgos externos responsables de la diferencia en nuestras creencias son distales e históricos, situados al final de una larga cadena causal. Los rasgos del *presente* no son relevantes: si sucede que yo estoy rodeado por XYZ circunstancias (que he sido teletransportado a la Tierra Gemela de la célebre teoría), mis creencias sobre el agua mojada se mantendrán debido a mi propia historia anterior. En estos casos, los rasgos externos relevantes son *pasivos*. Por su naturaleza distal, no juegan ningún papel en la dirección de los procesos cognitivos en el aquí y ahora. Esto se refleja en el hecho de que las acciones realizadas por mí y por mi gemelo son físicamente indiscernibles, a pesar de nuestras diferencias externas.

En los casos que describimos nosotros, en cambio, los factores externos relevantes son *activos*, tienen un papel crucial en el aquí y ahora. Dado que existe una complementación con el organismo humano, tienen un impacto directo en el organismo y en su conducta. En estos casos, las partes relevantes del mundo están por así decir *en el ajo*, no colgando al final de una larga cadena de causas. Al reflexionar sobre este tipo de complementariedad nos vemos abocados a un *externalismo activo*, opuesto al externalismo pasivo de Putnam y Burge.

Muchos autores objetan que incluso si Putnam y Burge tienen razón sobre la externalidad del contenido, no está claro que esos aspectos externos tengan un papel causal o explicativo en la generación de la acción. En ejemplos contrafácticos en los que la estructura interna se mantiene constante pero estos factores externos se cambian, la conducta se mantiene igual; así la estructura interna parece estar realizando la tarea crucial. No trataremos este asunto aquí, pero notamos que el externalismo

activo no se ve amenazado por este problema. Los factores externos en un sistema ensamblado tienen un papel que no puede obviarse —si mantenemos la estructura interna y cambiamos los factores externos, la conducta puede cambiar completamente. Los factores externos son en este caso tan causalmente relevantes como los factores internos del cerebro¹.

Al abrazar el externalismo activo, nos permitimos una explicación más natural de todo tipo de acciones. Podemos explicar mi elección de palabras en el juego Scrabble, por ejemplo, como el resultado de un proceso cognitivo extendido que incluye la reconfiguración de piezas en mi bandeja del juego. Por supuesto, siempre podemos explicar mi acción en términos de procesos internos y de una serie de “inputs” y “acciones”, pero esta explicación será innecesariamente compleja. Si un proceso isomórfico tuviera lugar dentro de mi cabeza, no sentiríamos la necesidad de describirlo de esta manera tan pintoresca². En un sentido muy real, la reconfiguración de piezas en mi bandeja no es parte de una acción, sino parte del *pensamiento*.

El punto de vista que defendemos aquí se refleja en un creciente conjunto de investigaciones en ciencia cognitiva. En áreas tan diversas como la teoría de la cognición de situación (Suchman 1987), los estudios sobre robótica y mundo real (Beer 1989), el enfoque dinámico del desarrollo infantil (Thelen y Smith 1994) y la investigación sobre las propiedades cognitivas de colectivos (Hutchins 1995), la cognición se contempla como un continuo con procesos en el entorno³. Así, al entender la cognición como un proceso extendido, no solamente tomamos una decisión terminológica, sino que establecemos una diferencia significativa en la metodología de la investigación científica. En efecto, los métodos explicativos que en su día parecían apropiados para estudiar únicamente los procesos “internos” hoy son adaptados para el estudio de los procesos exteriores, y con ello parece claro que nuestra comprensión de la cognición será más rica que hasta ahora.

Algunos autores encuentran este tipo de externalismo inaceptable. Una razón puede ser que muchos identifican lo cognitivo con lo consciente, y parece lejos de lo plausible que la consciencia se extienda más allá de la cabeza en estos casos. Pero no todos los procesos cognitivos, al menos en el uso común, son procesos conscientes. Es ampliamente aceptada la idea de que todo tipo de procesos más allá de los límites de la consciencia juegan un papel crucial en el procesamiento cognitivo: en

¹ La atracción que experimenta el externalismo de la mente en Filosofía puede provenir del atractivo intuitivo del externalismo activo. Los externalistas a menudo hacen analogías con sistemas pareados que implican factores externos, y apelan a la arbitrariedad de los límites entre cerebro y entorno. Pero estas intuiciones encajan mal con la teoría estricta del externalismo típico. En la mayoría de los ejemplos de Putnam y Burge, el entorno inmediato es irrelevante; solamente cuenta el entorno o contexto histórico. El debate se centró siempre en la cuestión de si la mente está dentro de la cabeza, pero quizás sea más relevante otra cuestión relacionada con estos ejemplos, como es: ¿pero está la mente en el presente?

² Herbert Simon (1981) sugería una vez que consideráramos la memoria interna como, en efecto, una fuente externa sobre la que operan procesos “realmente” internos. “La búsqueda memorística”, argumenta, “no es muy diferente a la búsqueda en el entorno externo”. La idea de Simon tiene al menos la virtud de tratar los procesos internos y externos con la paridad que merecen, pero sospechamos que en ella la mente se queda muy pequeña para la mayoría de los gustos humanos.

³ Teorías filosóficas de un similar espíritu se hallarán en Haugeland 1995, McClamrock 1985, Varela *et al.* 1991 y Wilson 1994.

la recuperación de recuerdos, en los procesos lingüísticos, en la adquisición de habilidades, por ejemplo. Así, el mero hecho de que los procesos externos sean externos cuando la consciencia es interna no es razón para negar que dichos procesos sean cognitivos.

Es más interesante el argumento de que lo que sostiene los procesos de cognición real en la cabeza es el requisito de que los procesos cognitivos sean *portátiles*. Aquí, nos vemos dentro de una visión de lo que podríamos llamar la Mente Desnuda: existe un conjunto de recursos y operaciones que podemos llevar con nosotros para que desempeñen una tarea cognitiva, independientemente del entorno local. En esta teoría, el problema ante sistemas ensamblados es que se *desensamblan* demasiado fácilmente. Los verdaderos procesos cognitivos son aquellos que se mantienen en el núcleo constante del sistema; cualquier otra cosa es un extra añadido.

Existe algo importante en esta objeción. El cerebro (o el cerebro y el cuerpo) comprenden un conjunto de recursos básicos, portátiles, que tienen interés en sí mismos. Estos recursos pueden incorporar acciones corporales en procesos cognitivos, como cuando usamos nuestros dedos como memoria de trabajo en un cálculo complicado, pero no incluirán factores más contingentes de nuestro entorno externo, como una calculadora de bolsillo. Aún más, la simple contingencia de establecer un sistema ensamblado no implica un estatus cognitivo. En un futuro lejano puede que podamos enchufar diversos módulos a nuestro cerebro para asistirnos: un módulo para una memoria a corto plazo extra cuando sea necesaria, por ejemplo. Cuando se conecta un módulo, los procesos implicados son tan cognitivos como si hubieran estado ahí todo el tiempo⁴.

Incluso aunque aceptáramos el giro del criterio de portabilidad, el externalismo activo no se vería minado. Contar con los dedos siempre se ha considerado algo aceptado en el proceso cognitivo, y es fácil llevar las cosas más allá. Pensemos en la vieja imagen del ingeniero que lleva en el cinturón su regla allá donde va. ¿Qué pasa si la gente siempre lleva consigo una calculadora, o se la implanta? El criterio verdadero de la intuición de la portabilidad es que en los sistemas ensamblados en el núcleo de la cognición, se requiere un ensamblaje del sistema *fiabile*. Ocurre que la mayoría de los sistemas complementarios tienen lugar dentro del cerebro, pero puede fácilmente aparecer un sistema ensamblado fiable ligado al entorno mismo. Si los recursos de mi calculadora o mi agenda Filofax están siempre disponibles cuando los necesito, se emparejan conmigo tan fiablemente como precise. En efecto, son parte de mi conjunto básico de recursos cognitivos que llevo encima en mi vida cotidiana. Estos sistemas no pueden ser impugnados simplemente sobre la idea de que podrían sufrir un daño concreto, dejar de funcionar o perderse, o porque con ello se produciría algún desensamblaje ocasional: el cerebro biológico también sufre estos

⁴ O considérese el siguiente pasaje de una novela reciente de ciencia ficción (McHugh 1992: 213): “Soy conducido al departamento de sistemas donde me conectan con el mismo. Todo lo que hago es conectarme y entonces el técnico instruye al sistema para que se adapte, y así lo hace. Me desconecto y pregunto la hora. Las 10 y 52. Aparece la información. Siempre antes de que pueda acceder a ella cuando estoy conectado, tengo la impresión de que sabía lo que iba a pensar y lo que el sistema me iba a decir, pero ahora, ¿cómo sé lo que es el sistema y lo que es Zhang?”

peligros, y en ocasiones pierde temporalmente capacidades en episodios de sueño, intoxicación, y emoción. Si las capacidades relevantes están generalmente donde se precisan, existe eficientemente un sistema ensamblado.

Además, puede que el cerebro biológico haya de hecho evolucionado y madurado de maneras que tienen que ver con la presencia fiable de un entorno externo manipulable. Parece ciertamente que la evolución ha favorecido capacidades de almacenamiento que han sido especialmente orientadas a parasitar el entorno local con la finalidad de reducir la carga de memoria interna, e incluso para transformar la naturaleza de los problemas computacionales mismos. Nuestros sistemas visuales han evolucionado para depender del entorno en diversas formas: aprovechan los hechos contingentes derivados de la estructura de las escenas naturales (*vid.* Ullman y Richards 1984), por ejemplo, y se benefician de los atajos computacionales para el movimiento y locomoción corporal (*vid.* Blake y Yuille 1992). Quizás hay otros casos en que la evolución ha encontrado ventajoso explotar la posibilidad de que el entorno esté integrado en la cognición. Si es así, entonces los sistemas ensamblados con factores externos son parte del conjunto básico de recursos cognitivos que siempre llevamos con nosotros en el mundo.

El lenguaje puede ser un ejemplo. El lenguaje parece ser un medio central mediante el cual los procesos cognitivos se extienden al mundo. Pensemos en un grupo de gente reunida en una mesa para generar ideas, o en un filósofo que piensa mejor escribiendo, desarrollando sus ideas conforme lo hace. Puede que el lenguaje evolucionara, en parte, para permitir estas extensiones de nuestros recursos cognitivos dentro de sistemas activos de complementariedad con el entorno.

Dentro del tiempo vital de un organismo, igualmente, el aprendizaje individual puede haber conformado el cerebro de modos que tienen que ver con las extensiones cognitivas que le han rodeado conforme iba aprendiendo. El lenguaje es de nuevo un ejemplo crucial, como lo son los variados artefactos físicos y digitales que rutinariamente se usan como extensiones cognitivas por los niños en los colegios y por quienes se forman en numerosas profesiones. En estos casos el cerebro se desarrolla de un modo que complementa las estructuras externas, y aprende a hacer su papel dentro de un sistema unificado y densamente ensamblado. Una vez reconocemos el papel crucial del entorno para forzar la evolución y el desarrollo de la cognición, vemos que la cognición extendida es un proceso cognitivo central, y no un elemento adicional extra.

Puede ayudarnos una analogía. La extraordinaria eficacia del pez, como dispositivo de nado, se debe en parte, según sabemos ahora, a una capacidad evolucionada de ensamblar o integrar su conducta natatoria con los entornos de energía cinética externa como los remolinos, turbulencias o vórtices del mundo acuático (*vid.* Triandafilú y G. Triandafilú 1995). Esos vórtices incluyen los que se generan naturalmente (por ejemplo cuando el agua golpea una roca) y los que el propio pez puede inducir (creándolos oportunamente con sus aletas). El pez nada construyendo esos procesos externos en el mismo corazón de sus rutinas locomotoras. El pez y los vórtices que le rodean constituyen en conjunto una máquina de natación unificada y notablemente eficaz.

Ahora consideremos un futuro fiable en el entorno humano, como es el mar de las palabras. El entorno lingüístico nos rodea desde que nacemos. En esas condiciones, el plástico cerebro humano llega a tratar esas estructuras como recurso fiable

para usarse en la conformación de las rutinas cognitivas más obvias. Donde las aletas del pez crean turbulencias y vórtices que luego aprovechará en el nado, nosotros intervenimos en múltiples medios lingüísticos, creando estructuras locales y turbulencias cuya fiable presencia conduce nuestros procesos internos en curso. Las palabras y los símbolos externos son por tanto miembros de pleno derecho entre los vórtices cognitivos que nos ayudan a constituir el pensamiento humano.

DE LA COGNICIÓN A LA MENTE

Hasta aquí hemos hablado largamente del “proceso cognitivo”, argumentando que se extiende al entorno. Algunos pueden pensar que la conclusión se ha alcanzado con muy poco esfuerzo. Quizás tenga lugar algún *proceso* en el entorno, pero ¿y en la *mente*? Todo cuanto hemos dicho hasta ahora es compatible con la perspectiva de que los verdaderos estados mentales —experiencias, creencias, deseos, emociones, y demás— están todos determinados por estados cerebrales. Quizás todo lo verdaderamente mental sea interno, después de todo, ¿o no?

Nos proponemos dar un paso más adelante. Mientras algunos estados mentales, como las experiencias, pueden estar determinados internamente, hay otros casos en que los factores externos contribuyen significativamente. En particular, argumentaremos que las *creencias* pueden constituirse en parte por factores del entorno, cuando dichos factores tienen el papel apropiado en la dirección de los procesos cognitivos. Si es así, la mente se extiende al mundo.

Primero, consideremos un caso normal de creencia implantada en la memoria. Inga oye decir a un amigo que hay una exposición en el Museo de Arte Moderno, y decide ir a visitarla. Piensa por un momento y recuerda que el museo está en la calle 53, así que va a pie a la calle 53 y entra en el museo. Parece claro que Inga cree que el museo está en la calle 53, y que esto lo creía incluso antes de consultar a su memoria. No se trata de una creencia sobrevenida, como tampoco lo son la mayoría de nuestras creencias. La creencia estaba alojada en algún lugar de la memoria, esperando para que se accediera a ella.

Ahora veamos el caso de Otto. Otto padece Alzheimer, y como otros enfermos de este mal, confía en la información del entorno para ayudarse a estructurar su vida. Otto lleva consigo un ordenador portátil allá donde vaya. Cuando aprende nueva información, la escribe. Cuando necesita recuperar información previa, la consulta en él. Para Otto, su portátil hace el papel que solía desempeñar su memoria biológica. Hoy Otto, al enterarse de la exposición que hay en el Museo de Arte Moderno, ha decidido ir a verla. Consulta su portátil, que le informa de que el museo está en la calle 53, de modo que va a pie a dicha calle y entra en el museo.

Claramente Otto ha paseado hasta la calle 53 porque deseaba visitar el museo y creía que ese museo estaba en la calle 53. Igual que Inga, quien tenía esa misma creencia incluso antes de consultar a su memoria, parece razonable afirmar que Otto cree que el museo está en la calle 53 incluso antes de consultar su ordenador. Pues en importantes aspectos los dos casos son enteramente análogos: el portátil tiene para Otto el mismo papel que la memoria juega para Inga. La información en el por-

tátil funciona igual que la información que forma parte de una creencia ordinaria no presente; la diferencia está en que esa información está más allá de la piel.

La alternativa es afirmar que Otto no tiene creencia hasta que no consulta su ordenador; como mucho, cree que el museo está ubicado en la dirección que el ordenador le indicará. Pero si seguimos a Otto durante algún tiempo, veremos lo antinatural que es esta manera de hablar. Otto usa constantemente su *notebook* para todos los asuntos. Es central en sus acciones en todo tipo de contextos, del mismo modo que la memoria ordinaria es central en la vida cotidiana. La misma información puede recuperarse una y otra vez, quizás modificándose ligeramente para cada ocasión, antes de volverse a guardar en los fondos de su memoria artificial. Decir que las creencias desaparecen cuando se cierra el ordenador parece igual de equivocado que afirmar que las creencias de Inga desaparecen en cuando no es ya consciente de ellas. En ambos casos la información es fiable cuando se necesita de ella, disponible para la consciencia y disponible para guiar la acción, del mismo modo en que esperamos que lo sea una creencia.

Ciertamente, en tanto las creencias y los deseos están caracterizados por su papel explicativo, los ejemplos de Otto y de Inga parecen ser análogos: la misma causalidad dinámica esencial en los dos casos los hace idénticos. Nos satisface explicar la acción de Inga en términos de su deseo presente de ir al museo y su creencia sólida en que el museo está en la calle 53, y también nos convence explicar la acción de Otto del mismo modo. La alternativa es explicar las acciones de Otto en términos de su deseo presente de ir al museo, su creencia sólida en que el Museo está en la dirección que está escrita en su portátil, y el hecho accesible de que el portátil afirma que el museo está en la calle 53; pero esto complica la explicación innecesariamente. Si tenemos que recurrir a explicar la acción de Otto de esta manera, entonces debemos hacer lo mismo con el infinito número de acciones en las que utiliza el ordenador igualmente; en cada una de las explicaciones, existirá un término extra relacionado con el portátil. Creemos que explicar las cosas de este modo es *llevarlas demasiado lejos*. Es innecesariamente complejo, del mismo modo en que es innecesariamente complejo explicar las acciones de Inga en términos de creencias sobre su memoria. El ordenador es una constante para Otto, del mismo modo que la memoria es una constante para Inga; si lo señalamos para explicar cada creencia/deseo será redundante. En una explicación, la sencillez es poder.

Si esto es cierto, podemos incluso construir el caso de un Otto gemelo, que es justamente igual que Otto excepto en que hace algún tiempo escribió en su ordenador erróneamente que el Museo de Arte Moderno estaba en la calle 51. Hoy, el Otto gemelo es un duplicado físico del otro, de piel para dentro, pero su portátil difiere. Consecuentemente, el Otto gemelo puede describirse con la creencia de que el museo está en la calle 51, mientras que Otto cree que está en la 53. En estos casos, la creencia no está simplemente en la cabeza.

Se trata de una conclusión idéntica a la de Putnam y Burge, pero de nuevo existen importantes diferencias. En el caso de Putnam y Burge, los factores externos que constituyen diferencias en la creencia, son distales e históricos, de modo que los gemelos en esos casos producen físicamente conductas no distinguibles entre sí. En los casos que describimos nosotros, los factores externos relevantes tienen un papel

activo en el aquí y ahora, y un impacto directo en la conducta. Donde Otto va a pie hasta la calle 53, el gemelo va a la 51. No hay una irrelevancia explicativa de este tipo de contenido de creencia externa; se introduce precisamente por el papel central explicativo que tienen en la conducta. Como los casos de Putnam y Burge, implican diferencias en la referencia y en las condiciones de verdad, pero también implican diferencias en la dinámica de la *cognición*⁵.

La conclusión es que cuando tratamos la creencia, no hay nada sagrado en el cráneo o la piel. Lo que hace que una información sea una creencia es el papel que juega, y no hay razón por la cual ese papel relevante solamente pueda ser desempeñado desde dentro del cuerpo.

Algunos se resistirán a esta conclusión. Una oponente puede salir a escena e insistir que cuando ella utiliza el término “creencia”, o quizás incluso de acuerdo con el uso típico, simplemente no puede decirse que encaje en lo que Otto hace cuando cree que el museo esté en la 53. No nos interesa debatir lo que constituye el uso típico; nuestra perspectiva amplia es la noción de creencia que *debería* usarse cuando Otto parece estar teniendo dicha creencia en cuestión. En todos los aspectos *importantes*, el caso de Otto es similar a un caso típico de creencia (no sobrevenida). Las diferencias entre el caso de Otto y el de Inga son llamativas, pero son superficiales. Al usar la noción de “creencia” en sentido amplio, resulta algo cercano a un tipo natural de la misma. La noción se vuelve más profunda y más unificada, y más útil en la explicación.

Para mostrar una resistencia sustancial, un opositor a la tesis debe mostrar que los casos de Otto y de Inga difieren en algún aspecto importante y relevante. Pero ¿en qué aspectos profundos son los dos casos diferentes? Argumentar únicamente en función de que la información está en la cabeza en un caso, pero no en el otro, no es sólido. Si esa referencia es relevante para causar una diferencia en la creencia, seguramente no es *primitivamente* relevante. Justificar el trato diferente requiere hallar alguna diferencia básica subyacente entre los dos.

Puede sugerirse que los casos muestran diferencias relevantes porque Inga tiene un acceso más *fiabile* a la información. Después de todo, alguien puede robarle a Otto su portátil, pero la memoria de Inga es más segura. No es implausible que la constancia es relevante: de hecho, el que Otto siempre utilice su *notebook* juega cierto papel en nuestra justificación de su estatus cognitivo. Si Otto estuviera consultando una guía en papel no le adscribiríamos la categoría de una creencia. Pero en el caso original, el acceso de Otto a su portátil es muy fiable, no absolutamente fiable, pero tampoco lo es el acceso de Inga a su propia memoria. Una cirugía podría limitar su cerebro, o más mundanamente, podría ser que hubiera bebido mucho. La mera posibilidad de tal limitación no basta para negar que ella tenga dicha creencia.

Uno puede pensar que el acceso de Otto a su ordenador es algo que *de hecho* viene y va. Por ejemplo, cuando se ducha no lo tiene consigo, y no puede leerlo

⁵ En la terminología de Chalmers “Los componentes del contenido” (léase): los gemelos en el caso de Putnam y Burge, difieren solamente en su contenido relacional, pero Otto y su gemelo difieren claramente en su contenido nocional, que es el tipo de contenido que gobierna la cognición. El contenido nocional es generalmente interno a un sistema cognitivo, pero en este caso el sistema cognitivo se ha extendido efectivamente para incluir un ordenador en él.

cuando está a oscuras. Seguramente su creencia no viene y va tan fácilmente, ¿no?. Para solucionar este problema podríamos redescibir la situación, pero en cualquier caso una desconexión temporal no amenaza nuestra idea. Después de todo, cuando Inga duerme, o cuando está intoxicada, no decimos que su creencia desaparece. Lo que realmente importa es que la información esté fácilmente disponible cuando el sujeto lo necesita, y que pueda satisfacerse su necesidad igualmente en ambos casos. Si el portátil de Otto estuviera frecuentemente inaccesible cuando su información fuera útil para él, esto sería un problema, pues la información no podría realizar el papel de guía de la acción que es central en la creencia; pero en las situaciones más importantes sí está disponible, de modo que la creencia no sufre peligro alguno.

Quizás una diferencia esté en que Inga tiene *mejor* acceso a la información que Otto. Los procesos “centrales” de Inga y su memoria probablemente tienen una conexión de banda ancha entre sí, comparada con la conexión de baja calidad que une a Otto a su portátil. Pero esto solamente no constituye una diferencia entre creer y no creer. Consideremos que una amiga de Inga, Lucy, que la acompaña al museo, tiene una conexión igualmente débil de su memoria biológica con sus sistemas centrales, bien sea por capacidades biológicas disminuidas o por alguna desventura en el pasado. El procesamiento en el caso de Lucy puede ser menos eficiente, pero en tanto la información relevante esté accesible, claramente Lucy cree que el museo está en la calle 53. Si la conexión fuera muy indirecta —si Lucy tuviera que esforzarse mucho en recuperar esa información, con resultados confusos, o si se necesitara la ayuda de un psicoterapeuta—, podríamos volvernos más reticentes a adscribir la creencia, pero estos casos están más allá de la situación de Otto, en la que la información es fácilmente accesible.

Otra sugerencia podría ser que Otto tiene acceso a la información relevante solo mediante la *percepción*, mientras que Inga accede directamente a ella mediante la introspección. En cierta forma, sin embargo, poner las cosas así implica negar las premisas mayores. Después de todo, nosotros abogamos efectivamente por un punto de vista en el que los procesos internos de Otto y su ordenador constituyen un solo sistema cognitivo. Desde el punto de vista de este sistema, el flujo de información entre ordenador y cerebro no es perceptual en absoluto; no implica el impacto de algo fuera del sistema. Es un proceso más cercano al propio flujo de información dentro de un cerebro. El único modo profundo en que este proceso es perceptual es en el caso de Otto, en que hay una fenomenología de la percepción distintiva asociada a la recuperación de información, mientras que en el caso de Inga no es así. Pero ¿por qué debería la naturaleza de una fenomenología asociada marcar una diferencia en el estatus de una creencia? La memoria de Inga puede tener una fenomenología perceptiva asociada, pero sigue siendo una creencia. La fenomenología no es visual, está claro. Pero para una fenomenología visual consideremos a Terminator, de la película de Arnold Schwarzenegger del mismo nombre. Cuando éste recupera cierta información de su memoria, ésta se “exhibe” ante él en su campo visual (presumiblemente él es consciente de ello, pues hay varias secuencias que muestran su propio punto de vista). El hecho de que recuerdos existentes se recuperen de este modo inusual no crea ninguna diferencia en su estatus de recuerdos existentes.

Estas variadas pequeñas diferencias entre los casos de Otto y de Inga son todas ellas *superficiales*. Fijarse en ellas sería perder el modo como para Otto, las entradas de su *notebook* tienen justamente el mismo papel que las creencias juegan en la guía de la mayoría de las vidas de la gente.

Quizás la intuición de que para Otto no se trata de una creencia verdaderamente proviene de un sentimiento residual de que las únicas creencias verdaderas son las que están presentes. Si tomamos este sentimiento en serio, tampoco la creencia de Inga sería auténtica, como muchas de las creencias que tenemos cada día. Ello sería una forma extrema, pero sin duda la más convincente, de negar la creencia de Otto. Incluso en una perspectiva menos extrema, la de que una creencia debe estar *disponible* para la consciencia, por ejemplo, las entradas del portátil de Otto encajan tan correctamente como los recuerdos de Inga. Una vez que aceptamos las creencias disponibles, es difícil resistirse a la conclusión de que el *notebook* de Otto tiene toda la disponibilidad relevante.

MÁS ALLÁ DE LOS LÍMITES EXTERIORES

Si esta tesis se acepta, ¿hasta dónde la llevaremos? Se nos plantean todo tipo de casos paradójicos. ¿Qué pasaría con los pueblerinos amnésicos de *Cien años de soledad*, que olvidaban los nombres de todo y por ello colocaban etiquetas a todo? ¿Y la información en mi Filofax, cuenta como parte de mi memoria? Si el *notebook* de Otto es manipulado, ¿creerá en la información que se instale en él? ¿Creo yo en los contenidos de la página que tengo delante conforme la leo? ¿Se expande mi estado cognitivo de algún modo en Internet?

No pensamos que haya respuestas categóricas a todas estas preguntas, y no las vamos a dar. Pero para ayudar a comprender lo que implica la asignación de creencias extendidas, podemos al menos examinar los rasgos de nuestro caso central que nos hizo aplicar la noción con tanta facilidad. Primero, el portátil es una constante en la vida de Otto, en las situaciones en que la información que contiene sea relevante, él rara vez emprendería una acción sin consultarlo previamente. En segundo lugar, la información en el ordenador está directamente disponible sin dificultad. Tercero, al recuperar información del mismo automáticamente él la asume. Cuarto, la información del ordenador ha sido asumida conscientemente en algún momento del pasado, y sin duda está ahí como consecuencia de dicha incorporación⁶. El estatus del cuarto rasgo como criterio para la creencia es discutible (¿quizá podemos adquirir creencias mediante la percepción subliminal, o mediante la manipulación de la memoria?) pero los tres primeros ciertamente tienen un papel crucial.

⁶ La constancia del criterio de la incorporación en el pasado puede sugerir que la historia es en parte constitutiva de la creencia. Puede reaccionarse a esto eliminando cualquier componente histórico (dando al criterio de la constancia una lectura puramente disposicional y eliminando el criterio de la incorporación pasada, por ejemplo), o bien puede permitirse dicho componente siempre que la carga esencial la tengan los factores del presente.

En tanto los exóticos casos paradójicos carezcan de estos rasgos, la aplicabilidad de la noción de “creencia” no podrá sostenerse. Si yo llevo a cabo muy rara vez una acción relevante sin consultar mi agenda Filofax, por ejemplo, su estatus en mi sistema cognitivo será como el del ordenador de Otto. Pero si yo actúo a menudo sin consultarla, por ejemplo, si yo a veces contesto preguntas relevantes que “no sé”, entonces la información que contiene es en menor medida parte de mi sistema de creencia. Internet normalmente no será fiable, a menos que yo confíe muchísimo en ella, esté adaptado a la tecnología, y además sea confiado por naturaleza, pero la información contenida en algunos de mis archivos del ordenador sí lo será. En casos intermedios, la cuestión de si una creencia está presente puede quedar indeterminada, o la respuesta depender de los variados criterios que aparecen en varios contextos en que se plantee esta cuestión. Pero cualquier indeterminación en esto no significa que en los casos centrales la respuesta no sea clara.

¿Y qué hay en el caso de la cognición socialmente extendida? ¿Pueden estar mis estados mentales en parte constituidos por los estados de otros pensadores? No vemos razón para que no sea así, en principio. En una pareja inusualmente interdependiente, es completamente posible que las creencias de uno de ellos tengan la misma función en el otro que el *notebook* tiene para Otto⁷. Lo esencial es el alto grado de confianza, fiabilidad y accesibilidad. En otras relaciones sociales estos criterios pueden no satisfacerse con tanta claridad, pero en cualquier caso sí pueden serlo en dominios específicos. Por ejemplo, el camarero de mi restaurante favorito puede actuar como un auténtico almacén de mis creencias sobre mis comidas favoritas (por manipulado que ello sea conforme a un deseo extendido). En otros casos, las creencias de uno pueden estar incorporadas por el secretario de uno, por su contable, por su colaborador⁸.

En cada uno de estos casos, el peso mayor del ensamblaje entre agentes lo tendrá el lenguaje. Sin el lenguaje, seríamos mucho más parecidos a la mentes “internas” discretas y cartesianas, en las que la cognición de alto nivel depende en gran medida de recursos internos. Pero la llegada del lenguaje nos ha permitido extender esa carga al mundo. El lenguaje, así entendido, no es un espejo de nuestros estados internos sino un complemento a los mismos. Sirve como una herramienta cuya misión es extender la cognición de forma que no pueden conseguirlo los dispositivos internos. Indudablemente, puede ser que la explosión intelectual de las épocas evolutivas recientes se deba en gran medida a esta extensión lingüística hecha posi-

⁷ Citado del *New York Times* del 30 de marzo de 1995, p. B7 en un artículo sobre un antiguo entrenador de baloncesto de UCLA John Wooden: “Wooden y su mujer asistieron a los Cuartos de Final del 36, y ella le servía de banco de memoria. Nell Wooden rara vez olvidaba un nombre, y su marido rara vez lo recordaba, así que en los pasillos de la final, ella reconocía a la gente para él”.

⁸ ¿Puede este tipo de razonamiento permitir algo como la “artritis” de las creencias de Burge? Después de todo, siempre puedo desviar a mi médico las acciones relevantes que emprender de acuerdo con mi enfermedad. Puede ser, pero hay algunas diferencias claras. Por ejemplo, cualquier creencia extendida debe basarse en una relación activa existente con el médico, más que en una relación histórica con una comunidad lingüística. Y en el análisis actual, mi delegación en el médico tenderá a depositar en él algo similar a una creencia auténtica de que tengo alguna otra enfermedad en mi muslo, y no la falsa creencia de que padezco artritis ahí. Por otra parte, si uso a expertos médicos solamente como consultores terminológicos, los resultados del análisis de Burge pueden repetirse.

ble como cualquier otro desarrollo independiente en nuestros recursos cognitivos interiores.

¿Y qué podemos decir, finalmente, del sí mismo? ¿Implica la mente extendida un sí mismo extendido? Así parece. La mayoría de nosotros ya acepta que el sí mismo excede los límites de la consciencia; mis creencias disposicionales, por ejemplo, constituyen en cierto sentido profundo parte de lo que soy. Si es así, entonces esos límites pueden también ir más allá de la piel. La información en el portátil de Otto, por ejemplo, es una parte central de su identidad como agente cognitivo. Ello implica que Otto *mismo* debe contemplarse como un sistema extendido, un ensamblaje de un organismo biológico y de recursos externos. Para oponernos consistentemente a esta conclusión, tendríamos que reducir el sí mismo a un simple montón de estados presentes, severamente amenazados en su continuidad psicológica profunda. Mucho mejor es tomar una visión más amplia, y ver a los agentes mismos extendidos en el mundo.

Como con cualquier reconcepción de nosotros mismos, esta visión deberá tener significativas consecuencias. Hay consecuencias obvias para las visiones filosóficas de la mente y para la metodología de investigación en ciencia cognitiva, pero también efectos en los campos moral y social. Puede que, por ejemplo, en algunos casos que interfirieran en el entorno de una persona, existan consecuencias morales como las que aparecen relacionadas con ella. Y si se toma verdaderamente en serio esta perspectiva, ciertas formas de actividad social deberían reconcebirse como menos próximas a la comunicación y la acción y más cercanas al pensamiento. En cualquier caso, una vez la hegemonía de la piel y el cráneo haya sido usurpada, podemos ser capaces de vernos verdaderamente como criaturas del mundo.

REFERENCIAS BIBLIOGRÁFICAS

- BEER, R. 1989. *Intelligence as Adaptive Behavior*. New York: Academic Press.
- BLAKE, A. & YUILLE, A. (eds) 1992. *Active Vision*. Cambridge, MA: MIT Press.
- BURGE, T. 1979. Individualism and the mental. *Midwest Studies in Philosophy* 4:73-122.
- CLARK, A. 1989. *Microcognition*. MIT Press.
- HAUGELAND, J. 1995. Mind embodied and embedded. In (Y. Houn and J. Ho, eds.), *Mind and Cognition*. Taipei: Academia Sinica.
- HUTCHINS, E. 1995. *Cognition in the Wild*. Cambridge, MA: MIT Press.
- KIRSH, D. 1995. The intelligent use of space. *Artificial Intelligence* 73:31-68.
- KIRSH, D. & MAGLIO, P. 1994. On distinguishing epistemic from pragmatic action. *Cognitive Science* 18:513-49.
- MCCLAMROCK, R. 1995. *Existential Cognition*. Chicago: University of Chicago Press.
- MCCLELLAND, J.L, D.E. RUMELHART, & G.E. HINTON 1986. The appeal of parallel distributed processing". In (McClelland & Rumelhart, eds) *Parallel Distributed Processing, Volume 2*. Cambridge, MA: MIT Press.
- MCHUGH, M. 1992. *China Mountain Zhang*. New York: Tom Doherty Associates.
- PUTNAM, H. 1975. The meaning of 'meaning'. In (K. Gunderson, ed) *Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press.

- SIMON, H. 1981. *The Sciences of the Artificial*. MIT Press.
- SUCHMAN, L. 1987. *Plans and Situated Actions*. Cambridge, UK: Cambridge University Press.
- THELEN, E. & SMITH, L. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*.
- TRIANAFYLLOU, M. & TRIANAFYLLOU, G. 1995. An Efficient Swimming Machine. *Scientific American* 272(3):64-70.
- ULLMAN, S. & RICHARDS, W. 1984. *Image Understanding*. Norwood, NJ: Ablex.
- VARELA, F., THOMPSON, E. & ROSCH, E. 1991. *The Embodied Mind*. Cambridge, MA: MIT Press.
- WILSON, R. 1994. Wide computationalism. *Mind* 103:351-72.

RESUMEN

¿Dónde acaba la mente y dónde empieza el resto del mundo?. La pregunta invita a dos respuestas típicas. Algunas personas aceptan las demarcaciones de la piel, y el cráneo, y afirman que lo que se halla fuera del cuerpo está también fuera de la mente. Otros sin embargo se inclinan por el argumento de que los significados de las palabras “simplemente no están en la cabeza” y sostienen este externalismo del significado, extendiéndolo a toda la mente. Nosotros queremos proponer una tercera posición. Abogamos por un tipo diferente de externalismo: un externalismo activo, basado en el papel activo que el entorno tiene en la consecución de los procesos cognitivos.

Palabras clave: Mente extendida, cognición, externalismo, tecnologías, mente distribuida

ABSTRACT

Where does the mind stop and the rest of the world begin? The question invites two standard replies. Some accept the intuitive demarcations of skin and skull, and say that what is outside the body is outside the mind. Others are impressed by the arguments of Putnam and Burge that the truth-conditions of our thoughts “just ain’t in the head”,[*] and hold that this externalism about meaning carries over into an externalism about mind. We propose to pursue a third position. We will advocate an externalism about mind, but one that is in no way grounded in the debatable role of truth-conditions and reference in fixing the contents of our mental states. Rather, we advocate an active externalism, based on the active role of the environment in driving cognitive processes.

Key words: Extended mind, distributed mind, cognition, technologies, externalism.

RÉSUMÉE

Où commence la pensée et où le reste du monde? La question nous invite à deux réponses communes. Quelques uns acceptent les démarcations intuitives de la peau et le crâne, en affirmant que ce que on voit dehors du corps est aussi dehors de la pensée. Des autres sont frappés par les arguments de Hillary Putnam et de Burge sur les conditions de vérité de nos pensées, que ne “sont pas dans la tête”, et soutiennent que cet externalisme du sens implique l’externalisme de la pensée. Nous proposons une troisième position, à faveur du externalisme de la pensée, mais pas fondé sur la discussion des conditions de vérité et de la référence dans la fixation des états mentaux. Au fait nous proposons un externalisme actif, fondé dans le rôle actif du environnement dans la direction des procès cognitifs.

Mots clé: Externalisme, pensée extendue, cognition, technologies, distribution de la pensée.