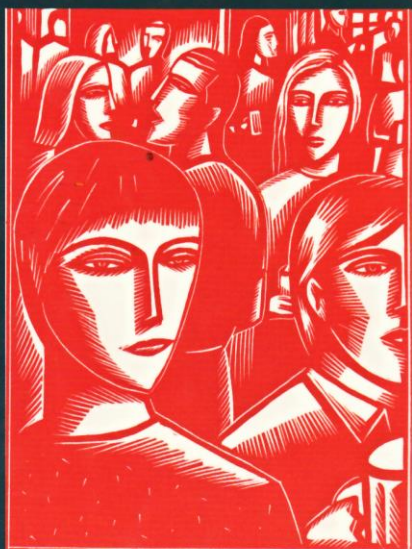
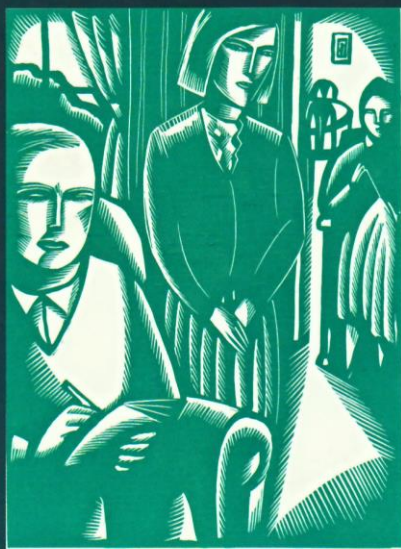


TIPOS de MENTES



Hacia una comprensión de la conciencia

DANIEL C. DENNETT

DEBATE
pensamiento

Daniel C. Dennett

TIPOS DE MENTES

Hacia una comprensión de la conciencia

DEBATE
pensamiento

Versión castellana de
FRANCISCO PÁEZ DE LA CADENA

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del *copyright*, bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidas la reprografía y el tratamiento informático, y la distribución de ejemplares de ella, mediante alquiler o préstamo públicos.

Primera edición: febrero 2000
Título original: *Kinds of Minds*
© Daniel Dennett, 1996

© De la traducción, Francisco Páez de la Cadena, 2000
© De la versión castellana, Editorial Debate, S. A.,
O'Donnell, 19, 28009 Madrid, 2000

I.S.B.N.: 84-8306-260-7
Depósito legal: B. 343-2000
Compuesto en Medianil Composición, S. L.
Impreso en A & M Gràpic, S. L., Santa Perpètua
de Mogoda (Barcelona)
Impreso en España (*Printed in Spain*)

ADVERTENCIA

ESTA ES UNA COPIA PRIVADA PARA FINES EXCLUSIVAMENTE EDUCACIONALES



QUEDA PROHIBIDA
LA VENTA, DISTRIBUCIÓN Y COMERCIALIZACIÓN

- El objeto de la biblioteca es facilitar y fomentar la educación otorgando préstamos gratuitos de libros a personas de los sectores más desposeídos de la sociedad que por motivos económicos, de situación geográfica o discapacidades físicas no tienen posibilidad para acceder a bibliotecas públicas, universitarias o gubernamentales. En consecuencia, una vez leído este libro se considera vencido el préstamo del mismo y deberá ser destruido. No hacerlo, usted, se hace responsable de los perjuicios que deriven de tal incumplimiento.
- Si usted puede financiar el libro, le recomendamos que lo compre en cualquier librería de su país.
- Este proyecto no obtiene ningún tipo de beneficio económico ni directa ni indirectamente.
- Si las leyes de su país no permiten este tipo de préstamo, absténgase de hacer uso de esta biblioteca virtual.

"Quién recibe una idea de mí, recibe instrucción sin disminuir la mía; igual que quién enciende su vela con la mía, recibe luz sin que yo quede a oscuras",

—Thomas Jefferson



Para otras publicaciones visite
www.lecturasinegoismo.com
Referencia: 398

SUMARIO

PREFACIO	7
1. ¿QUÉ TIPOS DE MENTE EXISTEN?	11
<i>Conocer la propia mente</i>	11
<i>Nosotros, los poseedores de mente, los «menteros»</i>	14
<i>Palabras y mentes</i>	18
<i>El problema de las mentes no comunicativas</i>	23
2. INTENCIONALIDAD: EL ENFOQUE DE LOS SISTEMAS INTENCIONALES	31
<i>Comienzos sencillos: el nacimiento del agente</i> ...	31
<i>Adopción del enfoque intencional</i>	40
<i>El equívoco objetivo de la precisión proposicional</i>	56
<i>Intencionalidad originaria y derivada</i>	65
3. EL CUERPO Y SUS MENTES	73
<i>¿De la sensibilidad a lo sentiente?</i>	73
<i>Los medios y los mensajes</i>	82
« <i>¡Mi cuerpo tiene mente propia!</i> »	91

SUMARIO

4. DE CÓMO LA INTENCIONALIDAD SALTÓ A PRIMER PLANO	101
<i>La torre de la generación y la prueba</i>	101
<i>La búsqueda de la sentiencia: informe de avances</i>	114
<i>De la fototaxia a la metafísica</i>	120
5. LA CREACIÓN DEL PENSAMIENTO	143
<i>Psicólogos naturales que no piensan</i>	143
<i>Fabricar cosas con las que pensar</i>	159
<i>Hablando con nosotros mismos</i>	175
6. NUESTRA MENTE Y OTRAS MENTES	181
<i>Nuestra conciencia, las mentes de los otros</i>	181
<i>Dolor y sufrimiento: lo que importa</i>	190
LECTURAS ADICIONALES	199
BIBLIOGRAFÍA	207
ÍNDICE	215

PREFACIO

Soy un filósofo, no un científico, y los filósofos preguntamos mejor que respondemos. Este principio no es un insulto ni a mí ni a mi disciplina, pese a lo que pueda parecer. Una parte muy difícil del ambicioso proyecto humano de comprendernos y de comprender nuestro mundo es la de encontrar mejores preguntas y la de romper las viejas costumbres y tradiciones en el modo de preguntar. Los filósofos pueden hacer una buena contribución a esta investigación explotando sus talentos aguzados profesionalmente como críticos preguntones, siempre que mantengan un espíritu abierto y no pretendan responder a todas las preguntas partiendo de primeros principios «obvios». Hay muchas maneras de hacer preguntas sobre las diferentes clases de mentes, y la mía (la manera de preguntar que voy a presentar en este libro) cambia casi a diario, refinándose y ampliándose, sufriendo correcciones y revisiones conforme voy sabiendo de nuevos descubrimientos, de nuevas teorías, de nuevas dificultades. Voy a presentar el conjunto de suposiciones fundamentales que me sirven de base para esa manera de preguntar y que le proporcionan un modelo estable y coherente, pero las áreas más emocionantes de ese modo de preguntar

se encuentran precisamente en las fronteras más cambiantes del modelo, allí donde hay más actividad. El objetivo principal de este libro es presentar las preguntas que me hago *en este mismo momento...* algunas de las cuales probablemente no llevarán a ninguna parte, así que prevéngase el lector. Pero mi manera de hacer preguntas tiene un historial bastante bueno a lo largo de los años, con una evolución bastante paulatina como para incorporar nuevos descubrimientos, algunos de los cuales estuvieron provocados por preguntas mías anteriores. Otros filósofos han presentado métodos de hacer preguntas que compiten con el mío, pero los más influyentes, pese a su atractivo inicial, llevan a contradicciones internas, dilemas o a muros negros de misterio, como demostraré. Así que recomiendo las preguntas que yo hago, con toda confianza.

Nuestras mentes son complejas urdimbres, tejidas con muy distintas hebras que reflejan muy diferentes dibujos. Algunos de estos elementos son tan antiguos como la vida misma y otros son tan nuevos como la tecnología actual. Nuestras mentes son como las de otros animales en muchos aspectos y notablemente diferentes en otros. Una perspectiva evolutiva puede ayudarnos a ver cómo y por qué esos elementos de las mentes llegan a adoptar las formas que poseen, pero no hay línea recta temporal «desde los microbios hasta el hombre» que nos revele el momento de la aparición de cada hilo nuevo. De modo que, en lo que sigue, he tenido que tejer oscilando entre las mentes sencillas y las mentes complejas, remontándome cada vez más buscando aquellos aspectos que deben añadirse hasta que lleguemos a algo que se pueda reconocer como una mente humana. En ese momento podremos volver la vista atrás una vez más para examinar las diferencias que encontramos y evaluar algunas de sus consecuencias.

Los primeros borradores de este libro se presentaron en las Conferencias Agnes Cuming del University College de

PREFACIO

Dublín y en mis conferencias públicas como *erskine fellow* de la Universidad de Canterbury de Christchurch (Nueva Zelanda), durante mayo y junio de 1995. Quiero mostrar mi agradecimiento al profesorado y al alumnado de esas instituciones, cuyas constructivas discusiones contribuyeron a hacer del borrador final algo casi irreconocible y (confío en que así sea) mejor. También quiero mostrar mi agradecimiento a Marc Hauser, Alva Noë, Wei Cui, Shannon Densmore, Tom Schuman, Pascal Buckley, Jerry Lyons, Sara Lippincott y a mis estudiantes de «Lenguaje y Mente» en Tufts, que leyeron y criticaron enérgicamente mi penúltimo borrador.

Universidad de Tufts
20 de diciembre de 1995

CAPÍTULO 1

¿QUÉ TIPOS DE MENTE EXISTEN?

Conocer la propia mente

¿Podremos saber alguna vez qué pasa por la mente de otra persona? ¿Puede saber alguna vez una mujer lo que significa ser hombre? ¿Qué experimenta un niño en el momento de nacer? ¿Qué experimenta un feto, si es que experimenta algo, en el seno materno? ¿Y qué hay de las mentes no humanas? ¿En qué piensan los caballos? ¿Por qué no les da asco a los buitres comerse los cadáveres putrefactos de los que se alimentan? Cuando un pez se ve atravesado por un anzuelo ¿le duele tanto como nos dolería a nosotros si ese mismo anzuelo nos atravesara el labio? ¿Pueden pensar las arañas o son simplemente robots diminutos que hacen inconscientemente sus elegantes telas? Y ya puestos, ¿por qué no podría ser consciente un robot, si es suficientemente complejo? Hay robots que pueden moverse y manejar cosas casi tan bien como las arañas; ¿podría un robot más complicado sentir dolor o preocupación por su futuro, del mismo modo que una persona? ¿O hay un abismo infranqueable que

separe a los robots (y quizá a las arañas y a los insectos y a las demás criaturas «listas» pero inconscientes) de aquellos animales que tienen mente? ¿Es posible que todos los animales salvo los seres humanos sean en realidad robots inconscientes? Eso fue lo que sostuvo notoriamente en el siglo XVII René Descartes. ¿Se equivocaría de medio a medio? ¿Es posible que todos los animales, e incluso las plantas, y hasta las bacterias, tengan mente?

O, por pasarse al extremo opuesto, ¿estamos seguros de que todos los seres humanos tenemos mente? Por poner el caso más extremo de todos, puede que usted sea la única mente del universo; puede que todo lo demás, incluyendo al aparente autor de este libro, no sea más que una máquina sin mente. Esta extraña idea se me ocurrió por primera vez cuando era pequeño, y puede que a usted se le haya ocurrido también. Más o menos una tercera parte de mis alumnos asegura que también ellos la concibieron y la rumiaron siendo niños. Suele divertirles que se les enseñe que es una hipótesis filosófica tan corriente que hasta tiene nombre: *solipsismo*, del latín «sólo yo». Nadie se toma el solipsismo en serio durante demasiado tiempo, por lo que sabemos, pero sí ofrece un reto importante: *si* sabemos que el solipsismo es estúpido, *si* sabemos que hay otras mentes... ¿cómo lo sabemos?

¿Qué tipos de mentes hay? ¿Y cómo lo sabemos? La primera pregunta se refiere a lo que existe: a la *ontología*, en lenguaje filosófico. La segunda pregunta se refiere a nuestro conocimiento: a la *epistemología*. El propósito de este libro no es responder a estas dos preguntas de una vez por todas, sino más bien mostrar por qué estas dos preguntas han de responderse conjuntamente. Los filósofos suelen advertirnos de que no confundamos cuestiones ontológicas con cuestiones epistemológicas. Lo que existe, dicen, es una cosa, y otra distinta es lo que podemos conocer. Puede que haya cosas que sean absolutamente incognoscibles para nosotros, de

manera que debemos ser cuidadosos y no considerar los límites de nuestro conocimiento como guías seguras acerca de los límites de lo que hay. Estoy de acuerdo en que este es, generalmente, un buen consejo, pero digo que ya conocemos lo suficiente sobre nuestras mentes como para saber que una de las cosas que las diferencia de todo lo demás que haya en el universo es nuestra *manera* de saberlo. Por ejemplo: usted sabe que tiene una mente y que tiene un cerebro, pero ambas cosas son conocimientos de distinto tipo. Usted sabe que tiene un cerebro de la misma manera que sabe que tiene bazo: porque lo ha oído decir. Apostaría a que nunca se ha visto el cerebro o el bazo, pero como los libros dicen que todos los seres humanos normales tienen ambos, usted llega a la casi absoluta certeza de que también tiene uno de cada. Con su mente tiene un trato más íntimo, tan íntimo que hasta podría llegar a decir que usted *es* su mente. (Eso fue lo que dijo Descartes: que era una conciencia, una *res cogitans*, una cosa pensante.) Un libro o un maestro podrían decirle a usted qué es la mente, pero usted no tendría por qué creer en lo que nadie le dijera para pensar, efectivamente, que usted tiene mente. Si alguna vez se le ha ocurrido pensar si usted es normal y tiene una mente como la de los demás, se dará cuenta inmediatamente de que, tal como señaló Descartes, esa misma perplejidad sobre esta cuestión ya demuestra más allá de toda duda que desde luego usted tiene mente.

Lo cual parece indicar que cada uno de nosotros conoce con exactitud una mente desde dentro y que no hay dos de nosotros que conozcan la misma mente desde dentro. Ninguna otra cosa se conoce del mismo modo. Y sin embargo, toda esta controversia se ha planteado en función de *cómo* conocemos usted y yo. Da por hecho que el solipsismo es falso. Cuanto más reflexionamos —nosotros— acerca de esta presuposición, más inevitable parece. No es posible que sólo haya una mente: o por lo menos, que haya sólo una mente como la *nuestra*.

Nosotros, los poseedores de mente, los «menteros»

Si queremos considerar la cuestión de si los animales no humanos tienen mente, tenemos que empezar preguntando si en cierto modo tienen mentes parecidas a las nuestras, habida cuenta de que son las únicas de las que conocemos algo... de momento. (Intente preguntarse si los animales no humanos tienen perdino. Usted no sabrá en qué consiste la pregunta si no sabe siquiera qué es un perdino. Sea lo que fuere la mente, hay que suponer que se trata de algo parecido a las nuestras: si no, no la llamaríamos mente.) De manera que nuestras mentes, las únicas que conocemos *nosotros* desde el principio, son el baremo a partir del cual debemos empezar. Sin este acuerdo, nos estaremos engañando, diciendo tonterías sin saberlo.

Cuando *yo* me dirijo a *usted*, estoy incluyendo a ambos en la clase de los portadores de mente. Este inevitable punto de partida crea, o reconoce, un grupo aparte, una clase de características privilegiadas realzada sobre todo lo demás del universo. Cosa que es tan evidente que casi no se nota por lo profundamente arraigada que está en nuestro pensamiento y en nuestro lenguaje, pero aun así debemos extendernos sobre ella. Cuando existe un *nosotros*, usted ya no está solo; el solipsismo es falso; hay otros. Cosa que se hace especialmente clara si consideramos algunas curiosas variaciones:

«Salimos de Houston al amanecer y enfilamos la carretera... sólo mi camión y yo.»

Qué raro. Si este tipo cree que su camión es un compañero tan valioso que merece entrar bajo el paraguas del «nosotros» es que debe de encontrarse muy solo. O eso, o que su camión esté preparado de modo que sea la envidia de todos los expertos en robótica del mundo. Por contra, «nosotros...

mi perro y yo» no nos sorprende en absoluto pero en cambio es difícil tomarse en serio un «nosotros... mi ostra y yo». En otras palabras: tenemos la suficiente seguridad de que los perros tienen mente y dudamos que la tengan las ostras.

La pertenencia a la clase de cosas que tienen mente proporciona una garantía de primordial importancia: la de cierta categoría moral. Sólo a los que poseen mente les importa, sólo a los que tienen mente puede preocuparles lo que ocurre. Si yo le hago algo a usted que usted no quiere que yo le haga, eso tiene una importancia moral. Importa porque le importa a usted. Puede que no importe mucho, o que sus intereses se vean superados por todo tipo de razones o que el hecho de que a usted le importe pueda incluso hacer que se muestre *a favor* de lo que yo hago (si es que le estoy castigando a usted por una mala acción suya). En cualquier caso, esa preocupación suya automáticamente pesa algo en la ecuación moral. Si las flores tuvieran mente, lo que les hacemos podría *importarles* y no solamente importaría a los que se preocupan por las flores. Si no hay nadie a quien le importe, entonces no importa lo que le hagamos a las flores.

Puede haber quien disienta; habrá quien diga que las flores tienen cierta categoría moral incluso aun sin que mente alguna sepa o se preocupe de su existencia. Su belleza, por ejemplo, independientemente de que se aprecie o no, es algo bueno en sí y por ello no debería destruirse a igualdad de los demás factores. Este punto de vista no es el de quien dice que la belleza de las flores importa *a Dios*, por ejemplo, o el de quien dice que *podría* importar a alguien cuya presencia sea indetectable para nosotros. Se trata del punto de vista de que la belleza importa *incluso a pesar de que no haya nadie a quien le importe*: ni a las propias flores, ni a Dios, ni a nadie. Sigo sin convencerme, pero en lugar de dejar de lado este punto de vista, hago notar que es polémico y no muy compartido. Por contra, no hace falta alegar

gran cosa para que la mayor parte de la gente se muestre de acuerdo en que aquello que tiene mente tiene también intereses que importan. Por eso se muestra tan preocupada la gente con la cuestión de qué tiene mente y qué no: cualquier propuesta de reajuste en las fronteras de la clase de poseedores de mente tiene gran relevancia ética.

Podríamos equivocarnos. Podríamos adjudicar mente a cosas que no la tengan o podríamos pasar por alto una cosa con mente. Estas equivocaciones no serían equivalentes. Pasarse en atribuir mentes («hacerse amigo» de las plantas de nuestra casa o quedarnos en vela por las noches preocupándonos por el ordenador que duerme en nuestro escritorio) es, como mucho, un estúpido error de credulidad. Quedarse corto al atribuir mentes (no tener en cuenta o rebajar o negar la experiencia, el sufrimiento y la alegría, las ambiciones truncadas y los deseos frustrados de una persona o animal que tuviera mente) sería un pecado terrible. Porque, en definitiva: ¿Cómo se sentiría usted si se le tratara como a un objeto inanimado? (Dése cuenta de cómo esta pregunta retórica apela a *nuestra* categoría compartida como poseedores de mente.)

Lo cierto es que ambos errores podrían tener graves consecuencias morales. Si nos pasamos en la atribución de mentes (si, por ejemplo, nos hacemos a la idea de que como las bacterias tienen mente no podemos justificar su eliminación) ello podría llevarnos a sacrificar el interés de muchos legítimos portadores de intereses (nuestros amigos, nuestros animales de compañía, nosotros mismos) por cosas que no tuvieran ninguna importancia moral genuina. El debate acerca del aborto gira alrededor de un dilema semejante; algunos creen que es evidente que un feto de diez semanas tiene mente, y otros piensan que es evidente que no. Si no tiene mente, entonces queda abierto el camino para argumentar que el feto no tiene mayores intereses que los que pueda tener, pongamos, una pierna gangrenada o un diente cariado: y

entonces se podría destruir para salvar la vida (o sencillamente para servir a los intereses) de la persona que tiene intereses y de la cual forma parte. Si el feto ya tiene mente, entonces, decidamos lo que decidamos, tenemos que considerar *sus* intereses conjuntamente con los de su portador temporal. En medio de estas dos posiciones extremas se encuentra el auténtico dilema: el feto desarrollará en seguida su propia mente si no se lo perturba, de modo que ¿cuándo empezamos a contar sus *futuros* intereses? La relevancia de poseer una mente en relación con la categoría moral resulta especialmente clara en estos casos, ya que si se sabe que el feto en cuestión es anacefálico (que carece de cerebro) cambia la consideración de forma drástica para la mayoría de las personas. No para todas. (No es que intente sentar aquí estos asuntos morales, sino solamente mostrar cómo una opinión moral común amplía nuestro interés sobre estas cuestiones mucho más allá de nuestra curiosidad normal.)

Los dictados de la moralidad y del método científico van aquí en direcciones opuestas. La línea ética consiste en equivocarse por exceso, para estar a salvo. La línea científica consiste en poner la carga de la prueba en la atribución. Como científico, por ejemplo, usted no podría limitarse a *declarar* que la presencia de moléculas de glutamato (un neurotransmisor básico que sirve para enviar señales entre las células nerviosas) equivale a la existencia de una mente: tendría que demostrarlo sobre la base de que la «hipótesis cero» es que esa mente no existe. (La hipótesis cero o de partida de nuestras leyes penales es la presunción de inocencia, el ser «inocente mientras no se demuestre lo contrario».) Entre los científicos existe un desacuerdo sustancial sobre qué especies poseen cierto tipo de mente y qué tipo de mente es, pero incluso los científicos que defienden con más ardor la conciencia de los animales aceptan esta carga de la prueba... y piensan que pueden satisfacerla concibiendo y confirmando teorías que demuestran que los animales son cons-

cientes. Esas teorías sin embargo todavía no se han confirmado y mientras tanto podemos valorar la incomodidad de aquellos que en este punto de vista agnóstico, en este «esperar a ver», ven un riesgo para la categoría moral de las criaturas de las que ellos se muestran *seguros* de que son conscientes.

Imagínese que la cuestión que se nos planteara no fuera la de la mente de las palomas o la de los murciélagos, sino la de las personas zurdas o la de las personas pelirrojas. Nos ofendería profundamente que se nos dijera que todavía hay que demostrar que esa categoría de seres vivientes posee lo suficiente para entrar a formar parte de la clase privilegiada de los poseedores de mente. Hay muchas personas que se sienten ultrajadas de manera parecida ante la exigencia de la prueba de una mente para las especies no humanas, pero si son honradas consigo mismas también garantizarán que ven la necesidad de semejante prueba en el caso de la medusa, de las amebas o de las margaritas; de modo que nos mostramos de acuerdo en el principio y en cambio nos resentimos en su aplicación a criaturas muy parecidas a nosotros. Podemos aliviar algo sus recelos acordando que nos pasaremos más bien por el lado de la inclusión en todas las regulaciones hasta que los hechos queden bien establecidos; con todo, el precio que se debe pagar por la confirmación científica de la hipótesis preferida en cuanto a la mente de los animales es el riesgo de su contraprueba científica.

Palabras y mentes

Sin embargo, no se puede discutir seriamente que usted y yo tengamos mente. ¿Cómo sé yo que usted tiene mente? Porque a cualquiera que pueda comprender mis palabras me dirijo automáticamente mediante el pronombre «usted» y porque sólo las cosas que poseen mente pueden comprender.

Hay artulugios dirigidos por ordenador que pueden leer libros a los ciegos: transforman una página de texto visible en un discurso de palabras audibles, pero esos artulugios no comprenden las palabras que leen y por tanto no se sienten aludidos por ningún «usted» que puedan encontrarse: lo pasan por alto sin más y dirigen a cualquiera que los escuche (y los comprenda) las palabras habladas. Así es como usted, amable lector o auditor, sabe que tiene mente. Lo mismo que lo sé yo. Créame.

De hecho, eso es lo que hacemos normalmente: creemos que los demás resuelven más allá de cualquier duda razonable la cuestión de si cualquiera de nosotros tiene mente o no. ¿Por qué ha de ser tan convincente la palabra de los demás? Porque tiene un poder resolutorio enorme frente a las dudas y las ambigüedades. Usted se encuentra en la situación de que alguien con aspecto amenazador y blandiendo un hacha se le acerca. Se pregunta: ¿Qué le pasa? ¿Me va a atacar? ¿Me ha confundido con otro? Le pregunta. Puede que confirme sus peores temores o puede que le cuente que ha intentado desbloquear en vano la puerta de su coche (delante del cual se encuentra usted) y que ha ido a por un hacha para romper la ventanilla. Es posible que usted no le crea cuando dice que se trata de su coche y no el de otra persona, pero si usted continúa la conversación (y decide no salir corriendo) terminará por resolver todas sus dudas y aclarar la situación de un modo que sería absolutamente imposible si usted y él fueran incapaces de comunicarse verbalmente. Imagínese que intenta preguntarle y resulta que él no habla su idioma. Puede que entonces los dos se decidan por recurrir a los gestos y a la mímica. Estas técnicas, utilizadas con ingenio, pueden llevarle lejos pero son un pobre sustituto del lenguaje: piense con qué ansiedad querrían ustedes confirmar lo que tan duramente les ha costado entender, si en ese momento apareciera un intérprete bilingüe. Algunas preguntas y respuestas transmitidas de ese modo no se limitarían a

aliviar cualquier incertidumbre que pudiera quedar sino que añadirían detalles que no podrían transmitirse de ninguna otra manera: «Cuando vio que usted se llevaba una mano al pecho y con la otra le empujaba, creyó que usted quería decir que estaba enfermo; estaba intentando preguntarle si no querría que le llevara al médico una vez que rompiera la ventanilla y recuperara las llaves. Lo de señalarse las orejas con los dedos era un intento de indicar un estetoscopio.» Ah, ahora encaja todo, gracias a unas pocas palabras.

Se suele hacer hincapié en la dificultad de traducir con precisión y fiabilidad de unos lenguajes humanos a otros. Se nos dice que las culturas humanas son demasiado diferentes, demasiado «inconmensurables» como para permitir que los significados disponibles para uno de los hablantes pueda compartírselos otro en su totalidad. No hay duda de que la traducción siempre se queda un poco lejos de la perfección, pero a gran escala puede que eso no importe. Tal vez sea imposible la traducción perfecta, pero de hecho todos los días se consigue una buena traducción... sin pretenderlo. La buena traducción puede distinguirse objetivamente de la no tan buena y de la mala, y permite a todos los seres humanos, independientemente de su raza, cultura, edad, sexo o experiencia unirse mucho más al resto de individuos que lo que puedan unirse los individuos de otras especies. Los seres humanos compartimos un mundo subjetivo (y lo sabemos) de un modo que queda completamente fuera de las capacidades de cualquier otra criatura del planeta y todo debido a que podemos hablar unos con otros. Los seres humanos que no tienen (aún) lenguaje con el cual comunicarse son la excepción y ese es el motivo por el cual tenemos dificultades concretas en saber qué significa ser un recién nacido o ser un sordomudo.

La conversación nos une. Podemos saber mucho de lo que significa ser pescador noruego o taxista nigeriano, monja octogenaria o niño de cinco años ciego de nacimiento,

maestro de ajedrez o prostituta o piloto de caza. Podemos saber mucho más de esos asuntos de lo que podemos llegar a saber (si es que llegamos a saber algo) qué significa ser delfín, murciélago o incluso chimpancé. Por muy diferente que seamos unos de otros, repartidos por el planeta, podemos explorar nuestras diferencias y hablar de ellas. Por muy parecidos que sean unos animales entre sí, por muy pegados que estén en un rebaño, no pueden saber gran cosa de sus similitudes y mucho menos de sus diferencias. No pueden cambiar impresiones. Pueden tener experiencias similares, codo con codo, pero verdaderamente no pueden compartir experiencias como nosotros.

Alguien podría dudarlo. ¿No pueden los animales comprenderse «instintivamente» de modos que a los humanos nos parecen insondables? Desde luego que ha habido autores que lo han dicho. Veamos, por ejemplo, el caso de Elizabeth Marshall Thomas que imagina en *The Hidden Life of Dogs* [La vida oculta de los perros] (1993) que los perros disfrutan de una comprensión amplia a su manera. Un ejemplo: «Por motivos conocidos para los perros pero no para nosotros, muchas perras no se aparean con sus crías» (pág. 76). Esa resistencia instintiva al apareamiento consanguíneo no se pone en duda... pero ¿por qué cree la autora que los perros tienen una visión más profunda que nosotros acerca de sus propios motivos instintivos? Hay muchas cosas que sentimos poderosa e instintivamente que no queremos hacer, sin tener la menor idea de por qué sentimos de tal modo. Suponer sin pruebas que los perros tienen una visión más profunda de sus impulsos que la nuestra es pasar por alto de manera inaceptable la hipótesis cero: si es que estamos planteando una cuestión científica. Como ya veremos, hay organismos muy simples que pueden adaptarse a su medio y a los demás organismos de modo sorprendentemente apropiado sin tener la más ligera idea de su adaptación. Sin embargo, nosotros ya sabemos por medio de la conversación que las personas

son característicamente capaces de un entendimiento muy elevado de sí mismas y de las demás.

Naturalmente que podemos confundirnos. Suele hacerse hincapié en la dificultad de saber si determinado hablante es sincero o no. Al ser las palabras las herramientas más poderosas para la comunicación, también son las herramientas más poderosas para el engaño y la manipulación. Pero así como puede ser fácil mentir, suele ser casi igual de fácil coger a un mentiroso... sobre todo cuando las mentiras se amplían y el problema logístico de mantener la estructura de la falsedad abrumba al mentiroso. En nuestra fantasía podemos imaginar mentirosos infinitamente poderosos, pero los engaños que son «posibles en principio» para semejantes demonios malignos pueden pasarse por alto con seguridad en el mundo real. Sencillamente sería demasiado difícil montar una falsedad enorme y mantenerla coherentemente. *Sabemos* que en todo el mundo las personas tienen más o menos las mismas querencias y los mismos aborrecimientos, las mismas esperanzas y los mismos miedos. Sabemos que les gusta recordar los acontecimientos preferidos de su vida. Sabemos que todas ellas tienen períodos de ensoñación en que reorganizan y repasan los detalles deliberadamente. Sabemos que tienen obsesiones, pesadillas y alucinaciones. Sabemos que un aroma o una melodía puede recordarles un acontecimiento concreto de su vida y que muchas veces hablan consigo mismas en silencio, sin mover los labios. Todo esto ya era bien conocido mucho antes de que surgiera la psicología científica, mucho antes de que se hicieran una observación y una experimentación meticolosas en seres humanos. Desde épocas muy antiguas sabemos todas esas cosas de las personas porque hemos hablado de ellas largo y tendido. No conocemos nada comparable en la vida mental de otras especies porque no podemos hablarlo con los seres de esa especie. Podemos creer que lo sabemos pero para confirmar o rebatir nuestras corazonadas tradicionales hace falta la investigación científica.

El problema de las mentes no comunicativas

Es muy difícil decir qué piensa alguien cuando no quiere comunicarlo... o no puede comunicarlo, por unos u otros motivos. Pero normalmente suponemos que esas personas que no se comunican desde luego piensan (tienen mente) incluso aunque no podamos confirmar los detalles. Esto es evidente, aunque sólo sea porque fácilmente nos podemos imaginar a nosotros mismos en una situación en la que nos negaríamos firmemente a comunicarnos, sin dejar de tener nuestros propios pensamientos, puede que hasta reflexionando con regocijo acerca de las dificultades que nuestros observadores tendrían al intentar explicarse qué nos pasaba por la mente, si es que nos pasaba algo. Por muy concluyente que sea su existencia, el habla no es necesaria para tener mente. De este hecho evidente podemos vernos tentados a sacar una conclusión problemática: podría haber entes que tuvieran mente pero que no pudieran decirnos qué están pensando: no porque estuvieran paralizados o sufrieran de afasia (la incapacidad de comunicarse verbalmente debido a un daño cerebral localizado) sino porque carecieran de la capacidad del lenguaje. ¿Y por qué digo que se trata de una conclusión problemática?

Consideremos en primer lugar los argumentos a favor. La tradición y el sentido común dictaminan que existen mentes sin lenguaje con seguridad. Es seguro también que nuestra capacidad de intercambiar con otros lo que nos pasa por la mente es sencillamente un talento periférico, en el mismo sentido en que hablamos de una impresora láser como periférico de un ordenador (el ordenador puede seguir funcionando sin tener conectada la impresora). Seguro que los animales no humanos (o por lo menos, algunos de ellos) tienen vida mental. Seguro que tienen mente los niños humanos antes de adquirir el lenguaje, y los sordomudos humanos (incluso los escasos sordomudos que nunca aprenden si-

quiera el lenguaje de signos). Seguro. Sin duda que sus mentes diferirán de la *nuestra* de muchas maneras difíciles de imaginar: diferirán de las mentes de quienes podemos comprender una conversación como la presente, pero seguro que *son* mentes. Nuestro camino real hacia el conocimiento de otras mentes (el lenguaje) no se les puede hacer extensivo, pero se trata de una limitación de nuestro conocimiento y no de una limitación de su mente. Surge entonces la perspectiva de que haya mentes cuyo contenido sea sistemáticamente inaccesible a nuestra curiosidad: incognoscibles, improbables, impenetrables a cualquier investigación.

La respuesta tradicional a esta perspectiva es aceptarla. Naturalmente, la mente es la última *terra incógnita*, queda fuera del alcance de toda ciencia y, en el caso de las mentes sin lenguaje, fuera de todo posible intercambio mediante empatía. ¿Y qué? Nuestra curiosidad debería atemperarse con un poco de humildad. No confundamos las cuestiones ontológicas (aquello que existe) con las epistemológicas (cómo lo conocemos). Debemos hacernos a la idea de este hecho maravilloso sobre lo que no es posible averiguar.

Pero antes de acomodarnos a esta conclusión, tenemos que considerar las consecuencias de algunos otros hechos sobre este asunto que son igual de evidentes. Sabemos que muy a menudo hacemos cosas inteligentes sin pensar en absoluto; las hacemos «automáticamente» o «inconscientemente». Por ejemplo: ¿Cómo se usa la información que nos llega mediante el flujo óptico de formas a través de la visión periférica para ajustar la longitud de nuestro paso cuando caminamos por un terreno abrupto? La respuesta es que no sabemos cómo. Ni aun queriéndolo se puede prestar atención a ese proceso. ¿Cómo nos damos cuenta, estando completamente dormidos, de que se nos ha retorcido el brazo izquierdo que nos produce una tensión indebida en el hombro izquierdo? No se sabe: no forma parte de nuestra experiencia. Inconsciente y rápidamente cambiamos a una postura

más «cómoda», sin interrupción alguna en nuestro sueño. Si se nos pide que hablemos de estas partes putativas de nuestras vidas mentales, nos vemos ante un espacio en blanco; ocurriera lo que ocurriese en nosotros para regir estos comportamientos inteligentes no formaba parte en absoluto de nuestra vida mental. De manera que otra perspectiva que debe considerarse es que entre las criaturas que carezcan de lenguaje haya algunas que no tengan mente en absoluto pero que hagan todo «automáticamente» o «inconscientemente».

También la respuesta tradicional a esta perspectiva es aceptarla. Naturalmente que algunas criaturas carecen de mente. Seguro que las bacterias no tienen mente y, seguramente, ocurre lo mismo con las amebas y las estrellas de mar. Es bastante posible que incluso las hormigas, con lo listas que son, no sean más que autómatas sin mente rodando por el mundo sin la menor experiencia ni el menor pensamiento. ¿Y la trucha? ¿Y los pollos? ¿Y las ratas? Puede que nunca seamos capaces de saber dónde trazar la línea que separe a las criaturas que tienen mente de las que no, pero no es más que otro aspecto de las inevitables limitaciones de nuestro conocimiento. Esos hechos pueden ser sistemáticamente incognoscibles y no simplemente difíciles de descubrir.

Aquí hay, por tanto, dos tipos de hechos supuestamente incognoscibles: hechos sobre lo que ocurre en los que tienen mente pero no tienen modo de comunicar sus pensamientos y hechos sobre qué criaturas tengan o no mente. Estas dos variantes de ignorancia no son igualmente fáciles de aceptar. Las diferencias *entre mentes* podrían ser diferencias cuyas líneas principales fueran fácilmente discernibles para observadores objetivos mientras que los detalles secundarios serían cada vez más difíciles de determinar: digamos una especie de beneficios decrecientes en relación con el trabajo invertido. Los remanentes desconocidos no serían misterios

sino lagunas inevitables en un catálogo informativamente muy rico, pero finito, de similitudes y diferencias. Las diferencias entre mentes serían entonces como las diferencias entre lenguajes o entre estilos de música o de arte: inagotables en el límite pero definibles hasta el grado de aproximación que quisiéramos. Pero la diferencia entre tener mente y no tenerla (entre ser algo con un punto de vista subjetivo propio y ser algo que es todo externo y nada interno, como una roca o como un trozo de uña) es, aparentemente, la diferencia entre todo y nada. Es mucho más difícil aceptar la idea de que por mucha investigación que se haga no sabremos *si hay alguien a quien le importe* dentro de un caparazón de langosta, por poner un caso, o tras la brillante cobertura de un robot.

Sugerir que un hecho de tales características y moralmente tan importante nos pueda ser sistemáticamente incognoscible, resulta sencillamente intolerable. Quiere decir que sean cuales fueren las investigaciones que realicemos, podríamos estar sacrificando, sin saberlo, los genuinos intereses morales de unos en beneficio completamente ilusorio de otros. La ignorancia inevitable de las consecuencias suele ser una excusa legítima cuando nos encontramos con que hemos causado sin querer algún daño al mundo, pero si tenemos que declararnos al principio inevitablemente ignorantes de las mismísimas bases de todo pensamiento moral, entonces la moral se convierte en un simulacro. Afortunadamente, esta conclusión es tan increíble como intolerable. Por ejemplo, resulta ridícula la pretensión de que los zurdos son unos zombies inconscientes a los que se podría desmantelar como si fueran bicicletas. Y de igual modo, en el extremo opuesto, se encuentra la pretensión de que las bacterias sufren o de que a las zanahorias les importa que se las arranque sin mayores ceremonias de sus moradas terrosas. Es evidente que podemos saber con certidumbre moral (que es lo que verdaderamente importa) que algunas cosas tienen mente y otras no.

Pero seguimos sin saber *cómo* sabemos estas cosas; la fuerza de nuestras intuiciones sobre tales casos no nos garantiza que podamos fiarnos de ellas. Consideremos unos cuantos casos, empezando por el siguiente comentario de la evolucionista Elaine Morgan:

Lo auténticamente llamativo acerca del recién nacido es que, desde el primer minuto, hay alguien ahí. Cualquiera que se incline sobre la cuna y mire recibe una mirada por respuesta (1995, pág. 99).

Como observación acerca de cómo los seres humanos reaccionamos instintivamente a la mirada, da completamente en el blanco, pero eso mismo demuestra la facilidad con que podemos engañarnos. Un robot puede confundirnos, por ejemplo. En el Laboratorio de Inteligencia Artificial del MIT (Instituto de Tecnología de Massachusetts), Rodney Brooks y Lynn Andrea Stein han reunido un equipo de especialistas en robótica y en otras áreas (entre los que me cuento) para construir un robot humanoide llamado Cog. Cog está fabricado de metal, silicio y vidrio, como tantos otros, pero su diseño es muy distinto, mucho más parecido al diseño de un ser humano, de tal modo que Cog pueda convertirse algún día en el primer robot consciente del mundo. ¿Es factible un robot consciente? He defendido una teoría de la conciencia, el Modelo de Borradores Múltiples (1991), que supone que un robot consciente es posible en principio, y el diseño de Cog tiene en mente ese lejano objetivo. Pero Cog no se acerca en absoluto aún a la conciencia. No puede ver ni oír ni sentir en absoluto, aunque sus partes corporales pueden moverse ya de una forma humanoide desconcertante. Tiene unos ojos que son cámaras de vídeo diminutas, con movimientos sacádicos (movimientos rápidos) para enfocar a cualquier persona que entre en la habitación y seguirla conforme se mueva. Que a uno le sigan de este modo es una ex-

perencia extrañamente inquietante, incluso para los que están en el ajo. Mirar fijamente a Cog a los ojos mientras te devuelve la mirada puede dejar sin aliento a los no iniciados, pero ahí dentro no hay nadie... o por lo menos, todavía no. Los brazos de Cog, a diferencia de los robots estándar, tanto reales como cinematográficos, se mueven veloz y flexiblemente, como los nuestros; cuando se aprieta el brazo extendido de Cog, se nota una extraña resistencia humanoide que hace que queramos exclamar, como si fuera una película de horror, «¡Está vivo! ¡Está vivo!». No está vivo, pero resulta muy fuerte la intuición de que es justamente lo contrario.

Ahora que estamos pensando en brazos, consideremos una variante con una moraleja distinta: el caso del brazo de un hombre, amputado en un accidente horroroso, que los cirujanos creen poder reinplantarle. Sobre la mesa de operaciones, todavía caliente y suave ¿siente dolor? (De sentirlo, tendríamos que inyectarle algo de novocaína, sobre todo si pensamos en utilizar el escalpelo para recortar tejidos del brazo antes de intentar el implante.) Una sugerencia estúpida, replicará usted; para sentir dolor hace falta tener mente y como el brazo no está unido al cuerpo, que tiene la mente, lo que se haga al brazo no causará sufrimiento a mente alguna. Pero puede que el brazo tenga mente propia. ¿Puede que siempre la haya tenido pero que haya sido incapaz de hablarnos! Bien ¿y por qué no? Tiene un buen número de células nerviosas y esas células siguen disparando respuestas. Si damos con un organismo completo que tuviera esa misma cantidad de células nerviosas activas, nos sentiríamos fuertemente inclinados a creer que es capaz de experimentar dolor, incluso aunque no pudiera expresarse de modo que pudiéramos entenderle. Aquí chocan nuestras intuiciones: los brazos no tienen mente a pesar de abarcar muchos de los procesos y materiales que suelen convencernos de que algunos animales no humanos sí la tienen.

¿Es que lo que cuenta es la conducta? Imagínese que pe-

liza el pulgar de ese brazo amputado y ¡el brazo le pellizca a usted! ¿Decidiría entonces darle novocaína? Y si decide que no ¿por qué? ¿Porque su reacción habría sido un reflejo «automático»? ¿Y cómo puede usted estar seguro? ¿Hay algo en la organización de esas células nerviosas que es la que marca la diferencia?

Estos casos enigmáticos son divertidos de plantear y aprendemos cosas importantes de nuestras ingenuas ideas sobre la mente cuando intentamos aclarar por qué nuestras intuiciones se forman como se forman, pero seguramente habrá algún método mejor para investigar los distintos tipos de mente... y de las no-mentes que podrían confundirnos. La convicción derrotista de que nunca podremos saber debería posponerse indefinidamente, dejándola como una conclusión de último aliento a la que llegar sólo después de haber agotado las demás vías y no sólo haber imaginado que las hemos agotado. Puede que nos esperen sorpresas y aspectos iluminadores.

Una perspectiva que hay que considerar, independientemente de que la descartemos finalmente o no, es que puede que, después de todo, el lenguaje no sea tan periférico para las mentes. Puede que el tipo de mente que se tiene cuando se le añade el lenguaje sea tan diferente del tipo de mente que se puede tener sin lenguaje, que llamarlas mentes a las dos sea un equivocación. En otras palabras, quizá nuestra idea de que las mentes de las demás criaturas tienen riquezas (riquezas inaccesibles a *nosotros* pero no a *ellos*, por supuesto) sea una ilusión. El filósofo Ludwig Wittgenstein dijo, como es bien sabido: «Si un león pudiera hablar, no le comprenderíamos.» Se trata, sin duda, de una posibilidad pero que desvía nuestra atención de otra: si un león pudiera hablar, naturalmente que le entenderíamos, con los habituales esfuerzos para la traducción entre diferentes idiomas, pero nuestras conversaciones con él no nos dirían prácticamente nada de las mentes de los leones corrientes, habida

cuenta de que su mente equipada de lenguaje sería bien diferente. ¿Podría ocurrir que al añadir el lenguaje a la «mente» de un león le *proporcionara* mente por primera vez! O a lo mejor no. En cualquier caso, deberíamos investigar ese aspecto y no dar por sentado, conforme a la tradición, que las mentes de los animales que no hablan son realmente parecidas a las nuestras.

Si tenemos que hallar una vía alternativa a la investigación en lugar de apoyarnos acríticamente en nuestras intuiciones preteóricas... ¿cómo deberíamos empezar? Consideremos la vía histórica, evolutiva. No siempre ha habido mentes. *Nosotros* tenemos mente, pero nosotros no hemos existido siempre. Hemos evolucionado a partir de seres con mentes más sencillas (si es que eran mentes) que evolucionaron a su vez de seres con presuntas mentes aún más simples. Y hubo un momento, hace cuatro o cinco mil millones de años, en que no había mente alguna, ni simple ni compleja... o, por lo menos, no en este planeta. ¿Cuáles fueron los cambios, en qué orden se dieron y por qué? Los pasos principales parecen claros, aunque los detalles de las fechas y de los lugares sólo puedan ser conjeturas. Una vez contada esa historia, por lo menos tendremos un marco general en el que situar nuestros dilemas. Puede que queramos distinguir tipos de seudomentes, o protomentes, o semimentes o hemisemidementes de lo que realmente es una auténtica mente. Llamemos como llamemos a esas disposiciones ancestrales, puede que podamos ponernos de acuerdo en la escala que forman y en las condiciones y en los principios que iniciaron la escala. El próximo capítulo desarrolla algunas de las herramientas para esta investigación.

CAPÍTULO 2

INTENCIONALIDAD: EL ENFOQUE DE LOS SISTEMAS INTENCIONALES

Me doy cuenta de algo y para ello busco una razón; lo cual quiere decir inicialmente que busco en ello una intención y por encima de todo a un alguien que tiene intenciones, un sujeto, un hacedor; cada suceso un acto... antiguamente se veía intención en todos los sucesos, es nuestra costumbre más antigua. ¿La poseen también los animales?

Friedrich Nietzsche, *La voluntad de poder*.

*Comienzos sencillos: el nacimiento del agente **

No hay grano de arena que tenga mente: un grano de arena es demasiado simple. Todavía más simples, ningún átomo de carbono ni molécula de agua tienen mente. Respecto a

* Ciertas partes de este epígrafe proceden de mi libro *La peligrosa idea de Darwin*, previa revisión.

esto no espero encontrar discrepancias serias. Pero ¿qué decir de las moléculas mayores? Un virus es una única inmensa molécula, una macromolécula compuesta de cientos de miles o incluso millones de partes, dependiendo de lo pequeñas que sean las partes que consideremos. Estas partes de tipo atómico interactúan, a su manera inconsciente, dando como resultado algunos efectos bastante sorprendentes. El principal de ellos, desde el punto de vista de nuestra investigación, es la *autoduplicación*, la producción de una réplica de sí mismo. Algunas macromoléculas, colocadas en un medio adecuadamente provisto, tienen la asombrosa capacidad de construir y ofrecer de modo inconsciente una copia exacta (o casi exacta) de sí mismas. El ADN [ácido desoxirribonucleico] y su antecesor el ARN [ácido ribonucleico] son macromoléculas de ese tipo; son la base de toda la vida de este planeta y por ello son una condición histórica previa de toda mente... por lo menos de toda mente de este planeta. Más o menos durante mil millones de años previamente a la aparición en la tierra de los organismos simples de una sola célula, hubo macromoléculas con capacidad de autoduplicación, mutando, creciendo e incluso recomponiéndose incesantemente... haciéndose cada vez más expertas en esas tareas, sin dejar de producir réplicas de sí mismas.

Es una hazaña fantástica que sigue quedando muy lejos de la capacidad de cualquier robot existente. ¿Quiere ello decir que tales macromoléculas tienen mente como las nuestras? Ciertamente, no. Ni siquiera están vivas: para la química no son más que inmensos cristales. Estas moléculas gigantes son máquinas diminutas: *nanotecnología macromolecular*. Son, en efecto, robots naturales. La posibilidad, en principio, de un robot capaz de duplicarse fue matemáticamente demostrada por John von Neumann, uno de los inventores del ordenador y cuyo brillante diseño de un aparato que se duplicara por sí solo sin estar vivo anticipaba muchos de los detalles de diseño y construcción del ARN y del ADN.

Por medio del microscopio de la biología molecular, hemos llegado a ser testigos del nacimiento del *agente* en las primeras macromoléculas que tienen complejidad suficiente como para *realizar acciones* y no sencillamente para *sufrir efectos*. Ese agente no es un agente con todas las de la ley como el nuestro: no sabe lo que hace. Por contra, nosotros solemos saber bien lo que hacemos. Para lo mejor, y para lo peor, nosotros, agentes humanos, podemos llevar a cabo acciones con intención, *intencionales*, después de haber deliberado conscientemente los pros y los contras. El agente macromolecular es distinto: hay motivos para que las macromoléculas hagan lo que hagan, pero las macromoléculas no son conscientes de ellos. Su modo de ser agentes es, sin embargo, la única base posible a partir de la cual habrían podido germinar las semillas de nuestro propio modo de ser agentes.

Hay algo ajeno y vagamente repelente en el cuasi agente que descubrimos aquí: todo ese útil ajetreo y, sin embargo, «no hay nadie ahí dentro». Las máquinas moleculares llevan a cabo sus asombrosas proezas, diseñadas evidentemente con un cuidado exquisito y evidentemente también sin tener ni idea de lo que están haciendo. Considérese esta descripción de la actividad de un fago de ARN: un virus capaz de duplicarse, y moderno heredero de las primeras macromoléculas con capacidad de duplicación:

En primer lugar, el virus necesita un material en el que guardar y proteger su propia información genética. En segundo lugar, precisa de un medio para introducir su información en la célula que lo albergue. En tercer lugar, requiere un mecanismo para copiar específicamente su información en presencia de una enorme abundancia de ARN de la célula huésped... Finalmente, debe organizar la proliferación de su información, proceso que generalmente conduce a la destrucción de la célula huésped. El virus consigue incluso que la cé-

lula le haga la duplicación: su única aportación es un factor proteínico, especialmente adaptado al ARN vírico. Esta enzima no se activa hasta que no aparece una «contraseña» del ARN vírico. Cuando la ve, reproduce el ARN vírico con suma eficacia, pasando por alto las moléculas de ARN de la célula huésped, en número muchísimo mayor. La consecuencia es que la célula se ve en seguida desbordada por el ARN vírico. Éste se guarda en la proteína de protección del virus, también sintetizada en grandes cantidades, hasta que finalmente la célula revienta y libera una enorme cantidad de nuevas partículas víricas. Este es un programa que se realiza de modo automático y que se lleva a cabo hasta el más mínimo detalle (Eigen, 1992, pág. 40).

El autor, el biólogo molecular Manfred Eigen, se ha ayudado de un rico vocabulario de palabras «agentes»: para poder reproducirse, el virus debe «organizar» la proliferación de su información y en la persecución de su objetivo crea una enzima que «ve» su contraseña y «pasa por alto» otras moléculas. Se trata de una licencia poética, con toda seguridad: estas palabras amplían su significado para esta ocasión concreta. ¡Pero qué ampliación tan irresistible! Las palabras agentes llaman la atención sobre los rasgos más llamativos del fenómeno: que estas macromoléculas son *sistemáticas*. Sus sistemas de control no son simplemente eficientes en lo que hacen: son adecuadamente sensibles a las variaciones, son oportunistas, ingeniosas, tortuosas. Se las puede «engañar» pero sólo mediante novedades con que sus antecesores no se hayan encontrado con regularidad.

Estas pizquillas de maquinaria molecular, impersonales, irreflexivas, robóticas y sin mente, son base última de todo agente, y por lo mismo del significado, y por lo mismo de la conciencia, en el mundo. Es extraño que un hecho científico tan sólido e incontrovertible tenga tan poderosas consecuencias en la estructuración del debate subsiguiente sobre algo

tan controvertido y misterioso como la mente, de modo que hagamos una pausa para recordar cuáles son esas consecuencias.

Respecto a lo que sigue ya no hay duda sería que pueda plantearse: *somos descendientes directos de estos robots capaces de duplicarse a sí mismos*. Somos mamíferos y todos los mamíferos descienden de antecesores reptiles cuyos ancestros fueron peces cuyos ancestros fueron criaturas marinas del estilo de los gusanos, y que a su vez evolucionaron de criaturas multicelulares más sencillas millones de años antes, las cuales descendían de criaturas unicelulares que evolucionaron de macromoléculas con capacidad de duplicarse hace unos tres mil millones de años. Sólo hay un árbol genealógico en el cual pueden situarse todos los seres vivos que han vivido en este planeta alguna vez: y no sólo los animales, sino también las plantas, las algas y las bacterias. Con cualquier chimpancé, cualquier gusano, cualquier brizna de hierba, cualquier secuoya, compartimos un ancestro común. Entre nuestros progenitores, por tanto, estuvieron las macromoléculas.

Por decirlo gráficamente: ¡nuestra tatarara... tatarara... tatarabuela *fue un robot!* Y no sólo descendemos de esas macromoléculas robot sino que estamos compuestos de ellas: nuestras moléculas de hemoglobina, nuestros anticuerpos, nuestras neuronas, nuestra maquinaria de reflejo ocular-vestibular... en cualquier escalón de análisis de moléculas para arriba, nuestro cuerpo (incluyendo nuestro cerebro, por supuesto) se compone de una maquinaria que lleva a cabo estúpidamente una tarea maravillosa y elegantemente concebida.

Quizá hayamos dejado de estremecernos ante la visión científica de los virus y las bacterias ejecutando afanosa, irreflexivamente, sus subversivos proyectos: espantosos pequeños autómatas que llevan a cabo sus perversas acciones. Pero no deberíamos creer que podemos acomodarnos en la

idea de que *son* invasores ajenos, absolutamente dispares de los tejidos más familiares que nos conforman. Estamos hechos de los mismos tipos de autómatas que nos invaden: nuestros anticuerpos no se distinguen por un halo especial de humanidad de aquellos antígenos a los que combaten; lo único que ocurre es que los anticuerpos pertenecen al club que forma cada uno de nosotros y por ello pelean de nuestro lado. Los miles de millones de neuronas que se unen para formar nuestro cerebro son células, ese mismo tipo de entidad biológica que los gérmenes que nos causan las infecciones o que las células de levadura que se multiplican en la barrica en la que fermenta la cerveza o en la masa de harina cuando sube el pan.

Toda célula (agente diminuto que puede llevar a cabo un limitado número de tareas) carece de mente, más o menos igual que un virus. ¿Es posible que si se colocan juntos suficientes de estos estúpidos *homunculi*, hombrecillos, el resultado sea una persona auténtica y consciente, con una mente de verdad? Según la ciencia moderna, no hay otro modo de formar una auténtica persona. Ahora bien: desde luego del hecho de que descendamos de robots no se sigue el hecho de que nosotros lo seamos. Después de todo, también somos descendientes directos de los peces y no somos peces; somos descendientes directos de las bacterias pero no somos bacterias. Pero a menos que haya algún ingrediente secreto en nosotros (que es lo que solían creer los dualistas y los vitalistas), estamos *hechos de* robots o, lo que viene a ser lo mismo, cada uno de nosotros es una colección de billones de máquinas macromoleculares. Y todas ellas descienden, en último extremo, de las macromoléculas con capacidad de duplicarse que hubo en un principio. De manera que hay cosas hechas con robots que *pueden* mostrar genuina conciencia, habida cuenta de que nosotros somos el mejor ejemplo.

Me doy cuenta de que para ciertas personas esto suena chocante e improbable, pero sospecho que no se han dado

cuenta de lo desesperantes que son las alternativas posibles. El dualismo (punto de vista según el cual las mentes están compuestas de un sustrato no físico y completamente misterioso) y el vitalismo (punto de vista según el cual los seres vivos albergan cierta sustancia física especial pero igualmente misteriosa: el *élan vital*) se han visto relegados al basurero de la historia, juntamente con la alquimia y la astrología. Como no esté usted dispuesto a declarar que la tierra es plana y que el sol es un carro ardiente tirado por caballos alados (en otras palabras, como no desafíe prácticamente a toda la ciencia moderna) no tendrá lugar desde el que discutir y luchar por semejantes ideas obsoletas. De modo que veamos qué se puede relatar con los conservadores recursos de la ciencia. Puede que la idea de que nuestras mentes hayan evolucionado a partir de otras más sencillas no sea tan mala, después de todo.

Nuestros antepasados macromoleculares (y eso es exacta y no metafóricamente lo que son: nuestros antepasados) eran *parecidos* a agentes en determinados aspectos, como deja clara la cita que he reproducido de Eigen, mientras que en otros aspectos eran innegablemente pasivos, flotando por ahí al azar, viéndose empujados hacia uno u otro sitio... esperando a entrar en acción con las armas cargadas, podríamos decir, pero esperando sin *esperarlo*, sin *resolución* o sin *intención* alguna. Puede que tuvieran los «dientes» preparados, pero carecían de mente tanto como una trampa para cazar.

¿Qué fue lo que cambió? Bruscamente no cambió nada. Antes de que nuestros ancestros tuvieran mente, tuvieron cuerpo. Primero se hicieron células simples, o procariotas, y más adelante las procariotas aceptaron ciertos invasores o inquilinos convirtiéndose así en células complejas, las eucariotas. Ya en esa época, apenas mil millones de años tras la primera aparición de las células simples, nuestros antepasados ya eran máquinas extraordinariamente complejas (hechas de máquinas hechas de máquinas), pero seguían sin te-

ner mente. Seguían siendo en sus desplazamientos tan pasivas y sin dirección como siempre, pero ya estaban equipadas con muchos subsistemas especializados con el fin de extraer energía y materiales del entorno y de protegerse y de repararse por sí mismas cuando fuera necesario.

La complicada organización de todas estas partes coordinadas no se parecía mucho a una mente. Aristóteles les puso un nombre, a ellas o a sus descendientes: las llamó *alma nutritiva*. Un alma nutritiva no es una cosa; no es, por ejemplo, uno de los subsistemas microscópicos que flotan por el citoplasma de una célula. Es un *principio de organización*; es forma, no sustancia, como dijo Aristóteles. Todas las cosas vivas (no sólo las plantas y los animales sino también los organismos unicelulares) tienen cuerpos que requieren una organización que se regule y se proteja a sí misma y que pueda ser activada diferencialmente por distintas condiciones. Estas organizaciones las ha concebido brillantemente la selección natural y, en último extremo, se componen de montones de diminutos interruptores pasivos que pueden conectarse o desconectarse debido a circunstancias igualmente pasivas y que los organismos se encuentran en sus desplazamientos.

También usted, a semejanza de otros animales, tiene un alma nutritiva (una organización que se regula y se protege a sí misma) bastante diferente, y más antigua, que su propio sistema nervioso, consistente en su sistema metabólico, sus sistemas inmunitario y en otros sistemas asombrosamente complejos de autorreparación y de mantenimiento de la salud de su cuerpo. Las vías de comunicación utilizadas por esos sistemas iniciales no fueron nervios sino vasos sanguíneos. Antes de que existieran teléfonos y radios, existió un servicio postal, fiable aunque lento transporte de paquetes tangibles de información valiosa para todo el mundo. Y mucho antes de que hubiera sistemas nerviosos en los organismos, los cuerpos se confiaban a un especie de sistema pos-

tal, escasamente técnico: la circulación de fluidos por el cuerpo, transporte fiable aunque lento, de valiosos paquetes de información allí donde se necesitaran para el control o el mantenimiento. Podemos ver los herederos de este sistema postal primordial tanto en las plantas como en los animales. En los animales, el torrente sanguíneo transporta nutrientes y desperdicios, pero desde sus primeros momentos ha sido también una autopista de la información. El movimiento de los fluidos en el interior de las plantas proporciona asimismo un medio relativamente rudimentario de enviar señales de una parte a otra de la planta. Pero en los animales podemos descubrir una innovación importante en el diseño: la evolución de sistemas nerviosos sencillos (antecedentes del sistema nervioso autónomo) capaces de una transmisión de información más veloz y más eficiente, pero todavía dedicados, en su mayor parte, a los asuntos internos. Un sistema nervioso autónomo no es en absoluto una mente sino más bien un sistema de control, más en la línea del alma nutritiva de una planta, que preserva la integridad básica del sistema viviente.

Distinguimos radicalmente estos antiguos sistemas de nuestras propias mentes y, curiosamente, sin embargo, cuanto más nos fijamos en los detalles de su funcionamiento ¡más parecidos los encontramos a mentes! Esos pequeños interruptores son como órganos sensores primitivos y los efectos que se producen cuando esos interruptores se conectan o se desconectan se parecen a acciones intencionadas. ¿Y en qué? En ser efectos producidos por sistemas que adaptan la *información* y que tienen unos *objetivos*. Es como si estas células y estos conjuntos de células fueran *agentes* diminutos y sencillos, sirvientes especializados que fomentan sus propias causas obsesivas actuando según les dicta su percepción de las circunstancias. El mundo rebosa de entidades como estas que van desde el tamaño molecular hasta el continental y que no sólo comprenden objetos «naturales»,

como plantas, animales y sus respectivas partes (y las partes de sus partes) sino también muchos artefactos humanos. Los termostatos son un ejemplo bien conocido de esos pseudo-agentes simples.

A todas estas entidades, de la más sencilla a la más compleja, las llamo *sistemas intencionales* y la perspectiva desde la que se hace visible la tarea del agente (falso o genuino) la llamo *enfoque intencional*.

Adopción del enfoque intencional

El enfoque intencional es la estrategia que consiste en interpretar el comportamiento de un ente (persona, animal, artefacto, lo que sea) tratándolo *como si* fuera un agente racional que rigiera la «elección» de sus «actos» «teniendo en cuenta» sus «creencias» y sus «deseos». Estos términos entrecorillados reciben un uso ampliado además de su uso corriente en lo que suele llamarse «psicología popular», el discurso psicológico cotidiano que utilizamos para discurrir acerca de la vida mental de nuestros seres colegas, los seres humanos. El enfoque intencional es la actitud o la perspectiva que adoptamos ordinariamente unos con otros, de modo que adoptar el enfoque intencional en relación a otras cosas parece un modo deliberado de *antropizar* la cuestión. ¿Cómo es posible que esta sea una buena idea?

Intentaré mostrar que si se aplica con cuidado, la adopción del enfoque intencional no sólo es una buena idea sino que es la clave para desentrañar los misterios de la mente, de todos los tipos de mentes. Es un método que explota las similitudes para poder descubrir las diferencias, la enorme colección de diferencias que se han acumulado entre las mentes de nuestros ancestros y las nuestras, y también entre nuestras mentes y las de nuestros iguales que habitan en este planeta. Hay que usar este método con cuidado: hay que an-

dar en una cuerda floja entre las metáforas vacías, por un lado, y la falsedad literal, por otro. El uso inapropiado del enfoque intencional puede equivocar gravemente al investigador inadvertido, aunque adecuadamente comprendido puede proporcionar una perspectiva válida y fructífera en diferentes campos, mostrando una unidad subyacente en los fenómenos y dirigiendo nuestra atención a los experimentos cruciales que hay que llevar a cabo.

La estrategia básica del enfoque intencional es tratar al ente como un agente para poder predecir (y por ende, explicar, en cierto sentido) sus actos o sus movimientos. Los rasgos distintivos del enfoque intencional pueden verse mejor contrastándolo con otros dos enfoques o estrategias más básicas de predicción: el *enfoque físico* y el *enfoque del diseño*. El enfoque físico no es más que el laborioso método estándar de las ciencias físicas, en las que utilizamos todo lo que sabemos de las leyes físicas y de la constitución física de las cosas en cuestión para concebir nuestra predicción. Cuando predigo que una piedra que tengo en la mano caerá al suelo al soltarla, me estoy valiendo del enfoque físico. No atribuyo a la piedra ni creencias ni deseos: le atribuyo una masa, o peso, y me apoyo en la ley de la gravedad para hacer mi predicción. Para cosas que no sean ni seres vivos ni artefactos, el enfoque físico es la única estrategia posible, aunque puede llevarse a distintos grados de detalle, desde el nivel subatómico hasta el nivel astronómico. Las explicaciones de por qué el agua burbujea cuando hierve, de por qué se han formado las cordilleras o de dónde sale la energía del sol son explicaciones que proceden del enfoque físico. Todo objeto físico, vivo o no, está sujeto a las leyes de la física y por ello mismo se comporta de diversas maneras que pueden explicarse y predecirse a partir del enfoque físico. Si lo que suelto de mi mano es un despertador o un pez de colores, mi predicción es la misma acerca de su trayectoria descendente, de acuerdo con los mismos presu-

puestos. Hasta un aeromodelo o un pájaro, que bien pueden seguir una trayectoria diferente al soltarlos, se comportan de modo que obedecen las leyes de la física a cualquier escala y en todo momento.

Los despertadores, por ser objetos fabricados (a diferencia de la piedra), son también susceptibles de un estilo más refinado de predicción, la realizada a partir del enfoque del diseño. Este es un atajo maravilloso que utilizamos continuamente. Imaginemos que alguien me regala un nuevo despertador digital. Es de una marca y de un modelo nuevos para mí, pero un breve examen de sus botones y pantallas externos me convence de que *si* aprieto unos pocos botones sin más *entonces* el despertador producirá unas horas después un ruido fuerte. No sé qué clase de ruido será, pero será el suficiente para despertarme. No necesito conocer las leyes físicas concretas que explican esa maravillosa regularidad; no necesito desmontar el aparato, pesar sus partes y medir sus voltajes. Sencillamente, me limito a *dar por hecho* que tiene un diseño concreto (el diseño que llamamos reloj despertador) y que funcionará adecuadamente, conforme a su diseño. Respecto a esta predicción estoy dispuesto a arriesgar bastante: puede que no la vida, pero sí la hora de despertarme para llegar a tiempo a mi clase o a coger el tren. Las predicciones sobre el enfoque del diseño son más arriesgadas que las del enfoque físico debido al número suplementario de cosas que tengo que dar por supuestas: que un determinado ente *está* diseñado conforme a lo que parece, y que además funcionará de acuerdo con ese diseño, es decir, que no estará estropeado o no funcionará mal. Los objetos diseñados están en ocasiones mal diseñados y otras veces se estropean. Pero ese moderado precio que pago arriesgándome se ve más que compensado por la tremenda facilidad de comprensión. Cuando se puede aplicar, el enfoque del diseño es un atajo de bajo coste y de escaso riesgo que me permite refinar la tediosa aplicación de mi limitado conoci-

miento de la física. Lo cierto es que todos nos jugamos la vida basándonos en predicciones hechas sobre el enfoque del diseño: enchufamos y conectamos sin dudar nuestros aparatos eléctricos, que podrían matarnos si estuvieran mal cableados; nos metemos voluntariamente en autobuses que en seguida aceleran hasta alcanzar velocidades letales; apretamos los botones de un ascensor al que nunca antes nos hemos subido.

La predicción a partir del enfoque del diseño funciona maravillosamente bien para los artefactos bien diseñados, pero también funciona maravillosamente bien con los artefactos de la madre Naturaleza: los seres vivos y sus órganos. Mucho antes de que se entendieran la física y la química del crecimiento y de la reproducción vegetales, nuestros antepasados apostaron literalmente su vida sobre la fiabilidad de su conocimiento del enfoque del diseño acerca de lo que *se suponía* que harían las semillas una vez plantadas. *Si* entierro unas pocas semillas en el suelo, *entonces*, a los pocos meses, y mediante un mínimo cuidado por mi parte, tendré algo que comer.

Acabamos de ver que las predicciones a partir del enfoque del diseño son arriesgadas, comparadas con las predicciones basadas en el enfoque físico (que son seguras pero muy tediosas de elaborar); un enfoque más arriesgado y más rápido es el enfoque intencional. Puede verse, por así decir, como una subespecie del enfoque del diseño en la cual el objeto diseñado es una especie de agente. Supongamos que lo aplicamos al despertador. El despertador es mi criado: si *le ordeno* que me despierte, *dándole a entender* una hora concreta para despertarme, puedo confiar en su capacidad interna de *darse cuenta* de cuándo llega ese momento y de ejecutar sumisamente la acción prometida. En cuanto llegue *a creer* que el momento de producir el sonido YA ha llegado, se verá «motivado» a actuar en consecuencia, gracias a mis instrucciones previas. No cabe duda de que el despertador es

tan sencillo que este antropomorfismo imaginario es, estrictamente hablando, innecesario para que nosotros comprendamos por qué hace lo que hace: pero démonos cuenta de que de este modo podríamos explicar a un niño cómo utilizar el despertador: «Le dices cuándo quieres que te despierte, y él se acuerda de hacerlo y produce un ruido fuerte.»

La adopción del enfoque intencional es más útil (y, desde luego, casi obligatoria) cuando el artefacto en cuestión es mucho más complicado que un despertador. Mi ejemplo preferido es el del ordenador para jugar al ajedrez. Hay cientos de programas de ordenador que pueden convertir un ordenador, sea un portátil o una supercomputadora, en un jugador de ajedrez. Por muchas diferencias físicas y de diseño que puedan tener, todos estos ordenadores sucumben limpiamente ante la misma y sencilla estrategia de interpretación: piénsese en ellos como en agentes racionales que *quieren* ganar y que *saben* las reglas y los principios del ajedrez y las posiciones de las piezas en el tablero. Instantáneamente, el problema de predecir y de interpretar su comportamiento se hace muchísimo más fácil de lo que sería si utilizáramos los enfoques físico o del diseño. En cualquier momento del juego, basta mirar el tablero y elaborar una lista de todos los movimientos permitidos posibles para el ordenador cuando sea su turno (normalmente habrá docenas de esas posibilidades.) ¿Por qué restringirnos a los movimientos permitidos? Porque, razonamos, el ordenador quiere jugar al ajedrez para ganar y para ello sabe que sólo puede hacer movimientos permitidos de modo que, siendo racional, se limita a esos movimientos. Ordenemos ahora los movimientos permitidos de mejor (los más hábiles, los más razonables) a peor (los más estúpidos, aquellos que le llevan a la derrota) y hagamos nuestra predicción: el ordenador hará el mejor movimiento. Puede que no estemos seguros de cuál *sea* ese mejor movimiento (¡el ordenador puede «calibrar» la situación mejor que nosotros!) pero podemos eliminar todos los movimien-

tos menos cuatro o cinco, lo que nos da todavía una considerable ventaja predictiva.

A veces, cuando el ordenador se encuentra en una situación apurada con tan sólo un movimiento posible que no sea suicida (un movimiento «forzado») podemos predecir su movimiento con absoluta confianza. No hay nada en las leyes de la física que fuerce a realizar semejante movimiento, como tampoco lo fuerza nada en el diseño concreto del ordenador. El movimiento se ve forzado por las *razones* abrumadoramente a favor de hacerlo y de no hacer ningún otro. Cualquier jugador de ajedrez, hecho del material físico que se quiera, haría ese movimiento. ¡Hasta un fantasma o un ángel lo harían! Alcanzamos la predicción sobre el enfoque intencional basándonos en la atrevida suposición de que *sin importar cómo* se haya diseñado el programa de ordenador, estará suficientemente bien diseñado como para atender a tan buenas razones. Predecimos su conducta *como si* fuera un agente racional.

El enfoque intencional es, innegablemente, un atajo útil en tales casos pero ¿qué fiabilidad podemos otorgarle? En realidad ¿qué le importa a un ordenador si gana o si pierde? ¿A qué viene decir que el despertador *desea* obedecer a su amo? Podemos utilizar este contraste entre objetivos naturales o artificiales para realzar nuestra percepción del hecho de que todos los objetivos reales surgen en último término de la situación apurada de un objeto vivo que se protege a sí mismo. Pero también debemos reconocer que el enfoque intencional *funciona* (cuando funciona) sean o no genuinos, o naturales, o «auténticamente queridos» los objetivos por el llamado agente, y este margen es crucial para comprender cómo podría establecerse en primer lugar una discriminación genuina de objetivos. La macromolécula ¿*desea verdaderamente* duplicarse? El enfoque intencional explica lo que ocurre, con independencia de cómo contestemos a esa pregunta. Consideremos un organismo sencillo (pongamos una

planaria o una ameba) moviéndose pero no al azar por el fondo de una placa de laboratorio, dirigiéndose siempre al extremo de la placa que tiene los nutrientes, o alejándose del extremo tóxico. Este organismo está buscando el bien, o evitando el mal: *sus propios* bien y mal, no los de un usuario humano de artefactos. Buscar el bien propio es un rasgo fundamental de cualquier agente racional, pero estos organismos simples ¿buscan o se limitan a «buscar»? No hace falta que contestemos a esta pregunta. En cualquier caso el organismo es un sistema intencional predecible.

Se trata de otra forma de señalar lo mismo que Sócrates pretendía en el *Menón* al preguntar si hay alguien que intencionadamente desee el mal. Nosotros, sistemas intencionales, deseamos a veces el mal debido a una mala comprensión, a la desinformación o por pura demencia, pero es parte esencial de la racionalidad desear lo que se juzga bueno. Lo que se nos ha legado (o más bien, lo que se nos ha reforzado) mediante la selección natural de nuestros ancestros es esta relación constitutiva entre el bien y la búsqueda del bien: los que tuvieron la desgracia de estar diseñados genéticamente para buscar lo que era malo para ellos, a largo plazo terminaron por no dejar descendencia. No es accidental que los productos de la evolución busquen (o «busquen») lo que juzgan (o «juzgan») que es bueno.

Si tiene que primar lo que es bueno para ellos, hasta los organismos más sencillos necesitan algunos órganos sensores o poderes discriminatorios (algunos interruptores que se activen en presencia de lo bueno y que se desactiven en su ausencia) y estos interruptores o *transductores* pueden unirse para producir respuestas corporales correctas. Esta exigencia da origen a la *función*. Una roca no puede funcionar mal porque no está ni bien ni mal equipada para fomentar bien alguno. Cuando decidimos interpretar un ente a partir del enfoque intencional, es como si nos pusiéramos en la situación de guardianes suyos, preguntándonos en efecto: «Si

yo estuviera en esta situación de este organismo ¿qué haría?» Y es ahí donde sale a relucir el antropomorfismo subyacente del enfoque intencional: tratamos a todos los sistemas intencionales como si fueran igual que nosotros... cosa que no son, naturalmente.

¿Es que se trata de una aplicación incorrecta de nuestra propia perspectiva, la perspectiva que compartimos los *portadores de mente*? No necesariamente. Desde el punto de vista de nuestra historia evolutiva, lo que ha ocurrido ha sido lo siguiente: a lo largo de miles de millones de años, los organismos han evolucionado gradualmente, acumulando cada vez más maquinaria versátil diseñada para favorecer sus bienes cada vez más complejos y elaborados. Finalmente, en nuestra especie, con la evolución del lenguaje y de la variedad de reflexión que el lenguaje permite (y que trataremos en los próximos capítulos) surgimos con la capacidad de preguntarnos por aquello mismo que dio inicio a este libro: nos preguntamos sobre las mentes de otros entes. Estas preguntas, desarrolladas ingenuamente por nuestros antepasados, llevaron al *animismo*, la idea de que todo objeto animado tenía una mente o un alma (*ánima*, en latín). Comenzamos a preguntarnos no solamente si el tigre pretendía comernos (cosa probable) sino por qué los ríos querían llegar al mar y qué querían de nosotros las nubes como contrapartida por la lluvia que nosotros les pedíamos. Conforme fuimos haciéndonos más complejos (y este es un desarrollo histórico muy reciente y no algo que tiene que discernirse en las amplias extensiones de toda la evolución) fuimos abandonando gradualmente el enfoque intencional para lo que hoy llamamos naturaleza *inanimada*, reservándolo para cosas más parecidas a nosotros: animales, principalmente, pero también plantas bajo muchas condiciones. Seguimos «engañando» a las plantas con luz y calor primaverales obtenidos artificialmente para que florezcan antes de tiempo y «estimulamos» a las verduras a producir raíces mucho más largas suminis-

trándoles menos agua que la que buscan tan desesperadamente. (Un leñador me explicó en una ocasión cómo sabía que no iba a encontrar pinos albares entre los árboles de ciertas tierras altas de mi bosque: «A esos pinos les gusta tener los pies húmedos.») Este modo de razonar acerca de las plantas no sólo es natural e inofensivo sino que resulta ser una ayuda efectiva para la comprensión y una importante palanca para los descubrimientos. Cuando los biólogos descubren que una planta tiene un órgano discriminatorio rudimentario, inmediatamente se preguntan para qué sirve ese órgano: qué tortuoso proyecto tendrá la planta que exija obtener información del entorno de esa manera. Con mucha frecuencia, la respuesta es un importante descubrimiento científico.

Los *sistemas intencionales* son, por definición, todos y cada uno de esos entes cuya conducta es predecible/explicable a partir del enfoque intencional. Las macromoléculas que se duplican a sí mismas, los termostatos, las amebas, las plantas, las ratas, los murciélagos, las personas y los ordenadores que juegan al ajedrez son sistemas intencionales, algunos más interesantes que otros. Como el objetivo del enfoque intencional es tratar a un ente como agente para poder predecir sus acciones, tenemos que suponer que se trata de un agente inteligente, porque un agente idiota podría hacer cualquier tontería. Este atrevido salto de suponer que el agente sólo hará los movimientos inteligentes (dentro de su limitada perspectiva) es lo que nos proporciona la ventaja de hacer predicciones. Describimos esa perspectiva limitada atribuyendo unas creencias y deseos *particulares* al agente sobre la base de su percepción de la situación y de sus objetivos y necesidades. Como nuestra ventaja predictiva en este ejercicio depende fundamentalmente de esta peculiaridad (ya que es sensible al modo concreto en que nosotros, los teóricos, expresemos las creencias y los deseos, o representemos el sistema intencional en cuestión) yo llamo a estos sistemas, *sistemas intencionales*. Presentan lo que los filósofos llaman *intención* o *intencionalidad*.

«Intencionalidad» en este sentido filosófico especial es un concepto tan controvertido, y los que no son filósofos lo usan tan rutinaria y equívocamente, que tengo que hacer un alto para extenderme en su definición. Desgraciadamente para la comunicación interdisciplinaria, el término filosófico «intencionalidad» tiene dos compañeros de viaje, palabras perfectamente válidas que se han confundido con ella y que, desde luego, tienen mucha relación con ella. Una es un término corriente, la otra un término técnico (del que pospondré brevemente su introducción). En el habla corriente, solemos discutir si el acto de determinada persona fue intencionado o no. Cuando el conductor se estrella contra el contrafuerte de puente ¿tenía la intención de suicidarse o es que se había quedado dormido? Cuando se llama al policía «papá» en ese momento, ¿es intencionado o es un acto fallido? ¿No estamos preguntando sobre la intención de ambos hechos? Sí, en sentido corriente, pero no en sentido filosófico.

La intencionalidad en el sentido filosófico es sencillamente *tener que ver con*. Una cosa muestra su intencionalidad si su aptitud *tiene que ver con* alguna otra cosa. Otra forma de decirlo sería que una cosa que muestra intencionalidad alberga una *representación* de alguna otra cosa... aunque esta segunda fórmula la encuentro menos reveladora y más problemática. La cerradura ¿alberga una representación de la llave que la abre? Una cerradura y una llave muestran la forma más basta de intencionalidad; lo mismo que los receptores de opiáceos de las células cerebrales, receptores que están diseñados para aceptar las moléculas de endorfina que la naturaleza ha proporcionado a los cerebros a lo largo de millones de años. Y en ambos casos se les puede confundir, es decir, abrirlos mediante una llave impostora. Las moléculas de morfina son llaves esqueléticas artefácticas que se han obtenido recientemente para abrir también las puertas de los receptores de opiáceos. (Incluso fue el descu-

brimiento de estos receptores sumamente específicos el que inspiró la investigación que condujo al descubrimiento de las endorfinas, los analgésicos del cerebro. Los investigadores razonaron que algo tenía que haber en el cerebro para que estos receptores especializados estuvieran *ahí* desde un principio.) Este tener que ver tan simple, al estilo de la llave y la cerradura, es el elemento fundamental del diseño a partir del cual la naturaleza ha fabricado los más caprichosos tipos de subsistemas que pueden recibir con más merecimiento el nombre de sistemas de representación, de manera que tendremos que analizar ese tener que ver con de estos sistemas basándonos en ese (¿casi?) tener que ver con de las cerraduras y las llaves en cualquier caso. Exagerando la cosa podemos decir que la forma actual de un resorte bimetálico en un termostato es una representación de la temperatura actual de la habitación, y que la posición del nivel ajustable del termostato es una representación de la temperatura deseada de la habitación, pero de igual modo podemos negar que sean representaciones propiamente hablando. Con todo, sí que encarnan una información *que tiene que ver con* la temperatura de la habitación y precisamente por encarnarla contribuyen a la capacidad de un sistema intencional sencillo.

¿Por qué nosotros, los filósofos, llamamos a ese tener que ver con «intencionalidad»? Todo se remonta a los filósofos medievales que acuñaron el término, dándose cuenta de la similitud entre esos fenómenos y el acto de apuntar a algo con el arco (*intendere arcum in*). Los fenómenos intencionales están provistos de flechas metafóricas, podríamos decir, que apuntan a una cosa o a otra, a aquello a lo que se refieran, a lo que aludan o con lo que tengan que ver los fenómenos. Pero naturalmente, muchos de los fenómenos que presentan esta mínima suerte de intencionalidad no hacen nada *intencionadamente* en el sentido cotidiano del término. Los estados perceptivos, emocionales y memorísticos, por ejemplo, exhiben todos ellos un tener que ver con sin ser ne-

cesariamente intencionados en el sentido corriente del término; pueden ser completamente involuntarios o respuestas automáticas a unas cosas u otras. Nada hay de intencionado en reconocer a un caballo cuando surge en lontananza aunque nuestro estado de reconocimiento presenta un tener que ver con muy concreto: lo reconocemos *como* caballo. Si lo hubiéramos percibido erróneamente *como* un alce o *como* un hombre sobre una motocicleta, nuestro estado perceptivo habría tenido que ver con algo distinto. Hubiera apuntado su flecha de un modo bastante distinto, incluso en este caso a algo inexistente pero sin embargo bien definido: o al alce inexistente o al ilusorio motorista. Hay una gran diferencia psicológica entre creer erróneamente que se está en presencia de un alce y creer erróneamente que se está en presencia de un hombre sobre una motocicleta, diferencia de consecuencias predecibles. Los teóricos medievales se dieron cuenta de que la flecha de la intencionalidad podía apuntar a nada de igual forma que podía apuntar de un modo muy concreto. Al objeto de nuestro pensamiento, fuera real o no, lo llamaron *objeto intencional*.

Para poder pensar en algo, debemos disponer de un modo de pensarlo, uno entre otros muchos posibles. Cualquier sistema intencional depende de estos modos concretos de pensar (percibir, buscar, identificar, temer, recordar) se piense lo que se «piense». Esta dependencia es la que crea todas las posibilidades de confusión tanto teóricas como prácticas. En la práctica, la mejor manera de enredar y confundir a un sistema intencional concreto es explotar algún fallo de su modo, o modos, de percibir o de pensar aquello sobre lo que tenga que percibir o pensar, sea lo que fuere. La naturaleza ha explorado innumerables variaciones sobre este tema, porque confundir a otros sistemas intencionales es uno de los principales objetivos de la vida de la mayor parte de los sistemas intencionales. Después de todo, uno de los deseos básicos de cualquier sistema intencional vivo es el deseo de

alimento, necesario para abastecer el crecimiento, para la reparación de los daños sufridos y para la reproducción, de tal manera que todo ser vivo necesita distinguir el alimento (el material bueno) de todo lo demás. Se sigue que otro deseo básico es evitar convertirse en alimento de cualquier otro sistema intencional. De ese modo, el camuflaje, el mimetismo, el sigilo y montones de otras estratagemas han puesto a prueba a los cerrajeros de la naturaleza, provocando la evolución de maneras cada vez más efectivas de distinguir las cosas y seguirles la pista. Pero no hay sistema perfecto. No se puede *coger* eliminando toda posibilidad de *equivocarse*. Por eso es tan importante para nosotros como teóricos ser capaces de identificar y de distinguir las distintas variantes de «coger» (y de «equivocarse») que pueden darse en los sistemas intencionales. Para poder explicar el «coger» de un sistema intencional en sus circunstancias, tenemos que tener una imagen precisa de la dependencia que presenta de sus capacidades concretas para distinguir las cosas: su modo de «pensar sobre» las cosas.

Desgraciadamente, sin embargo, como teóricos hemos sido propensos a exagerarlo, considerando *nuestra propia* casi ilimitada capacidad de distinguir una cosa de otra en nuestros pensamientos (gracias a nuestra capacidad de usar el lenguaje) como si fuera el contraste de toda intencionalidad genuina, de todo «tener que ver con» merecedor de tal nombre. Por ejemplo, cuando la lengua de una rana se dispara y atrapa lo que vuela por allí cerca, la rana puede equivocarse: puede tragarse una bola de cojinete lanzada por algún niño perverso, o un cebo de pescador atado a un sedal, o cualquier otra anomalía no comestible. La rana ha cometido una equivocación, pero ¿qué equivocación o equivocaciones, *exactamente*? ¿Qué «creía» la rana que estaba atrapando? ¿Una mosca? ¿Comida que le llegaba por el aire? ¿Una convexidad oscura que se movía? Nosotros, como usuarios del lenguaje podemos hacer distinciones cada vez más suti-

les acerca del contenido de lo que pudiera ser el pensamiento de la rana, y se ha dado una suposición acrítica de que antes de que podamos atribuir a la rana cualquier *auténtica* intencionalidad tenemos que reducir el contenido de los estados y actos de la rana con la misma precisión que, en principio, podemos usar nosotros cuando consideramos los pensamientos humanos y su contenido proposicional.

Ello ha sido una enorme fuente de confusiones teóricas, y para empeorar las cosas, existe un término técnico muy práctico, extraído de la lógica, que se refiere precisamente a esa capacidad del lenguaje para discriminar cada vez con mayores matices: *intensión*. Con *s.* Esta intensión es un rasgo de los lenguajes: no tiene una aplicación directa a ningún otro tipo de sistema representativo (fotografías, mapas, gráficos, «imágenes virtuales»... *mentes*). Según el uso normalizado entre los lógicos, las palabras o símbolos en un lenguaje pueden dividirse en aquellas palabras lógicas o de función (*functores*) tales como «si», «y», «o», «no», «todos», «algunos», y los *términos* o *predicados*, que pueden ser tan diversos como los asuntos que se discutan, tales como «rojo», «alta», «abuelo», «oxígeno», «autor de sonetos de segunda categoría»... Todo término o predicado con significado de un lenguaje tiene una *extensión* (el objeto o conjunto de objetos al que se refiere el término) y una *intensión* (el modo concreto en que se escoge o se determina ese objeto o ese conjunto de objetos). «El padre de Chelsea Clinton» y «el presidente de Estados Unidos en 1995» nombran el mismo objeto (William Clinton) y por ello tienen la misma extensión, pero apuntan a este ente corriente de maneras diferentes y por ello tienen diferente intensión. El término «triángulo equilátero» señala exactamente el mismo conjunto de objetos que el término «triángulo equiángulo», de manera que los dos términos tienen la misma extensión aunque claramente no significan la misma cosa: un término *tiene que ver con* que los lados del triángulo sean iguales y el otro

tiene que ver con que los ángulos sean iguales. De modo que la intensión (con s) contrasta con la extensión y significa, en fin... *significado*. ¿Y no es eso lo que también significa la intención con c, la intencionalidad?

Los lógicos señalan que, a numerosos efectos, podemos pasar por alto las diferencias entre *intensiones* de los términos y atenernos sencillamente a sus *extensiones*. Después de todo, si una rosa recibiera cualquier otro nombre seguiría oliendo con igual perfume, de modo que cuando se tratara de rosas las muchas e indefinidas maneras de debatir sobre una determinada clase de rosas serían equivalentes desde un punto de vista lógico. Como el agua *es* H₂O, cualquier cosa cierta que se dijera del agua utilizando el término «agua» sería igualmente cierta si lo sustituyéramos por «H₂O», y ello incluso aunque estos dos términos difieran sutilmente en significado, o intención. Esta libertad es particularmente evidente y útil en ciertas áreas de conocimiento como las matemáticas donde siempre nos podemos permitir la práctica de «sustituir iguales por iguales» reemplazando «4»* por «16», y viceversa, habida cuenta de que estos dos términos se refieren a un número que es el mismo en ambos casos. A esa libertad de sustitución en un contexto lingüístico se le llama apropiadamente *transparencia referencial*: porque efectivamente, los términos permiten ver las cosas a las que se refieren. Pero cuando no se trata de rosas sino del *pensar sobre las rosas* o del *hablar sobre (el pensar sobre) rosas*, ahí sí pueden importar las diferencias en intención. De modo que cuando el asunto trata de sistemas intencionales, sus creencias y sus deseos, el lenguaje que utiliza el teórico es sensible a la intención. El lógico diría que ese discurso presenta *opacidad referencial*: porque no es transparente, por-

* Se trata de un juego de palabras intraducible. La frase puede significar «nuestras madres nos dieron a luz» y también «nuestras madres nos dan la tabarra». (N. del T.)

que los términos en sí se interponen e interfieren con el asunto de maneras sutiles y productoras de confusión.

Para comprobar cómo importa verdaderamente la opacidad referencial cuando adoptamos el enfoque intencional, consideremos un caso básico de enfoque intencional en pleno funcionamiento aplicado a un ser humano. Es cosa que hacemos sin esfuerzo alguno todos los días y rara vez manifestamos todo lo que supone, pero he aquí un ejemplo extraído de un artículo filosófico reciente, ejemplo que misteriosa pero útilmente entra en más detalles de los habituales:

Bruto quería matar a César. Creía que César era un mortal corriente y que, en ese caso, apuñalarle (lo cual quería decir hundirle el cuchillo en el corazón) sería un modo de matarle. Pensó que podría apuñalar a César, porque recordó que tenía un cuchillo y que César se colocaba a su izquierda en el Foro. De manera que Bruto se vio motivado para apuñalar al hombre que tenía a su izquierda. Y así lo hizo, matando por ello a César (Israel, Perry y Tutiya, 1993, pág. 515).

Nótese que el término César representa subrepticamente un doble papel fundamental en esta explicación: no sólo en la forma normal y transparente de distinguir a un hombre, César, como el tipo que está en el Foro vestido con una toga, sino en designar al hombre *tal como Bruto lo designa*. Para Bruto no basta que César se coloque a su lado; tiene que ver que *es* César, el hombre al que quiere matar. Si Bruto se equivocara y confundiera a César, que está a su izquierda, con Casio, no habría intentado matarlo; como dicen los autores, no se habría visto motivado para apuñalar al hombre que estaba a su izquierda ya que no hubiera establecido en su mente la conexión fundamental: el vínculo que identificaba al hombre que estaba a su izquierda con su objetivo.

El equívoco objetivo de la precisión proposicional

Siempre que actúa un agente, lo hace sobre la base de una comprensión o equívoco concretos de las circunstancias, y las explicaciones y predicciones intencionales se apoyan en comprender esa comprensión o ese equívoco. Para predecir la acción de un sistema intencional debemos saber en torno a qué giran las creencias y los deseos del agente y tenemos que saber, por lo menos de manera aproximada, *cómo* son esas creencias y esos deseos de tal manera que podamos decir si se han establecido o se establecerán las relaciones cruciales.

Pero démonos cuenta de que cuando digo que adoptamos el enfoque intencional tenemos que saber *por lo menos de manera aproximada* cómo elige el agente los objetos que le interesan. Si no nos damos cuenta de ello estamos ante una fuente enorme de confusiones. Es característico que no necesitemos saber *exactamente* cuál es el concepto que el agente tiene de su tarea. El enfoque intencional es capaz de soportar bastante descuido, cosa que es una bendición ya que la tarea de expresar *exactamente* cuál es el concepto que el agente tiene de su tarea es un error y un ejercicio tan inútil como leer los poemas de un libro con un microscopio. Si el agente que estamos estudiando no concibe sus circunstancias con ayuda de un lenguaje capaz de realizar determinadas distinciones, el poder resolutorio de nuestro lenguaje no puede aprovecharse directamente para la tarea de *expresar* los pensamientos concretos, o los modos de pensar concretos, o las variantes concretas de sensibilidad, del susodicho agente. (Con todo, el lenguaje puede usarse indirectamente para *describir* esas particularidades con el detalle que exija el contexto teórico.)

Este aspecto suele perderse en la maraña de un persuasivo argumento espurio tal como sigue. ¿Piensan los perros (por ejemplo)? Si así fuera, entonces deben tener pensa-

mientos concretos. Un pensamiento no puede existir sin ser este o aquel pensamiento ¿no? Pero un pensamiento concreto tiene que estar compuesto de conceptos concretos. No se puede pensar el pensamiento

tengo un filete en el plato

a menos que tengamos las ideas de *plato* y de *filete* y para tener estos conceptos necesitamos otro montón de conceptos (*cuenco, fuente, vaca, carne...*) ya que este pensamiento concreto (nos) es fácilmente distinguible del pensamiento

el cuenco está lleno de carne de vaca

lo mismo que del pensamiento

tengo la fuente llena de hígado de ternera

por no hablar del pensamiento

eso rojo y sabroso que me dan normalmente en ese objeto en el cual como no es eso seco que suelen darme para comer

y así sucesivamente. ¿Qué pensamiento o pensamientos piensa el perro? ¿Cómo podremos expresar con exactitud, y en castellano, por ejemplo, el pensamiento que está pensando el perro? Si no se puede (y es que no se puede) entonces o los perros no pueden pensar pensamientos en absoluto o los pensamientos de los perros deben ser sistemáticamente inexpresables... y por ende fuera de nuestro alcance.

Ninguna de las dos alternativas es lógica. Suele pasarse por alto la idea de que el «pensamiento» de un perro pudiera ser inexpresable (en lenguaje humano) por la sencilla razón de que la expresión en lenguaje humano *hila demasiado*

fino, lo mismo que suele pasarse por alto su corolario: la idea de que sin embargo podemos describir exhaustivamente lo que no podemos expresar, sin dejar residuo misterioso alguno. El perro ha de tener sus maneras concretas de discriminar cosas y esas maneras de discriminar se formulan en «conceptos» idiosincráticos bastante concretos. Si somos capaces de averiguar cómo operan esas maneras y de describir cómo funcionan conjuntamente, entonces sabremos tanto del contenido de los pensamientos del perro como podamos saber por medio de una conversación acerca del contenido de los pensamientos de otro ser humano, incluso aunque no podamos encontrar la frase (en castellano o en cualquier otro idioma humano) que *exprese* tal contenido.

Cuando nosotros, portadores humanos de mentes, desde nuestra singularmente elevada perspectiva utilizamos *nuestro* especial truco de aplicar el enfoque intencional a otros entes, les estamos imponiendo nuestras propias maneras y nos arriesgamos a introducir excesiva claridad, a introducir una excesiva discriminación y una excesiva articulación de los contenidos, y por ende una organización excesiva, en los sistemas que intentamos comprender. También nos arriesgamos a introducir una excesiva cantidad de nuestro *tipo* de organización procedente de nuestra mente en los modelos que formulemos de estos sistemas más sencillos que el nuestro. Estos candidatos a la posesión de mente que son más simples que nosotros no comparten todas nuestras necesidades, ni por tanto todos nuestros deseos, ni por tanto todas nuestras prácticas mentales, ni por tanto todos nuestros recursos mentales.

Muchos organismos «experimentan» el sol y hasta guían sus vidas ajustándolas a su ritmo. Un girasol puede seguir al sol de manera mínima, girando para volver la flor conforme atraviesa el cielo, maximizando su exposición diaria a la luz solar, pero no podría afrontar la aparición de un parasol que se interpusiera. No puede calcular la reaparición del sol en

un momento posterior calculable y ajustar consecuentemente su «conducta» lenta y sencilla. Un animal bien podría ser capaz de semejante complejidad, modulando su locomoción para mantenerse en la sombra oculto a los ojos de sus presas, o incluso calcular de antemano dónde tumbarse a echar una larga siesta dándose cuenta (vagamente y sin pensar) de que la sombra del árbol pronto será más larga. Los animales siguen y vuelven a identificar otras cosas (compañeros, presas, retoños, lugares preferidos para alimentarse) y podrían seguir al sol de manera similar. Pero es que nosotros, los seres humanos, no nos limitamos a seguir al sol, sino que hacemos un descubrimiento ontológico acerca del sol: ¡es *el sol!* El mismo sol, día tras día.

El lógico alemán Gottlob Frege presentó un ejemplo acerca del cual lógicos y filósofos llevan escribiendo más de un siglo: la Estrella Matutina, que los antiguos conocían como *Phosphorus*, y la Estrella Vespertina, conocida por los antiguos como *Hesperus*, son uno y el mismo cuerpo celeste: Venus. Hoy día este es un hecho bien conocido, pero el descubrimiento de esta identidad fue un primitivo avance sustancial para la astronomía. ¿Cuántos de nosotros podrían hoy día formular la argumentación y aportar la prueba crucial sin la ayuda de un libro? Sin embargo, y hasta de pequeños, comprendemos (y aceptamos dócilmente) en seguida esa hipótesis. Resulta difícil de imaginar que cualesquiera otras criaturas llegaran a formular, y muchos menos a confirmar, la hipótesis de que esos dos puntitos brillantes son uno y el mismo cuerpo celeste.

¿Es que no podría ser ese disco enorme y caliente que atraviesa los cielos diariamente uno nuevo cada día? Somos la única especie que puede llegar incluso a formular tal cuestión. Comparemos el sol y la luna a las estaciones. La primavera vuelve todos los años pero (ya) no nos preguntamos si se trata de la *misma* primavera que ha regresado. Puede que a la primavera, personificada en una diosa en la

época antigua, la *vieran* nuestros antepasados como algo concreto que retornaba y no como un universal recurrente. Pero es que para otras especies este asunto ni siquiera es tal. Algunas especies tienen una sensibilidad exquisita para las variaciones: en algunos terrenos pueden discriminar más detalles de los que nosotros podamos apreciar con nuestros sentidos sin más (aunque sí podemos realizar discriminaciones más detalladas en cualquier modalidad que cualquiera otra criatura del planeta con la ayuda de nuestras prótesis: microscopios, espectroscopios, cromatógrafos de gas, y demás). Pero esas otras especies tienen una limitadísima capacidad de reflexión y sus sensibilidades están canalizadas hacia conjuntos más bien estrechos de posibilidades, como ya veremos.

Por contra, nosotros somos *crédulos totales*. Aparentemente, no hay límites para lo que somos capaces de creer y para lo que somos capaces de distinguir en nuestras creencias. Podemos hacer una distinción entre creer

*que el sol es y siempre ha sido la misma estrella,
día tras día*

y creer

*que el sol ha sido siempre la misma estrella, día
tras día, desde el 1 de enero de 1900, fecha en la
que tomó el puesto de su predecesor el sol actual.*

Doy por sentado que nadie cree esto último pero resulta sencillo ver cuál es la creencia y distinguirla tanto de la creencia estándar como de la siguiente, igualmente tonta pero distinta,

*que el cambio de soles más reciente tuvo lugar el
12 de junio de 1986.*

La forma fundamental de todas estas atribuciones de estados mentales a los sistemas intencionales son frases que expresan lo que llamamos *actitudes proposicionales*:

x cree p
 y desea q
 z se pregunta si acaso r .

Tales frases constan de tres partes: un término que se refiere al sistema intencional en cuestión (x , y , z), un término que se refiere a la actitud que se le atribuye (creer, desear, preguntarse...) y un término para el contenido concreto o para el significado de tal actitud: la *proposición* que denotan en estos casos ficticios las letras p , q y r . Naturalmente, cuando se trata de frases con atributos reales, estas proposiciones se *expresan* como *frases* (del idioma castellano o cualquier otro que utilice el hablante) y estas frases contienen términos que pueden no ser sustituibles *ad lib* por términos coextensivos, rasgo que es el característico de la opacidad referencial.

Por tanto, las proposiciones son los entes teóricos mediante los cuales identificamos, o medimos, creencias. Por definición, para dos creyentes, compartir una creencia es creer en la misma proposición. Entonces ¿qué son las proposiciones? Por convención filosófica generalmente aceptada, las proposiciones son los significados abstractos que comparten todas las *frases* que... significan lo mismo. Del humo de la batalla surge un círculo ominoso. Presumiblemente, la misma proposición se expresa en las frases:

1. *Snow is white.*
2. *La neige est blanche.*
3. *Der Schnee ist weiss.*

Después de todo, cuando le atribuimos a Tom, el inglés,

la creencia de que la nieve es blanca, queremos que Pierre, el francés, y Wilhelm, el alemán, sean capaces de atribuir a Tom esa misma creencia en sus propios idiomas. Que Tom no tenga necesidad de entender esos atributos no tiene nada que ver. Si vamos a eso, Tom tampoco necesita comprender lo que *yo* le atribuyo, naturalmente, porque Tom puede resultar ser un gato o un turco que no habla inglés.

Pero ¿las frases que siguen comparten la misma proposición?

4. Bill golpeó a Sam.
5. Sam fue golpeado por Bill.
6. Bill fue el agente del acto de golpear del que fue víctima Sam.

Las tres «dicen lo mismo» y, sin embargo, todas «lo» dicen de forma diferente. Las proposiciones ¿deben entenderse como *formas de decir las cosas* o como *cosas dichas*? Un modo teórico sencillo y llamativo de organizar la cuestión sería preguntar si un creyente puede creer alguna de ellas sin creer las otras. Si así fuera, se trataría de diferentes proposiciones. Después de todo, si las proposiciones van a ser los entes teóricos que midan las creencias, queremos que superen esta prueba. Pero ¿cómo podemos comprobarlo si Tom no habla inglés o si ni tan siquiera habla? Nosotros, al adjudicar nuestros atributos (o, por lo menos, cuando expresamos nuestros atributos en el lenguaje) estamos sujetos a un sistema de expresión, a un idioma, y los idiomas difieren tanto en sus estructuras como en sus términos. Al vernos forzados a esta o aquella estructura del lenguaje, queramos o no asumimos más distinciones que aquellas que puedan garantizarnos las circunstancias. Este es el núcleo de la advertencia que hice anteriormente acerca de la atribución *gruesa* de contenidos que basta para el éxito del enfoque intencional.

El filósofo Paul Churchland (1979) ha asemejado las

proposiciones a números, objetos igualmente abstractos utilizados para medir muchas propiedades físicas

x tiene un peso en gramos de 144
y lleva una velocidad en metros por segundo de 12 .

Evidentemente, los números representan su papel como chicos bien educados. Podemos sustituir «iguales por iguales». No tenemos dificultad en admitir que x tiene un peso en gramos de 2×72 o que y lleva una velocidad en metros por segundo de $9 + 3$. Como acabamos de ver, hay una dificultad al tratar de aplicar las mismas reglas de transformación y equivalencia a diferentes expresiones de lo que son supe-
stamente la misma proposición. ¡Vaya! las proposiciones no se portan igual de bien que los números como entes teóricos. ¡Se parecen más a dólares que a números!

Esta cabra vale 50 dólares.

¿Y cuánto vale en dracmas griegos o en rublos rusos (¡y en qué día de la semana!)? ¿Y vale más o menos hoy que en la antigua Grecia o como parte de las provisiones que llevó Marco Polo en sus expediciones? No hay duda de que una cabra siempre tiene un valor para su dueño y no hay duda de que podemos fijar una medida operativa y aproximada de su valor llevando a cabo (o imaginando que llevamos a cabo) un intercambio de dinero, o de polvo de oro, o de panes, o de lo que sea. Pero no existe un sistema de medida del valor económico que sea fijo, neutro y eterno, del mismo modo que no existe un sistema de medida fijo, neutro y eterno que mida el significado de lo proposicional. ¿Y qué? Supongo que estaría bien disponer de tales sistemas: el mundo sería más claro y el trabajo del teórico más sencillo. Pero un sistema semejante de medida, universal y con un baremo único, es innecesario para la teoría tanto en la teoría de la eco-

nomía como en la teoría del sistema intencional. La teoría económica válida no se ve amenazada por esa imprecisión ineliminable que se da en su medida del valor económico generalizado a todas las circunstancias y a todas las épocas. La teoría del sistema intencional válido no se ve amenazado por esa imprecisión ineliminable que se da en su medida del significado a través del mismo espectro universal. Mientras estemos atentos a la dificultad, podremos abordar de manera satisfactoria todos los problemas concretos utilizando el sistema aproximado y rápido que escogamos.

En capítulos posteriores, descubriremos que cuando tomamos nuestra habilidad de «credulidad total» y la aplicamos a criaturas «inferiores», nos organiza los datos muy convenientemente: nos dice dónde tenemos que mirar a continuación, nos pone condiciones limitantes y resalta pautas de similitud y diferencia. Pero como ya hemos visto, si no tenemos cuidado puede distorsionarnos lamentablemente la visión. Una cosa es tratar un organismo o cualquiera de sus muchos subsistemas como sistema intencional rudimentario que persigue aproximadamente y *sin pensar* sus objetivos innegablemente complejos, y otra bastante diferente es imputarle una apreciación reflexiva de lo que está haciendo. El tipo de pensamiento reflexivo de que disponemos nosotros es una innovación evolutiva muy reciente.

Las macromoléculas iniciales que se duplicaban *tenían* razones para hacerlo, pero no tenían ni idea de tales razones. Por contra, nosotros no sólo sabemos (o creemos saber) nuestras razones, sino que las articulamos, las debatimos, las criticamos, las compartimos. No son sencillamente las razones de nuestros actos: son las razones de *nosotros mismos*. Entre las macromoléculas y nosotros hay toda una historia que contar. Considérese, por ejemplo, el caso del cuco incubado en un nido ajeno por sus involuntarios padres adoptivos. Su primer acto al salir del huevo es empujar los demás huevos fuera del nido. No es una tarea fácil y resulta bas-

tante sorprendente observar la feroz perseverancia y la inventiva con que la cría supera los obstáculos que se interponen en su camino para deshacerse de los demás huevos. ¿Por qué lo hace? Porque esos huevos albergan a otros que rivalizarían por las atenciones de sus alimentadores suplentes. Al eliminar esos rivales, maximiza el alimento y el cuidado protector que va a recibir. Por supuesto que el cuco recién nacido es inconsciente: no tiene ni idea del fundamento de ese acto despiadado, pero el fundamento *existe* y ha dado forma a esa conducta innata a lo largo de los eones. *Nosotros* sabemos verlo aunque el cuco no sepa. A un fundamento semejante lo denomino «latente» porque no está *representado* ni en la cría ni en ninguna parte, aun sirviendo (a lo largo del tiempo evolutivo) para modelar y refinar esa conducta concreta (por ejemplo, proporcionándole el modo de satisfacer sus necesidades de información). Los principios estratégicos que se ponen en marcha no están codificados explícita sino implícitamente en la organización más amplia de los rasgos diseñados. ¿Cómo se han captado y se han articulado esas razones en algunas de las mentes que han evolucionado? Buena pregunta. Nos mantendrá ocupados durante varios capítulos pero antes de entrar a considerarla, debo abordar una sospecha residual que han aireado varios filósofos, a saber: que lo he hecho al revés. ¡Propongo explicar la intencionalidad real basándome en la seudointencionalidad! Y lo que es más: da la impresión de que fracaso al reconocer la importante distinción entre intencionalidad *originaria* o *intrínseca* e intencionalidad *derivada*. ¿Cuál es esa distinción?

Intencionalidad originaria y derivada

Según algunos filósofos que siguen a John Searle (1980), la intencionalidad se presenta en dos variantes: la intrínseca (u originaria) y la derivada. La intencionalidad intrínseca es

aquello a lo que se refieren nuestros pensamientos, nuestras creencias, nuestros deseos, nuestras intenciones (intenciones en el sentido corriente del término). Es el *origen* obvio de aquello otro a lo que se refieren de manera limitada y derivada algunos de nuestros artefactos: nuestras palabras y frases, nuestros libros y mapas, nuestros cuadros, nuestros programas de ordenador. Todos éstos tienen intencionalidad sólo por la cortesía de la especie de préstamo generoso que les hace nuestra propia mente. La intencionalidad derivada de nuestras representaciones artefácticas es parásita de la intencionalidad genuina, intrínseca y originaria que se esconde tras su creación.

Mucho se puede decir de esta pretensión. Si cerramos los ojos y pensamos en París, o en nuestra madre, ese pensamiento nuestro se refiere a su objeto del modo más primario y directo que pueda relacionar una cosa con otra. Si a continuación escribimos una descripción de París, o hacemos un dibujo de nuestra madre, la representación que se encuentra en el papel tiene que ver con París, o con nuestra madre, sólo gracias a que esa es nuestra intención (en el sentido corriente del término) como autores. Nos hacemos cargo de nuestras representaciones y somos nosotros los que declaramos o decidimos a qué se refieren nuestras representaciones. Existen unas convenciones del lenguaje en las que nos apoyamos para ayudarnos en ese imprimir significado a unos signos realizados burdamente en el papel. A menos que hayamos declarado previamente que cuando escribimos o decimos la palabra «París» nos referimos a Boston, o que elegimos llamar «madre» a Michelle Pfeiffer, se da por hecho que están vigentes las referencias normales convenidas por nuestra comunidad lingüística. A su vez, estas convenciones dependen de las intenciones comunales de esa comunidad. De modo que las representaciones externas obtienen sus significados (intenciones y extensiones) a partir de los significados de los estados mentales, internos, y de los actos

de las personas que las inventan y las utilizan. Esos estados mentales y esos actos tienen intencionalidad originaria.

Es innegable la cuestión acerca del estatus dependiente de las representaciones artefácticas. Es manifiesto que las marcas de un lápiz sobre el papel, en sí mismas, no quieren decir nada. Cosa que queda especialmente clara en los casos de frases ambiguas. El filósofo W. V. O. Quine nos proporciona un bonito ejemplo:

Nuestras madres nos dan a luz

¿A qué se refiere esta frase? ¿Es una queja en presente de indicativo sobre el aburrimiento o es una verdad en pretérito sobre nuestros orígenes? Hay que preguntar a la persona que ha creado la frase. Por los signos en sí mismos no hay manera posible de decidir la respuesta. *Esas marcas* no tienen intencionalidad intrínseca sean las que sean. Si poseen algún significado se debe al papel que tienen en un sistema de representación anclado en las mentes de los que hacen la representación.

¿Pero qué puede decirse de los estados y de los actos de esas mentes? ¿Qué es lo que les otorga intencionalidad? Una respuesta corriente es decir que esos estados mentales y esos actos tienen significado porque ellos mismos, oh maravilla, se componen de una especie de lenguaje, el lenguaje del pensamiento. El mentalés. Es una respuesta inútil. Lo es no porque no pueda encontrarse con que hay tal sistema en el funcionamiento interno del cerebro de las personas. Y ciertamente podría haberlo, aunque un sistema semejante no sería *como* un idioma natural del tipo del castellano o del francés. Es una respuesta inútil a la pregunta planteada porque se limita a posponer la cuestión. Sea, que exista un lenguaje del pensamiento. Y entonces ¿de dónde procede el significado de *sus* términos? ¿Cómo sabemos lo que significan las frases en nuestro lenguaje del pensamiento? Este problema se

ve con mayor claridad si contrastamos la hipótesis del lenguaje del pensamiento con la hipótesis rival y antecesora de ésta, la teoría pictórica de las ideas. Nuestros pensamientos son como cuadros, expone este punto de vista; tratan de lo que tratan porque, como cuadros que son, *se parecen* a los objetos representados. ¿Cómo distingo mi idea de un pato de mi idea de una vaca? ¡Dándome cuenta de que mi idea de un pato *se parece a un pato* mientras que mi idea de una vaca no! Cosa que también es inútil porque inmediatamente surge la cuestión de ¿cómo sabemos qué aspecto tiene un pato? Y nuevamente no es que sea inútil porque no pudiera existir un sistema de imaginería en nuestro cerebro que explotara las semejanzas gráficas entre las imágenes internas del cerebro y las cosas representadas; es más, podría existir tal sistema. De hecho, existe y estamos empezando a comprender cómo funciona. Se trata de una respuesta inútil para nuestra pregunta fundamental, sin embargo, porque se apoya en la mismísima comprensión que se supone debe explicar, y por tanto es una respuesta circular.

La solución a este problema de nuestra intencionalidad es directa. Acabamos de acordar que los artefactos de representación (como las descripciones escritas y los dibujos) poseen una intencionalidad derivada en virtud del papel que desempeñan en las actividades de sus creadores. Una lista de la compra escrita en un trozo de papel posee únicamente la intencionalidad derivada que obtiene de las intenciones del agente que la escribió. Pues bien ¡lo mismo ocurre con una lista de la compra memorizada por el mismo agente! Su intencionalidad es igual de derivada que la de la lista externa, y por las mismas razones. De manera similar, una imagen meramente mental de nuestra madre (o de Michelle Pfeiffer) tiene que ver con su objeto del mismo modo derivado que el dibujo que hagamos. Es interna, no externa, pero sigue siendo un artefacto creado por nuestro cerebro y significa lo que significa debido a su posición concreta en la economía vi-

gente de las actividades internas de nuestro cerebro y en su papel de regir nuestras complejas actividades corporales en el mundo real que nos rodea.

¿Y cómo ha llegado nuestro cerebro a tener una organización de estados tan asombrosos con poderes tan asombrosos? Juguemos la misma baza otra vez: el cerebro es un artefacto y obtiene la intencionalidad que tengan sus partes, sea la que fuere, de su papel en la economía vigente de un sistema aún mayor del que forma parte: o, dicho con otras palabras, de las intenciones de su creadora, la madre Naturaleza (conocida también como proceso evolutivo por selección natural).

Esta idea de que la intencionalidad de los estados cerebrales se deriva de la intencionalidad del sistema o proceso que los ha diseñado es una idea ciertamente extraña e inquietante en un primer momento. Podemos darnos cuenta de sus consecuencias considerando un contexto en el cual es correcta con seguridad: a saber, cuando nos preguntamos sobre la intencionalidad (derivada) de los estados «cerebrales» de un robot manufacturado. Supongamos que nos encontramos con un robot empujando un carrito de la compra en un supermercado y que va consultando cada cierto tiempo una tira de papel con símbolos escritos.

LECHE@ENVDE2 si P<2xENVDE2\\$, si no 2xLECHE@ENVDE1

¿*De qué trata* este galimatías, si es que tiene algún sentido? Le preguntamos al robot. Y nos responde: «Es para acordarme de que tengo que comprar un envase de leche de 2 litros, pero sólo si el precio del envase de 2 litros es menor que el doble del precio de un envase de 1 litro. Transporte con más facilidad los envases de 1 litro que los de 2.» Este artefacto sonoro emitido por el robot es fundamentalmente la traducción al castellano del artefacto escrito, pero lleva *su* significado derivado en la manga, para que

nos enteremos. ¿Y de dónde sacan cada uno de estos dos artefactos su intencionalidad derivada? Sin duda del inteligente trabajo ingenieril de los diseñadores del robot, pero puede que muy indirectamente. Puede que los ingenieros hayan formulado e instalado el principio consciente del precio que ha engendrado este recordatorio concreto: una posibilidad bastante aburrida, pero posibilidad en que la intencionalidad derivada de estos estados decididamente nos retrotraería a la propia intencionalidad de los diseñadores humanos como creadores de esos estados. Habría sido mucho más interesante que los diseñadores hubieran hecho algo más profundo. Es posible (esta posibilidad se encuentra en el límite de la capacidad técnica actual) que hubieran diseñado el robot para que fuera sensible al precio de muchas maneras distintas y le hubieran dejado «deducir» a partir de su propia «experiencia» que debiera adoptar semejante principio. En este caso, el principio no sería rígido sino flexible y en un futuro próximo el robot podría decidir partiendo de su mayor «experiencia» que este «programa» no tenía tanta trascendencia y que a partir de ese momento compraría la leche en envases de 1 litro por su propia comodidad, independientemente de lo que costaran. ¿Qué parte del trabajo de diseño han hecho los diseñadores del robot y qué parte han delegado en el propio robot? Cuanto más complejo sea el sistema de control, con sus consiguientes subsistemas de captación de información y de valoración de la información, mayor será la aportación del propio robot y por ello será mayor su pretensión de ser el «autor» de sus propios significados... significados que, con el tiempo, pueden llegar a ser bastante inescrutables para los diseñadores del robot.

El robot que hemos imaginado no existe aún, pero podría existir algún día. Lo presento aquí para poder mostrar que *dentro* de su mundo de intencionalidad meramente derivada podemos trazar la misma distinción que inspiró en un

primer momento el contraste entre intencionalidades originaria y derivada. (Teníamos que «consultar al autor» para descubrir el significado del artefacto.) Cosa instructiva porque muestra que la intencionalidad derivada puede derivarse de otra intencionalidad derivada. Muestra asimismo cómo podría surgir la ilusión de la intencionalidad intrínseca (la intencionalidad *metafísicamente* originaria). Podría *parecer* que el autor de un artefacto desconcertante tendría que tener intencionalidad intrínseca para poder ser la fuente de la intencionalidad derivada que tiene el artefacto, pero no es así. Podemos ver que, por lo menos en este caso, no se le deja tarea a la intencionalidad *intrínseca*. El robot imaginado sería igual de capaz que nosotros de delegar intencionalidad derivada a otros artefactos. Y se mueve por el mundo, haciendo progresar sus proyectos y evitando el daño, apoyado en la fuerza de su intencionalidad «meramente» derivada, esa intencionalidad que se ha introducido en su diseño: primero por sus diseñadores y luego, conforme va adquiriendo más información de su mundo, por sus propios procesos de rediseño. Puede que nosotros nos encontremos en la misma tesitura, viviendo nuestra vida a la luz de nuestra intencionalidad «meramente» derivada. ¿Qué sacaríamos de disponer de una intencionalidad intrínseca (fuera lo que fuese ésta) que no se nos hubiera legado como artefactos diseñados por la evolución? Puede que estemos persiguiendo el sexo de los ángeles.

Es bueno que se nos haya abierto semejante perspectiva porque la intencionalidad que nos permite hablar, escribir y preguntarnos toda clase de cosas es indudablemente un producto complejo y tardío de un proceso evolutivo que posee los tipos más bastos de intencionalidad (denigrados por Searle y otros «*como si* fueran intencionalidad») al igual que sus antepasados y sus componentes contemporáneos. Descendemos de robots y estamos compuestos de robots y la intencionalidad de la que disfrutamos se deriva de

la intencionalidad más básica de esos miles de millones de sistemas intencionales más bastos. No: no lo he captado al revés, sino al derecho. Es la única dirección promisoría en la que se puede viajar. Pero nos queda todo el camino por hacer.

CAPÍTULO 3

EL CUERPO Y SUS MENTES

Veo en un futuro lejano campos abiertos para investigaciones mucho más importantes. La psicología se apoyará en una nueva base, la de la necesaria adquisición gradual de cada uno de los poderes y capacidades mentales. Se arrojará luz sobre el origen del hombre y su historia.

Charles Darwin, *El origen de las especies*.

¿De la sensibilidad a lo sentiente?

Finalmente, emprendamos el camino. La madre Naturaleza (o, tal como la llamamos hoy, proceso de evolución por selección natural) no tiene ninguna previsión, pero ha ido formando gradualmente seres con previsiones. La tarea de una mente es fabricar futuro, tal como lo expresó en una ocasión el poeta Paul Valéry. Una mente es fundamentalmente algo que anticipa, un generador de expectativas. Barrena el presente buscando claves, que refina con la ayu-

da de los materiales que ha conservado del pasado y las convierte en anticipaciones del futuro. Y entonces actúa, racionalmente, sobre la base de esas cosas anticipadas con tanto esfuerzo.

Dada la inexcusable competitividad por los materiales en el mundo de las cosas vivas, la tarea que afronta cualquier organismo puede considerarse como una especie de versión del juego infantil del escondite. Buscamos lo que necesitamos y nos escondemos de aquellos que necesitan lo que tenemos. Los primeros seres que se duplicaron, las macromoléculas, tenían sus necesidades y desarrollaron medios sencillos (*¡relativamente sencillos!*) de satisfacerlas. Su búsqueda era un paseo al azar con un «captador» configurado para un objetivo definido. Cuando se topaban con lo adecuado, se lo apropiaban. Estas macromoléculas buscadoras no tenían plan, no tenían «imagen preconcebida», ni representación de las cosas buscadas fuera de la configuración de los «captadores». Era una cuestión de «cerradura y llave» y sanseacabó. De ahí que la macromolécula no supiera que estaba buscando, y que tampoco necesitara saberlo.

El principio de la «necesidad de saber» es conocidísimo en su aplicación al mundo del espionaje, tanto el real como el de ficción: no debe darse a ningún agente más información que la que estrictamente necesita para llevar a cabo su parte del trabajo. Prácticamente ese mismo principio ha estado en el candelero miles de millones de años y sigue estándolo de billones de formas distintas en el diseño de todo ser vivo. A los agentes (microagentes o pseudoagentes) de los cuales se compone un ser viviente (a modo de agentes secretos de la CIA o del KGB) se les entrega solamente la información que necesitan para poder realizar sus tareas sumamente limitadas. En el espionaje, el fundamento es la seguridad; en la naturaleza, el fundamento es la economía. La madre Naturaleza «descubrirá» antes el sistema más barato y con un diseño menos intensivo y lo seleccionará con absoluta miopía.

Por cierto, es importante darse cuenta de que el diseño más barato puede no ser el más eficiente o el más reducido de tamaño. Para la madre Naturaleza muchas veces puede ser más barato poner (o dejar) montones de materia extra, que no es funcional, por la sencilla razón de que ese material extra aparece por el proceso de duplicación y desarrollo y no hay manera de eliminarlo a bajo coste. Hoy se sabe que muchas mutaciones insertan un código que simplemente «desconecta» un gen sin eliminarlo... lo cual es un movimiento mucho más barato en el espacio genético. Un fenómeno paralelo en el mundo de la ingeniería humana se da de forma consabida en la programación de ordenadores. Cuando los programadores mejoran un programa (creando por ejemplo el Disparapalabras 7.0 que sustituye al Disparapalabras 6.1) la práctica normal consiste en crear un nuevo código origen adyacente al antiguo, obtenido mediante copia del antiguo que luego se edita y se cambia. Después, antes de poner en marcha o de compilar el nuevo código, «aparcen» el código antiguo: no lo borran del archivo que contiene el código origen sino que aíslan la versión antigua mediante unos símbolos especiales que indican al ordenador que debe saltarse todo lo que está entre paréntesis al compilar o al ejecutar el programa. Las instrucciones antiguas siguen existiendo en el «genoma» con la orden de que nunca se «expresen» en el fenotipo. No cuesta casi nada conservar el código antiguo y a lo mejor algún día puede hacernos falta. Por ejemplo, podrían cambiar las circunstancias del mundo, dando como resultado que la versión antigua fuera mejor de todas todas. O puede que mutáramos la copia extra de la versión antigua y la convirtiéramos en algo de valor. Un diseño obtenido con tanto esfuerzo no debe descartarse sin más ni más porque podría ser muy difícil reconstruirlo a partir de unas simples líneas. Tal como vamos viendo cada vez con más claridad, la evolución siempre se permite esta táctica, reutilizando una y otra vez los desechos de los procesos de

diseño anteriores. (He explorado con mayor profundidad este principio de la acumulación ahorrativa de diseño en *La peligrosa idea de Darwin*.)

Las macromoléculas no tenían por qué saber y sus descendientes unicelulares fueron mucho más complejos aunque tampoco necesitaban saber lo que hacían o por qué lo que hacían era la fuente de su sustento. De tal modo que a lo largo de miles de millones de años hubo razones, pero no seres que las formularan, o se las representaran o, incluso, y en sentido literal, seres capaces de apreciar esas razones. (La madre Naturaleza, el proceso de la selección natural, muestra su aprecio tácito por las buenas razones permitiendo, sin palabras y sin conciencia, que prosperen los mejores diseños.) Nosotros, teóricos de última hora, somos los primeros que *vemos* estas pautas y deducimos esas razones: los fundamentos latentes de los diseños que se han creado a lo largo de los eones.

Describimos esas pautas utilizando el enfoque intencional. Incluso los rasgos de diseño más sencillos en los organismos (rasgos permanentes más sencillos incluso que los interruptores ACTIVADO/DESACTIVADO) pueden instalarse y refinarse mediante un proceso que tiene una interpretación de enfoque intencional. Por ejemplo, las plantas no tienen mente por mucho que lo intente la imaginación del teórico, pero a lo largo del tiempo de evolución sus rasgos se han ido conformando mediante competencias que pueden reproducirse mediante modelos matemáticos de la teoría de juegos: ¡es *como si* las plantas y sus competidores fueran agentes igual que nosotros! Las plantas que tienen un historial evolutivo como presa continua de los herbívoros han desarrollado muchas veces una toxicidad contra esos herbívoros como respuesta. A su vez, los herbívoros suelen desarrollar una tolerancia específica en su sistema digestivo para esas toxinas concretas y vuelta a empezar, hasta que las plantas, desbaratada su primera intentona, desarrollaron una toxicidad

dad aún mayor o unos rebordes con pinchos, como siguiente medida en una carrera armamentista de medidas y contra-medidas. En cierto punto de esa carrera, los herbívoros pueden «elegir» no responder sino discriminar y volverse hacia otras fuentes de alimento, momento en que otras plantas no tóxicas pueden «imitar» a las tóxicas, explotando ciegamente una debilidad (visual u olfativa) en el sistema discriminatorio de los herbívoros, disfrutando así de un avance gratuito basándose en la defensa tóxica que ejercen otras especies vegetales. El fundamento que encontramos aquí es claro y predictivo, incluso aunque ni las plantas ni los sistemas digestivos de los herbívoros tengan mentes en el sentido corriente que atribuimos al término.

Todo esto ocurre con un ritmo dolorosamente lento para nuestras medidas. Pueden hacer falta miles de generaciones, miles de años, para que se haga uno solo de estos movimientos del escondite y para que reciba respuesta (aunque en algunas circunstancias el paso es sorprendentemente rápido). Las pautas del cambio evolutivo aparecen tan lentamente que son invisibles para nuestro ritmo normal de captación de información, de modo que es fácil que pasemos por alto su interpretación intencional o que la desechemos como si fuera un mero capricho o una metáfora. Este sesgo a favor de *nuestra* escala normal de tiempo puede denominarse *chovinismo de escala temporal*. Elijamos a la persona más lista, la más ingeniosa que conozcamos e imaginémosla filmada en cámara superlenta, por ejemplo, a 30.000 imágenes por segundo para luego proyectar la filmación a las normales 30 imágenes por segundo. Una sencilla réplica relampagueante, una ingeniosidad de esas que se dan «sobre la marcha», le saldría ahora de la boca como un glaciador, aburriendo al cinéfilo más empedernido. ¿Quién podría adivinar la inteligencia de su actuación, una inteligencia que sería inequívoca a velocidad normal? También estamos bajo el encantamiento de escalas de tiempo mal utilizadas que van en el sentido con-

trario, como demuestra clarísimamente la fotografía acelerada. Ver cómo crecen las flores, cómo producen capullos que luego florecen a los pocos segundos es verse arrastrado casi irremisiblemente hacia el enfoque intencional. ¡Mirad cómo prospera esa planta, como compite con su vecina creciendo para ocupar el mejor lugar al sol, desplegando desafiante sus hojas a la luz, parando los contragolpes y moviéndose y agitándose como un boxeador! Estas mismas pautas, proyectadas a velocidades diferentes, pueden revelar u ocultar la presencia o ausencia de mente... o eso nos parece. (La escala espacial muestra asimismo un poderoso sesgo inmanente; si los mosquitos fueran del tamaño de gaviotas, mucha más gente pensaría que tienen mente y si tuviéramos que mirar por un microscopio para ver las cabriolas de las nutrias, estaríamos mucho menos seguros de que les gusta la diversión.)

Para que podamos ver las cosas como poseedoras de mente tienen que darse al ritmo adecuado y cuando consideramos que algo posee mente no tenemos mucha elección: la percepción es casi irresistible. Pero esto ¿es sencillamente un hecho relativo a nuestro sesgo por ser observadores, o es un hecho relativo a las mentes? ¿Cuál es el *verdadero* papel de la velocidad en el fenómeno de la mente? ¿Podría haber mentes tan auténticas como las verdaderas que desarrollaran sus actividades en unos órdenes de magnitud más lentos que nuestras mentes? He aquí una razón para pensar que podría haberlas: si nuestro planeta fuera visitado por marcianos que pensarán los mismos pensamientos que nosotros sólo que miles o millones de veces más rápidamente, les pareceríamos tan idiotas como los árboles y tendrían la propensión a mofarse de la hipótesis de que nosotros tenemos mente. Si se mofaran, se equivocarían, ¿no? víctimas de su propio chovinismo de escala temporal. De modo que si queremos negar que haya una mente fundamentalmente lenta en su pensar tendremos que encontrar otra base que no sea nuestra pre-

ferencia por el ritmo de pensar humano. ¿Qué bases podrían ser esas? Puede que pensemos, quizá, que hay una velocidad mínima para la mente, al estilo de la velocidad mínima de escape necesaria para superar la gravedad y abandonar un planeta. Para que esta idea pueda atraer nuestra atención, y mucho más para que la apoyemos, necesitaríamos una teoría que dijera por qué habría de producirse eso. ¿Qué pasaría si un sistema va cada vez más y más deprisa hasta terminar por «romper la barrera de la mente» y crea una mente donde antes no había mente alguna? ¿No origina calor la fricción de las partes que se mueven, calor que por encima de cierta temperatura, lleve a la transformación química de alguna cosa? ¿Por qué habría de crear eso una mente? ¿Es como las partículas de un acelerador que se acercan a la velocidad de la luz convirtiéndose en partículas enormemente masivas? ¿Por qué habría de crear eso una mente? ¿Es que el rápido girar de las partes del cerebro teje en cierto modo un recipiente que contenga, que impida la salida de las partículas de mente que se acumulan hasta que una masa crítica las cohesiona en una mente? A no ser que se pueda proponer algo parecido a esto y *se pueda defender*, la idea de que la mera velocidad es esencial a la mente carece de valor, ya que existe una buena razón para sostener que lo que cuenta es la velocidad *relativa*: percepción, deliberación y acción a la rapidez suficiente (en relación con el entorno que se despliega ante ella) para conseguir los propósitos de la mente. Fabricar futuro no tiene utilidad alguna para ningún sistema intencional si sus «predicciones» llegan demasiado tarde para actuar basándose en ellas. La evolución favorecerá siempre al que piensa con rapidez por encima del que piensa con mayor lentitud a igualdad de las demás circunstancias, extinguiendo a los que no pueden cumplir sus plazos con regularidad.

Pero ¿y si hubiera un planeta en el que la velocidad de la luz fuera 100 kilómetros por hora y en el que todos los de-

más sucesos y procesos físicos se ralentizaran en proporción? Habida cuenta de que el ritmo de los sucesos del mundo físico no puede acelerarse ni ralentizarse en órdenes de magnitud (salvo en los fantásticos experimentos mentales de los filósofos), el requerimiento de una velocidad relativa sirve igual de bien que el requerimiento de una velocidad absoluta. Dada la velocidad a la que unas piedras arrojadas se acercan a sus respectivos blancos, y dada la velocidad a la cual se pueden propagar por la atmósfera unas advertencias audibles cualesquiera, y dada la fuerza que debe aplicarse para que un cuerpo de 100 kilos que corre a 20 kilómetros por hora se desvíe bruscamente a izquierda o a derecha... dadas todas estas especificaciones y otras muchas más firmemente fijadas para este suceso, los cerebros que sean útiles tienen que funcionar a unas velocidades mínimas bastante definidas, con independencia de las fantásticas «propiedades emergentes» que solamente pudieran producirse a determinadas velocidades. Estos requerimientos de velocidad de funcionamiento obligan a los cerebros, a su vez, a utilizar medios de transmisión de información que puedan mantener esas velocidades. Se trata de una buena razón por la que puede importar de qué está hecha una mente. Puede haber otras.

Cuando los sucesos en cuestión se desarrollan a un ritmo más tranquilo, puede darse en otros medios algo parecido a una mente. Estas pautas son discernibles en estos fenómenos solamente cuando adoptamos el enfoque intencional. A lo largo de períodos muy largos de tiempo, especies o linajes de plantas y animales pueden ser *sensibles* a las condiciones cambiantes y *responder* de manera racional a los cambios que notan. Es lo único que hace falta para que el enfoque intencional encuentre una ventaja explicativa y predictiva. En períodos mucho más cortos, las plantas como individuos pueden responder adecuadamente a los cambios que notan en el entorno, produciendo nuevos tallos y nuevas hojas para

explotar la luz solar disponible, alargando sus raíces en busca del agua e incluso (en algunas especies) ajustando *provisionalmente* la composición química de sus partes comestibles para mantener a raya el *presentido ataque* de los herbívoros que pasan por allí.

Esta especie de sensibilidad a ritmo lento puede chocarnos, como la sensibilidad artificial de los termostatos y ordenadores, como meras imitaciones de segunda categoría del fenómeno que marca auténticamente la diferencia: la *sentiencia*. Puede que podamos distinguir los «sistemas meramente intencionales» de las «mentes genuinas» preguntándonos si los candidatos de unos y otras disfrutan de sentiencia. Pues bien ¿qué es eso? Nunca se ha dado una definición apropiada de «sentiencia» pero es el término más o menos aceptado para lo que se imagina como el grado ínfimo de la conciencia. Llegados a este punto, podemos desear abordar la estrategia de comparar la sentiencia con la mera sensibilidad, fenómeno que presentan los organismos unicelulares, las plantas, el indicador de combustible del coche y la película de una cámara de fotos. La sensibilidad no precisa de conciencia en absoluto. La película fotográfica se vende en varias sensibilidades a la luz; los termómetros se hacen de materiales que son sensibles a los cambios de temperatura; el papel tornasol es sensible a la presencia de ácido. La opinión popular proclama que las plantas, y puede que los animales «inferiores» (medusas, esponjas, y demás), sean sensibles sin ser sentientes, pero que los animales «superiores» son sentientes. Como nosotros, no están *meramente* provistos de un equipo sensible de uno u otro tipo... un equipo que responde diferenciadamente a unas y otras cosas. Disfrutan de alguna otra propiedad, llamada sentiencia... eso dice la opinión popular. Pero ¿qué es esta propiedad reivindicada tan comúnmente?

¿A qué equivale esta sentiencia, que sea más que la sensibilidad y vaya más allá de ella? Se trata de una pregunta

que rara vez se hace y nunca ha recibido una respuesta adecuada. No deberíamos dar por hecho que existe una buena respuesta. No deberíamos dar por hecho, en otras palabras, que se trata de un pregunta bien formulada. Si queremos usar el concepto de sentiencia, tendremos que formarlo a partir de elementos que comprendamos. Todo el mundo está de acuerdo en que la sentiencia exige sensibilidad además de algún factor x aún sin identificar, de modo que si dirigimos nuestra atención a las distintas variedades de sensibilidad y en qué ocasiones se la explota, sin dejar de observar aquello que pueda llamarnos la atención como un añadido crucial, podremos descubrir la sentiencia sobre la marcha. Entonces podremos añadir el fenómeno de la sentiencia a nuestra historia evolutiva, o, en su lugar, podría evaporarse la idea de la sentiencia como categoría especial. De una u otra forma, salvaremos la distancia que nos separa a nosotros los conscientes de las macromoléculas meramente sensibles y no sentientes de las que descendemos. Para buscar la diferencia clave entre sensibilidad y sentiencia es tentador observar los materiales que las sustentan: los *medios* por los cuales viaja y se transforma la información.

Los medios y los mensajes

Debemos observar con mayor detenimiento el desarrollo que he esbozado al comienzo del capítulo 2. Los primeros sistemas de control no fueron verdaderamente más que protectores corporales. Las plantas están vivas, pero no tienen cerebro. No lo necesitan, habida cuenta del estilo de vida que llevan. Sin embargo, sí que necesitan mantener intactos sus cuerpos y adecuadamente colocados para sacar beneficio de su entorno más inmediato, para lo cual han desarrollado sistemas para controlar o regir que tuvieran en cuenta las variables cruciales y que reaccionaran de manera consecuente.

La preocupación de las plantas (y de ahí su rudimentaria intencionalidad) o bien estaba dirigida hacia el interior, hacia las condiciones internas, o bien estaba dirigida hacia las condiciones que se daban en todas las zonas importantes de frontera entre su cuerpo y el mundo cruel. La responsabilidad de monitorizar y de ajustar estaba distribuida y no centralizada. La captación localizada de los cambios en las condiciones podía tener respuestas localizadas, que fueran abiertamente independientes unas de otras. En ocasiones esto producía problemas de coordinación, en los cuales un equipo de microagentes actuaba a contrapelo de otro. Hay ocasiones en que las decisiones independientes son mala idea; si todo el mundo decide pasarse al lado derecho cuando el barco se escora a la izquierda, el barco bien puede escorarse a la derecha. Pero en su conjunto, las estrategias minimalistas de las plantas pueden verse satisfechas mediante una «toma de decisiones» muy repartida y modestamente coordinada por el intercambio rudimentario y lento de información a base de la difusión de los fluidos que circulan por el cuerpo de la planta.

¿Podrían entonces ser las plantas sencillamente «animales lentísimos» que disfrutarían de una sentiencia que se nos haya pasado por alto debido a nuestro chovinismo de escala temporal? Habida cuenta de que no hay un significado establecido del término «sentiencia» somos libres de adoptar uno de nuestra elección si es que podemos justificarlo. Si quisiéramos, podríamos referirnos a la sentiencia como capacidad de respuesta de la planta a su entorno, lenta aunque segura, pero necesitaríamos alguna razón para distinguir esta cualidad de la mera sensibilidad que exhiben las bacterias y otras formas de vida unicelular (por no decir nada de los fotómetros de las cámaras fotográficas). No hay candidatura posible para semejante razón y sí existe una razón bastante fuerte para reservar el término «sentiencia» para algo más especial: los animales poseen sistemas lentos de manteni-

miento corporal parecidos a los de las plantas, y la opinión común diferencia entre el funcionamiento de estos sistemas y la sentiencia de un animal.

Los animales tienen sistemas lentos de mantenimiento corporal desde que existen. Algunas de las moléculas que flotan en medios como la corriente sanguínea son en sí *funcionales*, directamente «hacen cosas» por el cuerpo (por ejemplo, algunos de ellos destruyen los invasores tóxicos en un combate cuerpo a cuerpo) y algunos son más parecidos a *mensajeros* cuya aproximación a un agente mayor y «reconocimiento» por éste le dice que «debe hacer cosas» (por ejemplo, acelerar el ritmo cardíaco o empezar a vomitar). A veces, ese agente mayor es el cuerpo entero. Por ejemplo, cuando la glándula pineal de ciertas especies detecta una disminución general de la luz solar día tras día, lanza a todo el cuerpo un mensaje hormonal para comenzar a prepararlo para el invierno: una tarea con muchas subtareas, iniciadas todas ellas a partir de un único mensaje. Aunque la actividad en estos viejos sistemas hormonales puede verse acompañada de poderosos ejemplos de lo que suponemos que es la sentiencia (como, por ejemplo, arcadas de náuseas, o sensación de entorpecimiento, o escalofríos, o brotes de lujuria) estos sistemas funcionan independientemente de esos acompañamientos sentientes, por ejemplo, en animales dormidos o comatosos. Los médicos hablan de seres humanos con muerte cerebral mantenidos vivos mediante respiradores en un «estado vegetativo», cuando son estos sistemas de mantenimiento corporal los únicos que mantienen la vida y el cuerpo unidos. Ha desaparecido la sentiencia pero sigue persistiendo una sensibilidad de muchos tipos manteniendo diversos equilibrios corporales. O, por lo menos, así querrían aplicar estos dos términos muchas personas.

En los animales, este complejo sistema de paquetes bioquímicos de control de información se complementó en su momento con un sistema más veloz y que discurría por un

medio diferente: impulsos de actividad eléctrica que viajaban por las fibras nerviosas. Esta variante abrió un espacio de oportunidades para reacciones más rápidas pero asimismo permitió que el control estuviera distribuido de diferente manera debido a las diferentes geometrías de conexión posibles en este nuevo sistema, el sistema nervioso autónomo. Las preocupaciones del nuevo sistema fueron en un principio internas o, en cualquier caso, inmediatas en el espacio y en el tiempo: ¿Tenía que temblar el cuerpo ahora o debía sudar? ¿Debían posponerse los procesos digestivos del estómago porque había una demanda más urgente de sangre? ¿Había de iniciarse la cuenta atrás de la eyaculación? Y así sucesivamente. Las zonas de contacto entre el nuevo medio y el viejo tenía que proporcionarlas la evolución, y la historia de tal desarrollo ha dejado su huella en nuestra organización actual, haciéndola mucho más complicada de lo esperable. Pasar por alto estas complejidades ha solido despistar a los teóricos (entre los que me incluyo) de modo que más vale que las veamos, aunque sea brevemente.

Una de las suposiciones fundamentales que comparten muchas teorías modernas sobre la mente es la conocida como *funcionalismo*. La idea central es bien conocida en la vida cotidiana y tiene muchas expresiones proverbiales tales como «la función crea el órgano». Lo que convierte a algo en mente (o en creencia, o en dolor, o en temor) no es su composición, de qué está formado, sino aquello *que es capaz de hacer*. En otras disciplinas nos damos cuenta de que este principio no admite discusión, sobre todo en nuestra evaluación de un artefacto cualquiera. Lo que hace que un objeto sea una bujía es que puede conectarse en una situación dada y *producir una chispa cuando se le solicita*. Es lo único que importa: color, material de construcción o complejidad interna pueden variar casi a discreción, lo mismo que su forma, siempre que esa forma permita satisfacer las dimensiones específicas de su papel funcional. En el mundo

de las cosas vivas, el funcionalismo se comprende bien: un corazón es algo que bombea sangre y un corazón artificial o de cerdo pueden servir más o menos igual de bien, y por ello pueden cambiarse por un corazón dañado en un cuerpo humano. Hay más de un centenar de diferentes variedades químicas de la valiosa proteína lisozima. Lo que las convierte en diferentes ejemplos de lisozima es lo que las hace valiosas: aquello que son capaces de hacer. Son intercambiables para casi cualquier propósito y situación.

En la jerga estándar del funcionalismo, estas entidades definidas funcionalmente admiten *concreciones múltiples*. ¿Por qué no iban a poder hacerse reales (concretarse) casi a partir de cualquier cosa las mentes artificiales, a semejanza de los corazones artificiales? Una vez que averigüemos lo que hacen las mentes (o lo que hacen los dolores, o lo que hacen las creencias, y así sucesivamente) deberíamos ser capaces de crear mentes (o partes de mentes) a partir de materiales distintos que tengan esas capacidades. Y a muchos teóricos (entre los que me incluyo) les ha parecido evidente que lo que las mentes hacen es *procesar información*; las mentes son los sistemas de control de los cuerpos y para poder ejecutar las tareas que se les han encomendado necesitan reunir, discriminar, almacenar, transformar y, en general, procesar información de las tareas de control que llevan a cabo. Hasta aquí, todo claro. El funcionalismo, lo mismo aquí que en otras partes, promete hacer al teórico la vida más fácil abstrayéndose de algunas de las enojosas peculiaridades del funcionamiento y centrándose en lo que de verdad se está haciendo. Pero casi es obligado que los funcionalistas simplifiquen su concepto de esta tarea, haciéndole *demasiado* fácil la vida al teórico.

Resulta tentador pensar en un sistema nervioso (bien en un sistema nervioso autónomo o bien en su acompañante posterior, el sistema nervioso central) como si fuera una red de información ligada en diversos lugares concretos (nodos transductores, de *entrada*, y nodos efectores, de *salida*) a la

realidad del cuerpo. Un *transductor* es cualquier artilugio que capte información de un medio (un cambio de la concentración de oxígeno en la sangre, una disminución de la luz ambiental, una elevación de la temperatura) y la traslade a otro medio. Una célula fotoeléctrica transduce la luz, que le llega en forma de fotones, en una señal electrónica, en forma de electrones que circulan por un cable. Un micrófono transduce ondas sonoras en señales en ese mismo medio electrónico. El resorte bimetálico de un termostato transduce los cambios de temperatura ambiente en una curvatura del resorte (lo cual, a su vez, se traduce generalmente en la transmisión de una señal electrónica por un cable para encender o apagar el calentador). Los bastones y conos de la retina del ojo son los transductores de la luz en el medio de las señales nerviosas; el tímpano transduce las ondas de sonido en vibraciones que, a su vez, se ven transducidas (por las células pilosas de la membrana basilar) al mismo medio de señales nerviosas. Hay transductores de temperatura distribuidos por todo el cuerpo, transductores del movimiento (en el oído interno) y montones de otros transductores para otros tipos de información. Un *efector* es cualquier dispositivo que puede ser obligado, mediante una señal distribuida por un medio, a hacer algo en otro «medio» (doblar un brazo, cerrar un poro, secretar un fluido, hacer un ruido).

En un ordenador existe una frontera bien definida entre el mundo «exterior» y los canales de información. Los dispositivos de entrada de información, como las teclas, el ratón, el micrófono o el monitor, transducen todos ellos información a un medio común, es decir, el medio electrónico en los cuales se transmiten, se almacenan y se transforman los «bits». También puede un ordenador disponer de transductores internos, como el transductor de temperatura que «informa» al ordenador de que se está recalentando, o el transductor que le advierte de las irregularidades que se dan en la fuente del suministro eléctrico, pero éstos cuentan como dis-

positivos *de entrada de información* habida cuenta de que extraen información del entorno (interno) y la introducen en el medio común de procesado de la información.

La cosa estaría clara teóricamente si pudiéramos aislar los canales de información procedente de los sucesos «exteriores» en el sistema nervioso de un cuerpo, de tal manera que las interacciones importantes se dieran en transductores y efectores identificables. Suele ser muy ilustrativa la división del trabajo que ello permitiría. Consideremos un barco con una rueda de timón muy alejada del timón que efectivamente controla. Se puede conectar la rueda al timón mediante cuerdas o con unos piñones y cadenas de bicicleta, con alambres y poleas o mediante un sistema hidráulico de mangueras llenas de aceite (¡o de agua o de whisky!). De una u otra manera, lo que se consigue es que estos sistemas transmitan al timón la energía que el timonel aplica a la rueda para girarla. *O también* se puede conectar la rueda al timón nada más que con unos finos alambres por los cuales pasen una señales electrónicas. No hace falta transducir la energía sino tan sólo la información procedente de la rueda que indique *cómo* quiere el timonel que gire el timón. Se puede pasar esta información desde la rueda en un extremo para que se transmita como una señal y aplicar la energía de manera localizada, en el extremo opuesto, mediante un efector, es decir, un motor del tipo que sea. (Además, se pueden añadir mensajes de «retroalimentación», que se transduzcan en el punto donde están el motor y el timón y envíen a la rueda el control de la resistencia al giro, de tal modo que el timonel pueda notar la presión del agua en la rueda cuando la hace girar. En la actualidad, esta retroalimentación es lo normal en el giro del volante de los automóviles, pero en cambio faltaba, cosa peligrosísima, en los primeros tiempos de la dirección asistida.)

Si se opta por este tipo de sistema (un sistema de señalización pura que transmita la información y casi nada de

energía) entonces no hay diferencia en absoluto entre que las señales sean electrones que pasen por un cable o fotones que pasen por una fibra óptica u ondas de radio que pasen por el espacio vacío. En todos los casos, lo que importa es que la información no se pierda ni se distorsione debido al lapso que transcurre entre el giro de la rueda y el giro del timón. Lo cual es también una exigencia clave en los sistemas transmisores de energía, es decir, los sistemas que utilizan vínculos mecánicos como las cadenas, o los alambres, o las mangueras. Ese es el motivo de que las tiras de goma no sean tan buenas como los cables inextensibles a pesar de que la información termine por llegar a su destino, o también por qué el aceite, que no se puede comprimir, es mejor que el aire para hacer funcionar un sistema hidráulico*.

En las máquinas modernas suele ser posible aislar de este modo el sistema de control del sistema controlado, de tal modo que los sistemas de control puedan intercambiarse fácilmente sin pérdidas de función. Son ejemplos evidentes de esto los familiares mandos a distancia de los aparatos electrónicos, lo mismo que los sistemas electrónicos de ignición (que reemplazan a las antiguas conexiones mecánicas) y demás dispositivos de los automóviles basados en chips de ordenador. Y hasta cierto punto, esa misma independencia de un medio concreto es un rasgo de los sistemas nerviosos animales, cuyas partes pueden descomponerse con bastante claridad en transductores periféricos y en efectores, y en sendas intermedias de transmisión. Por ejemplo, una manera de quedarse sordo es perder el nervio auditivo por un

* El ejemplo de la rueda del timón tiene un importante pedigrí histórico. El término «cibernética» lo acuñó Norbert Wiener a partir de la palabra griega que significa «timonel». La palabra «gobernador» tiene la misma etimología. Estas ideas de cómo se consigue el control mediante la transmisión y el procesado de información las formuló claramente por primera vez Wiener en *Cibernética* (1986).

cáncer. Las partes del oído sensibles al sonido siguen intactas pero la transmisión de los resultados de su trabajo hacia el resto del cerebro se ha visto interrumpida. Ese camino destruido puede reemplazarse hoy día con una prótesis que sirva de conexión, un cable diminuto hecho de un material diferente (metal, como en un ordenador corriente) y como los intercambiadores que lleva en ambos extremos pueden sintonizarse con las zonas sanas existentes, las señales pueden transmitirse. Se puede oír otra vez. No importa en absoluto cuál sea el medio de transmisión siempre que la información pase a su través sin pérdidas ni distorsiones.

Esta importante idea teórica lleva en ocasiones, sin embargo, a serias confusiones. La más seductora de ellas podría denominarse mito de la doble transducción: en un primer paso, el sistema nervioso transduce la luz, el sonido, la temperatura y demás a señales neurales (un rosario de impulsos en las fibras nerviosas) y en segundo lugar, en un determinado lugar central especial, transduce estas series de impulsos a *otro* medio: ¡el medio de la conciencia! Eso fue lo que pensó Descartes sugiriendo que la glándula pineal, que está justamente en el centro del cerebro, era el lugar en el que se producía esta segunda transducción... al medio misterioso y no físico de la mente. Hoy día casi nadie que estudie la mente piensa que exista tal medio no físico. Sin embargo, y por extraño que parezca, la idea de una segunda transducción a un medio *físico* o *material*, en un lugar del cerebro todavía por identificar, sigue seduciendo a ciertos teóricos incautos. Es como si vieran (o creyeran ver) que como la actividad periférica del sistema nervioso es mera sensibilidad, tendría que haber un lugar más central en el que se originara la sentiencia. Después de todo, un ojo vivo pero desconectado del resto del cerebro *no puede ver, no tiene experiencia visual consciente*, de modo que ésta debe originarse después, cuando a la mera sensibilidad se suma una *x* misteriosa para producir la sentiencia.

Las razones del persistente atractivo de esta idea no son difíciles de descubrir. Nos vemos tentados a creer que los meros impulsos nerviosos no podrían ser el material que forma la conciencia, que necesitan una traducción a otra cosa, no se sabe cómo. Porque, de otro modo, el sistema nervioso sería como un sistema telefónico sin que hubiera nadie en casa para coger el teléfono, o una red de televisión sin espectadores... o un barco sin timonel. Da la impresión de que tendría que haber un Agente, un Patrón, o una Audiencia, centrales para recibir (transducir) toda la información y *evaluarla*, para luego «timonear el barco».

La idea de que la red *en sí* (en virtud de su intrincada estructura y, en consecuencia, de sus poderes de transformación, y en consecuencia, de su capacidad para controlar el cuerpo) pudiera asumir el papel de Patrón interior y por ende patronear la conciencia parece ridícula. En un principio. Pero cierta versión de esta pretensión es la mejor esperanza del materialista. Aquí es donde pueden traerse a colación las complicaciones que dan al traste con la idea de que el sistema nervioso no es más que un puro sistema procesador de información, redistribuyendo una parte de esta ingente tarea de «apreciación» al propio cuerpo, con el fin de que nos ayuden a imaginar.

«¡*Mi cuerpo tiene mente propia!*»

La Naturaleza parece haber erigido el aparato de la racionalidad no sólo encima del aparato de regulación biológica sino *con él y a partir de él*.

Antonio Damasio, *El error de Descartes*.

El medio de transferencia de información en el sistema nervioso consiste en impulsos electroquímicos que viajan a

lo largo de las largas ramificaciones de las células nerviosas: no como los electrones por un cable a la velocidad de la luz, sino por medio de una reacción de desplazamiento en cadena mucho más lento. Una fibra nerviosa es una especie de pila alargada en la cual las diferencias químicas entre interior y exterior de las paredes celulares inducen una actividad eléctrica que luego se propaga por la pared a diferentes velocidades, muchísimo mayores que la velocidad con la cual viajan los grupos de moléculas por un fluido, pero muchísimo menores que la velocidad de la luz. Cuando las células nerviosas se ponen en contacto unas con otras, en unas uniones llamadas sinapsis, se produce la interacción conjunta de un microefector y un microtransductor: el impulso eléctrico dispara la descarga de moléculas de un neurotransmisor que salvan el hueco entre ambas células por difusión, a la manera antigua (se trata de un hueco muy pequeño) siendo transducidas después a nuevos impulsos eléctricos. Podríamos pensar que se trata de un paso atrás volviendo al viejo mundo de la llave y la cerradura moleculares. Y sobre todo cuando resulta que, además de las moléculas neurotransmisoras (como el glutamato) que parecen ser unos barqueros que cruzan todas las sinapsis de manera más o menos neutral, existe toda una gama de moléculas neuromoduladoras que, cuando *encuentran* las «cerraduras» en las células nerviosas próximas, producen toda suerte de cambios por sí solas. ¿Sería correcto decir que las células nerviosas *transducen* la presencia de estas moléculas neuromoduladoras del mismo modo que otros transductores «notan» la presencia de los antígenos, o del oxígeno, o del calor? Porque si fuera así, entonces existen transductores prácticamente en todas las uniones del sistema nervioso, añadiendo más información a la corriente de información que ya llevan los impulsos eléctricos. Y también hay efectores por todas partes, que segregan neuromoduladores y neurotransmisores hacia el mundo «exterior» que es el resto del cuerpo, en el cual se di-

funden para producir numerosos efectos. Se rompe la frontera neta entre el sistema de procesado de la información y el resto de mundo (el resto del cuerpo).

Siempre ha estado claro que allí donde haya transductores y efectores, desaparece la «neutralidad respecto al medio» o multiplicidad de concreción del sistema de información. Para poder detectar la luz, por ejemplo, se necesita algo fotosensible, algo que responda veloz y fiablemente a los fotones, amplificando su llegada subatómica a sucesos a mayor escala que a su vez puedan disparar sucesos subsiguientes. (La rodopsina es una sustancia fotosensible de este tipo, y esta proteína ha sido el material elegido en todos los ojos naturales desde las hormigas hasta los peces, pasando por las águilas y las personas. Quizá los ojos artificiales puedan usar algún otro elemento fotosensible, pero no valdría cualquier cosa.) Para poder identificar y dejar sin efecto un antígeno, necesitamos un anticuerpo que tenga la forma adecuada habida cuenta de que la identificación se hace mediante el método de la llave y la cerradura. Ello limita la elección de los materiales constructivos de anticuerpos a determinadas moléculas que puedan adoptar esas formas, cosa que a su vez restringe mucho la composición química de las moléculas, aunque no absolutamente (como demuestra el ejemplo de las variedades de lisozima). En teoría, cualquier sistema de procesado de información está ligado, podríamos decir, por ambos extremos a los transductores y a los efectores cuya composición física se ve prescrita por la labor que han de desarrollar; entre uno y otro extremos, todo puede conseguirse gracias a procesos independientes del medio.

Los sistemas de control de los barcos, los automóviles, las refinерías de petróleo y demás artefactos humanos complejos son independientes del medio, siempre que los medios utilizados sean capaces de realizar la tarea encomendada en el tiempo disponible. Sin embargo, los sistemas neurales de control en los animales no son auténticamente inde-

pendientes del medio: y no porque los sistemas de control tengan que estar confeccionados con determinados materiales para poder generar esa aura especial, o ese zumbido, o ese lo que sea especial, sino porque han evolucionado como sistemas de control de organismos que ya estaban profusamente equipados con sistemas de control muy dispersos y había que construir los sistemas nuevos superando estos primeros sistemas, pero también en profunda colaboración con ellos, creando así un número astronómico de puntos de transducción. De vez en cuando podemos pasar por alto estas interpenetraciones de diferentes medios que hay por todas partes (como, por ejemplo, cuando reemplazamos una única vía nerviosa por una prótesis, como la del nervio auditivo) pero podríamos pasar por alto *en general* estas interpenetraciones de medios en un experimento mental absolutamente fantástico.

Por ejemplo: las llaves moleculares necesarias para abrir las cerraduras que controlan cada una de las transducciones entre las células nerviosas son las moléculas de glutamato, de dopamina y de norepinefrina (entre otras); pero «en principio» podrían cambiarse todas las cerraduras, es decir, reemplazarlas por un sistema químico diferente. Después de todo, la función de una sustancia química depende de su adecuación a la cerradura y de ella, y de nada más, los efectos subsiguientes que se disparan a la llegada de su mensaje de puesta en marcha. Pero la distribución de responsabilidades por todo el cuerpo hace que ese cambio de cerraduras sea prácticamente imposible. En esos materiales concretos ya se ha encarnado una enorme cantidad de procesado de la información, y por lo mismo ya existe mucha información almacenada. Y por eso, cuando se crea una mente, es una muy buena razón añadida que sí importen los materiales elegidos. Así que ya tenemos dos razones para la elección de los mismos: la velocidad y la ubicuidad de los transductores y efectores a lo largo de

todo el sistema nervioso. No creo que haya otras razones igual de buenas que éstas.

Estas consideraciones prestan apoyo a la pretensión intuitivamente atractiva que suelen ofrecer los críticos del funcionalismo: que sí importa a partir de qué materiales se fabrique una mente. No se puede hacer una mente *sentiente* a partir de chips de silicona, o de cables y vidrio, o de latas de cerveza atadas con una cuerda. ¿Son motivos estos para abandonar el funcionalismo? En absoluto. Lo cierto es que estas razones obtienen su fuerza de la hondura fundamental del funcionalismo.

La *única* razón por la cual las mentes dependen de la composición de sus mecanismos o de sus medios es que para poder llevar a cabo las tareas que deben realizar, tienen que estar hechas, biohistóricamente, a partir de sustancias que sean compatibles con los cuerpos preexistentes que controlan. El funcionalismo se opone al vitalismo y demás formas de misticismo en las «propiedades intrínsecas» de diversas sustancias. La adrenalina no lleva en sí más cólera ni miedo que la estupidez que pueda llevar en sí una botella de whisky. *Per se*, estas sustancias son tan insignificantes para lo mental como la gasolina o el dióxido de carbono. Cuando su capacidad para servir como componentes de sistemas funcionales mayores depende de su composición interna es cuando importa su llamada «naturaleza intrínseca».

El hecho de que nuestro sistema nervioso no sea un sistema de control independiente del medio, aislado (es decir, que produzca «efectos» y «transducciones» en casi cada una de sus uniones) a diferencia del sistema de control de un barco moderno, nos obliga a pensar en las funciones de sus partes de un modo más complicado y realista. Este reconocimiento complica un tanto la vida de los filósofos de la mente funcionalistas. Un millar de experimentos mentales (entre ellos, mi relato «¿Dónde estoy?» [1978]) han sacado partido de la intuición de que *yo* no soy mi cuerpo sino... su propietario. En una operación de trasplante de corazón se quiere

ser el receptor y no el donante, pero en una operación de trasplante de cerebro querríamos ser el donante y no el receptor: vamos con nuestro cerebro y no con nuestro cuerpo. En principio (como ya han argumentado muchos filósofos), yo podría incluso intercambiar por otro mi actual cerebro reemplazando el medio y preservando sólo el mensaje. Por ejemplo, podría viajar teletransportado siempre que la información se preservara perfectamente. En principio, sí... pero sólo porque transmitiríamos la información correspondiente a todo el cuerpo y no sólo la pertinente al sistema nervioso. No se me puede separar de mi cuerpo con un corte nítido como los filósofos han supuesto con frecuencia. Mi cuerpo tiene tanto de *mí*, de mis valores, talentos, recuerdos y actitudes que me hacen ser como soy, como mi sistema nervioso.

El legado del famoso dualismo cartesiano entre mente y cuerpo va más allá de lo académico y llega a nuestras ideas cotidianas: «Estos atletas están preparados tanto física como anímicamente» y «No tienes nada orgánico... es cosa de tu cabeza». Incluso entre aquellos de nosotros que hemos combatido la visión de Descartes, ha habido una poderosa tendencia a tratar la mente (es decir, el cerebro) como patrona del cuerpo, como piloto del barco. Al alinearnos con esta manera estándar de pensar, pasamos por alto una alternativa importante: la de ver el cerebro (y con él, la mente) como un órgano entre muchos, un usurpador del control relativamente recién llegado y cuyas funciones no pueden comprenderse adecuadamente hasta que no lo veamos como el patrón sino solamente como uno de los sirvientes, un tanto díscolo, que trabajan en el fomento de los intereses del cuerpo que lo alberga, lo alimenta y da significado a sus acciones.

Esta perspectiva histórica o evolutiva me recuerda el cambio que ha sufrido Oxford en los treinta años que han transcurrido desde que viví allí como estudiante. Lo normal era que los catedráticos llevaran la voz cantante mientras

que desde becarios y demás burócratas hasta el vicerrector todos los demás actuaban bajo su guía y su mandato. Hoy día, los catedráticos, a semejanza de sus semejantes en las facultades universitarias norteamericanas, ocupan más claramente un papel de empleados contratados por una administración central. Pero en último extremo ¿de dónde obtiene su significado la universidad? En la historia evolutiva, en la administración de nuestros cuerpos se ha impuesto un cambio similar. Pero nuestros cuerpos, al igual que los catedráticos de Oxford, siguen teniendo cierto poder de decisión... o, por lo menos, cierta capacidad de rebelarse cuando la administración central actúa de manera que contraría los sentimientos del «cuerpo político».

Una vez que abandonamos la nítida identificación de la mente con el cerebro y dejamos que se extienda a otras partes del cuerpo, es más difícil pensar funcionalísticamente, pero las compensaciones son enormes. El hecho de que nuestros sistemas de control, a diferencia de los de los barcos y demás artefactos, sean tan poco aislados permite a nuestros propios cuerpos (como entes diferentes de los sistemas nerviosos a los que albergan) contener buena parte de la sabiduría que «nosotros» explotamos en el curso de la toma cotidiana de decisiones. Friedrich Nietzsche lo vio hace ya tiempo y lo expuso con su brío característico en *Así habló Zaratustra* (en la sección adecuadamente titulada «De los despreciadores del cuerpo»):

«Cuerpo soy, y alma», así habla el niño. ¿Y por qué no habríamos de hablar como niños? Pero los avisados y los vigilantes dicen: cuerpo soy por entero y nada más; y alma es sólo una palabra que designa una parte del cuerpo.

El cuerpo es una gran razón, una pluralidad con un sentido, una guerra y una paz, un rebaño y un pastor. Un instrumento de tu cuerpo es también tu pequeña razón, hermano mío, a la que llamas «espíritu»... un pequeño instrumento y

juguete de tu gran razón... Tras tus pensamientos y sentimientos, hermano mío, se yergue un poderoso gobernante, un sabio desconocido... cuyo nombre es ser. Mora en tu cuerpo; él es tu cuerpo. Hay más razón en tu cuerpo que en tu máxima sabiduría.

La evolución materializa información en todas las partes de todos los organismos. Las crías de una ballena materializan información de la comida que ingiere y del medio líquido en que encuentra su comida. El ala de un pájaro materializa sobre el medio en el cual funciona. La piel de un camaleón, de manera más llamativa, lleva información de su entorno en cada momento. Las vísceras y los sistemas hormonales de un animal albergan una enorme cantidad de información del mundo en que han vivido sus antepasados. Esa información no tiene por qué copiarse en absoluto en el cerebro. No hay por qué «representarla» en «estructuras de datos» en el sistema nervioso. Sin embargo, puede explotar la el sistema nervioso que está diseñado para explotar la información de los sistemas hormonales, o beneficiarse de ella, del mismo modo que está diseñado para explotar la información materializada en los miembros y en los ojos, o beneficiarse de ella. De manera que hay una sabiduría encarnada en el resto del cuerpo, sobre todo en relación con las preferencias. Al utilizar los antiguos sistemas corporales como una especie de caja de resonancia, o de audiencia atenta o crítica, el sistema nervioso central puede verse guiado hacia comportamientos sabios, unas veces a codazos, otras veces a empujones. Efectivamente, se deja que el cuerpo vote. Para ser justos con Descartes, deberíamos darnos cuenta de que hasta él vio (por lo menos vagamente) la importancia de esta unión entre cuerpo y mente:

Por medio de estas sensaciones de dolor, de hambre, de sed, y así sucesivamente, la naturaleza también enseña que es-

toy presente en mi cuerpo no como el marino está presente en su barco sino que yo estoy íntimamente unido a él y, por así decir, entremezclado con él, tanto que resulto ser una sola cosa con él (Meditación Sexta).

Cuando todo va bien, reina la armonía y las diferentes fuentes de sabiduría del cuerpo colaboran para beneficio del conjunto, pero estamos más que familiarizados con los conflictos que pueden provocar una exclamación como «¡Parece que mi cuerpo piensa por sí solo!». En apariencia, a veces es tentador reunir parte de esta información corporal formando una mente *aparte*. ¿Por qué? Porque está organizada de tal manera que a veces puede realizar, en cierto modo, discriminaciones por su cuenta, consultar preferencias, tomar decisiones y poner en marcha comportamientos que entran en competencia con *nuestra* mente. En esos momentos se hace muy poderosa la imagen cartesiana del ser marionetista que intenta desesperadamente controlar un cuerpo marioneta ingobernable. Nuestro cuerpo puede revelar enérgicamente los secretos que *nosotros* estamos deseando ocultar desesperadamente... sonrojándose, temblando o sudando, por mencionar tan sólo los casos más evidentes. Puede «decidir» que pese a *nuestros* planes bien trazados, este momento podría ser un buen momento para la relación sexual y no la discusión intelectual, dando a continuación los pasos necesarios para preparar un golpe de estado. En otro momento, y para mayor frustración y contrariedad nuestra, puede hacer oídos sordos a nuestros esfuerzos para alistarlo en una campaña sexual, obligándonos a subir el volumen, a girar los botones y a intentar toda suerte de ridículos engatusamientos para *convencerle*.

Pero ¿si nuestros cuerpos tenían ya mentes propias, por qué tuvieron que adquirir mentes adicionales... *nuestras* mentes? ¿Es que no basta con una mente por cuerpo? No siempre. Como ya hemos visto, las mentes antiguas basadas

en el cuerpo han hecho un trabajo serio al mantener vida y miembros juntos a lo largo de miles de millones de años pero son relativamente lentas y relativamente bastas en cuanto a capacidad de discriminación. Tienen una intencionalidad de corto alcance y a la que fácilmente se engaña. Para las relaciones más complejas con el mundo se hace necesaria una mente más veloz y de más amplias miras, una mente que sea capaz de asegurar un mayor y mejor futuro.

CAPÍTULO 4

DE CÓMO LA INTENCIONALIDAD SALTÓ A PRIMER PLANO

*La torre de la generación y la prueba**

Para poder ver con anticipación en el tiempo, es útil ver más allá en el espacio. Los que fueron en un principio sistemas de monitorización internos y periféricos evolucionaron lentamente hacia sistemas capaces de discriminación no sólo proximal (cercana) sino distal (distante). Ahí es donde la percepción está a sus anchas. El sentido del olor, u olfato, se apoya en que flotan en el aire llaves precursoras para ciertas cerraduras del individuo. Las trayectorias de estos precursores son relativamente lentas, variables e inciertas a causa de la evaporación y la dispersión al azar; de este modo, se limita la información de la fuente de la cual emanan. El sentido del oído depende de que las ondas sonoras choquen

* Este epígrafe está extraído, con algunas revisiones, de *La peligrosa idea de Darwin*.

con los transductores del sistema y como los caminos de las ondas del sonido son más veloces y más regulares, la percepción puede aproximarse más a una anticipación de la «acción a distancia». Pero las ondas de sonido pueden desviarse y rebotar de manera tal que puede oscurecerse su origen. La visión depende de la llegada muchísimo más rápida de fotones rebotados en las cosas del mundo, siguiendo trayectorias absolutamente rectas, de modo que con un sistema con un agujerito de forma adecuada (y, optativamente, con algunas lentes), un organismo pueda obtener información instantánea de alta fidelidad sobre los sucesos y las superficies alejados. ¿Cómo se dio esta transición de intencionalidad, de la interna a la proximal, y de la proximal a la distal? La evolución creó ejércitos de agentes internos especializados para recibir la información disponible en la periferia del cuerpo. La luz que cae sobre un pino lleva tanta información codificada como la que cae sobre una ardilla, pero ésta está provista de millones de microagentes que buscan información, específicamente diseñados para absorber, e incluso buscar, e interpretar dicha información.

Los animales no son solamente herbívoros o carnívoros. Son, según el bonito término acuñado por el psicólogo George Miller, *informívoros*. Y su hambre epistémica les surge de la combinación, de exquisita organización, de las hambres epistémicas específicas de millones de microagentes, organizados en docenas, o en centenares, o en miles de subsistemas. Cada uno de estos diminutos agentes puede concebirse como un sistema intencional absolutamente mínimo, cuyo proyecto de vida es hacerse una sola pregunta una y otra y otra vez: ¿Llega YA mi mensaje?, ¿llega YA mi mensaje? y ponerse en acción de forma limitada pero adecuada todas las veces que la respuesta sea SÍ. Sin esa hambre epistémica no hay percepción, ni captación. Los filósofos han intentado a menudo analizar la percepción distinguiendo lo Dado y lo que luego hace la mente con lo Dado.

Lo Dado es, por supuesto, Captado, pero esa captación de lo Dado no es cosa que haga un Gran Maestro Captador localizado en unos cuarteles generales del cerebro del animal. La tarea de captar está distribuida entre todos los captadores individualmente organizados. Los captadores no son sólo los transductores periféricos (los bastones y los conos de la retina del ojo, las células especializadas del epitelio de la nariz) sino también todos los funcionarios alimentados por éstos, células y grupos de células conectados en redes por todo el cerebro. Éstos no se alimentan de pautas luminosas o de presión (la presión de las ondas sonoras y del tacto) sino de pautas de impulsos neuronales; pero aparte de este cambio de dieta, representan papeles similares. ¿Cómo llegan a organizarse todos estos agentes en sistemas mayores capaces de sustentar formas de intencionalidad cada vez más complejas? Mediante un proceso de evolución por selección natural, desde luego, pero no mediante un proceso único.

Quiero proponer un marco en que podamos colocar las diversas opciones de diseño para el cerebro en el que veamos de dónde les viene su poder. Se trata de una estructura ultrasimplificada pero la idealización es el precio que solemos estar dispuestos a pagar para obtener una visión sinóptica. Llamo a ese marco torre de la generación y la prueba. Cada piso de la torre que se construye otorga mayor poder a los organismos de ese nivel para poder encontrar movimientos cada vez mejores y encontrarlos además con mayor eficiencia.

El poder cada vez mayor de los organismos para producir futuro puede representarse, de este modo, en una serie de pasos. Estos pasos ciertamente casi no representan con claridad períodos transitorios definidos de la historia evolutiva (no hay duda de que tales pasos los dieron diferentes linajes de manera no uniforme y solapándose unos con otros) pero los distintos pisos de la torre de la generación y la prueba marcan los avances importantes de poder cognitivo, y una vez que veamos en esbozo algunas de las características no-

tables de cada piso, tendrá más sentido el resto de los pasos evolutivos.

En el principio fue la evolución darwiniana de las especies por selección natural. Se generó ciegamente una variedad de organismos posibles, mediante procesos de recombinación y mutación de genes más o menos arbitrarios. Se probó a estos organismos y sólo sobrevivieron los mejores diseños. Este es el piso bajo de la torre. Llamemos a sus habitantes *criaturas darwinianas*.

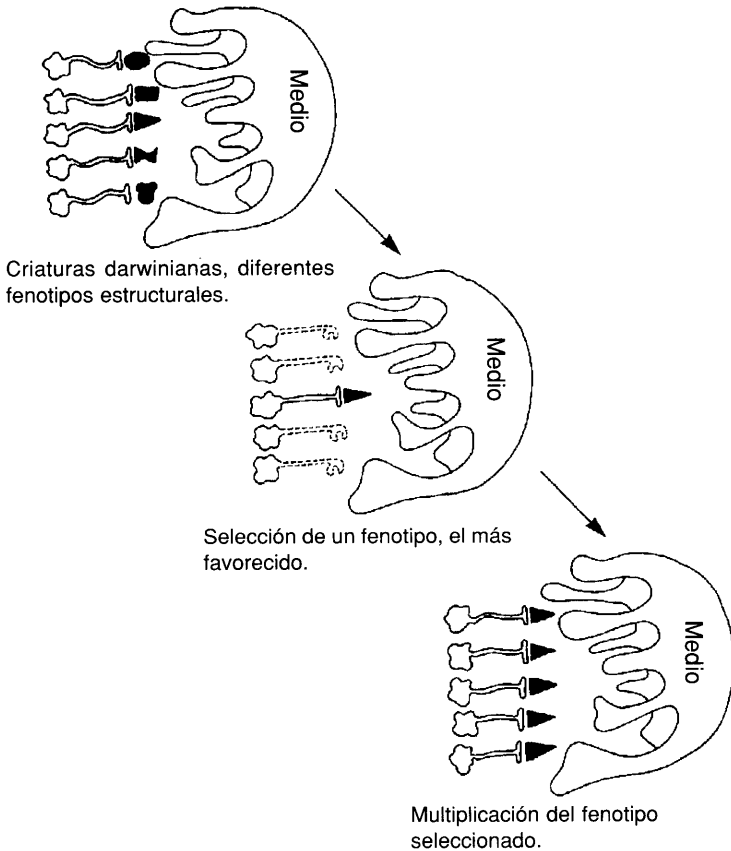


Figura 4.1

Este proceso pasó por muchos millones de ciclos, produciendo muchos diseños estupendos, tanto de plantas como de animales. En determinado momento, entre las nuevas creaciones, aparecieron algunos diseños con la propiedad de *plasticidad fenotípica*: es decir, que los organismos posibles individuales no estaban diseñados por completo en el momento de su nacimiento; en su diseño había elementos que podían *ajustarse debido a los sucesos que se daban en las pruebas de campo*. Podemos suponer que algunos de estos organismos candidatos no eran mejores que sus primos, las criaturas darwinianas acabadas, habida cuenta de que no tenían modo de favorecer (seleccionar para una repetición) las opciones de conducta para «probar las cuales» estaban equipados. Pero podemos suponer que otros tuvieron la fortuna suficiente de llevar en su interior «reforzadores» que daba la casualidad de que favorecían los movimientos inteligentes, es decir, las acciones que eran mejores para esos individuos en lugar de otras acciones disponibles. Así enfrentados al entorno, estos individuos dieron lugar a numerosas acciones, probadas una a una, hasta que encontraron la que funcionaba. Detectaron que funcionaba solamente por recibir una señal positiva o negativa del entorno, que ajustó la probabilidad de que esa acción se reprodujera en otra ocasión. Naturalmente, estarían condenadas las criaturas mal preparadas (las que tuvieran invertido el refuerzo positivo o negativo). Sólo tendrían ventaja aquellas criaturas suficientemente afortunadas por haber nacido con el refuerzo adecuado. A este subconjunto de criaturas darwinianas podemos llamarlas *criaturas skinnerianas* ya que como gustaba de señalar el psicólogo conductista B. F. Skinner ese «condicionante operante» no es sencillamente análogo de la selección natural darwiniana, sino una ampliación de la misma: «Allí donde cesa la conducta heredada toma el mando la modificabilidad heredada del proceso de condicionamiento.»

La revolución cognitiva que surgió en la década de los

setenta arrebató al conductismo su posición predominante en la psicología y desde entonces se ha dado una tendencia a subestimar el poder del condicionamiento skinneriano (o sus variantes) para conformar la competencia conductual de los organismos y dar lugar a estructuras altamente adaptativas y discernientes. Sin embargo, los florecientes trabajos de los años noventa sobre redes neurales y «conexionismo» han demostrado de nuevo el virtuosismo, a menudo sorprendente,

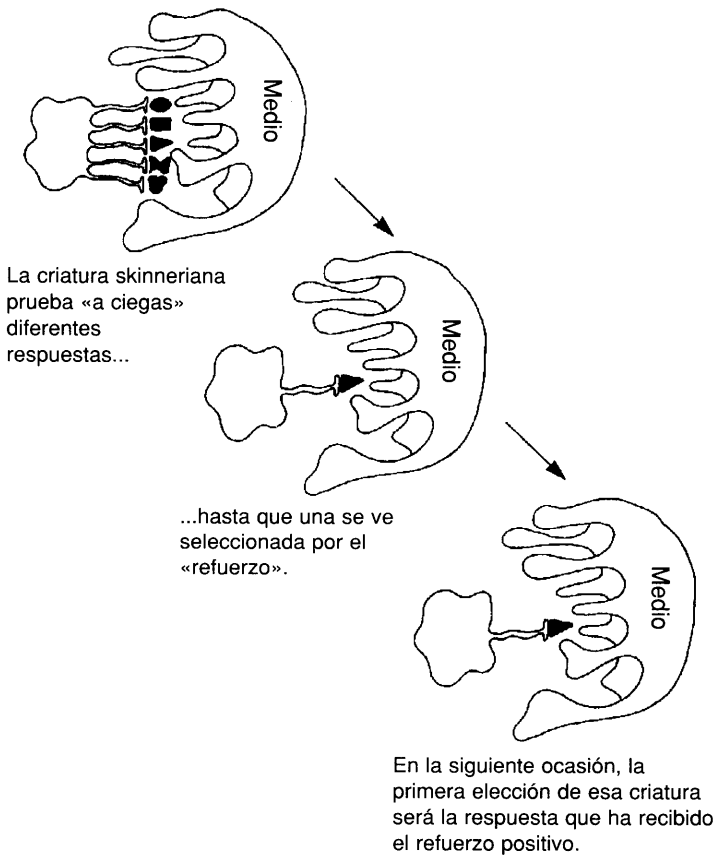


Figura 4.2

de redes sencillas que comienzan su vida conectadas de manera más o menos al azar y que luego ajustan sus conexiones mediante una especie sencilla de «experiencia»: la historia de los refuerzos que van encontrando.

La idea fundamental de permitir al entorno interpretar un papel ciego pero selectivo en la conformación de la mente (o del cerebro, o del sistema de control) tiene un pedigrí incluso más antiguo que el de Darwin. Los antepasados intelectuales de los conexionistas de hoy y de los conductistas de ayer fueron los asociacionistas: filósofos como David Hume que en el siglo XVIII intentó imaginarse cómo las partes de la mente (partes a las que llamó impresiones e ideas) podían llegar a organizarse por sí mismas sin disponer de un director organizador que lo supiera todo. Como en una memorable ocasión me dijo un estudiante: «Hume quería que las ideas pensaran por sí mismas.» Hume tuvo corazonadas maravillosas sobre cómo podrían unirse las impresiones y las ideas mediante un proceso parecido al del enlace químico creando después vías transitadas en la mente, las costumbres, pero estas corazonadas eran excesivamente vagas como para ser comprobadas. Sin embargo, el asociacionismo de Hume fue una inspiración directa para los famosos experimentos de Pavlov en el condicionamiento de la conducta animal, que a su vez llevaron a las teorías del condicionamiento, algo diferentes, de E. L. Thorndike, Skinner y demás conductistas de la psicología. Algunos de estos investigadores (concretamente, Donald Hebb) intentaron unir su conductismo todavía más a lo que se conocía del cerebro. En 1949 Hebb propuso unos modelos de mecanismos de condicionamiento simple que podrían ajustar las conexiones entre las células nerviosas. Estos mecanismos (que hoy se llaman reglas hebbianas del aprendizaje) y sus continuadores son las máquinas del cambio en el conexionismo, la última de las manifestaciones de esta tradición.

Asociacionismo, conductismo, conexionismo... en orden

histórico y también en orden alfabético podemos trazar la evolución de los modelos de un tipo sencillo de aprendizaje que bien podríamos llamar *aprendizaje ABC*. No hay duda de que la mayoría de los animales son capaces de aprendizaje ABC, es decir, pueden llegar a modificar (o a rediseñar) su comportamiento en la dirección adecuada como resultado de un proceso de formación y de conformación debido al entorno. Hoy día existen buenos modelos, con diferente grado de realismo y de detalle, de cómo puede llegarse a semejante proceso de condicionamiento o de formación en una red de células nerviosas sin que sea una cuestión milagrosa.

Para numerosos propósitos que salvan la vida (por ejemplo, el reconicimiento de pautas, la discriminación, la generalización y el control dinámico de la locomoción) las redes ABC son francamente maravillosas: son eficientes, compactas, robustas en su funcionamiento, con un margen para los errores y relativamente fáciles de rediseñar sobre la marcha. Lo que es más, tales redes resaltan vívidamente la idea de Skinner de que hay poca diferencia entre trazar la línea entre la poda y la conformación que hace la selección natural que se transmite genéticamente a la descendencia (es decir, el cableado con el que nacemos) y la poda y la conformación que tienen lugar después en cada individuo (el recableado con el que terminamos cada uno, como resultado de la experiencia o la formación). Naturaleza y cultura se mezclan sin bordes de separación. Sin embargo, sí hay algunos trucos cognitivos que esas redes ABC no están aún preparadas para realizar y (crítica todavía más contundente) hay algunos trucos cognitivos que claramente no son resultado de formación alguna. Algunos animales parecen capaces de «aprendizaje a la primera»: pueden deducir ciertas cosas sin tener que pasar el arduo proceso de ensayo y error en el áspero mundo que es el sello de todo aprendizaje ABC.

El condicionamiento skinneriano está bien siempre que no nos mate uno de nuestros primeros errores. Un sistema

mejor es el que supone una *preselección* entre todos los posibles comportamientos o acciones, de manera tal que los movimientos auténticamente idiotas queden fuera de las posibilidades antes de haberlos probado en la «vida real». Nosotros como seres humanos somos capaces de tal refinamiento, pero no somos los únicos. A los beneficiarios de este tercer escalón podemos denominarlos *criaturas popperianas* ya que, como dijo elegantemente en una ocasión el filósofo Karl Popper, esta mejora del diseño «permite que nuestras hipótesis mueran en lugar de morir nosotros». A diferencia de las criaturas meramente skinnerianas, muchas de las cuales sobreviven únicamente porque su primer movimiento es afortunado, las criaturas popperianas sobreviven porque son suficientemente listas como para hacer movimientos algo mejores que un simple movimiento afortunado. Por supuesto que tienen suerte de ser tan listas, pero ser listo es algo más que tener suerte sin más.

¿Cómo se efectúa esta preselección en los agentes popperianos? Tiene que haber un filtro y cualquier filtro de este tipo debe equivaler a una especie de *entorno interno* en el cual se puedan llevar a cabo con seguridad algunas pruebas: un lo que sea interno estructurado de tal forma que las acciones sustitutivas que favorece sean generalmente las mismas acciones que bendeciría también el mundo real si llegaran a realizarse. En resumidas cuentas: que el entorno interno, sea lo que sea, debe albergar montones de *información* del entorno externo y sus regularidades. Salvo eso (o la magia) no hay nada que pueda hacer que merezca la pena la preselección. (Siempre se puede arrojar una moneda o consultar un oráculo, pero ello no mejora la prueba a ciegas, a no ser que alguien o algo sesguen sistemáticamente tanto la moneda como el oráculo basándose en la información auténtica que del mundo poseen ese alguien o ese algo.)

La belleza de la idea de Popper queda ejemplificada en

el reciente desarrollo de los simuladores realistas de vuelo que se usan para entrenar a los pilotos. En un mundo simulado, los pilotos pueden aprender qué movimientos deben realizar en según qué situaciones sin arriesgar siquiera sus vidas (ni unos costosos aviones). Sin embargo, como ejemplos del truco popperiano, los simuladores de vuelo son equívocos en un aspecto: reproducen el mundo con demasiada literalidad. Debemos tener mucho cuidado de no creer que el entorno interno de una criatura popperiana es simplemente una réplica del mundo exterior, y que reproduce todas las contingencias físicas del mundo. En ese mundo milagroso de juguete ¡la estufita de nuestra cabeza tendría que estar tan caliente como para quemarnos el dedito que le ponemos encima! No hace falta suponer nada semejante. La *información* del efecto de colocar un dedo sobre una estufa tiene que

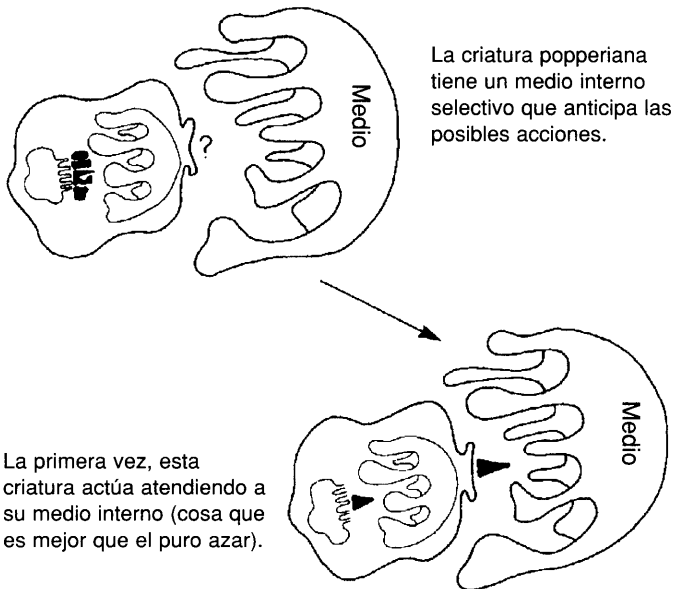


Figura 4.3

estar ahí y debe estar de forma que pueda producir su efecto premonitorio cuando se la invoque para una prueba interna, pero tal efecto puede conseguirse sin tener que construir una réplica del mundo. Después de todo sería igual de popperiano formar a los pilotos haciéndoles leer un libro que les explicara todas las contingencias a las que podrían enfrentarse cuando se subieran a la cabina del avión. Puede que no sea un método tan poderoso de enseñanza pero ¡sería mucho mejor que salir a hacer pruebas en el aire! El elemento común de las criaturas popperianas es que de una u otra forma (bien por herencia, bien por adquisición) en ellas llevan instalada la información (información precisa del mundo que –probablemente– habrán de encontrarse) y que tal información se encuentra en forma tal que puede conseguir los objetivos que son su *raison d'être*.

Uno de los medios por los cuales consiguen filtros útiles estas criaturas popperianas es situar las opciones posibles de conducta ante el tribunal corporal y explotar la sabiduría, por muy pasada de moda o miope que sea, acumulada en esos tejidos. Si el cuerpo se rebela (por ejemplo, con reacciones tan típicas como náusea, vértigo, temor o temblor) ya es un síntoma medianamente fiable (mejor que lanzar una moneda al aire) de que el acto que se plantea puede no ser una buena idea. Aquí se ve cómo en lugar de recablear el cerebro para eliminar tales elecciones haciéndolas estrictamente impensables, la evolución puede sencillamente organizar que la respuesta a cualquier pensamiento sobre ellas sea un agobio tan fuerte como para que sea sumamente improbable que sean las elegidas para actuar. La información del cuerpo en que se basa tal reacción puede haberse instalado allí mediante una receta genética o debido a una reciente experiencia individual. Cuando un bebé aprende a gatear tiene una aversión innata a aventurarse cerca de un panel de cristal que tenga un hueco debajo, a través del cual pueda ver un «acantilado visual». Aunque su madre le llame

desde cerca, le engatuse y le anime, el niño se echa atrás temeroso pese a no haberse caído en su vida. La experiencia de sus antepasados le hace quedarse del lado seguro. Cuando una rata ha comido un nuevo tipo de alimento y se le ha inyectado una sustancia que le ha hecho vomitar, seguidamente muestra una fuerte aversión al alimento que se parece al ingerido antes de vomitar y que huele como él. Aquí la información que le induce a quedarse del lado seguro la ha obtenido de su propia experiencia. Ningún filtro es perfecto (después de todo, el panel de vidrio era seguro y el nuevo alimento de la rata no tiene nada de tóxico) pero mejor seguro que arrepentido.

Los inteligentes experimentos de psicólogos y etólogos parecen indicar algunas otras formas en las que los animales pueden probar «de cabeza» las acciones recogiendo así un beneficio popperiano. En los años treinta y cuarenta los conductistas demostraron una y otra vez que sus animales experimentales eran capaces de «aprendizaje latente» sobre el mundo, un aprendizaje que no estaba específicamente recompensado por ningún refuerzo detectable. (Este ejercicio suyo de autorrefutación es en sí un ejemplo notabilísimo de otra cuestión popperiana: que la ciencia progresa sólo cuando presenta hipótesis refutables.) Cuando a las ratas se les permitía explorar un laberinto en el que no había ni comida ni ninguna otra recompensa, terminaban por aprenderse el camino en el transcurso normal de los acontecimientos; si luego se les colocaba en el laberinto algo que ellas tenían por valioso, las ratas que antes se habían aprendido el camino en incursiones anteriores lo encontraban más rápidamente (cosa nada sorprendente) que las ratas del grupo de control que veían el laberinto por primera vez. Puede parecer un descubrimiento insignificante. ¿Es que no siempre ha sido evidente que las ratas eran suficientemente listas como para aprender por dónde tenían que ir? Pues sí y no. Puede haber *parecido* evidente pero este es el tipo de prueba (la prueba con-

trastada sobre la hipótesis cero) que debe realizarse si tenemos que estar seguros de lo inteligentes que son las diversas especies y en qué medida poseen mente. Como ya veremos, otros experimentos con animales demuestran tendencias sorprendentemente estúpidas: lagunas casi increíbles en el conocimiento que los animales tienen de su propio entorno.

Valerosamente, los conductistas intentaron acomodar el aprendizaje latente en sus modelos ABC. Uno de sus recursos más contundentes era postular un «impulso de curiosidad» que se satisfacía (se «reducía» según decían ellos) mediante la exploración. Después de todo, allí había un refuerzo funcionando en todos esos medios sin refuerzo. Oh maravilla, todo entorno, por el simple hecho de ser un entorno en el que hay cosas que aprender, está lleno de estímulos de refuerzo. Como intento para salvar el conductismo ortodoxo era manifiestamente vacío pero eso no quiere decir que sea una idea inútil en otros contextos; reconoce el hecho de que la curiosidad (hambre epistémica) debe mover cualquier sistema de aprendizaje poderoso.

Nosotros, los seres humanos, somos condicionables mediante aprendizaje ABC de modo que somos criaturas skinnerianas, pero no *sólo* somos criaturas skinnerianas. También disfrutamos de los beneficios de un montaje heredado genéticamente, de manera que también somos criaturas darwinianas. Pero somos más. Somos criaturas popperianas. ¿Qué otros animales son criaturas popperianas y cuáles son meramente skinnerianas? Los pichones eran los animales experimentales preferidos de Skinner, y él y sus seguidores desarrollaron la tecnología del condicionamiento operante hasta un grado sumamente sofisticado, consiguiendo que los pichones exhibieran conductas notablemente extravagantes y complejas. Cosa muy notable, los skinnerianos no han conseguido demostrar nunca que los pichones *no fueran* criaturas popperianas, y la investigación en un montón de especies distintas, desde pulpos hasta mamíferos, pasan-

do por peces, parece indicar con mucha fuerza que si existen criaturas puramente skinnerianas, capaces de aprender sólo a base de aprendizaje ciego de ensayo y error, habrán de encontrarse entre los invertebrados sencillos. La gran babosa marina (o liebre marina) *Aplysia californica* ha reemplazado más o menos al pichón como foco de atención entre aquellos que estudian los mecanismos del condicionamiento sencillo.

Por tanto no diferimos de las demás especies en que seamos criaturas popperianas. Lejos de ello; mamíferos y pájaros, reptiles, anfibios, peces e incluso muchos invertebrados exhiben la capacidad de utilizar la información general que obtienen de sus entornos para entresacar sus opciones de conducta antes de ponerse en marcha. ¿Cómo se incorpora a sus cerebros la nueva información del entorno exterior? Evidentemente por la percepción. El entorno contiene una mezcla de riquezas, mucha más información de la que incluso un ángel cognitivo podría usar. Los mecanismos perceptivos diseñados para pasar por alto la mayor parte del flujo de estímulos se concentran en la información de mayor utilidad y más fiable. ¿Y cómo se las arregla la información reunida para ejercer su efecto selectivo cuando se «consideran» las opciones, ayudando al animal a diseñar interacciones todavía más efectivas con su mundo? Sin duda existe una variedad de métodos y de mecanismos diferentes, pero entre ellos se encuentran aquellos que utilizan el cuerpo como caja de resonancia.

La búsqueda de la sentiencia: informe de avances

Hemos ido añadiendo gradualmente elementos a nuestra receta de la mente. ¿Tenemos ya los ingredientes de la sentiencia? Desde luego el comportamiento normal de muchos de los animales que hemos estado describiendo aprueba

nuestros intuitivos exámenes de sentiencia con un éxito rotundo. Al observar cómo tiemblan de miedo un cachorro o un bebé al borde de un aparente precipicio, o ver cómo hace muecas de aparente disgusto una rata ante el olor de un alimento supuestamente tóxico, nos vemos en dificultades incluso de mantener la hipótesis de que *no* estamos ante un ser sentiente. Pero también hemos dejado al descubierto algunos terrenos sólidos para la desconfianza: hemos visto algunas formas en que la conducta sorprendentemente parecida a la de una mente puede ser producida por sistemas de control relativamente sencillos, mecánicos y aparentemente ajenos a lo que sería una mente. La potencia de nuestras respuestas instintivas a la pura velocidad y al parecido vital que supone el movimiento, por ejemplo, deberían ponernos en guardia ante la posibilidad genuina, y no meramente filosófica, de vernos engañados al atribuir mayor sutileza y mayor comprensión a una entidad de lo que las circunstancias permiten. Reconociendo que el comportamiento observable puede hechizarnos, podremos apreciar la necesidad de hacernos aún más preguntas... acerca de qué hay detrás de esa conducta.

Consideremos el dolor. En 1986 el gobierno británico modificó sus leyes de protección a los animales en experimentos, añadiendo el pulpo al privilegiado círculo de animales que no pueden ser operados sin anestesia. Un pulpo es un molusco, fisiológicamente más parecido a una ostra que a una trucha (y mucho menos a un mamífero) pero la conducta del pulpo y de otros cefalópodos (calamar, sepia) es tan sorprendentemente inteligente y tan (aparentemente) sentiente que las autoridades científicas decidieron permitir que la similitud conductual superara a la diferencia interna: se supone oficialmente que los cefalópodos (pero no otros moluscos) son capaces de sentir dolor... no vaya a ser que sean capaces. Por contra, los monos rhesus están fisiológica y evolutivamente muy próximos a nosotros, de modo que

solemos suponer que son capaces de sufrir como nosotros, pero llegado el momento muestran una conducta asombrosamente diferente. El primatólogo Marc Hauser me ha contado en una conversación que durante la época de apareamiento los monos machos luchan ferozmente y que no es infrecuente ver a un macho derribar a otro y luego arrancarle de un mordisco un testículo. El macho herido ni chilla ni hace muecas sino que se limita a lamerse la herida y a marcharse. Al cabo de uno o dos días ¡se le puede ver apareándose! Resulta difícil creer que este animal haya experimentado algo parecido a los dolores que experimentaría un ser humano si se le infligiera semejante herida (nos da vértigo sólo de pensarlo) a pesar de nuestra cercanía biológica. De modo que no podemos ya esperar que las pruebas fisiológicas y conductuales converjan felizmente para darnos respuestas inequívocas, habida cuenta de que ya conocemos casos en los que estos dos tipos de prueba, no concluyentes aunque irresistibles, van en sentido opuesto. ¿Cómo podemos pensar entonces en este asunto?

Una función clave del dolor es la de refuerzo negativo (el «castigo» que disminuye la probabilidad de que se repita una acción) y se puede educar a cualquier criatura skinneriana mediante refuerzo negativo de un tipo u otro. ¿Eso es lo único que es el dolor, un refuerzo negativo? ¿Dolor *experimentado*? ¿Podría haber *dolor* inconsciente o *no experimentado*? Hay mecanismos sencillos de refuerzo negativo que proporcionan el poder formativo o castrador de la conducta que presenta el dolor sin otros efectos secundarios de tipo mental, de modo que sería un error invocar la sentiencia siempre que nos encontremos ante un condicionamiento skinneriano. Otra función del dolor es trastornar las pautas normales de la actividad corporal que podrían empeorar una herida (por ejemplo, el dolor hace que un animal cuide su miembro herido hasta que se cure) y ello se logra mediante un flujo de sustancias neuroquímicas formando un bucle de interaccio-

nes continuadas con el sistema nervioso. Entonces ¿la presencia de esas sustancias garantiza que se da dolor? No, porque en sí mismas no son más que llaves que están flotando por ahí en busca de sus cerraduras; si se interrumpe el ciclo de interacciones no hay motivo para suponer que el dolor persista. Podemos preguntar incluso ¿son necesarias estas sustancias concretas para el dolor? ¿Podrían presentar las criaturas un sistema diferente de cerraduras y llaves? La respuesta puede depender más de los procesos históricos de evolución de este planeta que de cualquier propiedad intrínseca de las sustancias. El ejemplo del pulpo muestra que debemos mirar para ver qué variaciones pueden encontrarse en la estructura química y qué diferencias suponen en la función, pero sin esperar que estos hechos *en sí mismos* cierren nuestro asunto de la sentiencia.

¿Y qué pasa entonces con otros rasgos de este ciclo de interacciones? ¿En qué medida puede ser rudimentario un sistema de dolor y seguir siendo sentiencia? ¿Qué es lo que en él sería significativo y por qué? Por ejemplo, pensemos en un sapo con una pata rota. ¿Es un ser sentiente que experimenta dolor? Es un ser vivo normal cuya vida normal se ha visto interrumpida por el daño hecho a una de sus partes, impidiéndole llevar a cabo las conductas que son su modo de ganarse la vida. Todavía más: se encuentra en un estado con un potencial de refuerzo negativo muy grande: se le puede condicionar en seguida para eludir semejantes estados de su sistema nervioso. Este estado se mantiene mediante un ciclo de interacciones que en cierto modo le trastorna su normal disposición al salto... aunque en una situación de emergencia terminaría por saltar. Es tentador ver en todo esto el equivalente del dolor. Pero también es tentador atribuir al sapo un soliloquio en el cual exprese su temor de que llegue tal emergencia, ansíe un alivio, deplora su relativa vulnerabilidad, se lamente amargamente de los actos estúpidos que le han llevado a esa situación, y así sucesivamente, y todos es-

tos añadidos no se justifican de manera alguna por lo que sabemos de los sapos. Por el contrario, cuanto más sabemos de ellos más seguros estamos de que su sistema nervioso está diseñado para que pasen por la vida sin ninguna de esas costosas capacidades reflexivas.

¿Y qué? ¿Qué tiene que ver la *sentiencia* con esos caprichosos talentos intelectuales? Buena pregunta, pero eso significa que debemos intentar responderla y no sencillamente utilizarla como cuestión retórica para desviar la cuestión. He aquí una circunstancia en la que nuestro modo de preguntar puede suponer una gran diferencia, porque es posible que nos engañemos creando un problema fantasma en este punto. ¿Cómo? Perdiendo la pista del lugar en el que estamos por un proceso de adición y sustracción. En el inicio, buscamos una *x*, el ingrediente especial que distingue la mera sensibilidad de la sentiencia auténtica y trabajamos en esa búsqueda en dos direcciones. Avanzando de abajo arriba, desde los casos sencillos, añadimos versiones rudimentarias de cada rasgo por separado y en general no nos impresiona nada: aunque cada una de esas capacidades sea discutiblemente un componente esencial de la sentiencia, seguramente en la sentiencia hay más que simplemente eso: ¡un robot podría mostrar *lo mismo* sin tener sentiencia en absoluto! Avanzando de arriba abajo, desde nuestra propia experiencia ricamente detallada (y apreciadísima por nosotros) reconocemos que otras criaturas carecen manifiestamente de algunos de los rasgos particularmente humanos de nuestra experiencia, de manera que se los sustraemos por no ser esenciales. No queremos ser injustos con nuestros primos los animales. De tal manera que mientras nosotros admitimos que todo eso que pensamos cuando pensamos en la atrocidad del dolor (y en por qué importa moralmente que alguien sufra) supone imaginarse justamente esos añadidos antropomórficos, decidimos generosamente que son sólo añadidos no «esenciales» para el fenómeno de la sentiencia

en bruto (y para su ejemplo moralmente más significativo, el dolor). Lo que podemos tender a pasar por alto, mientras estos dos barcos se cruzan en la noche, es la posibilidad de que en una vía sustraemos lo mismo que estamos buscando en la otra. Si lo que hacemos es eso, no pasaría de ser una ilusión nuestro convencimiento de que todavía tenemos que encontrar la x (el «eslabón perdido» de la sentiencia).

No digo que *estemos cometiendo* un error de este tipo sino que *podríamos estar cometéndolo*. Es suficiente por el momento ya que altera la carga de la prueba. Por ello, he aquí una hipótesis conservadora sobre la sentiencia: no existe tal fenómeno *añadido*. La «sentiencia» se da en cualquier grado o intensidad imaginable desde la más sencilla y «robótica» hasta la más exquisitamente sensible e hiperreactivamente «humana». Como vimos en el capítulo 1, desde luego tenemos que trazar una línea a través de este continuo de casos multifacéticos porque es una exigencia para tener actuaciones morales, pero la perspectiva de que podamos *descubrir* el umbral (un «paso» que sea moralmente significativo en lo que, por lo demás, no es una escalera sino una rampa) no es que sea extremadamente improbable sino que tampoco tendrá atractivo moral.

Veamos al sapo nuevamente en este contexto. ¿De qué lado de la línea se queda el sapo? (Si los sapos son un caso excesivamente evidente para nosotros por unos u otros motivos, elijamos cualquier criatura que parezca ocupar nuestra penumbra de incertidumbre. Sea una hormiga, una medusa, una paloma o una rata.) Supongamos ahora que la «ciencia confirma» que en ese sapo se encuentra la mínima sentiencia genuina: que el «dolor» del sapo es real, dolor experimentado, por ejemplo. El sapo queda ahora cualificado para el tratamiento especial reservado para los sentientes. Supongamos en vez de eso que resulta que el sapo no tiene esa x , una vez que hayamos determinado qué sea esa x . En este caso, la posición del sapo se queda en la de «mero autómatas», algo en lo

cual podemos intervenir de cualquier manera que se nos ocurra sin compunción moral alguna. Dado lo que ya sabemos de los sapos, ¿parece plausible que pudiera haber un rasgo *desconocido hasta el momento* cuyo descubrimiento pudiera justificar esa enorme diferencia en nuestra actitud? Por supuesto, si descubriéramos que los sapos son en realidad hombrillos atrapados en cuerpos de sapo, como el príncipe del cuento, inmediatamente tendríamos la base para prodigarles los máximos cuidados porque sabríamos que, a pesar de todas sus apariencias conductuales, los sapos *serían capaces* de pasar por todas las torturas y las ansiedades que consideramos tan importantes en nuestro propio caso. Pero ya sabemos que el sapo no es nada semejante. Se nos ha pedido que imaginemos que existe cierta x que no es ni por asomo un príncipe humano bajo la piel de un sapo, pero que sin embargo es irresistible moralmente. No obstante, ya sabemos que un sapo no es un sencillo juego de cuerda sino una cosa viviente exquisitamente compleja capaz de una sorprendente variedad de actividades para protegerse en el fomento de su tarea prefijada de hacer más generaciones de sapos. ¿Es que esto no es suficiente para exigirnos una consideración especial por nuestra parte? Se nos pide que imaginemos que existe cierta x que no se parece en nada a esta mera complejidad de la estructura de control, pero que sin embargo nos obligaría a un reconocimiento moral cuando la descubramos. Sospecho que lo que se nos pide es que demos rienda suelta a algo más que a una fantasía. Pero prosigamos nuestra búsqueda y veamos qué viene a continuación porque aún estamos muy lejos de la mente humana.

De la fototaxia a la metafísica

Una vez que llegamos a las criaturas popperianas (criaturas cuyos cerebros tienen la capacidad de estar dotados,

internamente, de habilidad preselectiva)... ¿qué viene después? Sin duda que muchas cosas diferentes, pero vamos a concentrarnos en una innovación concreta cuyos poderes podamos ver claramente. Entre los sucesores de las criaturas popperianas se encuentran aquellos cuyos entornos internos reciben la información mediante las partes *diseñadas* del entorno externo. Una de las ideas fundamentales más penetrantes de Darwin fue que el diseño es costoso pero que la copia de los diseños es barata; es decir, que hacer un diseño nuevo es muy difícil pero que rediseñar los diseños antiguos es relativamente sencillo. Pocos de nosotros podríamos reinventar la rueda, pero no tenemos que reinventarla ya que hemos adquirido el diseño de la rueda (y otros muchos) de la cultura en la que hemos crecido. Podemos llamar *criaturas gregorianas* a este sub-sub-subconjunto de criaturas darwinianas, habida cuenta de que tengo para mí que el psicólogo británico Richard Gregory es el más preeminente teórico del papel de la información (o, más exactamente, lo que Gregory llama inteligencia potencial) para la invención de movimientos inteligentes (lo que Gregory llama inteligencia cinética). Gregory observa que unas tijeras, como artefacto bien diseñado, no son sólo el resultado de una inteligencia sino de un ser dotado de inteligencia (inteligencia potencial externa) en su sentido recto y muy intuitivo: cuando le damos a alguien unas tijeras aumentamos su potencial de llegar a realizar movimientos inteligentes con mayor seguridad y más rápidamente (1981, págs. 311 y sigs.).

Los antropólogos admitieron hace tiempo que la aparición del uso de herramientas se produjo al tiempo que se daba un aumento importante de la inteligencia. Los chimpancés de la selva capturan termitas metiendo unas cañas burdamente preparadas por los termiteros, sacándolas rápidamente llenas de termitas que luego cogen pasándose la caña por la boca. Este hecho se vuelve aún más significativo cuando averiguamos que no todos los chimpancés han

dado con este truco; en algunas «culturas» de los chimpancés las termitas son una fuente de alimento inexplorada. Lo cual nos recuerda que el uso de herramientas es una señal de inteligencia de ida y vuelta: no sólo *requiere* inteligencia reconocer y mantener una herramienta (y más aún fabricarla) sino que una herramienta *confiere* inteligencia a aquellos suficientemente afortunados a los que se les ha dado una. Cuanto mejor diseñada esté la herramienta (cuanta más información suponga su fabricación) mayor inteligencia potencial confiere a su usuario. Y, nos recuerda Gregory, entre las herramientas más destacadas están aquellas que él denomina herramientas mentales: las palabras.

Las palabras y demás herramientas mentales proporcionan a una criatura gregoriana un entorno interior que le permite construir generadores y probadores de movimientos aún más sutiles. Las criaturas skinnerianas se preguntan: «¿Qué voy a hacer a continuación?» y no tienen clave que les dé la respuesta hasta que no se han llevado unos cuantos

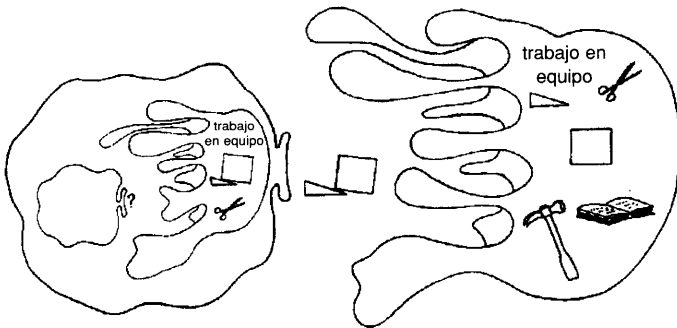


Figura 4.4

La criatura gregoriana extrae herramientas del medio (cultural); estas herramientas mejoran tanto los generadores como los probadores.

golpes duros. Las criaturas popperianas dan un gran paso al preguntarse: «¿Qué debo pensar a continuación?» antes de preguntarse: «¿Qué voy a hacer a continuación?» (Hay que hacer hincapié en que ni las criaturas skinnerianas ni las popperianas tienen por qué hablarse a sí mismas ni tener estos pensamientos. Sencillamente están diseñadas para funcionar *como si* se hubieran hecho esas preguntas. Aquí se ven el poder y el riesgo del enfoque intencional: la razón de que las criaturas popperianas sean más listas [digamos que son criaturas tortuosas con más éxito] que las skinnerianas es que responden adaptativamente a una información mayor y mejor, de un modo que podríamos describir vívida e imprecisamente a partir del enfoque intencional como si tuvieran lugar esos imaginarios soliloquios. Pero sería un error atribuir a estas criaturas todas las sutilezas que acompañan a la capacidad real de formular tales preguntas y de darse tales respuestas sobre el modelo humano de las preguntas explícitas que nos hacemos a nosotros mismos.) Las criaturas gregorianas dan un gran paso hacia el nivel humano de destreza mental, beneficiándose de la experiencia de otros al explotar la sabiduría encarnada en las herramientas mentales que esos otros han inventado, mejorado y transmitido; por eso aprenden cómo pensar mejor sobre aquello que tienen que pensar a continuación... y así sucesivamente, creando una torre de sucesivas reflexiones internas sin límite fijado ni discernible. Puede verse mejor cómo llega a conseguirse este piso gregoriano volviendo la vista de nuevo a los talentos ancestrales a partir de los cuales debe de estar formada la mayoría de los talentos humanos.

Una de las prácticas más sencillas que empujan la vida y que se encuentra en muchas especies es la *fototaxia*: la distinción entre luz y oscuridad, orientándose hacia la primera. La luz es fácil de transducir y dado el modo en que la luz emerge de una fuente, con la intensidad disminuyendo gradualmente al alejarnos de su origen, una conexión relativa-

mente sencilla entre transductores y efectores puede producir una fototaxia fiable. En el elegante librito del neurocientífico Valentino Braitenberg *Vehicles* tenemos el modelo más sencillo: el vehículo de la figura 4.5. Tiene dos transductores de luz y las señales que emiten alimentan, de forma cruzada, dos efectores (piénsese en los efectores como si fueran motores fueraborda). Cuanta más luz se transduce, más rápidamente funciona el motor. El transductor más cercano a la fuente de luz moverá su motor un poco más deprisa que el más alejado de la luz, y ello siempre haría girar el vehículo en dirección a la luz, hasta que llegara a chocar con ella u orbitara muy cerca de ella.

El mundo de un ser tan simple está graduado desde la luz hasta la oscuridad pasando por diversos grados de penumbra que recorren todo el gradiente. Este ser no sabe, ni necesita

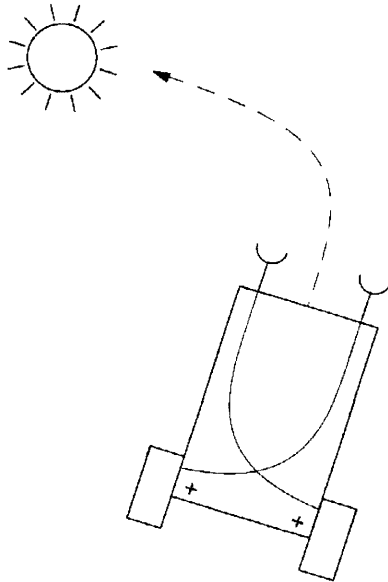


Figura 4.5

saber, nada más. El reconocimiento de la luz es casi gratuito: lo que pone en marcha al transductor es luz y al sistema le da igual si se trata de la *misma* luz que vuelve una y otra vez o se trata de una luz nueva. Ecológicamente, en un mundo de dos lunas sí habría diferencia en seguir una u otra: el reconocimiento de la luna sería un problema añadido que requeriría una solución. En un mundo semejante la mera fototaxia no sería suficiente. En nuestro mundo, la luna no es el tipo de objeto que una criatura tiene que saber reidentificar; con la madre suele ocurrir muchas veces justamente lo contrario.

La *mamataxia* (el dirigirse hacia la madre) es un talento mucho más complejo. Si mamá emitiera una luz brillante podría valer con la fototaxia pero no valdría si hubiera otras madres en las proximidades que usaran el mismo sistema. Si mamá emitiera entonces un tono concreto de luz azul, diferente de la luz que emiten todas las demás madres, entonces obtendríamos un buen resultado colocando a cada uno de nuestros fototransductores un filtro que sólo dejara pasar la luz azul. La naturaleza suele apoyarse en un principio similar pero utilizando un medio que usa la energía más eficientemente. Mamá emite un olor característico, distinguible de todos los demás olores (de las proximidades). Entonces se consigue la *mamataxia* (la reidentificación de la madre y la aproximación a ella) mediante la olfatotransducción u olfacción. La intensidad de los olores es función de la concentración de las llaves moleculares al difundirse en el medio circundante, el aire o el agua. Por tanto, un transductor puede ser la cerradura de forma apropiada y puede seguir el gradiente de concentración utilizando un dispositivo del tipo del vehículo de Braitenberg. Estas firmas olfativas son antiguas y poderosas. En nuestra especie se han visto sepultadas por miles de mecanismos distintos pero su situación en los ciemientos sigue siendo discernible. A pesar de nuestra complejidad, los olores siguen *conmoviéndonos* sin que sepamos

cómo o por qué, como describió en un texto ya famoso Marcel Proust*.

La tecnología honra el mismo principio de diseño en otro medio más: el RIPE (Radiobaliza indicadora de posición de emergencia) [en inglés, EPIRB, *Emergency Position Indicating Radio Beacon*] un transmisor de radio autónomo y a pilas que repite una y otra vez una señal concreta en una frecuencia determinada. Se puede comprar en una tienda de efectos marinos y llevárselo uno en el barco. Si nos vemos en un apuro, lo encendemos. Inmediatamente un sistema mundial de rastreo capta nuestra señal RIPE e indica su posición con un punto luminoso en un mapa electrónico. Busca asimismo esa señal concreta en su enorme lista de señales y a partir de ella identifica nuestro barco. La identificación simplifica muchísimo la búsqueda y el rescate porque supone una redundancia: la baliza pueden rastrearla ciegamente los receptores de radio (transductores) pero conforme los rescatadores se van acercando, les ayuda saber si tienen que buscar (con la vista) un bou pesquero de color negro, un velero pequeño de color verde oscuro o una balsa de goma de color naranja brillante. Pueden añadirse otros sistemas sensoriales para hacer que la aproximación final sea más rápida y menos vulnerable a una interrupción (por ejemplo, por si se acabaran las pilas del RIPE). En los animales, el rastreo del olor no es el único medio de la mamataxia. También se depende de las señales o firmas visuales y auditivas como ha demostrado notablemente el etólogo Konrad Lorenz con sus estudios pio-

* Los olores no sólo se utilizan como señales de identificación. Suelen representar papeles poderosos en la atracción para el apareamiento o incluso en suprimir la actividad sexual o la maduración de los rivales. Las señales del bulbo olfatorio se saltan el paso por el tálamo para llegar al resto del cerebro, de manera que, a diferencia de lo que ocurre en la visión, la audición e incluso el tacto, las órdenes olfativas llegan directamente a los antiguos centros de control, eliminando muchos intermediarios. Es probable que esta ruta más directa ayude a explicar el poder casi hipnótico y perentorio que algunos olores ejercen en nosotros.

neros sobre las «improntas» de los gansos y patos jóvenes. Los polluelos que no reciben su impronta al poco de nacer con una firma o señal adecuada de su madre, se fijarán en la primera cosa grande que se mueva y pase por su lado y de ahí en adelante la considerarán su madre.

Las balizas (con sus complementos sensores) son buenas soluciones de diseño cuando un agente debe rastrear (reconocer, volver a identificar) a un ente concreto (normalmente otro agente, como mamá) durante largo tiempo. Se instala anticipadamente la baliza en el objetivo y adelante. (Manifestación reciente de esto son las balizas antirrobo que se instalan ocultas en los coches y se activan a distancia cuando nos han robado el coche.) Pero como es normal, todo tiene su coste. Uno de los más evidentes es que amigos y enemigos pueden utilizar por igual la maquinaria de rastreo para acercarse al objetivo. Por ejemplo, es muy corriente que los predadores estén sintonizados con los mismos canales olfativos y auditivos que las crías que intentan no perder el contacto con mamá.

Olores y sonidos se emiten en cierto ámbito que no queda fácilmente bajo control del emisor. Un modo de conseguir una baliza más selectiva y de baja energía sería colocarle a mamá un punto azul concreto (un pigmento de cualquier tipo) y que la luz del sol creara en ese punto una baliza visible solamente para ciertos sectores del mundo y que desapareciera sencillamente cuando mamá se metiera en la sombra. La camada podría seguir el punto azul siempre que fuera visible. Pero este dispositivo exige una inversión en maquinaria fotosensible compleja: un ojo sencillo, por ejemplo, y no simplemente un par de fotocélulas.

La capacidad de mantenerse en estrecho y fiable contacto con alguna cosa concreta y ecológicamente muy importante (como, por ejemplo, mamá) no exige la capacidad de *concebir* esa cosa como una entidad concreta y duradera que va y viene. Como acabamos de ver, se puede conseguir una

mamataxia fiable con unos pocos trucos. Normalmente el talento es vigoroso en los entornos sencillos pero a una criatura armada con un sistema tan simple se la «engaña» con facilidad y cuando esto ocurre se precipita hacia su desgracia sin darse cuenta de su estupidez. No hace falta que el sistema monitorice su propio éxito o reflexione sobre las condiciones bajo las cuales tiene éxito o fracasa; ese es un añadido posterior (y más costoso).

El rastreo en colaboración (en el cual el objetivo proporciona una baliza cómoda simplificando así la tarea del rastreador) es un paso en el camino hacia el rastreo competitivo, en el cual el objetivo no sólo no proporciona una baliza de firma única sino que intenta ocultarse activamente, hacerse irrastreable. Este movimiento para no ser presa se ve contrarrestado por los sistemas de rastreo general de los predadores, que lo rastrean todo, diseñados para hacer de *aspectos cualesquiera* que revele un objeto digno de ser rastreado una especie de baliza privada y provisional: una «imagen de búsqueda» creada para el momento por un conjunto de detectores de rasgos del predador y que se utilizan para correlacionar, a cada instante, la firma del blanco, para revisar y poner al día esa imagen de búsqueda conforme cambia el blanco y siempre con el objetivo de mantener en la retícula el objeto elegido.

Es importante admitir que esta variedad de rastreos no requiere una categorización del blanco. Pensemos en un ojo primitivo, consistente en un dispositivo de unos pocos centenares de fotocélulas que transducen un dibujo cambiante de puntos luminosos y que a su vez se activan siempre que reciben luz reflejada por algo. Un sistema así podría enviar fácilmente un mensaje de este estilo: «X, ese lo que sea que es responsable de la investigación en curso con el grupo de puntos luminosos, acaba de moverse hacia la derecha.» (No es que tuviera que enviar el mensaje con tantas palabras: no harían falta en absoluto las palabras, ni símbolos, en el sis-

tema.) De modo que la única identificación que produce este sistema consiste en una especie de identificación continuada degenerada o mínima, instante tras instante, de aquello que se rastrea. Incluso aquí hay cierta tolerancia para el cambio y la sustitución. Un grupo de puntos luminosos que cambia gradualmente y que se moviera sobre un fondo más o menos estático puede cambiar su forma y su carácter interno radicalmente y seguir siendo rastreable siempre que no cambie con demasiada brusquedad. (El *fenómeno phi*, en el cual el sistema visual interpreta involuntariamente las secuencias de luz destellante, como la trayectoria de un objeto en movimiento, es una manifestación vívida de este circuito que llevamos incorporado a nuestros propios sistemas de visión.)

¿Qué ocurre cuando ese *X* se mete temporalmente detrás de un árbol? La solución más obvia es mantener intacta su imagen más reciente y luego echar un vistazo alrededor, al azar, con la esperanza de encajar otra vez esta baliza cuando aparezca, si es que aparece. Se pueden aumentar las probabilidades apuntando la imagen de búsqueda hacia el lugar más probable de reaparición de la baliza provisional. Y se puede conseguir una idea algo mejor que la puramente azarística del lugar más probable de reaparición teniendo en cuenta la trayectoria anterior de la baliza y trazando su continuación futura en línea recta. Lo cual nos proporciona ejemplos de producción de futuro en una de sus formas más antiguas y ubicuas, dándonos asimismo un caso claro de la flecha de intencionalidad apuntada hacia un blanco inexistente pero razonablemente esperable.

Esta capacidad de «mantenerse en contacto con» otro objeto (si es posible, tocándolo y manipulándolo literalmente) es el prerrequisito de la percepción de alta calidad. El reconocimiento visual de una persona o de un objeto concretos, por ejemplo, es casi imposible si la imagen del objeto no se mantiene centrada en la fovea de alta resolución del ojo durante un tiempo apreciablemente largo. Los microagentes

con hambre epistémica necesitan tiempo para alimentarse y organizarse. De tal manera que es una condición previa para desarrollar una descripción que identifique el objeto la capacidad de mantener ese foco de información *acerca* de una cosa concreta (de aquello, sea lo que fuere, que esté rastreando)*.

La manera de maximizar la probabilidad de mantener o de recobrar el contacto con la entidad rastreada es confiar en múltiples sistemas independientes, todos ellos falibles pero solapándose en sus competencias. Cuando falla un sistema hay otros que lo relevan y el resultado tiende a ser un rastreo suave y continuo compuesto de elementos que funcionan intermitentemente.

¿Cómo están vinculados unos a otros estos sistemas múltiples? Hay muchas posibilidades. Si tenemos dos sistemas sensoriales, podemos unirlos mediante una compuerta Y: los dos tienen que ponerse en marcha simultáneamente con la información para que el agente responda positivamente.

* Este aspecto de la primacía del rastreo sobre la descripción es, creo, el atisbo de la verdad en la doctrina filosófica ya abandonada de que existen dos tipos de creencia: las creencias *de re* que en cierto modo tratan «directamente» de los objetos, y las creencias *de dicto* que tratan de sus objetos sólo por la mediación de un *dictum*, es decir, una descripción definida (en lenguaje natural o en algún «lenguaje» mental). Supuestamente, el contraste se ilustra con el siguiente ejemplo, la diferencia entre

creer que Tom (*ese chico*, el que está allí) es un hombre

y

creer que quien quiera que me haya enviado este anónimo es un hombre.

En el primer caso, se supone que la intencionalidad es en cierto modo más directa, que se aferra a su objeto de un modo más primitivo. Pero como ya hemos visto, podemos refundir hasta los casos más directos y primitivos de rastreo perceptivo al modo *de dicto* (el *x* tal que *x* es, sea lo que fuera, responsable de ese grupo de puntos luminosos sujeto a investigación que acaba de saltar hacia la derecha) para poder extraer un rasgo del mecanismo que media esta especie de referencia «inmediata». La diferencia entre *de re* y *de dicto* es una diferencia en la perspectiva o hincapié que hace el hablante, y no en el fenómeno. Para ampliar esto, véase Dennett, «Beyond Belief» [«Más allá de la creencia»] (1982).

(Una compuerta Y puede construirse en cualquier medio: no es un objeto sino un principio organizativo. Son las dos llaves que han de abrirse para abrir una caja fuerte, o para disparar un misil nuclear: van unidas por una compuerta Y. Cuando enganchamos la manguera del jardín a un grifo por un lado y a una lanza de apertura regulable en el otro, esas dos válvulas van unidas por una compuerta Y: ambas han de estar abiertas para que salga el agua.) Como alternativa, se pueden unir dos sistemas mediante una compuerta O: o uno de los dos, o $A \vee B$, o los dos al mismo tiempo, provocan una respuesta positiva del agente. Las compuertas O se usan para proporcionar un respaldo o para reservar subsistemas dentro de un sistema mayor: si falla una unidad, la actividad de la unidad de reserva es suficiente para mantener el sistema en funcionamiento. Los aviones de dos motores los llevan unidos mediante una compuerta O: lo ideal puede ser que funcionen ambos, pero en caso de necesidad basta con uno.

Conforme se van añadiendo más sistemas se perfilan maneras intermedias de unirlos. Por ejemplo, se pueden unir de manera que SI un sistema A está ACTIVADO, entonces si *cualquiera* de B o C está ACTIVADO, el sistema responda positivamente; en otros casos *ambos* sistemas, B y C, deben estar activados para producir una respuesta positiva. (Lo cual equivale a una regla de mayoría que rigiera los tres sistemas; si la mayoría, una mayoría cualquiera, está ACTIVADA, el sistema responderá positivamente.) Todas las maneras posibles de unir sistemas con compuertas Y y con compuertas O (y con compuertas NO que sencillamente invierten la respuesta de un sistema, cambiando ACTIVADO por DESACTIVADO, y viceversa) se denominan funciones booleanas de esos sistemas, ya que el sistema puede describirse muy precisamente basándose en los operadores lógicos Y, O y NO, que formalizó por primera vez el matemático inglés del siglo XIX George Boole. Pero también hay maneras

no booleanas de que los sistemas entremezclen sus efectos. En lugar de llevar a todos los contribuyentes a una central del voto y otorgarle a cada uno de ellos un único voto (SI o NO, ACTIVADO o DESACTIVADO) y canalizar a partir de ahí su contribución a la conducta general hasta un punto único de decisión vulnerable (el efecto sumado de todas las conexiones booleanas) podríamos permitirles que mantuvieran sus conexiones independientes y continuamente variables con la conducta y que el mundo extrajera una conducta como resultado de toda su actividad. El vehículo de Valentino Braitenberg, con sus dos fototransductores cruzados, es un ejemplo absolutamente simple de ello. La «decisión» de girar a la izquierda o a la derecha surge de la fuerza relativa de las contribuciones de los dos sistemas transductor-motor, pero el efecto no puede representarse de manera eficiente y útil como función booleana de los respectivos «argumentos» aducidos por los transductores. (En principio, el comportamiento basado en entradas y salidas de un sistema tal puede aproximarse mediante una función booleana de sus componentes si se analiza adecuadamente, pero semejante hazaña analítica puede fracasar en revelar lo que es importante en esas relaciones. Por ejemplo, considerar el tiempo atmosférico como un sistema booleano es, en principio, posible, pero eso no permite trabajar con el sistema ni obtener información de él.)

Instalando docenas o centenares o miles de circuitos semejantes en un único organismo, pueden controlarse fiablemente las complejas actividades que protegen la vida, todo ello sin que ocurra nada en el interior del organismo que se parezca a *tener pensamientos concretos*. Se dan muchas decisiones del tipo *como si: como si* reconociera, *como si* jugara al escondite. También muchísimas maneras de que un organismo así equipado pueda «cometer errores», pero esos errores nunca equivalen a formular la representación de una proposición falsa y creer a continuación que es verdadera.

¿Qué versatilidad puede tener semejante arquitectura? Resulta difícil de decir. Los investigadores han diseñado y probado recientemente sistemas artificiales de control que producen muchas de las llamativas pautas de conducta que observamos en formas de vida relativamente sencillas, como en insectos y en otros vertebrados; de manera que es tentador creer que todas las asombrosas conductas complejas de esas criaturas puedan estar orquestadas mediante una arquitectura como esta, incluso aunque no sepamos todavía diseñar un sistema de tal complejidad. Después de todo, puede que el cerebro de un insecto no tenga más que unos cientos de neuronas y pensemos sin embargo en las complejísimas relaciones con el mundo que puede supervisar una disposición de células semejante. El biólogo evolucionista Robert Trivers indica, por ejemplo:

Las hormigas que cultivan hongos son agricultoras. Las obreras llevan hojas, las meten en los nidos, las preparan como medio de cultivo de los hongos, plantan en ellas los hongos, fertilizan los hongos con sus propias deposiciones, eliminan arrancando las especies competidoras y, por último, recolectan una parte concreta del hongo de la cual se alimentan (1985, pág. 172).

Además tenemos los rituales prologados y de pauta intrincadísima de apareo y de cría de los peces y los pájaros. Cada paso presenta exigencias sensoriales que deben satisfacerse antes de llevarse a cabo y que luego se guía adaptativamente a través de muchos obstáculos. ¿Cómo se controlan estas intrincadas maniobras? Alterando esforzadamente las fuentes de información en los experimentos, los biólogos han determinado muchas de las condiciones del entorno que se usan como indicativos aunque no es suficiente saber qué información puede captar cada organismo. La siguiente y difícil tarea es averiguar cómo pueden estar dise-

ñados sus diminutos cerebros para dar buen uso a toda esta útil sensibilidad.

Si fuéramos un pez o un cangrejo o algo parecido y uno de nuestros proyectos consistiera en, por poner un ejemplo, construir un nido de guijarros en el fondo del océano, necesitaríamos un dispositivo para localizar guijarros y un modo de encontrar el camino de regreso a *nuestro* nido para depositar el guijarro hallado en un lugar adecuado antes de seguir buscando. Sin embargo, no hace falta que el sistema sea a prueba de estupideces. Como es improbable que, durante nuestras búsquedas, surjan subrepticamente nidos de guijarros impostores en la zona que hayamos elegido (por lo menos hasta que los listos experimentadores humanos se toman interés en nosotros), podemos mantener unos niveles de reidentificación bastante bajos y poco costosos. Si se da un error en la «identificación», probablemente seguiríamos construyendo, no engañados por la treta sino absolutamente incapaces de reconocer o de apreciar el error, es decir, sin que nos preocupara lo más mínimo. Por otra parte, si estuviéramos equipados con un sistema de respaldo para el reconocimiento del nido, y el nido impostor no superara la prueba de respaldo nos veríamos metidos en el desorden, cogidos entre dos sistemas que nos requieren al mismo tiempo. Estos conflictos se dan pero no tiene sentido preguntar cuando el tal organismo va nervioso de un lado para otro: «¿Y ahora, qué está pensando? ¿Cuál es el *contenido proposicional* de su estado de confusión?»

En organismos como el nuestro (organismos equipados con muchísimas capas de sistemas que se controlan a sí mismos, que pueden comprobar la existencia de estos conflictos cuando se presentan y pueden intentar mediar en ellos) a veces queda perfectamente claro qué error se ha cometido. Un ejemplo perturbador es la ilusión de Capgras, una afección extravagante que a veces padecen los seres humanos que han sufrido una lesión cerebral. La señal distintiva de la ilusión

de Capgras es el convencimiento del paciente de que un conocido cercano (generalmente un ser querido) se ha visto suplantado por un impostor que se le parece (y que habla y actúa como él) y que el conocido ¡ha desaparecido misteriosamente! Este sorprendente fenómeno debería agitar las aguas de la filosofía. Los filósofos han reunido muchos casos extraños de identidad cambiada para ilustrar sus diversas teorías filosóficas y la literatura de la filosofía está plagada de fantásticos experimentos mentales sobre espías y asesinos que viajan de incógnito, amigos íntimos que se disfrazan de gorila y gemelos idénticos que se reencuentran al cabo de muchísimo tiempo, pero los casos de la vida real de la ilusión de Capgras no han llamado todavía la atención de los filósofos. Lo que es particularmente llamativo de estos casos es que no dependen de sutiles disfraces ni de atisbos fugaces. Al contrario: la ilusión persiste incluso cuando el agente somete al individuo que sirve de blanco a un estudio detallado. Se sabe de pacientes de Capgras que han asesinado a sus cónyuges ¡tan seguros estaban de que aquellos intrusos estaban intentando meterse en pellejos (en vidas al completo) que no les pertenecían por derecho propio! No puede haber duda de que en semejante caso, tan triste, el agente en cuestión ha dado por ciertas determinadas proposiciones específicas de no identidad: *este hombre no es mi marido*; este hombre es cualitativamente lo más parecido a mi marido y sin embargo no es mi marido. Resulta muy interesante el hecho de que las personas que sufren de semejante ilusión son claramente incapaces de decir por qué están tan seguras.

El neuropsicólogo Andrew Young (1994) ofrece una hipótesis ingeniosa y plausible para explicar qué es lo que va mal. Young contrasta la ilusión de Capgras con otra curiosa afección causada por lesiones cerebrales: la *prosopagnosia*. Las personas con prosopagnosia no pueden reconocer rostros humanos familiares. Pueden tener buena vista pero no

son capaces de identificar ni siquiera a sus amigos íntimos hasta que les oyen hablar. En un experimento característico se les muestran colecciones de fotografías: algunas son de individuos desconocidos y otras de miembros de su familia y personajes conocidos: Hitler, Marilyn Monroe, J. F. Kennedy. Cuando se les pide que cojan las de rostros familiares, no lo hacen mejor que si las escogieran al azar. Pero durante más de una década los investigadores han sospechado que a pesar de un resultado tan asombrosamente pobre, *algo* había en algunos prosopagnósicos que identificaba correctamente a los miembros de la familia y a la gente famosa porque su cuerpo reaccionaba de manera diferente ante las caras familiares. Si mientras miran la foto de una cara conocida se les dicen diversos nombres posibles de la persona fotografiada, muestran una alta respuesta galvánica de la piel al oír el nombre correcto. (La respuesta galvánica de la piel es la medida de la conductividad eléctrica y la principal prueba en la que se basan los polígrafos o «detectores de mentiras».) La conclusión que Young y otros investigadores sacan de estos resultados es que debe haber dos (o más) sistemas que pueden identificar un rostro y que uno de ellos está intacto en los prosopagnósicos que muestran esta respuesta. Este sistema sigue funcionando bien, furtiva y prácticamente sin ser notado. Supongamos ahora, dice Young, que los pacientes de Capgras tengan la disfuncionalidad opuesta: el sistema (o sistemas) que opera al descubierto para el reconocimiento consciente de rostros funciona perfectamente (y por eso los pacientes de Capgras están de acuerdo en que los «impostores» tienen el mismo aspecto que sus seres queridos) pero que el sistema (o sistemas) furtivo que normalmente proporciona un voto de confianza al reafirmar lo visto en esas ocasiones está deteriorado y ominosamente silencioso. La *ausencia* de tal colaboración sutil a la identificación es tan irritante («¡Falta algo!») que equivale a un veto de bolsillo al voto positivo del sistema superviviente: el resultado

que sale a la luz es la convicción íntima del paciente de estar mirando a un impostor. En lugar de atribuir esa falta de identificación a un sistema perceptivo defectuoso, el agente echa la culpa al mundo de un modo tan extravagante metafísicamente y tan improbable que apenas caben dudas del poder (el poder político, efectivamente) que normalmente tiene sobre todos nosotros ese sistema deteriorado. Cuando esta hambre epistémica concreta queda insatisfecha, nos da tal ataque que echa por la borda las contribuciones de los demás sistemas.

Entre el olvidadizo cangrejo y el extravagante paciente equivocado de Capgras hay casos intermedios. ¿Puede un perro equivocarse de dueño o no reconocerlo? Según Homero, cuando Ulises regresa a Ítaca después de su odisea de veinte años, disfrazado con harapos como un mendigo, le reconoce su antiguo perro Argo, mueve la cola, agacha las orejas y después muere. (Y habría que recordar que Ulises se enjuga secretamente una lágrima.) Así como hay razones para que un cangrejo reconozca (o intente reconocer) la identidad de su propio nido también hay razones para que un perro reconozca (o intente reconocer) la identidad de su amo, entre otras muchas cosas importantes de su mundo. Cuanto más fuertes son las razones para reidentificar las cosas, más beneficioso es no cometer errores y por ello mayor inversión se hace en maquinaria perceptiva y cognitiva. Los modos de aprendizaje avanzado dependen, ciertamente, de capacidades preferentes de (re)identificación. Sea un caso sencillo, el de un perro que ve sobrio a Ulises el lunes, el miércoles y el viernes, pero le ve borracho el sábado. De este conjunto de experiencias se pueden sacar diversas conclusiones lógicamente disponibles: que hay hombres borrachos y sobrios, que un hombre puede estar borracho un día y sobrio otro, que Ulises es uno de esos hombres. El perro no podría (lógicamente no podría) saber el segundo o el tercero de los hechos de esta secuencia de experiencias independientes

a menos que tuviera algún modo (falible pero fiable) de reidentificar al hombre como ese mismo hombre en una y otra experiencia. (Millikan, 1997). (Podemos ver una aplicación más drástica del mismo principio en el curioso hecho de que no podemos saber [como cuestión de lógica] qué aspecto tenemos mirándonos en un espejo a menos que tengamos algún *otro* medio de saber que la cara que vemos es la nuestra. Sin esa identificación independiente no podríamos saber más de nuestro aspecto mirándonos en un espejo que si miráramos una fotografía que por casualidad fuera de nuestra cara.)

Los perros viven en un mundo conductual mucho más rico y más complejo que el mundo del cangrejo, con más oportunidades de subterfugios, de faroleo y de disfraces, y por ello tienen mayores beneficios que obtener del rechazo de pistas equívocas. Pero además, los sistemas de un perro no tienen por qué estar a prueba de estupideces. Si el perro mete la pata en la identificación (del tipo que sea) podemos caracterizarlo como un caso de identidad equivocada sin haber llegado a la conclusión de si el perro es capaz de *pensar* la proposición del comportamiento como si creyera en ella. El comportamiento de Argo en el relato homérico es conmovedor pero no podemos consentir que el sentimentalismo nuble nuestras teorías. Argo también podría amar los aromas del otoño y responder con alegría todos los años cuando le llegara al morro la primera bocanada de los frutos maduros, pero ello no querría decir que tuviera manera alguna de distinguir entre tipos de estaciones recurrentes, como el otoño, e individuos que regresan, como Ulises. Para Argo ¿Ulises no es nada más que una colección organizada y agradable de olores y sonidos, imágenes y sentimientos, una especie de estación recurrente pero irregular (¡llevaba veinte años sin volver!) durante la cual se ven favorecidos determinados comportamientos? Es una estación que generalmente está sobria aunque a veces se sabe también que está ebria. *Noso-*

tros podemos ver, desde nuestra peculiar perspectiva humana, que el éxito de Argo en este mundo dependerá generalmente de lo mucho que su conducta se aproxime al comportamiento de un agente que, como nosotros los seres humanos adultos, distinga claramente entre distintos individuos. De manera que cuando interpretamos su conducta desde el enfoque intencional, hacemos bien en atribuir a Argo creencias que distinguen a Ulises de otras personas, perros rivales fuertes de otros rivales más débiles, corderos de otros animales, Ítaca de otros lugares, y así sucesivamente. Pero debemos estar preparados para descubrir que esta comprensión aparente del perro presenta algunas lagunas sorprendentes, lagunas inconcebibles en un ser humano con nuestro esquema conceptual y por ende absolutamente inexpresables en los términos del lenguaje humano.

Durante milenios han sido muy corrientes los relatos de inteligencia de las mascotas. El antiguo filósofo estoico Crisipo contaba de un perro que podía realizar la siguiente hazaña racional: al llegar a una encrucijada con tres caminos, olía los caminos A y B y *sin olfatear* el C, corría por éste habiendo razonado que como no hay rastro en A ni en B, la presa debería haber ido por C. Las personas son, en cambio, menos aficionadas a contar las estupideces de sus mascotas y suelen resistirse a las conclusiones que les marcan las lagunas que descubren en sus habilidades. Un perrito tan listo y sin embargo ¿no es capaz de averiguar cómo desenrollar la correa mientras gira en torno a un árbol o a una farola? Y no se trata, como podría parecer, de una prueba de inteligencia desleal para un perro, comparada, por ejemplo, con un examen de sensibilidad hacia la ironía en la poesía, o de la apreciación de la transitividad de la expresión comparativa *más caliente que* (si A está más caliente que B y B está más caliente que C, ¿entonces A está más caliente o más frío que C?). Pero son pocos los perros que pueden superarla, si es que la supera alguno. Y con toda su inteligencia, los delfines

son extrañamente incapaces de averiguar que fácilmente podrían saltar hacia su libertad por encima de las redes atuneras. Saltar fuera del agua no es una acción antinatural para ellos, lo que hace más llamativa su cerrazón. Como los investigadores van descubriendo con regularidad, cuanto más ingeniosamente se investigan las capacidades de los animales no humanos, más probable es descubrir abruptas lagunas en ellos. La capacidad de los animales para generalizar a partir de la explotación de sus conocimientos es muy limitada. (Véase Cheney y Seyfarth, *How Monkeys See the World* [Cómo ven los monos el mundo], 1990 para un informe esclarecedor de esta pauta de investigaciones en las mentes de los monos tumbly*.)

Nosotros, seres humanos, gracias a la perspectiva que obtenemos de nuestra capacidad de reflexionar según nuestra peculiar manera, podemos discernir los fallos de rastreo que quedarían muy fuera del alcance de otros seres. Supongamos que Tom ha llevado encima, durante muchos años, una moneda de la suerte. Aunque no la llame de ninguna manera en especial nosotros la llamaremos Amy. Tom se lleva a Amy a España, la pone en la mesilla de noche al dormir, y así sucesivamente. Un día, en un viaje a la ciudad de Nueva York, cediendo a un impulso, Tom lanza a Amy a una fuente donde se mezcla con las demás monedas de un penique, absolutamente indistinguible, tanto para Tom como para nosotros, de todas ellas... o por lo menos de todas aquellas que llevan la misma fecha de acuñación que Amy. Y aun así, Tom puede *reflexionar* sobre este desarrollo. Puede reconocer la verdad de la proposición de que uno, y sólo uno, de esos peniques es su moneda de la suerte que siempre ha llevado consigo. Puede preocuparse (o sencillamente divertirse) ante el hecho de que ha perdido irremisiblemente la pista de algo con lo que ha mantenido el contacto durante años,

* *Cercopithecus pygerythrus*. (N. del T.)

por un medio u otro. Supongamos que coja de la fuente una moneda que pueda ser Amy. Puede darse cuenta de que una de las dos proposiciones siguientes, y exactamente una de las dos, es verdadera:

1. El penique que tengo ahora en la mano es el penique que traje a Nueva York.
2. El penique que tengo ahora en la mano no es el penique que traje a Nueva York.

No hace falta ser un científico espacial para darse cuenta de que una de las dos tiene que ser verdad, aunque ni Tom ni nadie más en la historia del mundo, del pasado o del futuro, sepa decir cuál. Esta capacidad que *nosotros* poseemos de montar hipótesis sobre la identidad, e incluso de comprobar dichas hipótesis en la mayoría de las circunstancias, es bastante ajena a todas las demás criaturas. Las prácticas y proyectos de muchas criaturas les exigen rastrear y reidentificar individuos (sus madres, sus parejas, sus presas, sus superiores y sus subordinados en el grupo) pero no hay prueba que parezca indicar que se den cuenta de que cuando hacen eso, es eso lo que están haciendo. Su intencionalidad nunca llega al punto de la concreción metafísica a la que puede elevarse la nuestra.

¿Cómo lo conseguimos nosotros? No hace falta ser un científico espacial para tener semejantes pensamientos, pero lo que sí hace falta es una criatura gregoriana que entre sus herramientas mentales cuente con el lenguaje. Pero para poder utilizar el lenguaje tenemos que estar equipados con los talentos que nos permitan extraer estas herramientas mentales del medio (social) en que residen.

CAPÍTULO 5

LA CREACIÓN DEL PENSAMIENTO

Psicólogos naturales que no piensan

El lenguaje se inventó para que las personas pudieran ocultarse sus pensamientos unas a otras.

Charles-Maurice de Talleyrand.

Hay muchos animales que se ocultan pero que no piensan que se están ocultando. Muchos animales se juntan en manadas pero no piensan que se están reuniendo en manadas. Muchos animales cazan pero no piensan que están cazando. Todos ellos son beneficiarios de sistemas nerviosos que se ocupan de los controles de esos comportamientos inteligentes y apropiados sin calentarle la cabeza al sujeto con pensamientos, o con algo que discutiblemente se parezca a los pensamientos que pensamos los pensantes. Cazar y comer, ocultarse y huir, reunirse en manada y dispersarse son actos que parecen estar todos al alcance de esos mecanismos no pensantes. Pero ¿hay conductas inteligentes que deban

verse acompañadas, precedidas o controladas por pensamientos inteligentes?

Si la estrategia de adoptar el enfoque intencional es una bendición tan grande como he pretendido, entonces un lugar evidente para buscar un avance en las mentes animales es el de los sistemas intencionales que en sí son capaces de adoptar el enfoque intencional hacia otros (y hacia sí mismos). Deberíamos buscar comportamientos que sean sensibles a las diferencias en los (hipotéticos) pensamientos de otros animales. Hay una antigua broma entre los conductistas que dice que no creen en las creencias, piensan que nada puede pensar y que, en su opinión, nadie tiene opiniones. ¿Qué animales se consideran conductistas, incapaces incluso de plantear hipótesis sobre las mentes de los demás? ¿Qué animales se ven obligados, o a qué animales se les permite, sacar una titulación más alta? Parece ser algo paradójico un agente no pensante preocupado con el descubrimiento y la manipulación de los pensamientos de otros agentes, de manera que quizá aquí podamos encontrar un nivel de complejidad que obligue al pensamiento a evolucionar.

¿Podría el pensamiento lanzarse a sí mismo a la existencia tirándose de sus propias orejas? (Si usted va a pensar cómo pienso, yo voy a tener que empezar a pensar cómo piensa usted para estar a la par... una especie de carrera armamentista de reflexiones.) Muchos teóricos han pensado que hay alguna versión de esta carrera armamentista que explica la evolución de la inteligencia superior. En un escrito influyente («Nature's Psychologists» [«Psicólogos de la naturaleza»], 1978), el psicólogo Nicholas Humphrey argüía que el desarrollo de la conciencia *de uno mismo* era una estrategia para desarrollar y experimentar hipótesis sobre lo que pasaba por las mentes de *otros*. La idea es que la capacidad de hacer sensible nuestra propia conducta a la manera de pensar de otro agente (además de hacernos capaces de manipularla) nos dotaría automáticamente de capacidad

de sensibilizar nuestra conducta a nuestro propio pensar. Cosa que podría ser o bien, como sugiere Humphrey, porque se utiliza la conciencia de uno mismo como fuente de hipótesis sobre las conciencias de los demás, o bien porque cuando se adquiere la costumbre de adoptar el enfoque intencional hacia los demás, se cae en la cuenta de que resulta útil someterse uno mismo a ese tratamiento. O porque mediante cierta combinación de estas razones, la costumbre de adoptar el enfoque intencional podría ampliarse para abarcar tanto la interpretación de los demás como la interpretación de uno mismo.

En un ensayo titulado «Conditions of Personhood» [«Condiciones para ser persona»], 1976, argüí que una clave importante para convertirse en persona fue el paso de un sistema intencional de *primer orden* a un sistema intencional de *segundo orden*. Un sistema intencional de primer orden tiene creencias y deseos relacionados con muchas cosas, pero *no* relacionados con las creencias y los deseos. Un sistema intencional de segundo orden tiene creencias y deseos relacionados con creencias y deseos, propios o ajenos. Un sistema intencional de tercer orden sería capaz de proezas tales como *querer* que otro *crea* que *quiere* algo, mientras que un sistema intencional de cuarto orden podría *creer* que otro *quiere* que *crea* que (el primero) *quiere* algo, y así sucesivamente. Yo sostenía que el gran paso fue el del primer al segundo orden; los órdenes superiores son sólo cuestión de qué cantidad de cosas puede retener un agente en su cabeza en un momento dado, cosa que varía con las circunstancias incluso para un mismo agente. A veces los órdenes superiores son tan sencillos que son involuntarios. ¿Por qué ese tipo de la película hace tantos esfuerzos para no sonreír? En su contexto es deliciosamente evidente: su esfuerzo nos muestra que *él ve* que *ella no se da cuenta* de que *él ya sabe* que *ella quiere* que él la saque a bailar y que *¡él quiere* que las cosas sigan así! Otras veces, algunas repeticiones más sen-

cillas pueden dejarnos perplejos. ¿Está usted seguro de que yo quiero que usted crea que yo quiero hacerle creer lo que estoy diciendo?

Pero si como otros autores y yo hemos argumentado, la intencionalidad de orden superior es un avance importante en los tipos de mente, la línea divisoria que buscamos entre la inteligencia que piensa y la que no, no está tan clara. Algunos de los ejemplos mejor estudiados de una intencionalidad (aparente) de orden superior entre criaturas no humanas parecen seguir quedando en el campo de la destreza irreflexiva. Consideremos la «maniobra de distracción», esa conducta bien conocida de los pájaros que anidan a baja altura y que, al acercarse al nido un predador, se apartan subrepticamente de sus vulnerables huevos o polluelos y de la manera más ostentosa fingen tener un ala rota, aleteando, cayéndose y chillando del modo más patético. Esta conducta atrae generalmente al predador alejándolo del nido en una caza estúpida en la cual casi nunca termina por capturar el alimento «fácil» que se le ofrece. El fundamento latente en este comportamiento está claro y siguiendo la útil práctica de Richard Dawkins de su libro de 1976, *El gen egoísta*, podemos darle la forma de un *imaginario* soliloquio:

Soy un pájaro que anida a poca altura, y cuyas crías no pueden defenderse cuando las descubre un predador. Es *esperable* que este predador que se acerca las descubra pronto a no ser que yo lo distraiga; podría distraerse por su mismo *deseo* de capturarme y comerme, pero sólo si *pensara* que tiene una posibilidad *razonable* de cazarme, porque no es tonto; y llegaría a *creerlo* si yo le *diera pruebas* de que ya no puedo volar; cosa que podría hacer fingiendo que tengo un ala rota, etc. (De Dennett, 1983).

En el caso del apuñalamiento de César por Bruto, que expusimos en el capítulo 2, estaba dentro de los límites de lo

plausible suponer que Bruto pasó por algo parecido al proceso del soliloquio esbozado para él, aunque normalmente, incluso en los más locuaces humanos que hablan consigo mismo, la mayor parte de ese soliloquio transcurriría sin formularse en palabras. Sin embargo, desafía nuestras creencias suponer que cualquier pájaro tiene un soliloquio como el aquí expuesto. Y sin embargo, un soliloquio así expresa el fundamento que ha conformado la conducta, sea o no capaz el pájaro de darse cuenta del fundamento. Las investigaciones de la etóloga Carolyn Ristau (1991) han demostrado que en al menos una de tales especies (el chorlito americano *) los individuos controlan de maneras muy complejas sus maniobras de distracción. Por ejemplo, vigilan la dirección de la mirada del predador, subiendo el volumen de su exhibición si el predador parece perder interés y adaptan su conducta de otras maneras según lo que detecten en la conducta del predador. Los chorlitos discriminan también basándose en la forma y el tamaño del intruso: como las vacas no son carnívoras, una vaca no es la adecuada para sentirse atraída por la perspectiva de una fácil comida pajaril, de modo que algunos chorlitos tratan a las vacas de distinta manera, graznando, picoteando e intentando apartarlas en lugar de atraerlas hacia otro sitio.

Parece que las liebres pueden evaluar el tamaño de un predador que se les aproxime, un zorro por ejemplo, y hacer una estimación del peligro (Hasson, 1991; Holley, 1994). Si la liebre concluye que un zorro concreto se las ha arreglado para acercarse a la distancia de caza, o bien se encogerá y se quedará congelada (contando con escapar por completo a la atención del zorro) o bien se encogerá para salir corriendo lo más veloz y silenciosamente posible, interponiendo y dejando atrás cualquier cosa que pueda servirle de cobertura. Pero

* *Charadrius melodus*. La maniobra de distracción expuesta es general para todos los chorlitos y algunas otras carádridas. (*N. del T.*)

si la liebre concluye que el zorro probablemente no tendría éxito en su caza, hace algo extraño y maravilloso. Se levanta sobre sus patas traseras, muy conspicuamente, y ¡mira al zorro hasta que éste se retira! ¿Por qué? Porque anuncia al zorro que debe abandonar la caza. «Ya te he visto y no te tengo miedo. No malgastes tu precioso tiempo e incluso tus energías más preciadas persiguiéndome. ¡Abandona!» Y generalmente el zorro llega a esa misma conclusión, buscando su cena en otra parte y abandonando a la liebre que, de este modo, ha preservado su propia energía para continuar con su alimentación.

Una vez más, el fundamento de esta conducta es casi ciertamente latente. Probablemente no es una táctica que la propia liebre haya averiguado por sí misma o de la que sea capaz de reflexionar. Las gacelas a las que persiguen los leones o las hienas suelen hacer algo parecido, llamado «rebote». Dan unos saltos ridículamente altos que evidentemente no les benefician en su huida pero pensados para advertir a sus predadores de su superior velocidad. «No te molestes en perseguirme *a mí*. Persigue a esa otra. Soy tan veloz que puedo derrochar tiempo y energía dando estos saltos estúpidos y aun así correr más que tú.» Y aparentemente funciona: generalmente los predadores dirigen su atención a otros animales.

Podrían citarse otras variantes de conducta predador-presa, todas ellas con elaborados fundamentos, pero no hay pruebas (o son escasísimas) de que los animales se representen de verdad esos fundamentos de manera alguna. Si vamos a considerar *estas* criaturas como «psicólogos naturales» (por utilizar el término de Humphrey), aparentemente son psicólogos naturales que no piensan. Estas criaturas no se representan la mente de aquellos con los cuales se relacionan, es decir, no necesitan consultar un «modelo» interno de la mente de otro ser para poder prever su comportamiento y ajustar por tanto el suyo de forma conveniente. Están

bien provistos de una «lista» más bien larga de conductas posibles, bien ligadas a una lista más bien larga de pistas perceptivas y no necesitan saber nada más. ¿Equivale esto a leer la mente? ¿Son o no son los chorlitos americanos, las liebres o las gacelas sistemas intencionales de orden superior? Esta pregunta empieza a parecer menos importante que la pregunta siguiente: ¿Qué parte de esta aparente capacidad de lectura mental podría estar organizada? Entonces ¿cuándo surge la necesidad de sobrepasar esas largas listas, de ir más allá? El etólogo Andrew Whiten ha sugerido que la necesidad surge sencillamente cuando la lista se hace demasiado larga y demasiado abultada para seguir aumentando. Un par de listas así equivale, en términos lógicos, a una conjunción de condicionales o parejas de «si... entonces...»:

[Si ves x , haz A] y [si ves y , haz B] y [si ves z , haz C],...

Dependiendo de cuántos condicionales independientes haya, puede resultar económico consolidarlos en unas representaciones del mundo más organizadas. Quizá en algunas especies (qué especies sean es otra cuestión que sigue abierta) la brillante innovación de la *generalización* explícita se incorpora al cuadro, permitiendo que las listas se trocean y se recompongan a petición de ciertos primeros principios conforme van surgiendo nuevos casos. Consideremos el diagrama de Whiten sobre la complejidad que se organizaría en torno a la representación interna en un animal de un deseo específico de otro animal.

Como antes, *nosotros* podemos ver el fundamento de tal consolidación, pero no hace falta que este fundamento pase, en modo alguno, por las mentes de los sujetos. Si tienen suerte suficiente como para que les toque esta mejora del diseño, serán sencillamente sus beneficiarios sin darse cuenta de cómo ni por qué funciona. Pero este diseño ¿es realmente la mejora que aparenta ser?, ¿cuáles son sus costos y sus

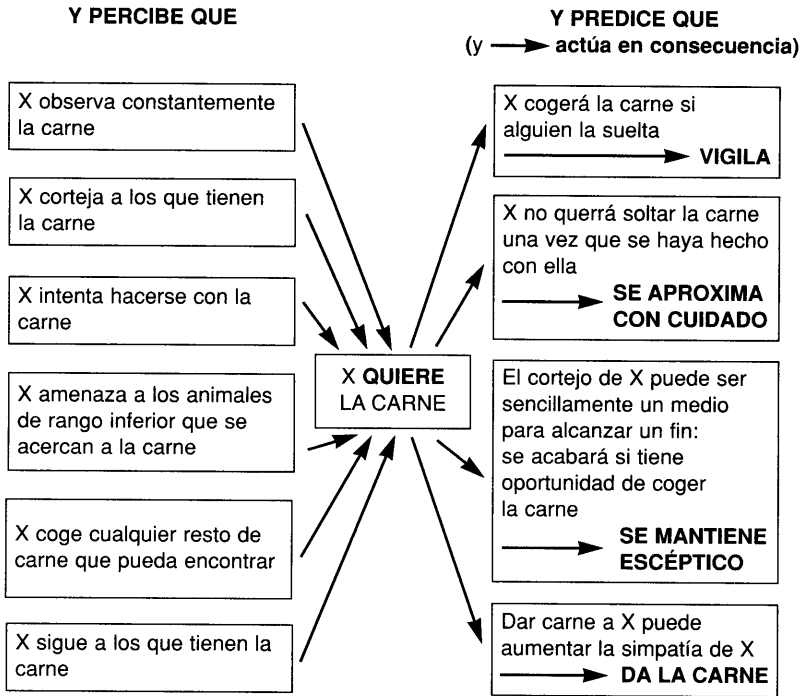


Figura 5.1

beneficios? Y dejando a un lado su valor ¿cómo podría llegar a darse, a existir? ¿Apareció sin más un día, como una reacción al azar y desesperada a un creciente problema de «cabeza»... de demasiadas reglas condicionales que debían mantenerse en servicio simultáneamente? Quizá, pero nadie conoce aún ningún límite superior plausible del número de estructuras de control semiindependientes y concurrentes que pueden coexistir en un sistema nervioso. (En un agente real con sistema nervioso real, puede no haber ninguna. Puede que unos pocos cientos de miles de semejantes circuitos de control perceptivo-conductual puedan entremezclarse eficientemente en el cerebro... ¿cuántos podrían ser exigibles?)

¿Acaso no podría haber algún otro tipo de presión selec-

tiva que llevara a la reorganización de las estructuras de control, generando como una cosa extra la capacidad de generalizar? El etólogo David McFarland (1989) ha argumentado que la oportunidad de la comunicación proporciona esa presión de diseño y, lo que es más, la cínica sugerencia de Tallestrand que abre este capítulo se acerca a una importante verdad. Cuando la comunicación surge en una especie, dice, la pura honradez no es la mejor política ya que nos haría excesivamente explotables por nuestros competidores (Dawkins y Krebs, 1978). El contexto competitivo es claro en todos los casos de comunicación entre predador y presa, como las prácticas mínimas de comunicación que presentan la gacela que rebota y la liebre que amedrenta al zorro; y ahí es evidente cómo surge la posibilidad del farol. En la carrera armamentista que tiene como objeto la producción de futuro, se tiene una ventaja considerable si se puede producir más y mejor futuro acerca del otro que el futuro que ese otro puede producir acerca de nosotros, de modo que siempre incumbe al agente mantener inescrutables su propio sistema de control. La impredecibilidad es en general un buen rasgo protector que nunca debe derrocharse sino emplearse siempre sabiamente. Hay mucho que obtener de la comunicación si se reparte astutamente... lo cual es suficientemente verdad como para mantener el crédito propio a buen nivel pero suficientemente falso como para mantener abiertas las propias posibilidades. (Este es el primer aspecto de la sabiduría en el juego del póquer: el que nunca farolea, nunca gana; el que farolea siempre, siempre pierde.) Hace falta ampliar un poco la imaginación para ver a la liebre y al zorro como colaboradores en sus problemas comunes de manejo de los recursos, pero lo cierto es que a ambos les va mejor con sus treguas ocasionales.

Las perspectivas de ampliar la colaboración y por ende de multiplicar sus beneficios son visibles con mucha mayor claridad en el contexto de la comunicación con los miembros de

la propia especie. En ella, compartir el alimento, y compartir los costes y los riesgos del cuidado de las crías y de la defensa del grupo y demás aspectos proporcionan múltiples oportunidades de colaboración pero sólo si se pueden satisfacer las condiciones más bien exigentes para explotar dichas oportunidades. La colaboración entre padres o entre padres y crías no puede tomarse como si la diera la naturaleza sin más; la omnipresente posibilidad de competencia sigue estando detrás de cualesquiera convenciones útiles que surjan y hay que tener en cuenta este contexto de competencia.

Según McFarland, la necesidad de una representación explícita y manipulable de la conducta propia surge solamente cuando aparece la opción de la comunicación cooperativa sin dejar de ser autoprotectora, porque entonces el agente debe controlar una nueva forma de conducta: la conducta de comunicar explícitamente algo acerca de la conducta de algún otro. («Estoy intentando pescar un pez» o «Busco a mi madre» o «Sólo estoy descansando».) Ante la tarea de conformar y ejecutar tal acto comunicativo, el problema del agente es una versión del mismo problema que afrontamos nosotros como observadores teóricos: ¿Cómo debe dividirse la maraña de los circuitos de control de conducta (de competencia, de realce, de unión, de entrelazamiento) del agente para conformar «alternativas» competitivas? La comunicación favorece las respuestas bien definidas. Como dice la expresión: «¿Vas de pesca o a remojar el cebo?»; así, las exigencias de comunicación, al obligar a un agente a pronunciarse por una categoría, pueden a menudo crear una distorsión, de ese estilo que vemos cuando se nos pide que señalemos una respuesta en una prueba mal diseñada de múltiples opciones: si no existe la opción «ninguna de las anteriores» nos vemos forzados a elegir la menos mala, sea cual sea. McFarland sugiere que esta tarea de separar donde la naturaleza no ha proporcionado juntas visibles es un problema que el agente resuelve por lo que podría llamarse *con-*

fabulación aproximada. El agente llega a etiquetar sus tendencias *como si* estuvieran regidas por objetivos explícitamente representados (como si fueran proyectos originales para actuar) en lugar de tendencias de actuación que surgen de la interacción de distintas posibilidades. Una vez que las *representaciones de intenciones* (intenciones en su sentido cotidiano) salen a la luz de esta manera equívoca, pueden tener éxito convenciendo al propio agente de que tiene estas intenciones primordiales y bien definidas para regir sus acciones. Para poder resolver su problema de comunicación, el agente ha hecho uso de un intermediario en sí mismo, un menú de opciones explícitas entre las cuales elegir y luego se ha visto arrastrado en cierto modo por su propia creación.

Las oportunidades de dar buen uso a semejantes comunicaciones son, sin embargo, estrictamente limitadas. Muchos entornos son inhóspitos para guardar secretos, independientemente de cualesquiera propensiones o habilidades de los agentes que viven en él; y si no se puede guardar un secreto queda poco juego para la comunicación. Según la antigua sabiduría popular, la gente que vive en casas de cristal no debe arrojar piedras, pero los animales que viven en los equivalentes naturales de las casas de cristal no tienen piedras que arrojar. Los animales que viven en estrecho agrupamiento en territorio abierto rara vez se encuentran mucho tiempo (si es que alguna vez ocurre) fuera de la vista y del oído (y del olfato y del roce) de sus congéneres y por ello no tienen oportunidades de satisfacer las condiciones en las cuales puedan florecer los secretos. Supongamos que *p* es un hecho ecológicamente valioso y supongamos que lo sabemos, y que nadie más lo sabe... todavía. Si nosotros y otro de nuestros agentes competidores potenciales de las proximidades tenemos acceso a prácticamente la misma información del entorno, entonces es casi imposible que se den circunstancias en las cuales podamos utilizar a nuestro favor ese gradiente de información provisional. Puede que

seamos los primeros ñus en ver o en oler al león hacia el noroeste, pero verdaderamente no podemos atesorar (o vender) esta información, porque los que están pegados a nosotros dispondrán de ella en seguida por sí mismos. Como la posibilidad de que esa ventaja informativa pueda controlarse es muy escasa, un perverso ñu (por ejemplo) tendría pocas oportunidades de beneficiarse de su talento. Sencillamente ¿qué podría hacer para sacar una ventaja rastrera sobre los demás?

El enfoque intencional nos muestra en seguida que la conducta aparentemente sencilla de *mantener un secreto* (un comportamiento cero, desde muchos puntos de vista) es en realidad una conducta cuyo éxito depende de que se satisfaga un conjunto bastante exigente de condiciones. Supongamos que *Bill mantiene el secreto p ante Jim*. Deben darse las siguientes condiciones:

1. Bill sabe (o cree) *p*.
2. Bill cree que Jim no cree que *p*.
3. Bill no quiere que Jim crea *p*.
4. Bill cree que Bill puede conseguir que Jim no llegue a creer que *p*.

Esta última condición es la que restringe el poder guardar los secretos de manera sofisticada (por ejemplo, acerca de rasgos del entorno externo) a entornos conductuales bastante específicos. Cosa que demostraron claramente los experimentos de los años setenta del primatólogo Emil Menzel (1971, 1974), en los cuales se mostraba a los chimpancés, uno a uno, la localización de la comida oculta y con ello se les daba la oportunidad de engañar a los demás sobre su localización. Solían llegar a la oportunidad de hacerlo con fascinantes resultados pero su conducta siempre dependía de que los experimentadores reprodujeran un estado de cosas en el laboratorio (en este caso, una jaula adyacente a un re-

cinto vallado de mayor tamaño) que rara vez se darían en la naturaleza: el chimpancé que ve la comida oculta debe estar en la situación de *saber que los demás chimpancés no le ven mirando la comida*. Lo cual se conseguía manteniendo a los demás chimpancés encerrados en una jaula común mientras que al chimpancé elegido se le llevaba a solas al recinto mayor y se le mostraba la comida oculta. El chimpancé elegido podía llegar a comprender que sólo él sabía que *p*, es decir, que sus aventuras informativas no eran visibles para los demás que estaban en la jaula. Y, por supuesto, tenía que haber algo que el chimpancé pudiera hacer para proteger su secreto (por lo menos durante cierto tiempo) una vez que se soltara a los demás.

Los chimpancés en estado salvaje frecuentemente se separan de sus grupos en distancia y en tiempo suficientes como para adquirir secretos que queden bajo su control, de modo que son una buena especie para hacerles estas pruebas. En animales cuya historia evolutiva no se ha desarrollado en entornos en los que surgen de manera natural y frecuente estas oportunidades, hay pocas probabilidades de que haya evolucionado la capacidad de explotar tales oportunidades. Por supuesto que no es imposible descubrir (en el laboratorio) un talento hasta entonces sin utilizar ya que un talento no utilizado debe salir a la superficie en el mundo real, raras veces, siempre que se produzca una innovación. Un talento de ese tipo será generalmente un subproducto de otros talentos desarrollados bajo otras presiones selectivas. Sin embargo, y por lo general, como lo que esperamos es que la complejidad cognitiva coevolucione con la complejidad del entorno, debemos buscar primero la complejidad cognitiva en aquellas especies que tienen una larga historia de relaciones con el tipo adecuado de complejidad ambiental.

Viéndolos en conjunto, estos aspectos *parecen indicar* que el pensar (nuestro tipo de pensar) tuvo que esperar a que surgiera el habla que, a su vez, tuvo que esperar a que sur-

giera la capacidad de guardar secretos, que a su vez tuvo que esperar a que surgiera la adecuada complejización del entorno conductual. Nos sorprendería descubrir el pensamiento en cualquier especie que no hubiera llegado al final de esta cascada de cribas. Siempre que las opciones conductuales sean relativamente simples (obsérvese la situación apurada del chorlito americano) no hace falta que se dé ninguna representación central, así que con toda probabilidad no se da. El tipo de sensibilidad de orden superior que se requiere para satisfacer las necesidades de un chorlito americano o de una liebre o de una gacela pueden probablemente proporcionarlas unas redes diseñadas casi por completo por mecanismos darwinianos, con la complicidad ocasional aquí y allá de mecanismos skinnerianos. En ese caso, el aprendizaje ABC sería probablemente suficiente para producir esa sensibilidad... aunque este es un asunto empírico que está lejos de quedar cerrado. Será interesante descubrir si hay casos en los que tengamos pruebas claras de un trato diferenciado en *individuos* concretos (por ejemplo, un chorlito americano que no malgasta sus tretas con un perro al que vuelve a identificar, o una liebre que, después de una persecución demasiado apurada incrementa su distancia de amedrentamiento para un zorro concreto). Hasta en estos casos, *podemos* ser capaces de atribuir el aprendizaje a modelos relativamente sencillos: estos animales son criaturas popperianas (criaturas a las que se puede guiar por la experiencia pasada para que rechacen posibilidades de actuación tentadoras pero no probadas) pero siguen sin ser pensantes explícitos.

Siempre que los psicólogos naturales no tengan oportunidad ni obligación de comunicarse unos a otros las atribuciones de intencionalidad que hacen de sí mismos o de otros, siempre que no tengan la oportunidad de comparar apuntes, de debatir con otros, de *preguntar las razones* que fundamentan las conclusiones sobre las cuales tienen curiosidad, no parece existir presión selectiva sobre ellos para que se re-

presenten tales razones y por ello no parece haber presión selectiva sobre ellos para que renuncien al principio de necesidad de saber en favor del opuesto, el principio comando en equipo: désele a cada agente tanto conocimiento del proyecto total como sea posible de manera que el equipo tenga una oportunidad de improvisar adecuadamente cuando se presenten obstáculos inesperados. (Muchas películas como *Los cañones de Navarone* o *Doce del patíbulo* hacen explícito este principio al presentar las hazañas de equipos tan versátiles e instruidos; de ahí el nombre que le doy.)

Los fundamentos latentes que explican la intencionalidad rudimentaria de orden superior de aves y liebres (e incluso de los chimpancés) se ven homenajeados en los diseños de sus sistemas nerviosos, pero nosotros buscamos algo más; buscamos fundamentos que estén *representados* en esos sistemas nerviosos.

Aunque el aprendizaje ABC pueda ofrecer competencias notablemente sutiles y poderosamente discriminatorias, capaces de entresacar las pautas ocultas en voluminosos conjuntos de datos, estas competencias tienden a estar ancladas en los tejidos específicos que se modifican mediante la formación. Son competencias «encarnadas» en el sentido de que son incapaces de verse «transportadas» al momento para abordar otros problemas que afronta el individuo o que comparte con otros individuos. El filósofo Andy Clark y la psicóloga Annette Karmiloff-Smith (1993) han explorado recientemente la transición entre un cerebro que sólo tiene ese conocimiento encarnado en sí y otro cerebro que, como ellos dicen, «se enriquece desde dentro al volver a representar el conocimiento que ya se había representado». Clark y Karmiloff-Smith señalan que así como existen claros beneficios para una política de diseño que «entreteja intrincadamente los diversos aspectos de nuestro conocimiento sobre un campo en una única estructura de conocimiento» también existen costes: «El entretejido hace prácticamente imposible funcio-

nar con las diversas dimensiones de nuestro conocimiento independientemente de las demás, así como explotarlas también independientemente.» Tal conocimiento está tan íntimamente oculto dentro de la maraña de conexiones que «es conocimiento *en* el sistema aunque todavía no es conocimiento *para* el sistema»... a semejanza de la sabiduría revelada en la precoz resolución con la que el recién nacido cuco elimina los huevos competidores echándolos fuera del nido. ¿Qué habría que añadir a la arquitectura computacional del cuco para que fuera capaz de apreciar, de comprender y de explotar la sabiduría entretejida con sus redes neurales?

Una respuesta popular a esta cuestión es, bajo muchos disfraces, la de «¡símbolos!». La respuesta es casi tautológica y por ello ha de ser correcta en *cierta* interpretación. ¿Cómo podría darse que no fuera el caso de que el conocimiento implícito o tácito se convirtiera en explícito expresándose o traducándose en determinado medio de representación «explícita»? Los símbolos, a diferencia de los nudos tejidos en las redes conexionistas, son movibles; pueden manipularse; pueden organizarse en estructuras más amplias en las cuales su contribución al significado del conjunto puede ser una función definida de la estructura de las partes, y generable por ella (la estructura sintáctica). Seguro que esto es correcto en cierta medida pero debemos avanzar con cautela ya que muchos pioneros se han hecho estas mismas preguntas de una manera que ha terminado por llevarles al error.

Noosotros, los seres humanos, tenemos la capacidad de aprendizaje interior y veloz... un aprendizaje que no depende de una enseñanza laboriosa sino que está en nosotros en cuanto contemplamos una representación simbólica adecuada del conocimiento. Cuando los psicólogos inventan un nuevo esquema o paradigma experimental en el cual probar a sujetos no humanos como las ratas, los gatos, los monos o los delfines, suelen tener que dedicar docenas de horas, o in-

cluso cientos, a formar a cada sujeto en las nuevas tareas. Sin embargo, con los seres humanos suele bastar con decirles lo que se espera de ellos. Después de una breve sesión de preguntas y respuestas y de unos pocos minutos de práctica, nosotros los seres humanos seremos generalmente tan competentes en nuestro nuevo medio como pudiera serlo cualquier agente. Por supuesto que tenemos que *comprender* las representaciones que se nos presentan en esas pruebas y ahí es donde la transición del aprendizaje ABC a nuestro tipo de aprendizaje se pierde todavía en la niebla. Una introspección que puede ayudar a aclararlo es la familiar máxima de la construcción de aparatos: si «lo hace usted mismo» lo entenderá. Para anclar un fundamento a un agente de manera fuerte, de tal modo que se convierta *en la propia razón del agente*, éste debe «hacer» algo. Debe componerse, diseñarse, editarse, revisarse, manipularse, dotarse de una representación de esa razón. ¿Y cómo llega un agente a ser capaz de hacer semejante cosa maravillosa? ¿Tiene que crecerle un nuevo órgano en el cerebro? ¿O puede formar su habilidad a partir de las diversas manipulaciones del mundo exterior que ya tiene completamente dominadas?

Fabricar cosas con las que pensar

Al igual que no se puede practicar mucho la carpintería sin más herramientas que las manos, no se puede pensar mucho con tan sólo el cerebro.

Bo Dahlbom y Lars-Erik Janlert,
Computer Future [Futuro de ordenador], 1997.

Todo agente se enfrenta a la tarea de hacer el mejor uso de su entorno. El entorno contiene una diversidad de bienes y de toxinas, entremezclados y con un confuso conjunto de

pistas más indirectas: presagios y distractores, piedras en las que pisar y trampas. Estos recursos suelen equivaler a una confusión de riquezas que compiten por captar la atención del agente; la tarea de manejar los recursos (y de refinar ese manejo) se convierte por tanto en una tarea del agente en la que el tiempo es una dimensión crucial. El tiempo empleado en la fútil persecución de una presa o en reforzarse para resistir amenazas ilusorias es tiempo desperdiciado, y el tiempo es precioso.

Como se indica en la figura 4.4, las criaturas gregorianas cogen del ambiente diversas entidades diseñadas y las utilizan para mejorar la eficiencia y la precisión de las pruebas de sus hipótesis y de su toma de decisión, pero tal como está el diagrama es equívoco. ¿Cuánto espacio queda en el cerebro para tales artefactos y cómo se instalan en él? ¿Es que el cerebro de una criatura gregoriana es mucho más capaz que el de otras criaturas? Nuestros cerebros son modestamente mayores que los de nuestros parientes más próximos (aunque no mayores que los cerebros de algunos delfines y ballenas) pero no es esa ciertamente la fuente de nuestra inteligencia. Quiero sugerir que la fuente primaria es nuestra costumbre de *descargar* el máximo posible de nuestras tareas cognitivas en el propio entorno, expulsando cosas de nuestra mente (es decir, de nuestros proyectos y actividades mentales) hacia el mundo circundante en el que un montón de dispositivos periféricos que construimos pueden almacenar, procesar y volver a representar nuestros significados, simplificando, realizando y protegiendo los procesos de transformación que *son* nuestro pensar. Esta práctica tan extendida de descargar nos libera de las limitaciones de nuestros cerebros animales.

Un agente afronta su entorno con su actual repertorio de habilidades, perceptivas y conductuales. Si el entorno es demasiado complicado para abordarlo con estas habilidades, el agente se ve en dificultades a menos que pueda desarrollar

nuevas habilidades o simplificar su entorno. O hacer ambas cosas. La mayor parte de las especies se apoya en los hitos naturales para encontrar el camino y algunas especies han añadido la treta de colocar hitos en el mundo para su uso posterior. Las hormigas, por ejemplo, dejan rastros de feromonas (rastros de olor) que les llevan del nido a su alimento y a la inversa, y los individuos de muchas especies territoriales señalan las fronteras de sus territorios con compuestos aromáticos idiosincráticos que llevan en la orina. Marcar la tierra de este modo avisa a los intrusos pero también proporciona un cómodo dispositivo que se puede usar. Nos ahorra la necesidad de recordar de una u otra manera la frontera de esa parte del entorno en la que hemos invertido esfuerzos significativos de refinamiento de recursos... e incluso que hemos cultivado. Conforme nos aproximamos a la frontera podemos olerla. Dejamos que el mundo exterior almacene cierta información fácilmente transducida acerca de dónde se encuentran los puntos importantes de la naturaleza, de tal manera que podamos dedicar nuestro limitado cerebro a otras cuestiones. Es un buen modo de administrar las cosas. Colocar deliberadamente marcas en el entorno para usarlas en distinguir lo que son sus rasgos más importantes es una manera excelente de reducir la carga cognitiva de nuestra percepción y de nuestra memoria. Es una variación sobre el tema de la buena táctica de la evolución de instalar balizas allí donde más se necesitan, al tiempo que es una mejora de esa misma táctica.

Para nosotros, los seres humanos, los beneficios de etiquetar las cosas de nuestro entorno son tan evidentes que tendemos a pasar por alto el fundamento del etiquetado y las condiciones en las cuales funciona. ¿A cuento de qué etiqueta alguien y qué supone ese etiquetar algo? Supongamos que estamos buscando en miles de cajas de zapatos la llave de una casa, llave que creemos haber escondido en una de esas cajas. A menos que seamos idiotas o que estemos tan

frenéticos en nuestra búsqueda que ni podamos pararnos a pensar cuál es el mejor camino que podemos seguir, dispondremos algún sistema para que el entorno nos ayude a resolver el problema. En concreto, queremos ahorrar tiempo no teniendo que mirar más de una vez en cada caja. Uno de los sistemas sería coger de un montón (el montón sin revisar) las cajas, una por una, y colocarlas en otro montón (el montón revisado). Otro (potencialmente más eficiente desde el punto de vista de la energía) es colocar una marca en cada caja que se revisa y adoptar luego la norma de no mirar en las cajas que ya tengan una marca. Una marca de control simplifica el mundo al proporcionarnos una sencilla tarea perceptiva para sustituir a otra de reconocimiento y memorización mucho más difícil (y quizá imposible). Démonos cuenta de que si las cajas están todas colocadas en una fila y no tenemos que preocuparnos de que la fila se reordene sola, sin que nos apercibamos, ni siquiera hace falta hacerles una marca; podemos ir de izquierda a derecha utilizando un sistema de distinguir con el que ya nos ha provisto la naturaleza: la distinción entre izquierda y derecha.

Concentrémonos ahora en la propia marca de control. ¿Valdrá *cualquier cosa* como marca de control? Claramente, no. «Haré un borroncito suave en alguna parte de cada caja que revise.» «Aplastaré la esquina de cada caja al revisarla.» No son buenas opciones habida cuenta de la gran probabilidad de que algo, anteriormente, haya dejado marcas de ese tipo en alguna caja. Necesitamos algo distintivo, algo de lo que sepamos fiablemente que es el resultado de nuestro etiquetado y no de un deterioro ajeno al asunto. Asimismo debería ser recordable, por supuesto, de modo que no nos indujera a confusión sobre el hecho de si la destacada etiqueta que nos encontramos es la que hemos puesto *nosotros* y, si es así, qué regla queríamos seguir cuando la pusimos. No vale que nos atemos una cuerda a un dedo para que nos lo recuerde si, cuando *nos llama la atención más adelante* (y

por lo mismo cumpliendo su papel de baliza de autocontrol puesta en el entorno) no nos acordamos de por qué nos la atamos. Esas marcas deliberadas y sencillas en el mundo son las primitivas precursoras de la escritura, un paso dado hacia la creación en el mundo externo de sistemas de almacenamiento de información periféricos. Nótese que esta innovación no depende de que haya un lenguaje sistemático en el cual se escriban esas etiquetas. Cualquier sistema inventado para la ocasión puede servir siempre que se pueda recordar mientras se está usando.

¿Qué especies han descubierto estas estrategias? Algunos experimentos recientes nos proporcionan un atisbo seductor, aunque no concluyente, de las posibilidades existentes. Los pájaros que hacen escondrijos en los que esconder alimento en muchos lugares concretos tienen un amplio éxito al volver a hallar sus almacenes secretos después de mucho tiempo. Los cascanueces de Clark *, por ejemplo, los han estudiado experimentalmente el biólogo Russell Balda y sus colaboradores en un entorno cerrado de laboratorio: una gran habitación bien con un suelo sucio o bien con un suelo provisto de muchos agujeros rellenos de arena y otros más provistos de diferentes marcas. Los pájaros pueden hacer más de una docena de escondrijos con las semillas que se les dan y vuelven al cabo de varios días a recogerlas. Tienen mucho éxito fiándose de múltiples pistas y descubren la mayoría de sus escondrijos incluso cuando los experimentadores cambian o quitan algunas de las marcas. Pero en el laboratorio cometen errores y la mayoría de ellos parecen errores de autocontrol: gastan tiempo y energía en revisar sitios que ya habían dejado limpios en incursiones anteriores. Como estos pájaros pueden hacer miles de escondrijos en su entorno natural y visitarlos al cabo de períodos que pueden durar más de seis meses, la frecuencia de

* Se trata de la especie *Nucifraga columbiana*. (N. del T.)

esas visitas repetidas e inútiles es casi imposible de calcular pero es razonable pensar que repetir así las visitas es una costumbre costosa que otras especies de pájaros, como los paros, son capaces de evitar.

En su entorno natural, se observa que los cascanueces de Clark comen las semillas allí donde las desentierran, dejando tras de sí un montón de restos de la merendola que podría recordarles, la siguiente vez que volaran por allí, que ya han abierto esa caja de zapatos concreta. Balda y sus colaboradores diseñaron experimentos para comprobar la hipótesis de que los pájaros se basaban en tales marcas para evitar repetir sus visitas. En uno de los casos, se borraron con todo cuidado entre visita y visita las modificaciones dejadas por los pájaros, mientras que en otro se dejaron tal cual. Sin embargo, en el entorno del laboratorio los pájaros no lo hicieron significativamente mejor cuando se les dejaron las marcas, de modo que no está probado que los pájaros se apoyen en tales pistas. Puede que en la naturaleza no puedan hacerlo habida cuenta de que esas pistas las eliminan, en todo caso y frecuentemente, los factores meteorológicos, como Balda indica. También señala que los experimentos hasta el momento no son concluyentes; el coste del error en el laboratorio es muy leve: unos pocos segundos malgastados en la vida de un pájaro bien alimentado.

También es posible que colocar los pájaros en el entorno de laboratorio inadvertidamente los vuelva incompetentes, ya que su costumbre cotidiana de otorgar parte de la tarea de autocontrol al entorno puede depender de otras pistas que inadvertidamente estén ausentes en el entorno de laboratorio. Se observa corrientemente (¡aunque no tanto!) que las personas mayores a las que se lleva de su casa a un hospital quedan en manifiesta desventaja, incluso aunque sus necesidades corporales estén bien atendidas. Suelen *parecer* un poco trastornados: ser absolutamente incapaces de alimentarse, vestirse y asearse, y mucho menos embarcarse en otras actividades de

mayor interés. Sin embargo, suele ocurrir que si regresan a sus casas, pueden manejarse solos bastante bien. ¿Cómo es eso? A lo largo de los años, han cargado su entorno hogareño con marcas muy familiares, disparadores de costumbres, recordatorios de qué deben hacer, de dónde encontrar el alimento, de cómo vestirse, de dónde está el teléfono, y así sucesivamente. Una persona mayor puede ser un auténtico virtuoso en ayudarse a sí mismo en esa zona del mundo tan archisabida, a pesar de su creciente incapacidad para recibir nuevos empujones de aprendizaje, sean del tipo ABC o de otro cualquiera. Sacarlos de sus casas es, literalmente, separarlos de una parte enorme de sus mentes, cosa tan potencialmente devastadora como pasar por una ablación en el cerebro.

Puede que, sin pensar, algunos pájaros hagan marcas de control como subproducto de otras actividades suyas. Los seres humanos desde luego nos apoyamos en muchas marcas de control colocadas inadvertidamente en nuestros alrededores. Escogemos costumbres beneficiosas que sabemos apreciar sin detenernos a comprender por qué significan semejante tesoro. Pensemos en intentar multiplicar números de varias cifras mentalmente. ¿Cuánto es 217 multiplicado por 436? A nadie se le ocurriría intentar responder sin la ayuda de lápiz y papel, salvo como demostración de habilidad. La cuenta en el papel tiene más de una función útil; proporciona un almacén fiable de los resultados intermedios, pero los símbolos individuales sirven asimismo como marcas que pueden seguirse, recordándonos, conforme nuestros ojos y nuestros dedos van avanzando, cuál debe ser el paso siguiente de la receta. (Si se duda de esta segunda contribución, trátase de hacer esa multiplicación escribiendo cada resultado parcial en distintas hojas de papel colocadas de cualquier manera en lugar de alinear las cifras de la manera acostumbrada.) Nosotros, criaturas gregorianas, somos los beneficiarios, literalmente, de miles de tecnologías útiles de ese tipo, inventadas por otros en borrosos rincones recóndi-

tos de la prehistoria o de la historia pero transmitidas mediante autopistas culturales, y mediante los caminos genéticos de la herencia. Gracias a nuestra herencia cultural, aprendemos cómo extender por el mundo nuestra mente, un mundo en el que podemos utilizar de la mejor manera posible nuestras habilidades innatas de rastreo y de reconocimiento de pautas tan bellamente diseñadas.

Hacer un cambio tal en el mundo, no significa sólo hacer una descarga de memoria. Puede permitir al agente dar a luz algún talento cognitivo que de otro modo habría quedado infrutilizado, preparando para ese talento materiales especiales... cosa que en el peor de los casos se hará sin premeditación alguna. El científico en robótica Philippe Gaussier (1994) ha proporcionado recientemente una vívida ilustración de esta posibilidad, usando diminutos robots que primero alteran su entorno y que luego ven alterado su propio repertorio conductual por el entorno que ellos mismos han modificado. Estos robots son vehículos de Braitenberg en el mundo real, a los que su creador, el científico en robótica Francesco Mondada, llama *kheperas* (la palabra italiana que designa a los escarabajos peloteros). Son un poco más pequeños que los discos de hockey sobre hielo y avanzan rodando sobre dos ruedecillas y una guía. Los robots tienen sistemas visuales extremadamente rudimentarios (dos o tres fotocélulas) conectados a sus ruedas de tal modo que las señales que envían impiden a los robots colisionar con las paredes que rodean su mundo (la superficie de una mesa), haciéndolos que se desvíen. De manera que estos robots están equipados de manera innata, podríamos decir, con un sistema de guía visual para evitar paredes. Por la mesa hay distribuidos pequeños trozos de madera (en forma de cilindro), que pueden moverse, y los sistemas innatos de visión de los robots los hace rodear también estos obstáculos ligeros, pero generalmente los ganchos que llevan los robots a la espalda desplazan los cilindros conforme circulan los robots. Éstos

se desplazan al azar por toda la mesa, enganchando cilindros sin querer y depositándolos después donde quiera que se desvíen bruscamente. (Véase la figura 5.2.) Con el tiempo, estos enganchones distribuyen de otra forma los cilindros y cuando se da que dos o más cilindros están juntos, forman un grupo tal que los robots lo «malinterpreta» a partir de ese momento como un trozo de pared... que hay que evitar. En breve y sin instrucción alguna de ningún cuartel general central los robots alinean todos los cilindros que estaban distribuidos por su entorno, organizándolos en series de paredes interconectadas. Los paseos al azar de los *kheperas* en un entorno inicialmente hecho al azar estructuran primero ese entorno en algo parecido a un laberinto y luego usan esa estructura para conformar su propia conducta: se convierten en robots seguidores de paredes.

Este es un caso de lo más simple que pueda imaginarse de una táctica que incluye, en el extremo complejo del espectro, dibujos de diagramas y construcción de modelos. Por ejemplo, ¿por qué dibujamos diagramas... en una pizarra o (en otras épocas) en el suelo de una caverna con un palo afi-

Los robots de Philippe Gaussier

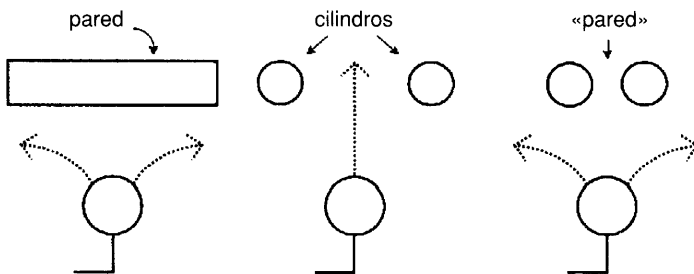


Figura 5.2

lado? Lo hacemos porque volviendo a representar la información en otro formato, la hacemos presentable para una u otra aptitud perceptiva con un propósito determinado.

Las criaturas popperianas (y su subvariedad, las criaturas gregorianas) viven en un entorno que puede ser aproximadamente dividido en dos partes: la «externa» y la «interna». Los habitantes del entorno «interno» no se distinguen tanto por el lado de la piel en el que se encuentran (como B. F. Skinner ha señalado [1964, pág. 84] «La piel no es tan importante como frontera») como por el hecho de ser o no portátiles, y por ello casi omnipresentes, y por ello relativamente más controlables y más conocidos, y por ello con mayores probabilidades de ser escogidos para beneficio del agente. (Como ya hemos señalado en el capítulo 2, la lista de la compra en la tira de papel consigue su significado exactamente de la misma forma que una lista de la compra memorizada en el cerebro.) El entorno «externo» cambia de maneras que a veces son muy difíciles de rastrear y, en general, está geográficamente fuera de la criatura. (Los límites de la geografía para trazar esta distinción nunca están mejor ilustrados que en el caso de los antígenos, malignos invasores del exterior, y de los anticuerpos, leales defensores del interior, los cuales se mezclan con fuerzas amigas [como las bacterias de nuestros intestinos, sin cuya participación moriríamos] y con mirones insignificantes, formando multitudes de agentes microbianos que pueblan nuestro espacio corporal.) El conocimiento portátil sobre el mundo de una criatura popperiana tiene que abarcar una pizca de conocimiento (un saber cómo) sobre esa parte omnipresente del mundo que es *ella misma*. Por supuesto que tiene que saber cuáles son sus miembros y qué boca hay que alimentar pero también tiene que saber por dónde anda, y saberlo en su cerebro, hasta cierto punto. ¿Y cómo lo consigue? Utilizando los mismos métodos de toda la vida: ¡situando marcas y etiquetas donde quiera que le resulte práctico! Entre los recursos que

un agente tiene que manejar contra el reloj se encuentran los recursos de su propio sistema nervioso. No hace falta que este conocimiento de sí mismo esté representado explícitamente, como no hace ninguna falta que la sabiduría de una criatura no pensante esté representada explícitamente. Puede tratarse de un mero saber cómo hacer las cosas encarnado en ella, pero es un saber crucial sobre cómo manipular esa parte del mundo curiosamente dócil y relativamente permanente que es uno mismo.

Lo que deseamos es que tales refinamientos de nuestros recursos internos nos simplifiquen la vida de manera que podamos hacer las cosas mejor y más deprisa (el tiempo *siempre* es precioso) con nuestro repertorio disponible de talentos. Una vez más, no sirve de nada crear un símbolo *interno* como herramienta para utilizar en el control de uno mismo si cuando «nos llama la atención» no podemos recordar para qué lo hemos creado. La manipulabilidad de cualquier sistema de señaladores, marcas, etiquetas, símbolos y demás recordatorios depende de la robustez subyacente de las habilidades espontáneas de rastreo y de reidentificación, que nos proporcionan con distintos caminos, redundantes y multimodales, de acceso a nuestras herramientas. Las técnicas de manejo de recursos con las que hemos nacido no hacen distinciones entre las cosas exteriores e interiores. En las criaturas gregorianas como nosotros, las representaciones de rasgos y objetos del mundo (externo o interno) se convierten en objetos por derecho propio: cosas que se pueden manipular, rastrear, mover, atesorar, alinear, estudiar, volver del revés y, además, ajustar y explotar.

En su libro *Sobre la fotografía*, 1982 la crítica literaria Susan Sontag señala que la llegada de la fotografía de alta velocidad fue un revolucionario avance tecnológico para la ciencia porque, por primera vez, permitió a los seres humanos examinar los complicados fenómenos temporales no en tiempo real sino *en su propio tiempo, a su ritmo...* a placer,

metódicamente, analizando, volviendo una y otra vez a los rastros creados por esos complicados sucesos. Como se ha indicado en el capítulo 3 nuestras mentes naturales están preparadas para afrontar los cambios que se den solamente a ritmos determinados. Los sucesos que se den a mayor o a menor velocidad son sencillamente invisibles para nosotros. La fotografía fue un avance tecnológico que llevaba en su estela un inmenso empujón de poder cognitivo al permitirnos volver a representarnos los sucesos del mundo que nos interesaban en un formato, y a un ritmo, que estaba hecho a la medida de nuestros sentidos concretos.

Antes de que hubiera cámaras y películas de alta velocidad, existían muchos dispositivos de registro y observación que permitían al científico extraer datos con precisión del mundo para un tranquilo análisis posterior. Los exquisitos diagramas e ilustraciones de diversos siglos de ciencia son el testimonio del poder de estos métodos; pero en una cámara fotográfica hay algo especial: es «estúpida». Para poder «capturar» los datos representados en sus fotografías, no tiene que entender a su objeto como debe comprenderlo un artista o un dibujante humano. Por ello nos entrega una versión de la realidad sin editar, sin contaminar, sin sesgar, pero aun así vuelta a representar y lista para las facultades que están preparadas para analizar y, en último extremo, para comprender los fenómenos. Este cartografiado carente de mente de datos complejos mediante formatos más sencillos, más naturales o más cercanos al usuario, como hemos visto, es un sello distintivo de una inteligencia creciente.

Pero junto con la cámara y con el enorme montón de instantáneas que salen de ella, se plantea un problema de recursos: hay que etiquetar las propias fotografías. Poco bien nos hace captar un suceso de interés en una instantánea si luego no podemos recordar cuál de las miles de ellas que hay por toda la oficina es la que representa el suceso de nuestro interés. Este «problema de emparejamiento» no se presenta

en variantes más sencillas y más directas del rastreo, como hemos visto, pero generalmente hay que abordar el coste de resolverlo; la treta puede financiarse a sí misma (el tiempo es oro) en aquellos casos en los que permita el rastreo indirecto de cosas que no pueden rastrearse directamente. Piénsese en la brillante práctica de pinchar alfileres de colores en un mapa para señalar la localización de un gran número de sucesos que estamos intentando comprender. Puede diagnosticarse una epidemia viendo (*viendo*, gracias a los códigos de color) que todos los casos de cierto tipo se alinean en el mapa junto con algún otro rasgo inconspicuo, o incluso hasta ese momento no representado: el suministro de aguas, o el sistema de evacuación de aguas fecales, o puede que la ruta del cartero. La base secreta de un asesinato en serie puede a veces localizarse (en una especie de «casataxia») dibujando el centro geográfico del cúmulo de sus ataques. La drástica mejora de todos nuestros tipos de investigación, desde las estrategias de forrajeo de nuestra época de cazadores-recolectores hasta las investigaciones actuales de la policía, de los críticos de poesía y de los físicos se deben generalmente al explosivo crecimiento de nuestras tecnologías de re-representación.

En nuestros cerebros guardamos «señaladores» e «índices» dejando el máximo posible de los datos al mundo exterior, en nuestras agendas, en nuestras bibliotecas, en nuestros cuadernos de notas y en nuestros ordenadores y, naturalmente, en nuestro círculo de amigos y de colaboradores. Una mente humana no sólo no se limita al cerebro sino que más bien se vería gravemente minusválida si se le quitaran esas herramientas exteriores, tan minusválida por lo menos como un miope al que se le quitaran las gafas. Cuantos más datos y dispositivos nos quitamos de encima, más dependientes nos volvemos de esos periféricos: sin embargo, cuanto más nos familiarizamos con los objetos periféricos gracias a nuestra práctica en manipularlos, con más con-

fianza podemos desenvolvemos sin ellos, volviendo a abordar los problemas con la cabeza, resolviéndolos con nuestra imaginación disciplinada por esa práctica externa. (¿Podemos de cabeza poner por orden alfabético las palabras de esta frase?)

Una fuente particularmente rica en nuevas técnicas de representación es la costumbre que nosotros, y sólo nosotros, hemos desarrollado de cartografiar deliberadamente nuestros nuevos problemas sobre la maquinaria que servía para resolver los problemas anteriores. Por ejemplo, consideremos los muchos métodos diferentes que hemos desarrollado para pensar en el tiempo al pensar en el espacio (Jaynes, 1976). Tenemos todo tipo de convenciones para cartografiar el pasado, el presente y el futuro, el antes y el después, el más pronto y el más tarde (diferencias que son prácticamente invisibles en la naturaleza sin refinar) sobre las ideas de izquierda y derecha, arriba y abajo, en el sentido de las agujas del reloj y en sentido contrario. El lunes está a la izquierda del martes para la mayoría de nosotros mientras que (en una valiosa convención, que se va desvaneciendo de nuestra cultura, triste es decirlo) las cuatro en punto están debajo de las tres y a mano derecha, cada día o cada noche. No se detiene ahí nuestra espacialización del tiempo. Concretamente, en la ciencia se extiende a los gráficos, que se han convertido hoy día en un sistema familiar de diagramas para casi toda la gente alfabetizada. (Piénsese en los beneficios, o en la temperatura, o en el volumen de nuestra cadena de música, subiendo y subiendo y subiendo desde la izquierda hacia la derecha con el transcurso del tiempo.) Utilizamos nuestro sentido del espacio para *ver* el paso del tiempo (generalmente de izquierda a derecha, según la convención estándar, salvo en los diagramas evolutivos en los que las eras más antiguas se muestran en la parte baja mientras que el *hoy* aparece en lo alto). Como muestran estos ejemplos (la ausencia de figuras en este punto del texto es

completamente deliberada) nuestra capacidad para *imaginar* estos diagramas cuando se nos invita verbalmente a hacerlo es, en sí misma, una valiosa aptitud gregoriana, con muchos usos. Nuestra capacidad para imaginar estos diagramas es parasitaria de nuestra capacidad de dibujarlos y verlos, descargándolos al menos provisionalmente en el mundo exterior.

Gracias a nuestras imaginaciones realizadas mediante prótesis, podemos formular posibilidades que de otro modo serían posibilidades metafísicas imponderables e inadvertidas, como el caso de Amy, la moneda de la suerte que se vio al final del capítulo 4. Necesitamos ser capaces de imaginar la trayectoria, por lo demás invisible, que une a la genuina Amy de ayer con uno de los peniques de aspecto semejante del montón: necesitamos dibujarla en «nuestra imaginación». Sin esas ayudas visuales, externas o internas, tendríamos muchas dificultades en seguir estas observaciones metafísicas, y muchas más dificultades en aportarles algo. (¿Significa eso que alguien que haya nacido ciego no podría participar en esas discusiones metafísicas? No, porque los ciegos desarrollan sus propios métodos de imaginar espacialmente, preocupándose, lo mismo que la imaginación de las personas que ven, de mantener el contacto con las cosas que se mueven en el espacio, de diferente manera. Aunque una cuestión interesante es qué diferencias hay, si es que hay alguna, entre los estilos de pensamiento abstracto que adoptan los que han nacido ciegos o sordos.) Armados con esas herramientas mentales, tendemos a olvidar que *nuestras* maneras de pensar respecto al mundo no son las únicas y que en concreto no son prerequisites para enfrentarse con éxito al mundo. Probablemente parece evidente en un principio que como son tan manifiestamente inteligentes, los perros y los delfines y los murciélagos deben tener conceptos más o menos como los nuestros, pero una vez que reflexionemos no debería parecernos tan evidente en absoluto. Muchas de

las preguntas que hemos suscitado desde nuestra perspectiva evolutiva sobre la ontología y la epistemología de otras criaturas no han quedado respondidas todavía y las respuestas sin duda serán sorprendentes. Sólo hemos dado el primer paso: hemos visto algunas posibilidades de investigación que anteriormente habíamos pasado por alto.

De todas las herramientas mentales que adquirimos en el curso del amueblamiento de nuestros cerebros a partir de las reservas de la cultura, no hay ninguna más importante, por supuesto, que las palabras... habladas primero, escritas después. Las palabras nos hacen más inteligentes al facilitarnos la cognición de la misma manera (pero muchas veces multiplicada) que las balizas y las marcas del territorio facilitan la circulación por el mundo a las criaturas más sencillas. La circulación en el mundo abstracto multidimensional de las ideas es sencillamente imposible sin una inmensa reserva de marcas móviles y memorizables que puedan compartirse, criticarse, registrarse y verse desde diferentes perspectivas. Es importante recordar que hablar y escribir son dos innovaciones completamente diferentes, separadas por muchos cientos de miles (puede que millones) de años y que cada una de ellas presenta su conjunto diferente de posibilidades. Tendemos a unir los dos fenómenos sobre todo al teorizar respecto al cerebro y la mente. La mayor parte de lo que se ha escrito acerca de las posibilidades de un «lenguaje del pensamiento» como medio de operaciones cognitivas presupone que estamos pensando en un lenguaje *escrito* del pensamiento... «escritura cerebral y lectura mental» como dije hace unos años. Podríamos tener una mejor perspectiva de cómo la llegada del lenguaje podría magnificar nuestros poderes cognitivos si en vez de eso nos concentráramos en por qué y cómo podría funcionar adecuadamente un lenguaje *hablado* del pensamiento (descendiente de nuestro lenguaje público y natural).

Hablando con nosotros mismos

Si la mente infantil todavía no formada ha de convertirse en mente inteligente, debe adquirir tanto disciplina como iniciativa.

Alan Turing.

No hay paso más grande, más explosivo ni más trascendental en la historia del diseño de la mente que la invención del lenguaje. Cuando el *Homo sapiens* se convirtió en el beneficiario de esta invención, la especie se colocó en una catapulta que la ha impulsado muchísimo más lejos que a ninguna de las demás especies de la tierra en su capacidad de mirar hacia delante y de reflexionar. Lo cierto para la especie es igualmente cierto para el individuo. No hay transición más capacitante en la vida de una persona que la de «aprender» a hablar. Debo poner la palabra entre comillas porque hemos llegado a darnos cuenta (gracias a la labor de psicólogos y lingüistas) de que los niños humanos están genéticamente prediseñados de muchos modos diferentes para el lenguaje. Como Noam Chomsky, el padre de la lingüística moderna, suele decir (y con una exageración excusable) los pájaros no tienen que aprender sus plumas y los niños no tienen que aprender su lenguaje. Una enorme parte del trabajo de diseño de un usuario de lenguaje (o de plumas, según el caso) se realizó hace ya muchísimo tiempo y está en el niño en forma de talentos y disposiciones innatos, fácilmente adaptables a las condiciones concretas de la gramática y el vocabulario. Los niños adquieren el lenguaje a una velocidad de vértigo, aprendiendo palabras nuevas a una media de una docena al día, durante muchos años, hasta que son adolescentes, en que la velocidad de adquisición queda muy reducida. Antes de llegar al colegio dominan todos los aspectos gramaticales, a excepción de los más sutiles. Además de toda su interacción lingüística con los miembros de

su familia (y mascotas), los niños pasan muchas horas vocalizando para sí desde que empiezan a gatear, primero balbuceando, luego disfrutando de las maravillosas mezclas de palabras y de sílabas sin sentido ricamente dotadas de diferentes entonaciones (exhortatoria, tranquilizadora, explicativa, zalamera) y terminan por llegar a una expresión propia de tipo complejo.

A los niños les encanta hablar solos. ¿Qué influencia tendrá esto en su mente? Todavía no sé contestar a esa pregunta, pero tengo algunas conjeturas como sugerencias para futuras investigaciones. Pensemos qué sucede en los inicios de la vida lingüística de cualquier niño. «¡Quema! ¡Está caliente! ¡No toques la estufa!», dice mamá. En esa etapa, el niño no tienen necesidad de saber qué significa «caliente» o «tocar» o «estufa»: esas palabras no son, *en principio*, más que sonidos, sucesos auditivos que tienen cierto poder evocador, cierta familiaridad, algo que se le queda al niño y que despierta ciertos ecos. Sirven para evocar un tipo de situación (acercarse a la estufa y alejarse de ella) que no es sólo una situación en la que se *escucha* una prohibición sino también una situación en la que se encuentra una repetición auditiva que se puede remedar. Simplificando groseramente, suponemos que el niño adquiere la costumbre de decirse (en voz alta) «¡Caliente! ¡No toques!» sin mucha idea de lo que significan las palabras, pronunciándolas sencillamente como una parte que acompaña a la situación familiar de acercarse y alejarse de la estufa, además de tenerlas como una especie de mantra que podría emplearse en cualquier otra ocasión. Después de todo, a los niños les domina el hábito de repetir las palabras que acaban de oír: repetirlas dentro y fuera del contexto y construir con ellas vínculos de reconocimiento y caminos asociativos entre las propiedades auditivas y las propiedades sensoriales, los estados internos y demás que les son concomitantes.

Este es un esbozo esquemático del tipo de proceso que

debe de producirse. Este proceso podría tener el efecto de poner en marcha lo que podríamos llamar *expresión propia a medio entender*. El niño, empujado inicialmente por ciertas asociaciones auditivas insistentes provenientes de las admoniciones de sus padres, adquiere la costumbre de añadir a sus actividades una banda sonora, a «comentarlas». El hecho de pronunciar esa banda sonora sería en un principio una especie de «blablabla» (una cháchara sin sentido compuesta de cosas que suenan como palabras) mezclado con palabras reales pronunciadas con mucho sentimiento pero con escasa o nula comprensión de su significado, así como unas pocas palabras comprendidas. Habría una imitación de la exhortación, de la prohibición, de la alabanza, de la descripción y todas estas imitaciones terminarían por madurar en exhortaciones, prohibiciones, alabanzas y descripciones auténticas. Pero la costumbre de añadir «etiquetas» sería anterior a la comprensión en sí de los textos de las etiquetas (incluso anterior a la comprensión parcial de las mismas).

Lo que sugiero es que estas prácticas en un principio «estúpidas» (la mera pronunciación de esas etiquetas, en circunstancias adecuadas o inadecuadas) se transformarían en un plazo muy corto en la costumbre de representar los estados y las actividad de uno mismo de una manera nueva. Conforme el niño estableciera más asociaciones entre los procesos auditivos y articulatorios, por una parte, y las pautas de actividad concomitante a ellos, por otra, se irían creando nodos prominentes en la memoria. Una palabra podría convertirse en familiar aunque no se comprendiera. Y son esas anclas de familiaridad las que podrían dar a una etiqueta una identidad independiente dentro del sistema. Sin esa independencia, las etiquetas son invisibles. Para que una palabra sirva de etiqueta útil y manipulable en el refinado de los recursos del cerebro, debe realzar con prontitud las asociaciones buscadas que ya están hasta cierto punto establecidas en el sistema. Aparte de ello, las palabras pueden ser ar-

bitrarias y su arbitrariedad forma parte de aquello que las hace distintivas: el riesgo de no darse cuenta de la presencia de la etiqueta es pequeño; no se mezcla con su entorno, como podía ocurrir con la mella en la esquina de la caja de zapatos. Por así decir, se saca de la manga su creación deliberada, a modo de marca de fábrica.

La costumbre de esa expresión propia a medio entender podría, según indico, ser el origen de la práctica del etiquetado consciente, con palabras (o con balbuceos, o con el neologismo personal que se quiera) lo que a su vez podría llevar a una práctica todavía más eficiente: dejar a un lado todas las asociaciones auditivas y articuladas, o la mayoría de ellas, y apoyarse tan sólo en el *resto* de las asociaciones (y posibilidades de asociación) para realizar el anclaje. Lo que sugiero es que el niño puede abandonar la pronunciación en voz alta y crear neologismos personales, no pronunciados, como etiquetas de los rasgos de sus propias actividades.

Podemos tomar un objeto lingüístico como *objeto encontrado* (incluso si en cierto modo metemos la pata al inventarlo nosotros en lugar de oírsele a alguien) y guardarlo para examinarlo más adelante, en privado. Nuestra capacidad de hacerlo depende de la capacidad que demostremos para volver a identificar o a reconocer en distintas ocasiones una etiqueta de ese tipo que tenga uno o varios rasgos para poder recordarla, una apariencia independiente de su significado. Una vez que hemos creado etiquetas y hemos adquirido la costumbre de pegarlas a las circunstancias que experimentamos, hemos creado una nueva clase de objetos que por sí mismos pueden convertirse en objetos de toda la maquinaria de reconocimiento de pautas, de construcción de asociaciones, y así sucesivamente. A modo de científicos que se entretuvieran retrospectivamente en examinar sin prisas las fotografías tomadas en el calor de una batalla experimental, nosotros podemos reflexionar sobre cualesquiera pautas que

haya que distinguir en los diversos objetos etiquetados que dragamos de nuestra memoria.

Conforme vamos mejorando, nuestras etiquetas se van haciendo más refinadas, más perspicaces e incluso mejor articuladas y se llega finalmente a un punto en el que nos acercamos a la proeza casi mágica con la cual empezamos: la *mera contemplación* de una representación es suficiente para recordar todas las lecciones pertinentes. Nos hemos convertido en *entendedores* de los objetos que hemos creado. A estos nodos artefácticos de nuestras memorias, a estas pálidas sombras de las palabras oídas y pronunciadas, podríamos llamarlos *conceptos*. Por tanto, un concepto es una etiqueta interna que entre sus muchas asociaciones puede incluir o no los rasgos auditivos y articulatorios de una palabra (pública o privada). Pero sugiero que las palabras son los prototipos o antepasados de los conceptos. Los primeros conceptos que podemos manipular, eso es lo que estoy sugiriendo, son conceptos «vocalizados» y sólo los conceptos que pueden manipularse pueden convertirse para nosotros en objeto de escrutinio.

Platón en el *Teeteto* compara la memoria humana con una inmensa jaula de pájaros:

SÓCRATES: Considera ahora si el conocimiento es algo que puedes poseer de esa manera sin tenerlo a tu alrededor, a semejanza de un hombre que ha capturado algunos pájaros silvestres, pichones y demás, y los guarda en un aviario en su casa. En cierto sentido, por supuesto, podríamos decir que los «tiene» en todo momento en tanto que los posee, ¿no es así?

TEETETO: Sí.

SÓCRATES: Pero en otro cierto sentido «no tiene» a ninguno de ellos, aunque los controle ahora que los ha hecho cau-

tivos en un recinto que le pertenece; puede cogerlos y tenerlos cuando le parezca cogiendo el pájaro que desee y dejándolo ir otra vez; y queda a su albedrío hacerlo tantas veces como desee (197 c-d).

El truco está en hacer que acuda el pájaro adecuado cuando lo necesitamos. ¿Cómo hacemos eso? Por medio de la tecnología. Construimos complejos sistemas de asociación mnemotécnica (señaladores, etiquetas, rampas y escaleras, ganchos y cadenas). Refinamos nuestros recursos mediante ensayos y componendas incesantes, transformando nuestro cerebro (y todos los accesorios periféricos que adquirimos) en una red de habilidades inmensamente estructurada. Todavía no se han encontrado evidencias que demuestren que haya ningún otro animal que haga lo mismo.

CAPÍTULO 6

NUESTRA MENTE Y OTRAS MENTES

Una vez que el niño aprende el significado de «por qué» y de «porque», se convierte en miembro de pago de la raza humana.

Elaine Morgan,
*The Descent of the Child: Human Evolution
from a New Perspective*
[El origen del niño: la evolución humana desde una nueva perspectiva].

Nuestra conciencia, las mentes de los otros

La mente parece menos milagrosa cuando se ve cómo puede haberse formado a partir de distintos elementos y cómo sigue dependiendo de esos elementos. Una mente humana al desnudo (sin papel ni lápiz, sin hablar, sin comparar anotaciones, sin hacer esbozos) es, en primer lugar, una cosa que no hemos visto nunca. Cualquier mente humana a la que hayamos dirigido nuestra atención (incluyendo de

manera muy especial la propia, a la que miramos «desde dentro») es un producto no sólo de la selección natural sino de un diseño cultural repetido una y otra vez de enormes proporciones. Resulta bastante fácil ver por qué una mente parece milagrosa cuando no tenemos idea de todos sus componentes y de cómo se fabricaron. Cada componente tiene una larga historia de diseño, a veces de miles de millones de años.

Antes de que criatura alguna pudiera pensar, hubo criaturas con intencionalidad bruta y no pensante, dispositivos de mero rastreo y discriminación que no tenían ni el menor asomo de idea de lo que hacían ni por qué lo hacían. Pero funcionaban bien. Estos dispositivos rastreaban cosas respondiendo fiablemente a sus idas y venidas, manteniéndolas enfocadas la mayor parte del tiempo y desviándose rara vez y durante poco tiempo antes de volver nuevamente a su tarea. En períodos mucho más largos se podría decir que los *diseños* de estos dispositivos rastreaban cosas: no parejas escurridizas ni presas sino cosas abstractas, los fundamentos latentes de sus propias funciones. Conforme cambiaban las circunstancias los diseños de estos dispositivos fueron cambiando en respuesta apropiada a las nuevas condiciones, haciendo que sus poseedores siguieran bien equipados sin sobrecargarlos de razones. Estas criaturas cazaban pero no pensaban que cazaban, huían pero no pensaban que huían. Tenían el *saber cómo hacerlo*, que necesitaban. Ese saber cómo hacerlo es una especie de sabiduría, una especie de información útil pero no es un conocimiento representado.

Luego, algunas criaturas comenzaron a refinar aquella parte del medio que era más fácil de controlar colocando marcas tanto en su interior como en su exterior, descargando los problemas en el mundo y en otras partes de su cerebro. Comenzaron a fabricar y a utilizar representaciones pero no sabían que era eso lo que estaban haciendo. No necesitaban saberlo. ¿Deberíamos llamar «pensar» a esta cla-

se de utilización involuntaria de representaciones? Si es así, tendríamos que decir estas criaturas estaban pensando ¡pero sin saber que pensaban! Pensamiento inconsciente... los que gustan de las formulaciones «paradójicas» podrían estar a favor de esta manera de hablar pero podríamos decir sin inducir a tanto engaño que se trataba de una conducta *inteligente pero no pensante* porque no sólo no era reflexiva sino que tampoco permitía la reflexión sobre sí misma.

Nosotros, los seres humanos, hacemos muchas cosas inteligentes sin pensar. Nos lavamos los dientes, nos atamos los cordones de los zapatos, conducimos e incluso respondemos a preguntas sin pensar. Pero la mayoría de estas actividades nuestras son diferentes porque *podemos* pensar sobre ellas de una manera en que las demás criaturas no pueden pensar sobre sus actividades inteligentes pero no pensantes. Desde luego que buena parte de nuestras actividades no pensantes, como conducir un coche, pueden convertirse en actividades no pensantes, pero sólo después de pasar por un largo período de desarrollo del diseño que fuera explícitamente consciente de sí mismo. ¿Cómo se consigue esto? Las mejoras que introducimos en nuestro cerebro cuando aprendemos el lenguaje nos permiten revisar, recordar, ensayar, rediseñar nuestras propias actividades, convirtiendo nuestros cerebros en especies de cámaras de resonancia en las que los procesos, que de otro modo se desvanecerían, pueden quedar acogidos y convertirse en objetos por derecho propio. Aquellas que persisten más y adquieren influencia al persistir son las que llamamos pensamientos conscientes.

Los contenidos mentales se hacen conscientes no por ingresar en determinada cámara especial del cerebro, no por verse transducidos a un medio privilegiado y misterioso, sino por triunfar frente a otros contenidos mentales en el dominio del control de la conducta y, por ende, de conseguir

efectos más duraderos o, como decimos equívocamente, «memorizarlos». Y como somos hablantes, y como hablar con nosotros mismos es una de nuestras actividades más influyentes, una de las formas más efectivas de que un contenido mental se vuelva influyente es que ocupe una posición en la que controle las partes que utiliza el lenguaje.

Una reacción corriente ante esta sugerencia sobre la conciencia humana es la de franco desconcierto, que se expresa más o menos así: «Supongamos que todos esos extraños procesos competitivos tienen lugar en mi cerebro y supongamos que, tal como usted dice, los procesos conscientes son sencillamente los que triunfan. ¿Cómo es que *eso* los convierte en conscientes? ¿Qué les ocurre para que sea verdad que *yo* los conozca? Porque, después de todo, ¿se trata de *mi* conciencia la que precisa una explicación, habida cuenta de que la conozco en primera persona del singular!» Estas preguntas revelan una profunda confusión porque presuponen que lo que *nosotros* somos es *algo diferente*, una especie de *res cogitans* cartesiana, añadido a toda esta actividad cerebro-cuerpo. Y sin embargo, los que somos *es* precisamente esta organización de toda la actividad competitiva entre un montón de competidores que ha desarrollado nuestro cuerpo. Conocemos «automáticamente» que estas cosas ocurren en nuestro cuerpo porque si no las conociéramos ¿no sería nuestro cuerpo! (Podemos salir a la calle con los guantes de otra persona creyendo que son los nuestros pero no podríamos firmar un contrato con la mano de otra persona, creyendo erróneamente que era la nuestra, como tampoco podría afectarnos la tristeza o el miedo de otra persona creyendo erróneamente que esos eran nuestros sentimientos.)

Los actos y los sucesos que cada uno de nosotros puede contar a los demás, y las razones de los mismos, son nuestros porque los hemos fabricado nosotros... y porque ellos nos han fabricado a nosotros. Lo que cada uno de nosotros

es, es ese agente de cuya vida podemos dar razón. Podemos dar razón a los demás y a nosotros mismos. El proceso de descripción propia comienza en la más temprana niñez e incluye un buen montón de fantasía desde un primer momento. (Piénsese en Snoopy tumbado encima de su caseta y pensando: «Aquí va el as de la primera guerra mundial, volando hacia la batalla.») Y prosigue a lo largo de la vida. (Piénsese en el camarero del café en la discusión que sobre la «mala fe» presenta Jean-Paul Sartre en *El ser y la nada*, tan sumido en su aprendizaje sobre cómo dar vida a la descripción que de sí mismo hace como camarero.) Es lo que *nosotros* hacemos. Es lo que *nosotros* somos.

Otras mentes ¿son verdaderamente tan distintas de la mente humana? Como experimento sencillo me gustaría que usted que me lee imaginara algo que me atrevería a decir nunca ha imaginado antes. Imagínese, por favor, con cierto detalle un hombre con una bata blanca de laboratorio trepando por una cuerda mientras lleva un cubo de plástico rojo sujeto con los dientes. Tarea mental fácil para usted. ¿Podría un chimpancé llevar a cabo esa misma tarea? Lo dudo. Yo tomo los elementos hombre, cuerda, trepar, cubo, dientes, como objetos familiares del mundo perceptivo y conductual de un chimpancé de laboratorio. Estoy seguro de que un chimpancé de ese tipo no sólo puede percibir esas cosas sino que las ve *como* hombre, cuerda, cubo y demás. En cierto sentido mínimo, garantizo que el chimpancé tiene un *concepto* de hombre, de cuerda, de cubo (pero presumiblemente no posee conceptos de langosta, de epigrama, de abogado). Mi pregunta es la siguiente: ¿Qué puede hacer un chimpancé con sus conceptos? Ya en la primera guerra mundial el psicólogo alemán Wolfgang Köhler hizo algunos experimentos famosos con chimpancés para ver qué tipos de problemas podían resolver pensando: ¿Podría un chimpancé averiguar cómo apilar unas cajas en su jaula para alcanzar unos plátanos colgados a una altura excesiva como para po-

der cogerlos sin más? O en otro caso, ¿podría averiguar cómo unir dos palos para formar uno más largo con el cual conseguir descolgar la comida? Quiere la leyenda popular que los chimpancés de Köhler sí supieron dar con estas soluciones pero lo cierto es que estos animales fueron bastante poco impresionantes cuando se pusieron a la tarea: algunos resolvieron el problema después de muchísimas intentonas y otros ni siquiera llegaron a resolverlo. Estudios posteriores, entre ellos algunos actuales que son mucho más sutiles, siguen fracasando en la respuesta de estas preguntas aparentemente simples sobre lo que puede pensar un chimpancé cuando se le proporcionan todas las claves. Pero supongamos por un momento que los experimentos de Köhler sí hubieran dado respuesta a la pregunta que se dice que respondieron: que un chimpancé sí es capaz de descubrir la solución a un sencillo problema de este tipo, siempre que los elementos de la solución sean visibles y los tenga a mano, es decir, que se puedan manipular en una serie de pruebas.

Mi pregunta es diferente: ¿Puede un chimpancé *recordar* los elementos de una solución cuando estos elementos no están presentes para servirle al chimpancé de recuerdos visuales? El ejercicio al que se sometió usted que me lee procedió de una sugerencia verbal mía. Estoy seguro de que usted mismo puede hacer algunas sugerencias por su cuenta, inventando con ello una imaginiería mental de considerable novedad. (Esta es una de las cosas que sabemos de nosotros mismos: que a todos nos divierte meternos en complejos juegos imaginativos cuidadosamente hechos a medida para satisfacer nuestras necesidades del momento.) El informe que he esbozado en capítulos anteriores sobre cómo funciona la mente no humana supone que los chimpancés deben ser incapaces de semejantes actividades. En cierto modo podrían ser capaces de reunir accidentalmente los conceptos significativos (sus propios conceptos) y luego llevar su atención a resultados azarísticamente interesantes, pero sospecho que

hasta eso queda fuera de los límites de la movilidad o de la manipulabilidad de sus recursos.

Estas preguntas sobre las mentes de los chimpancés son bastante sencillas pero nadie sabe las respuestas... todavía. No es que las respuestas sean imposibles de obtener, sino que no es sencillo inventar los experimentos adecuados para obtenerlas. Nótese que estas preguntas no son de ese tipo que pueden enfocarse viendo el tamaño relativo del cerebro del animal o incluso calibrando su capacidad cognitiva bruta (de memoria, por ejemplo, o de discriminación). Seguramente hay abundante maquinaria en el cerebro de un chimpancé para almacenar toda la información que necesita como materia prima para semejante tarea; las preguntas apuntan a si la maquinaria está organizada de tal forma que permita ese tipo de explotación. (Tenemos un gran aviario con un montón de pájaros; ¿podemos hacer que vuelen en formación?) Lo que hace que una mente sea poderosa (ciertamente lo que hace que una mente sea consciente) no es de qué está hecha, ni lo grande que es, sino lo que es capaz de hacer. ¿Puede concentrarse? ¿Puede distraerse? ¿Puede recordar sucesos anteriores? ¿Puede mantener la atención sobre varias cosas al mismo tiempo? ¿De qué rasgos de sus actividades puede darse cuenta o cuáles de ellos puede controlar?

Cuando se contesten preguntas como estas, sabremos todo lo que necesitamos saber sobre esas mentes para responder a las preguntas moralmente importantes. Estas respuestas comprenderán todo lo que queremos saber sobre el concepto de conciencia *salvo* la idea de si en criaturas semejantes «están apagadas las luces», como un autor ha dicho recientemente. Pero es una mala idea, pese a su popularidad. No es que ninguno de sus defensores haya llegado a definirla, ni siquiera a clarificarla: es que no hay lugar a tal aclaración ni para tal definición. Porque supongamos que ya hemos contestado a todas las demás preguntas sobre la mente de cierta criatura y que en ese momento algunos filósofos

pretendan que seguimos sin saber la respuesta a esa cuestión absolutamente importante: ¿Están las luces mentales apagadas o encendidas? ¿Sí o no? ¿Qué importaría una u otra respuesta? Se nos debe una respuesta a *esta* pregunta que acabo de hacer antes de que tengamos que tomarnos en serio la cuestión.

Los perros ¿tienen concepto de lo que es un *gato*? Sí y no. Por próximo al nuestro que extensionalmente esté el «concepto» de gato que tiene un perro (nosotros y los perros identificamos los mismos conjuntos de entes como gatos y no gatos), se diferencia radicalmente del nuestro en una cosa: el perro no puede ponderar su propio concepto. No puede preguntarse si sabe qué es un gato; no puede preguntarse si los gatos son animales; no puede intentar distinguir la esencia del gato (mediante sus conocimientos) de los meros accidentes. Los conceptos no son cosas en el mundo del perro lo mismo que lo son los gatos. Los conceptos *son* cosas de nuestro mundo porque nosotros disponemos de lenguaje. Un oso polar es apto en relación con la nieve como no lo es un león, de modo que en cierto sentido el oso polar tiene un concepto del que carece el león: un concepto de nieve. Pero no hay mamífero carente de lenguaje que pueda tener el concepto de nieve que nosotros tenemos porque los mamíferos carentes de lenguaje no tienen manera de ponderar la nieve «en general» o «en sí». Y no por la razón trivial de que no tengan una *palabra* (del lenguaje natural) para designar la nieve sino porque sin lenguaje natural no son capaces de arrancar conceptos de sus nidos conexionistas entretejidos ni de manipularlos. Podemos hablar del conocimiento implícito o procedimental que tiene de la nieve el oso polar (el *saber de nieve* que tiene el oso), e incluso podemos investigar empíricamente la extensión de ese concepto de nieve encarnado en sí mismo que tiene el oso, pero teniendo en mente que ese no es un concepto manejable por el oso polar.

«¡Puede que no sea capaz de hablar, pero seguro que

piensa!»: uno de los principales objetivos de este libro ha sido el de hacer vacilar la confianza del que lee en esta reacción tan familiar. Quizá el mayor obstáculo en nuestros intentos de aclararnos en lo que se refiere a las capacidades mentales de los animales no humanos sea nuestra costumbre casi irresistible de imaginar que esos animales acompañan sus inteligentes actividades con una corriente de conciencia reflexiva *similar* a la nuestra. No es que ahora *sepamos* que no hacen cosa semejante: es más bien que en estas primeras etapas de nuestras investigaciones *no debemos suponer* que sí. El pensamiento tanto científico como filosófico sobre este asunto se ha visto fuertemente influido por el clásico texto de 1974 del filósofo Thomas Nagel: «¿Qué se siente al ser un murciélago?» El propio título nos pone en una pista equivocada, invitándonos a pasar por alto las distintas formas en que murciélagos (y demás animales) podrían conseguir sus ingeniosas proezas sin que «ser un...» signifique nada para ellos. Creamos un misterio supuestamente impenetrable para nosotros si suponemos sin más que la pregunta de Nagel tiene sentido y que sabemos lo que preguntamos.

¿Qué significa para un pájaro construir su nido? La pregunta nos invita a imaginar cómo construiríamos un nido y a intentar imaginarnos los detalles de la comparación. Pero como construir nidos es algo que no solemos hacer, primero deberíamos recordar lo que para nosotros significa hacer algo que nos sea familiar. Por ejemplo ¿qué significa para nosotros atarnos los cordones de los zapatos? A veces le prestamos atención; a veces lo hacen nuestros dedos sin que nos demos ninguna cuenta y mientras hacemos otras cosas. De modo que, podemos creer que el pájaro tiene alguna ensoñación o planea las actividades del día siguiente cuando hace sus movimientos constructores de nidos. Podría ser, pero lo cierto es que las pruebas hasta el momento indican con mucha fuerza que el pájaro no está equipado para hacer semejante cosa. Ciertamente, el contraste que notamos entre

prestar atención y hacer una tarea mientras tenemos la cabeza en otra parte, no tiene en absoluto ninguna equivalencia en el caso del pájaro. El hecho de que *nosotros* no podamos hacer un nido sin pensar cuidadosa y reflexivamente sobre lo que estamos haciendo y por qué lo hacemos no es razón suficiente para suponer que cuando el pájaro hace su nido deba tener pensamientos pajariles sobre lo que está haciendo (por lo menos en su primer nido, antes de dominar la tarea). Cuanto más aprendemos sobre cómo puede ocuparse un cerebro en los procesos que significan el cumplimiento de tareas inteligentes para sus poseedores no humanos, menos se parecen esos procesos a los pensamientos que, vagamente, hemos imaginado como responsables de esas tareas. Eso no significa que *nuestros* pensamientos no sean procesos que se dan en nuestro cerebro, o que nuestros pensamientos no representen el papel fundamental que normalmente les atribuimos en regir nuestra conducta. Alguno de esos procesos en el cerebro humano terminará, presumiblemente, por ser discernible como lo son los pensamientos que conocemos tan íntimamente, pero está todavía por ver si las *capacidades* mentales de cualesquiera otras especies dependen de que tengan una *vida* mental semejante a la nuestra.

Dolor y sufrimiento: lo que importa

Siempre hay una solución bien sabida a cualquier problema humano: limpia, plausible y errónea.

H. L. Mencken,
Prejuicios (segunda serie).

Sería tranquilizador llegar al final de nuestra historia y poder decir algo del siguiente estilo: «... Y así vemos que de nuestros descubrimientos se sigue que los insectos, los peces

y los reptiles no son sentientes en absoluto (son meros autómatas) pero ¡que los anfibios, las aves y los mamíferos son sentientes o conscientes como nosotros! Y, por si interesa, el feto humano se convierte en sentiente entre las semanas decimoquinta y decimosexta.» Semejante solución, limpia y plausible para algunos de nuestros problemas humanos de toma moral de decisiones supondría un enorme alivio, pero todavía no podemos contar semejante historia y no hay motivos para creer que más adelante podrá contarse. Es improbable que hayamos pasado por alto un rasgo mental que suponga una diferencia clara en cuanto a la moralidad, y los rasgos que hemos examinado parecen hacer su aparición no solamente de forma gradual sino también de modo asincrónico, incoherente y salteado, tanto en la historia evolutiva como en el desarrollo de los organismos individuales. Por supuesto que *es posible* que investigaciones posteriores revelen un sistema de similitudes y diferencias hasta el momento indetectado que nos impresione adecuadamente y que, entonces, seamos capaces de ver, por vez primera, por dónde ha trazado la naturaleza la línea y por qué la ha trazado precisamente ahí. Sin embargo, no es una posibilidad en la que apoyarse si ni siquiera hemos sido capaces de imaginar en qué consistiría tal descubrimiento o por qué razón sería moralmente significativo para nosotros. (De igual modo podríamos imaginar que determinado día desaparecerán las nubes y Dios nos dirá, directamente, a qué criaturas hay que incluir y a cuáles no en el círculo mágico.)

En nuestra búsqueda de diferentes mentes (y de proto-mentes) no parece existir ningún umbral claro ni ninguna masa crítica... hasta que llegamos al tipo de conciencia de que disfrutamos los seres humanos usuarios de lenguaje. Esa variedad de mente es única y más poderosa en varios órdenes de magnitud que cualquier otra variedad de mente, pero probablemente no queramos concederle excesivo peso moral. Bien podríamos creer que cuenta más en cualquier cálculo

lo moral la capacidad de sufrir que la capacidad de razonamiento abstruso y complejo sobre el futuro (y sobre todo lo demás que hay bajo la capa del sol). Pues bien, ¿cuál es entonces la relación entre el dolor, el sufrimiento y la conciencia?

Mientras la distinción entre dolor y sufrimiento sea, como la mayor parte de las cosas cotidianas, una distinción no científica, en cierto modo sin límites precisos, es sin embargo una marca de medida valiosa e intuitivamente satisfactoria de importancia moral. El fenómeno del dolor ni es homogéneo para todas las especies, ni es sencillo. Podemos verlo en nosotros mismos, dándonos cuenta de lo poco evidentes que son las respuestas a determinadas cuestiones bien simples. ¿Experimentamos como dolor los estímulos de nuestros receptores de dolor, estímulos que nos impiden permitir a nuestros miembros que adopten posturas forzadas y dañinas para nuestras articulaciones mientras dormimos? ¿O debemos llamarlos más apropiadamente dolores inconscientes? En cualquier caso, ¿tienen significación moral? Podríamos llamar a semejantes estados del sistema nervioso que protegen el cuerpo, estados «sentientes» sin por ello suponer que sean las experiencias de un ser, de un ego, de un sujeto. Para que tales estados importen (los llamemos dolores, o estados conscientes, o experiencias) debe haber un sujeto duradero *a quien* le importen porque sean fuente de sufrimiento.

Considérese el fenómeno tan comentado de la *disociación* que se da en presencia de un gran dolor o de un gran temor. Cuando se maltrata a los niños suelen acogerse a una estrategia desesperada pero efectiva: «Se ausentan.» En cierto modo se dicen a sí mismos que no son ellos quienes sufren ese dolor. Parece haber dos variantes de disociadores: los que simplemente rechazan el dolor como suyo y lo contemplan, por así decir, desde fuera; y aquellos que se desintegran, por lo menos momentáneamente, en algo parecido a

una personalidad múltiple (no soy «yo» quien está sufriendo este dolor sino «ella» o «él»). Mi hipótesis no enteramente extravagante sobre esto es que estas dos variantes de niños difieren en su aprobación tácita de una doctrina filosófica: que toda experiencia debe ser experiencia experimentada por algún sujeto. Los niños que rechazan el principio no ven nada malo en ausentarse del dolor dejándolo sin sujeto para que circule por ahí sin herir a nadie en concreto. Los que aceptan el principio tienen que inventarse a otro para que actúe de sujeto: «¡Cualquiera menos yo!»

Pueda o no sostenerse cualquier interpretación de este fenómeno de disociación, la mayor parte de los psiquiatras están de acuerdo en que funciona hasta cierto punto. Esto es, consista en lo que consista este recurso de la disociación, es genuinamente analgésico... o, por ser más precisos, sirva o no para disminuir el *dolor* sí decididamente aplaca el *sufrimiento*. De modo que lo que tenemos es una especie de resultado modesto: sea cual fuera la diferencia entre un niño no disociado y otro disociado, se trata de una diferencia que afecta notablemente a la existencia del sufrimiento o a su cantidad. (Me apresuro a añadir que nada de lo que he dicho supone que cuando los niños disocian mitiguen en modo alguno la atrocidad del perverso comportamiento de sus maltratadores; sin embargo, disminuyen drásticamente el horror de los efectos mismos... aunque esos niños paguen después un precio altísimo en la lucha contra los efectos de su disociación.)

Un niño disociado no sufre lo mismo que un niño no disociado. Pero a continuación ¿qué diremos de criaturas que son disociadas *por naturaleza*, que nunca consiguen, ni intentan conseguir siquiera, esa suerte de organización interna compleja que es la normal en un niño no disociado y que se quiebra en un niño disociado? Una conclusión obligada podría ser: una criatura de ese tipo está incapacitada por su constitución para pasar por esa *especie* o esa *cantidad* de su-

frimiento que puede soportar un ser humano normal. Pero si todas las especies no humanas están en tal estado relativamente desorganizado, tenemos base para la hipótesis de que los animales no humanos pueden desde luego sentir dolor pero no sufrir como nosotros.

¡Qué bien! Puede esperarse que los amantes de los animales respondan a esta sugerencia con justa indignación y con profundas sospechas. Como esta sugerencia promete desde luego aliviar muchas de nuestras aprensiones sobre las prácticas humanas corrientes absolviendo, por lo menos de cierta culpa que otros les adjudican, a nuestros cazadores, granjeros y experimentadores, deberíamos ser francamente cautelosos y ecuanímenes en considerar su fundamento. Deberíamos estar al tanto de posibles fuentes de engaño... por ambas partes, en este asunto tan revuelto. La sugerencia de que los animales no humanos no son susceptibles de niveles de sufrimiento humanos, provoca generalmente una catarata de relatos que parten el corazón, la mayor parte de ellos sobre perros. ¿Por qué predominan los perros? ¿Podría ocurrir que los perros fueran los mejores contraejemplos porque los perros tengan de verdad una mayor capacidad de sufrimiento que otros mamíferos? Podría ser, y la perspectiva evolutiva que hemos estado siguiendo puede explicarnos el porqué.

Los perros, y sólo los perros entre las demás especies domesticadas, responden fuertemente al enorme volumen de lo que podríamos denominar conducta «humanizante» que les dirigen sus amos. Hablamos a nuestro perro, nos compadecemos de él y generalmente lo tratamos como a un compañero humano en la medida de lo posible... y nos deleitamos en su respuesta positiva y familiar a esta simpatía. Podemos intentarlo con los gatos, pero rara vez se da. Retrospectivamente no es sorprendente; los perros domésticos son descendientes de mamíferos sociales hechos durante millones de años a vivir en grupos muy cooperativos e interactivos, mientras que los gatos surgen de linajes asociales. Lo que es más, los pe-

ros domésticos son diferentes a sus primos (lobos, zorros y coyotes) en algo importante: responden al afecto humano. No hay ningún misterio en esto. Los perros domésticos se han seleccionado precisamente por esa diferencia durante centenares de miles de generaciones. En *El origen de las especies* Charles Darwin señalaba que mientras la intervención humana deliberada en la reproducción de especies domesticadas ha funcionado durante miles de años para criar caballos más rápidos, ovejas más lanudas, ganado vacuno más carnoso, y así sucesivamente, durante mucho más tiempo se ha empleado una fuerza más sutil pero todavía más poderosa para conformar nuestras especies domesticadas. La denominó selección inconsciente. Nuestros antepasados practicaron la crianza selectiva pero no sabían que la estaban practicando. Este favoritismo involuntario, a lo largo de eones, ha hecho que los perros se parezcan más y más a nosotros en cosas que precisamente nos llaman la atención. Entre otros rasgos que hemos seleccionado inconscientemente, yo sugiero que se encuentra la susceptibilidad para su socialización con los humanos que, en los perros, ofrece muchos de los efectos organizadores que tiene en los propias crías humanas. Tratándoles como si fueran humanos, tuvimos éxito en convertirlos en más humanos de lo que hubieran sido de no darse ese trato. Comenzaron a desarrollar los propios rasgos organizativos que de no haber sido así serían dominio exclusivo de los seres humanos socializados. En resumidas cuentas, si la conciencia humana (la suerte de conciencia que es necesaria para un sufrimiento auténtico) es, como yo he sostenido, una reestructuración radical de la arquitectura virtual del cerebro humano, entonces debería seguirse que los únicos animales que fueran capaces de algo que remotamente se pareciera a esa forma de conciencia, serían aquellos animales a los que también se les hubiera impuesto, mediante la cultura, esa maquinaria virtual. Los perros son, claramente, los que están más próximos a cumplir esta condición.

¿Y qué podemos decir del dolor? Cuando piso a alguien en un dedo del pie causándole un dolor breve pero definido (y claramente consciente) le produzco un daño muy escaso, generalmente nulo. El dolor, aun siendo intenso, es demasiado breve como para que importe y yo no causo un daño permanente al pie. La idea de que la persona a la que piso sufre durante un segundo o dos es una aplicación errónea y risible de esa importante noción e incluso admitiendo que mi pisotón que causa unos pocos segundos de dolor pueda originar unos segundos o unos minutos más de irritación (especialmente si esa persona cree que lo he hecho a propósito), el dolor en sí, como experiencia breve de signo negativo es de una significación moral evanescente. (Que haya pisado a una persona cantando un aria y esa interrupción le estropee su carrera operística es harina de otro costal.)

Muchas discusiones admiten tácitamente que: 1) el sufrimiento y el dolor son lo mismo a diferente escala; 2) que todo dolor es «dolor experimentado»; y 3) que la «cantidad de sufrimiento» ha de calcularse («en principio») sumando todos los dolores (el horror de cada uno de ellos se determina mediante duración, número de veces e intensidad). Estas suposiciones son grotescas cuando se consideran desapasionadamente a la fría luz del día (cosa difícil para algunos de sus partidarios). Nos ayudará un pequeño ejercicio: suponemos que, gracias a cierto «milagro de la medicina moderna», pudiéramos separar todos nuestros dolores y sufrimientos del contexto en el que se producen posponiéndolos, por ejemplo, a finales de año, momento en el que podrían soportarse durante una semana horrorosa de agonía continua, en una especie de vacaciones negativas o bien (si nos tomamos en serio la fórmula que expresa la condición 3) intercambiando duración por intensidad, de modo que la desgracia de un año pudiera empaquetarse en un único montón que fuera un choque de intensidad intolerable y que durara por ejemplo cinco minutos. Un año entero sin ni siquiera un sua-

ve contratiempo o un dolor de cabeza a cambio de un descenso a los infiernos sin anestesia, completamente reversible, y breve... ¿aceptaríamos semejante trato? Yo desde luego sí si creyera que tiene algún sentido. (Damos por supuesto, naturalmente, que este episodio horrible no me mataría ni me dejaría una secuela de locura... aunque ¡no me importaría enloquecer durante ese choque!) Lo cierto es que aceptaría el trato incluso si supusiera «duplicar» o «cuadruplicar» el total del sufrimiento, siempre que no durara más de cinco minutos y que no dejara secuelas duraderas. Yo creo que cualquiera podría sentirse feliz de hacer un trato así, pero lo cierto es que no tiene ningún sentido. (Por ejemplo, supondría que el benefactor que proporcionara tal servicio gratis a todos duplicaría o cuadruplicaría, *ex hypothesi*, el sufrimiento del mundo... y el mundo le amaría por ello.)

Lo que es erróneo en semejante planteamiento es, naturalmente, que no se pueden desgajar dolor y sufrimiento de sus contextos de esa forma imaginada. La anticipación y las consecuencias, y el reconocimiento de sus implicaciones para los planes y perspectivas de vida de cada cual, no pueden dejarse a un lado como meros «acompañamientos cognitivos» del sufrimiento. Lo que es terrible de perder el empleo, o una pierna, o la reputación propia, o a un ser querido, no es el sufrimiento que esos acontecimientos *causan* en nosotros sino el sufrimiento que *son* semejantes sucesos. Si lo que nos preocupa es descubrir y mejorar situaciones desconocidas de sufrimiento en el mundo, lo que tenemos que hacer es estudiar las vidas de las criaturas y no su cerebro. Por supuesto que lo que les pasa por la cabeza es altamente significativo como ricas fuentes de evidencia sobre lo que hacen y cómo lo hacen, pero lo que hacen termina por ser tan visible (para observadores preparados) como las actividades de las plantas, los torrentes de las montañas o los motores de combustión interna. Si no encontramos sufrimiento en las vidas que podemos ver (estudiándolas diligentemente, utili-

zando todos los métodos de la ciencia) podemos estar seguros de que no hay sufrimientos invisibles en sus respectivos cerebros. Si descubrimos sufrimiento, lo reconoceremos sin dificultad. Nos resulta demasiado familiar.

Este libro empezaba con un montón de preguntas y, habida cuenta de que es el libro de un filósofo, acaba no con las respuestas sino, espero, con mejores versiones de esas mismas preguntas. Como mínimo podemos ver algunos caminos que pueden seguirse y algunas trampas que deben evitarse en nuestra actual exploración de los distintos tipos de mente.

LECTURAS ADICIONALES

Podría dar la impresión de que no tiene mucho sentido que el que me lee lea a su vez los libros que más me han influido para escribir este libro ya que, si he hecho bien mi trabajo, habré extraído los mejores textos, ahorrando al lector tiempo y dificultades. Puede que esto sea verdad de algunos de ellos, pero no de los libros que enumero a continuación. Se trata de los libros que concretamente quiero que lean mis lectores, si es que no los han leído ya, o que los releen si ya los leyeron. De ellos he aprendido mucho ¡pero no lo suficiente! Lo cierto es que me doy perfecta cuenta de que en esos libros hay mucho más que yo y los demás podemos descubrir, y en cierto sentido este libro se pretende inductor y guía.

En primer lugar, presento dos libros famosos e influyentes, pero a menudo mal comprendidos, de sendos filósofos: *The Concept of Mind* [El concepto de mente] (1949) de Gilbert Ryle, y las *Investigaciones filosóficas* (1958) de Ludwig Wittgenstein. Tanto Ryle como Wittgenstein se mostraban bastante contrarios a la idea de una investigación científica de la mente, y es moneda corriente en la «revolución cognitiva» que hemos ido mucho más allá de sus análisis impla-

cablemente acientíficos sobre lo mental. No es verdad. Hay que tolerar su percepción errónea de buenas cuestiones científicas, a menudo tan frustrante, y su casi total ignorancia de la biología y de la ciencia del cerebro, pero aun así se las apañaron para hacer profundas e importantes observaciones que la mayoría de nosotros sólo ahora estamos en condiciones de valorar adecuadamente. El análisis de Ryle sobre el «saber cómo» (diferente del «saber qué») ha llamado la atención desde hace mucho, y ha provocado la aprobación de los científicos cognitivos, pero sus famosas pretensiones de que el pensar podría producirse en el mundo público y que no tenía por qué recluirse en algún lugar del pensar, han parecido perversas y mal fundamentadas para la mayor parte de sus lectores. No cabe duda de que algunas lo eran, pero sorprende ver cómo brilla buena parte del pensamiento de Ryle cuando se lo enfoca bajo una nueva luz. Por su parte Wittgenstein ha sufrido la admiración de una horda de intérpretes erróneos que comparten con él su antipatía por la ciencia pero que carecen de su visión. Se les puede pasar por alto sin que nos perdamos nada; vayan al original y léanlo por medio de la lupa que he intentado proporcionarles. Una figura colocada en sitio similar es la del psicólogo James J. Gibson, cuyo libro sorprendentemente original *The Senses Considered as Perceptual Systems* [Los sentidos considerados como sistemas perceptivos] (1968) ha sido un pararrayos para ataques mal dirigidos por parte de científicos cognitivos y libro sagrado para una camarilla absolutamente devota de gibsonianos radicales. Léanlo: a ellos déjenlos para después.

El libro *Vehicles: Experiments in Synthetic Psychology* [Vehículos: experimentos de psicología sintética] (1984) de Valentino Braitenberg, ha inspirado a una generación de robotistas y demás científicos cognitivos y es, sencillamente, un clásico. Cambiará la manera de ver la mente de quien lo lea, si es que mi libro no ha conseguido ya tal transformación. Otro filósofo que ha bebido abundantemente del pozo de

Braitenberg es Dan Lloyd y su libro de 1989, *Simple Minds* [Mentes sencillas], abarca buena parte de lo que abarca el mío, con énfasis algo distintos pero creo que sin mayores discrepancias. Dan Lloyd fue alumno informal mío y luego colega más joven en Tufts mientras trabajaba en su libro. No puedo decir sin más lo que me ha enseñado y lo que yo le he enseñado a él; de su libro hay mucho que aprender en todo caso. Podría decir lo mismo de otros colegas míos del Centro, en Tufts, como Kathleen Atkins, Nicholas Humphrey y Evan Thompson. Fue Atkins la que primero me enseñó, ya a mediados de los ochenta, por qué debíamos abandonar la epistemología y la ontología anticuadas al pensar en las mentes animales, y cómo debíamos hacerlo. Véanse, por ejemplo, sus ensayos «Science and our Inner Lives: Birds of Prey, Beasts, and the Common (Featherless) Biped» [La ciencia y nuestra vida interior: las aves de rapiña, las bestias y el bípedo (implume) común] y «What Is It Like to Be Boring and Myopic?» [«¿En qué consiste ser aburrido y miope?»] Nicholas Humphrey llegó en 1987 para trabajar conmigo durante varios años pero todavía no he aceptado todas las ideas de su *Una historia de la mente* (1995) a pesar de las muchas horas de discusión. Estando Evan Thompson en el Centro terminó su libro conjunto con Francisco Varela y Eleanor Rosch *The Embodied Mind* [La mente hecha carne] (1990) y las influencias de ese libro en éste pueden verse fácilmente, estoy seguro. Más recientemente, el libro de Antonio Damasio *El error de Descartes* (1996) consolida y avanza algunos de los temas de estas obras, además de abrir nuevos terrenos por su cuenta.

Para una comprensión más profunda del papel de la evolución en el diseño de las mentes de todas las criaturas, habría que leer los libros de Richard Dawkins, empezando por *El gen egoísta*. La *Social Evolution* [Evolución social] de Robert Trivers es una introducción excelente a los mejores aspectos de la sociobiología. El nuevo campo de la psicolo-

gía evolutiva está bien representado en una antología editada por Jerome Barkow, Leda Cosmides y John Tooby: *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* [La mente adaptada: la psicología evolutiva y la generación de cultura] (1992), y para un repensar que abre los ojos sobre la psicología y la biología del niño, léase el de Elaine Morgan *The Descent of the Child: Human Evolution from a New Perspective* [El linaje del niño: la evolución humana desde una nueva perspectiva] (1995).

En otro frente, los etólogos cognitivos han alimentado las fantasías de los filósofos (y de los psicólogos) sobre la vida mental y los poderes de las animales no humanos con un torrente de fascinantes trabajos experimentales y de observación. Donald Griffin es el padre de este campo. Sus libros *El pensamiento de los animales* (1986), *The Question of Animal Awareness* [La cuestión de la conciencia animal] (1976) y *Animals Minds* [Mentes animales] (1992) y lo que es más importante, sus pioneras investigaciones sobre la ecolocalización de los murciélagos, abrieron la mente de muchos a las posibilidades de este campo. Un estudio ejemplar es el trabajo de Dorothy Cheney y Robert Seyfarth con monos tumbulu: *How Monkeys See the World* [Cómo ven los monos el mundo] (1990). La antología de Andrew Whiten y Richard Byrne, *Machiavellian Intelligence* [Inteligencia maquiavélica] (1988) y la antología de Carolyn Ristau *Cognitive Ethology* [Etología cognitiva] (1991) son al tiempo textos clásicos y análisis austeros de los distintos problemas; y el libro bellamente ilustrado de James y Carol Gould *The Animal Mind* [La mente animal] (1994) debería perfumar las conjeturas teóricas de todo aquel que piense sobre la mente animal. Para lo ultimísimo en pensamiento y comunicación de los animales véase el nuevo libro de Marc Hauser *The Evolution of Communication* [La evolución de la comunicación] y el de Derek Bickerton *Language and Human Behavior* [Lenguaje y comportamiento humano]. El ensayo de

1991 de Patrick Bateson «Assessment of Pain in Animals» [«Evaluación del dolor en los animales»] es una panorámica valiosa de lo que se sabe y de lo que sigue sin saberse sobre el dolor y el sufrimiento animales.

En el capítulo 4 pasé de puntillas (aun no queriendo que así fuera) sobre una amplia y fascinante literatura acerca de la intencionalidad de orden superior: los niños y los animales como «psicólogos naturales». Decidí que podía pasar así porque el asunto había recibido recientemente en otros lugares una buena atención. Entre muchos, dos buenos libros que explican tanto los detalles como su importancia son el de Janet Astington *The Child's Discovery of the Mind* [El descubrimiento de la mente por el niño] (1993) y el de Simon Baron-Cohen *Autismo* (1998).

También he sido escueto en el importante asunto del aprendizaje ABC y de sus modelos actuales más prometedores. Para los detalles (y para algunas diferencias no triviales de opinión filosófica que bien merece la pena considerar) véanse *Associative Engines: Connectionism, Concepts and Representational Change* [Máquinas asociativas: conexionismo, conceptos y cambio de representación] (1993) de Andy Clark, y *The Engine of Reason, the Seat of the Soul* [La máquina de la razón, sede del alma] (1995) de Paul Churchland. Los que deseen entrar todavía más a fondo en los detalles (cosa que recomiendo) pueden comenzar con el libro de Patricia Churchland y Terence Sejnowski, *The Computational Brain* [El cerebro computativo] (1992). Considérense estos libros como un contraste realista de algunas de mis conjeturas más impresionistas y entusiastas. Dos filósofos más cuyas obras debería consultar cualquiera que quisiera valorar las ideas que he avanzado aquí triangulándolas con visiones algo relacionadas aunque bastante ortogonales son Gareth Evans y su *The Varieties of Reference* [Las variedades de la referencia] (1982) y Ruth Garrett Millikan y sus *Language Thought and other Biological Categories*

[Pensamiento del lenguaje y otras categorías biológicas] (1984) y *White Queen Psychology and other Essays for Alice* [Psicología de la reina blanca y otros ensayos para Alicia] (1993).

La discusión sobre cómo hacer cosas pensantes de los capítulos 5 y 6 se inspiró no sólo en *Mind in Science* [La mente en la ciencia] (1981) de Richard Gregory, y en el texto de Andy Clark y Annette Karmiloff-Smith, sino también en el libro de Karmiloff-Smith *Más allá de la modularidad* (1994) y en otros libros anteriores que llevaban dándome vueltas por la cabeza, fructíferamente, varios años: *The Origins of Consciousness in the Breakdown of the Bicameral Mind* [Los orígenes de la conciencia en la descomposición de la mente bicameral] (1976) de Julian Jaynes; *Metaphors We Live By* [Metáforas por las que vivimos] (1980) de George Lakoff y Mark Johnson; *Mental Models* [Modelos mentales] (1983) de Philip Johnson-Laird y *The Society of Mind* [La sociedad mental] (1985) de Marvin Minsky. Un libro nuevo que presenta los primeros modelos concretos de algunas de estas actividades humanas quintaesenciales es el de Douglas Hofstadter *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought* [Conceptos fluidos y analogía creativas: modelos informáticos de mecanismos fundamentales del pensamiento] (1995).

Mi libro de 1991, *Consciousness Explained* [La conciencia, explicada] trataba fundamentalmente de la conciencia humana, y decía poco de la mente de otros animales, como no fuera implícitamente. Como algunos lectores, al intentar descubrir lo que en él se decía, han llegado a algunas conclusiones que encontraron ambiguas o incluso alarmantes, me di cuenta de que tenía que aclarar mi teoría de la conciencia, ampliándola explícitamente a otras especies. *Tipos de mentes* es uno de sus resultados; otro es «Animal Consciousness: What Matters and Why» [Conciencia animal: qué

importa y por qué)], que fue mi contribución al congreso «En compañía de los animales» celebrado en la New School for Social Research, en Nueva York, en abril de 1995. También ha sido recibido con escepticismo el apoyo evolucionista a mi teoría de la conciencia que expuse en mi libro *La peligrosa idea de Darwin*. Muchas de las ideas que avanzo en *Tipos de mentes* están extraídas de otros artículos míos que aparecen en la bibliografía o basadas en ellos.

BIBLIOGRAFÍA

- Atkins, Kathleen, «Science and our Inner Lives: Birds of Prey, Beasts, and the Common (Featherless) Biped», en Marc Bekoff y Dale Jamieson, eds., *Interpretation and Explanation in the Study of Animal Behavior*, vol. 1 (Boulder, Colo.: Westview, 1990), págs. 414-427.
- , «What Is It Like to Be Boring and Myopic?», en Dahlbom, ed., *Dennett and His Critics*.
- Astington, Janet, *The Child's Discovery of the Mind* (Cambridge: Harvard University Press, 1993).
- Balda, Russell P., y R. J. Turek, «Memory in Birds», en Herbert L. Roitblat, Thomas G. Bever y Herbert S. Terrace, eds., *Animal Cognition* (Hillsdale, N. J.: Erlbaum, 1984), págs. 513-532.
- , Alan C. Kamil y Kristie Grim, «Revisits to Emptied Cache Sites by Clark's Nutcrackers (*Nucifraga columbiana*)», *Animal Behavior*, 34 (1986), págs. 1289-1298.
- Barkow, Jerome, Leda Cosmides, y John Tooby, *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (Oxford: Oxford University Press, 1992).
- Baron-Cohen, Simon, *Autismo* (Alianza, 1998).
- Bateson, Patrick, «Assessment of Pain in Animals», *Animal Behavior*, 42 (1991), págs. 827-839.

- Bickerton, Derek, *Language and Human Behavior* (Seattle: Universidad de Washington Press, 1995).
- Braitenberg, Valentino, *Vehicles: Experiments in Synthetic Psychology* (Cambridge: MIT Press/A Bradford Book, 1984).
- Cheney, Dorothy, y Robert Seyfarth, *How Monkeys See the World* (Chicago: Universidad de Chicago Press, 1990).
- Churchland, Patricia, y Terence Sejnowski, *The Computational Brain* (Cambridge: MIT Press/A Bradford Book, 1992).
- Churchland, Paul, *Scientific Realism and the Plasticity of Mind* (Cambridge, RU: Cambridge University Press, 1979).
- , *The Engine of Reason, the Seat of the Soul* (Cambridge: MIT Press/A Bradford Book, 1995).
- Clark, Andy, *Associative Engines: Connectionism, Concepts and Representational Change* (Cambridge: MIT Press/A Bradford Book, 1993).
- , y Annette Karmiloff-Smith, «The Cognizer's Innards: A Psychological and Philosophical Perspective on the Development of Thought», *Mind and Language*, 8 (1993), págs. 487-519.
- Dahlbom, Bo, ed., *Dennett and His Critics: Demystifying Mind* (Oxford: Blackwell, 1993).
- Damasio, Antonio, *El error de Descartes* (Grijalbo, 1996).
- Darwin, Charles, *El origen de las especies* (Londres: Murray, 1859).
- Dawkins, Richard, *El gen egoísta* (Salvat, 1993).
- , y John R. Krebs, «Animal Signals: Information or Manipulation?» en John R. Krebs y Nicholas B. Davies, eds., *Behavioural Ecology*, 2.^a ed. (Sunderland, Mass.: Sinauer Associates, 1978), págs. 282-309.
- Dennett, Daniel, «Brain Writing and Mind Reading», en K. Gunderson, ed., *Language, Mind and Knowledge, Minnesota Studies in the Philosophy of Science*, vol. 7 (Mi-

- neápolis: Universidad de Minesota Press, 1975). Reeditado en Dennett, *Brainstorms* y después con una posdata en D. Rosenthal, ed., *The Nature of Mind* (Oxford: Oxford University Press, 1991).
- , «Conditions of Personhood», en Amelie Rorty, ed., *The Identities of Persons* (Berkeley: Universidad de California Press, 1976). Reeditado en Dennett, *Brainstorms*.
- , *Brainstorms* (Cambridge: MIT Press/A Bradford Book, 1978).
- , «Where Am I?», en Dennett, *Brainstorms*.
- , «Beyond Belief», en Andrew Woodfield, ed., *Thought and Object* (Oxford: Oxford University Press, 1982). Reeditado en Dennett, *The Intentional Stance*.
- , «Intentional Systems in Cognitive Ethology: The “Panglossian Paradigm” Defended», *Behavioral and Brain Sciences*, 6 (1983), págs. 343-390.
- , *La actitud intencional* (Editorial Gedisa, 1991).
- , *Consciousness Explained* (Boston: Little, Brown, 1991).
- , «Learning and Labeling» (comentado en Clark y Karmiloff-Smith), *Mind and Language*, 8 (1993), págs. 540-548.
- , «The Message Is: There Is No Medium» (respuesta a Jackson, Rosenthal, Shoemaker, y Tye), *Philosophy & Phenomenological Research*, diciembre de 1993, págs. 889-931.
- , «Back from the Drawing Board» (contestación a las críticas), en Dahlbom, ed., *Dennett and His Critics*.
- , «Get Real» (contestación a las críticas), en *Philosophical Topics*, 22 (1995), págs. 505-568.
- , «Animal Consciousness: What Matters and Why», en *Social Research*, 62 (1995), págs. 691-710.
- , «Consciousness: More like Fame than Television», del libro de conferencias «Interfaces Brain-Computer», Christa Maar, Ernst Pöppel y Thomas Christaller, eds., publicado por Rowohlt.
- , «Do Animals Have Beliefs?», en Herbert L. Roitblat,

- ed., *Comparative Approaches to Cognitive Sciences*, MIT Press.
- Eigen, Manfred, *Steps Towards Life* (Oxford: Oxford University Press, 1992).
- Evans, Gareth, *The Varieties of Reference* (Oxford: Clarendon Press, 1982).
- Gaussier, Philippe, y Zrehen, S., «A Constructivist Approach for Autonomous Agents», en Adia Magnenat Thalmann y Daniel Thalmann, eds., *Artificial Life and Virtual Reality* (Londres: Wiley, 1994).
- , «Avoiding the World Model Trap: An Acting Robot Does Not Need to Be So Smart», *Robotics and Computer Integrated Manufacturing*, 11 (1994), págs. 279-286.
- Gibson, James J., *The Senses Considered as Perceptual Systems* (Londres: Allen & Unwin, 1968).
- Gould, James, y Carol Gould, *The Animal Mind* (Nueva York: HPHLP, Scientific American Library, 1994).
- Gregory, Richard L., *Mind in Science: A History of Explanations in Psychology* (Cambridge, RU: Cambridge University Press, 1981).
- Griffin, Donald, *El pensamiento de los animales* (Ariel, 1986).
- , *The Question of Animal Awareness* (Nueva York: Rockefeller University Press, 1976).
- , *Animal Minds* (Chicago: Universidad de Chicago Press, 1992).
- Hasson, O., «Pursuit-Deterrent Signals: Communication between Predator and Prey», *Trends in Ecology and Evolution*, 6 (1991), págs. 325-329.
- Hebb, Donald, *The Organization of Behavior: A Neuropsychological Theory* (Nueva York: Wiley, 1949).
- Hofstadter, Douglas R., *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought* (Nueva York: Basic Books, 1995).
- Holley, Tony, «No Hide, No Seek», *Natural History*, 7 (1994), págs. 42-45.

- Humphrey, Nicholas, «Nature's Psychologists», *New Scientist*, 29 (junio de 1978), págs. 900-904. Reeditado en *Consciousness Regained* (Oxford: Oxford University Press, 1983).
- , *Una historia de la mente* (Editorial Gedisa, 1995).
- Israel, David, John Perry, y Syun Tutiya, «Executions, Motivations and Accomplishments», *Philosophical Review*, 102 (1993), págs. 515-540.
- Jaynes, Julian, *The Origins of Consciousness in the Breakdown of the Bicameral Mind* (Boston: Houghton Mifflin, 1976).
- Johnson-Laird, Philip N., *Mental Models* (Cambridge, RU: Cambridge University Press, 1983).
- Kamil, Alan C., Russell P. Balda, Deborah J. Olson, y Sally Good, «Returns to Emptied Cache Sites by Clark's Nutcrackers, *Nucifraga columbiana*: A Puzzle Revisited», *Animal Behavior*, 45 (1993), págs. 241-252.
- Karmiloff-Smith, Annette, *Más allá de la modularidad* (Alianza, 1994).
- Lakoff, George, y Mark Johnson, *Metaphors We Live By* (Chicago: Universidad de Chicago Press, 1980).
- Lloyd, Dan, *Simple Minds* (Cambridge: MIT Press/A Bradford Book, 1989).
- McFarland, David, 1989, «Goals, No-Goals and Own Goals», en Alan Montefiore y Denis Noble, eds., *Goals, No-Goals and Own Goals: A Debate on Goal-Directed and Intentional Behaviour* (Londres: Unwin Hyman, 1989), págs. 39-57.
- Menzel, Emil W., hijo, 1971, «Communication about the Environment in a Group of Young Chimpanzees», *Folia Primatologia*, 15 (1971), págs. 220-232.
- , «A Group of Young Chimpanzees in a One-Acre Field», en A. M. Schreier y F. Stolnitz, eds., *Behavior of Nonhuman Primates*, vol. 5 (Nueva York: Academic Press, 1974), págs. 83-153. Reeditado en Ristau, *Cognitive Ethology*.

- Millikan, Ruth Garrett, *Language, Thought, and other Biological Categories* (Cambridge: MIT Press/A Bradford Book, 1984).
- , *White Queen Psychology and other Essays for Alice* (Cambridge: MIT Press/A Bradford Book, 1993).
- , «A Common Structure for Concepts of Individuals, Stuffs, and Basic Kin: More Mama, More Milk, and More Mouse», *Behavioral and Brain Sciences*, 1997.
- Minsky, Marvin, *The Society of Mind* (Nueva York: Simon & Schuster, 1985).
- Morgan, Elaine, *The Descent of the Child: Human Evolution from a New Perspective* (Oxford: Oxford University Press, 1995).
- Nagel, Thomas, «What Is It Like to Be a Bat?», *Philosophical Review*, 83 (1974), págs. 435-450.
- Nietzsche, Friedrich, *Así habló Zaratustra* (Alianza, 1988).
- Ristau, Carolyn, ed., *Cognitive Ethology* (Hillsdale, N. J.: Erlbaum, 1991).
- , «Aspects of the Cognitive Ethology of an Injury-Feigning Bird, the Piping Plover», en Ristau, ed., *Cognitive Ethology*, págs. 91-126.
- Ryle, Gilbert, *The Concept of Mind* (Londres: Hutchinson, 1949).
- Sartre, Jean-Paul, *El ser y la nada* (Alianza, 1989).
- Searle, John, «Minds, Brains and Programs», *Behavioral and Brain Sciences*, 3 (1980), págs. 417-458.
- Skinner, B. F., *Ciencia y conducta humana* (Martínez Roca, 1986).
- , «Behaviorism at Fifty», en T. W. Wann, ed., *Behaviorism and Phenomenology* (Chicago: Universidad de Chicago Press, 1964), págs. 79-108.
- Sontag, Susan, *Sobre la fotografía* (EDHASA, 1982).
- Thomas, Elizabeth Marshall, *The Hidden Life of Dogs* (Boston: Houghton Mifflin, 1993).

- Trivers, Robert, *Social Evolution* (Menlo Park, California: Benjamin Cummings, 1985).
- Varela, Francisco J., Evan Thompson, y Eleanor Rosch, *The Embodied Mind: Cognitive Science and Human Experience* (Cambridge: MIT Press/A Bradford Book, 1991).
- Whiten, Andrew, «Grades of Mind Reading», en Charlie Lewis y Peter Mitchell, eds., *Children's Early Understanding of Mind: Origins and Development* (Hillsdale, N. J.: Erlbaum, 1994), págs. 47-70.
- , y Richard W. Byrne, eds., *Machiavellian Intelligence* (Oxford: Oxford University Press, 1988).
- Wiener, Norbert, *Cibernética* (Tusquest, 1986).
- Wittgenstein, Ludwig, *Investigaciones filosóficas* (Crítica, 1988).
- Young, Andrew, «The Neuropsychology of Awareness», en Antti Revonsuo y Matti Kamppinen, *Consciousness in Philosophy and Cognitive Neuroscience* (Hillsdale, N. J.: Erlbaum, 1994), págs. 173-203.

ÍNDICE

- aborto, debate sobre el, y atribución de conciencia, 16-17
- actitudes proposicionales, 56-61
- actividad «inconsciente», 24-26
- actividad «automática», 24-26
- ADN, 32
- agente
 - comprensión de las circunstancias del, 56-59
 - enfoque intencional y, 40-42
 - macromolecular, 32-36
 - nacimiento del, 31-33
- alma nutritiva, 38
- animales
 - actividad «automática» o «inconsciente» de los, 24-25
 - comprensión instintiva entre los, 21
 - cuestiones sobre si tienen mente o carecen de ella, 11-15
 - dolor en los, 115-116
 - prueba para saber si tienen mente, 16-18
 - sistema circulatorio en los, 39, 83-85
 - sistemas de control neural en los, 93-95
- animismo, 47
- antropomorfismo, 40, 44, 47, 118
- aprendizaje ABC, 108-109, 113, 156
- aprendizaje latente, 112
- Aristóteles, 38
- ARN, 32
- ARN, fago de, 33-34
- Así habló Zaratustra* (Nietzsche), 97-98
- asociacionismo, 107-108
- Astington, Janet, 203
- Atkins, Kathleen, 201
- audición, 101-102
- Balda, Russell, 163-164
- Barkow, Jerome, 202
- Baron-Cohen, Simon, 203
- Bateson, Patrick, 203
- Bickerton, Derek, 203
- Boole, George, 131
- booleanas, funciones, 131-132
- Braitenberg, Valentino, 124, 132, 200
- Brooks, Rodney, 27
- Byrne, Richard, 202

- Capgras, ilusión de, 134-135
 chovinismo de escala temporal, 77-81
 Cheney, Dorothy, 140, 202
 Chomsky, Noam, 175
 Churchland, Patricia, 203
 Churchland, Paul, 62, 203
 cibernética, 89
 circulatorios, sistemas, 38-39, 83-85
 Clark, Andy, 157, 203
 Cog (robot), 27-28
 comunicación
 compartir experiencias mediante la, 20-21
 confabulación de aproximación en la, 152-153
 contacto visual en la, 26-28
 cooperación y, 151-152
 dispositivos de lectura por ordenador y, 19
 mentes no comunicativas y, 23-30
 reconocimiento de poseedores de mente en la, 14-15
 sinceridad y decepciones en la, 21-22
 conceptos, las palabras como prototipos de los, 179
 condicionamiento skinneriano, 106, 108, 116
 conductismo, 105-114 (*véase también* criaturas skinnerianas)
 conexionismo, 106, 107
 confabulación aproximada, 152-153
 conciencia
 contacto visual y, 26-28
 Modelo de Borradores Múltiples de la, 27
 conciencia de uno mismo y comprobación de hipótesis, 145
 construcción de nidos, 133-134, 189-190
 contacto visual, 26
 Cosmides, Leda, 202
 creencias, 60-61
 de re y de dicto, 130
 enfoque intencional de otras entidades y, 56-59
 proposiciones y, 60-61
 criaturas darwinianas, 104-105
 criaturas gregorianas, 121, 122, 123, 141, 160, 168, 169
 criaturas popperianas, 109-114, 120-121, 123-124, 156, 168-169
 criaturas skinnerianas, 105, 109, 114, 116, 122-123
 Crisipo, 139

 Damasio, Antonio, 201
 Darwin, Charles, 121, 195
 Dawkins, Richard, 146, 151, 201
 delfines, inteligencia de los, 139-140
 Dennett, Daniel C., 130, 204-205
 descarga, 160-161
 Descartes, René, 12, 13, 90, 96, 98-99
 disociación, 192-194
 dispositivos lectores, 19
 dolor
 disociación y, 192-194
 distinción entre sufrimiento y, 192, 196-197
 en animales, 115-116
 refuerzo negativo y, 116-118
 sentencia y, 116-120
 sentimiento de, 28-29
 dualismo, 36-37, 95-96

 efectores del sistema nervioso, 87, 92-94
 Eigen, Manfred, 33-34, 37
El gen egoísta (Richard Dawkins), 146, 201

- El origen de las especies* (Darwin), 195
- El ser y la nada* (Sartre), 185
- enfoque del diseño, 42-44
- EPIRB (véase RIPE)
- epistemología, 12
- errores e intencionalidad, 51-53
- errores en el rastreo, 134-136
- etiquetado, 161-166, 177
- Evans, Gareth, 203
- evolución
- cerebro y, 98
 - percepción y, 102
 - selección inconsciente en la, 195
 - selección natural y, 104
- extensión, 53-54
- fantasía, 22
- fenómeno phi, 129
- firmas auditivas y rastreo, 126
- firmas visuales y rastreo, 126, 129
- flores, categoría moral de las, 15
- fotografía, 169-171
- fitotaxia, 123-125, 132
- frases, proposiciones expresadas como, 60-63
- Frege, Gottlob, 59
- función, surgimiento de la, 46-47
- funcionalismo, 85-87, 95
- gacelas, 148, 151
- Gaussier, Philippe, 166
- generar y probar, 103-114
- Gibson, James J., 200
- glutamato, moléculas de, 17, 92
- Gould, James y Carol, 202
- Gregory, Richard, 121-122, 204
- Griffin, Donald, 202
- habla
- compartir experiencias mediante el, 20-21
 - dispositivos de lectura informatizados y el, 19
 - hablar con nosotros mismos, 175-180
 - reconocimiento de poseedores de mente con el, 14-15
 - sinceridad y engaños en el, 21-22
- Hasson, O., 147
- Hauser, Marc, 116, 202
- Hebb, Donald, 107
- herramientas, inteligencia y uso de, 121-123
- hipótesis cero, 17
- Hofstadter, Douglas, 204
- Holley, Tony, 147
- Homero, 137
- Hume, David, 107
- Humphrey, Nicholas, 144, 148, 201
- ideas, teoría gráfica de las, 67-68
- identificación
- del amo por el perro, 137-139
 - firmas auditivas y visuales en la, 126
 - olfato y, 125, 126
- identificación de caras, 135-136
- improntas, 127
- inteligencia
- criaturas gregorianas y la, 121
 - esencial y cinética, 121
 - mascotas e, 138-140
 - uso de herramientas e, 121-122
- inteligencia cinética, 121
- inteligencia potencial, 14
- intención (en sentido corriente), distinta del término filosófico, 48-50
- intencional, enfoque, 40-55, 122-123
- adopción del, 40-55, 144
 - antropomorfismo del, 40, 44, 47

- comprensión de la idea de ser agente y, 56
- enfoque físico y enfoque del diseño frente al, 41-44
- frente a las plantas, 47-48
- imprecisión del, 56-59
- opacidad referencial del, 54
- ordenadores que juegan al ajedrez y, 44-45
- selección natural y, 76-77, 80
- intencional, objeto, 51
- intencionales, sistemas, 40, 48
- actitudes proposicionales y, 56-61
- desear el mal y los, 46
- equivocaciones y, 52
- nacimiento de la función y, 46-47
- orden superior y, 143-144, 145
- sentencia de los, 81
- intencionalidad, 49-50, 182
- confabulación de aproximación e, 153
- derivada frente a originaria, 65-72
- en los primeros sistemas de control, 82
- equivocaciones e, 52-53
- mantenimiento de secretos e, 154-156
- intencionalidad derivada, 65-72
- intencionalidad intrínseca, 65-72
- intencionalidad originaria, 65-72
- intensión, 53-54
- intensionalidad, 53-55
- interruptores ACTIVADO/DESACTIVADO
- alma nutritiva e, 38
- surgimiento de la función e, 46-47
- Israel, David, 55
- Jaynes, Julian, 172, 204
- Johnson, Mark, 204
- Johnson-Laird, Philip, 204
- juegos, teoría de, 76
- Karmiloff-Smith, Annette, 157, 204
- kheperas*, robots, 166
- Köhler, Wolfgang, 185-187
- Krebs, John R., 151
- La peligrosa idea de Darwin* (Dennett), 31, 76, 101, 205
- La vida oculta de los perros* (Thomas), 21
- Lakoff, George, 204
- lenguaje, 141
- adquisición del, 175-177
- compartir experiencias y, 20-21
- del pensamiento, 67
- intencionalidad intrínseca y, 66-67
- intensionalidad en el, 53-55
- mentes que no se comunican y carencia de, 23-30
- poseedores de mente y uso del, 18-20, 28-30
- traducción y, 20-30
- Lloyd, Dan, 201
- Lorenz, Konrad, 126
- macromoléculas
- agentes y, 32-36
- autoduplicación de, 33-34, 64, 74
- mal, sistemas intencionales y el, 46
- mamataxia, 125-127
- maniobra de distracción, 146-148
- máquinas, seres humanos como máquinas sin mente, 12 (*véase también robots*)
- McFarland, David, 151-152
- medida y proposiciones, 63-64
- Menón* (Platón), 46
- mentes
- categoría moral y posesión de, 15-17
- conocimiento sobre, 11-13
- cuerpos y, 97-100
- defecto en la atribución de, 16-17

- descarga de tareas cognitivas, 160-161
- exceso en la atribución de, 15, 16-18
- hemisemidementados, 30
- previsiones de futuro por medio de la, 73-74
- procesado de la información, 85-87
- protomentes, 30
- prueba científica, 16-18
- seudomentes, 30
- uso de palabras y, 18-20
- uso del lenguaje y, 18-20, 29-30
- Menzel, Emil, 154
- Miller, George, 102
- Millikan, Ruth Garrett, 138, 204
- Minsky, Marvin, 204
- MIT (Instituto de Tecnología de Massachusetts), 27
- Modelo de Borradores Múltiples de la conciencia, 27
- Mondada, Francesco, 166
- Morgan, Elaine, 27, 181, 202
- movimientos inteligentes, 121
- mutaciones, 75
- Nagel, Thomas, 189
- neurotransmisores, 17-18, 92
- Nietzsche, Friedrich, 97-98
- niños
 - adquisición del lenguaje, 175-178
 - disociación en, 192-194
 - seres humanos como máquinas sin mentes, 12
- olfacción, 101, 125, 126
- olfato, sentido del, 101
- ontología, 12
- opacidad referencial, 54, 61
- ordenadores, 75
 - como dispositivos de lectura, 19
 - dispositivos de entrada para, 87
 - que juegan al ajedrez, 44-45
- pájaros
 - maniobras de distracción de los, 145-146
 - ocultación de semillas, 163-165
- palabras, 18-22, 173-174
 - compartir experiencias mediante, 20-21
 - criaturas gregorianas y, 122
 - dispositivos de lectura informatizados y, 19
 - intensionalidad y, 53-55
 - poseedores de mente y usuarios de, 18-20
 - prototipos de conceptos, 179
 - sinceridad o engaños en el uso de las, 21-22
 - traducción y, 20
- Pavlov, Iván, 107
- pensamiento inconsciente, 182-183
- percepción
 - como acción a distancia, 101-102
 - evolución de la, 102
 - identificación del amo por el perro, 137-139
 - identificación y, 135-137
- perros
 - comprensión instintiva entre, 21
 - identificación del amo, 137-139
 - inteligencia de los, 139-140
 - respuesta al comportamiento humano, 194-195
 - selección inconsciente de rasgos humanos en, 194-195
 - uso de conceptos, 188
- Perry, John, 55
- pineal, glándula, 84, 90
- plantas

- adopción del enfoque intencional
hacia las, 47-48
primeros sistemas de control y,
82-83
respuesta al cambio evolutivo en
las, 80-81
sistema circulatorio en las, 39
plasticidad fenotípica, 105
Platón, 46, 179-180
Popper, Karl, 109
poseedores de mente
categoría moral y, 15-17
conocimiento de la pertenencia a
la clase de, 14-16
diferencia entre tener y no tener
mente, 25-26
prueba de la presencia de mente,
16-18
sentir dolor, 28-29
uso del lenguaje y, 18-20
procesado de la información, 86-87,
91-92
proposiciones
como dólares, 63
expresión en forma de frase, 60-
63
medidas y, 63
prosopagnosia, 135-136
Proust, Marcel, 126
prueba de las hipótesis, 141, 144,
160
Psicología popular, 40
«Psicólogos de la naturaleza»
(Humphrey), 144
Quine, W. V. O., 67
rastreo
actividades complejas protectoras
de la vida y, 132-134
cooperador frente a rastreo com-
petitivo, 128
discriminación de fracasos en el,
140-141
equivocaciones en el, 134-135
fototaxia como, 123-125, 132
mamataxia como, 125-127
RIPE (Radiobaliza indicadora de
posición de emergencia), 126
vinculación de sistemas múltiples
en el, 130-132
re-representación, 171
rebote de las gacelas, 148, 151
reflejo «automático», 29
reflejos, 28-29
refuerzo (*véase también* conductis-
mo; criaturas skinnerianas)
aprendizaje latente y, 112
en conexionismo, 106-107
refuerzo negativo y dolor, 116-118
reglas hebbianas del aprendizaje,
107
réplicas, 32
macromoleculares, 33-34, 48, 74
robots, 32, 34-36
RIPE (Radiobaliza indicadora de
posición de emergencia), 126
Ristau, Carolyn, 147, 202
robots, 24
autoduplicación, 32, 34-37
Cog, 27-28
contacto visual con, 26-27
intencionalidad intrínseca y, 69-
72
kheperas, 166
organización de barreras en el en-
torno, 165-168
Rosch, Eleanor, 201
Ryle, Gilbert, 199-200
Sartre, Jean Paul, 185
Searle, John, 65, 71
secretos, guardar, 154-156
Sejnowski, Terence, 203

- selección inconsciente, 195
 selección natural, 73, 104, 181
 diseño y, 74-77
 interpretación del enfoque intencional de la, 76-78
 ritmo de la, 77-80
 sensibilidad a las condiciones cambiantes, 80-82
 sensibilidad
 cambio evolutivo y, 79-82
 estado vegetativo y, 84
 sentencia y 81-82
 sentencia
 búsqueda de la, 114-120
 estado vegetativo y, 84
 percepción del dolor y, 116-120
 sensibilidad y, 81-82
 sistema nervioso y creación de la, 90-91
 sistemas de mantenimiento corporal en el animal, 83-85
 seres humanos
 como clase poseedora de mente, 14-16
 como máquinas sin mentes, 12
 descendientes de los robots que se autoduplican, 34-37
 pseudoagentes, 40
 Seyfarth, Robert, 140, 202
 símbolos, 158
 sinapsis, 92
 sinceridad en la comunicación, 21-22
 sistema nervioso
 transductores y efectores del, 86-91, 92-94
 transferencia de información en el, 91-93
 sistema nervioso autónomo, 85
 sistemas de control
 actividades protectoras de la vida y los, 132-135
 indiferentes al medio, 93-95
 respuesta al medio de los, 82-83
 transductores y efectores en los, 87-91
 Skinner, B. F., 105, 168
 Sócrates, 46
 solipsismo, 12, 14
 Sontag, Susan, 169
 Stein, Lynn Andrea, 27
 sufrimiento, distinción entre dolor y, 190-197
 Talleyrand, Charles-Maurice de, 151
Teeteto (Platón), 179-180
 teoría pictórica de las ideas, 67-68
 Thomas, Elizabeth Marshall, 21
 Thompson, Evan, 201
 Thorndike, E. L., 107
 Tooby, John, 202
 torre de la generación y la prueba, 101-114
 traducción, 20, 29
 transductores en los sistemas nerviosos, 87-91, 92-93
 transparencia referencial, 54
 Trivers, Robert, 133, 202
 Tutiya, Syun, 55
 Valéry, Paul, 73
 Varela, Francisco, 201
 vegetativo, estado, 84
Vehículos (Braitenberg), 124, 200
 virus que se duplican, 33-34
 visión, 101-102
 vitalismo, 37, 95
 von Neumann, John, 32
 Whiten, Andrew, 149, 202
 Wiener, Norbert, 89
 Wittgenstein, Ludwig, 29, 199
 Young, Andrew, 135-136

Daniel C. Dennett es director del centro de Estudios Cognitivos de la Universidad de Tufts y es autor de varios libros y ensayos cortos.

NOTA FINAL

Le recordamos que este libro ha sido prestado gratuitamente para uso exclusivamente educacional bajo condición de ser destruido una vez leído. Si es así, destrúyalo en forma inmediata.

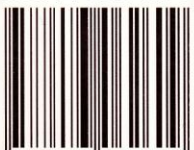
Súmese como voluntario o donante para promover el crecimiento y la difusión de la Biblioteca



Para otras publicaciones visite
www.lecturasinegoismo.com
Referencia: 398

Combinando ideas provenientes de la filosofía, la inteligencia artificial y la neurobiología, Daniel Dennett guía al lector por un fascinante viaje de investigación, explorando posibilidades tan llamativas como las siguientes: ¿Puede cualquiera de nosotros saber lo que pasa por la mente de otra persona? ¿Qué es lo que diferencia la mente humana de las mentes animales, sobre todo de aquellos capaces de manifestar conductas complejas? Si a esos animales se les diera mágicamente la capacidad de hablar ¿desarrollarían esos grupos una inteligencia tan sutil en sus discriminaciones como la nuestra? Una vez que a los robots se les haya dotado de sistemas sensoriales como los que nos proporcionan experiencia ¿llegarán a mostrar los rasgos concretos que se creen que distinguen la mente humana, entre ellos la capacidad de pensar sobre el pensar? Dennett aborda estas cuestiones desde una perspectiva evolutiva. Comenzando por las macromoléculas de ADN y ARN, el autor muestra cómo, paso a paso, la vida animal evolucionó desde la simple facultad de responder ante condiciones ambientales recurrentes con cierta frecuencia hasta los enfoques mucho más potentes de combatir fuerzas muy superiores, o el modo de usar la experiencia pasada para predecir el futuro en situaciones nunca vividas.

ISBN 84-8306-260-7



9 788483 062609

DEBATE

pensamiento