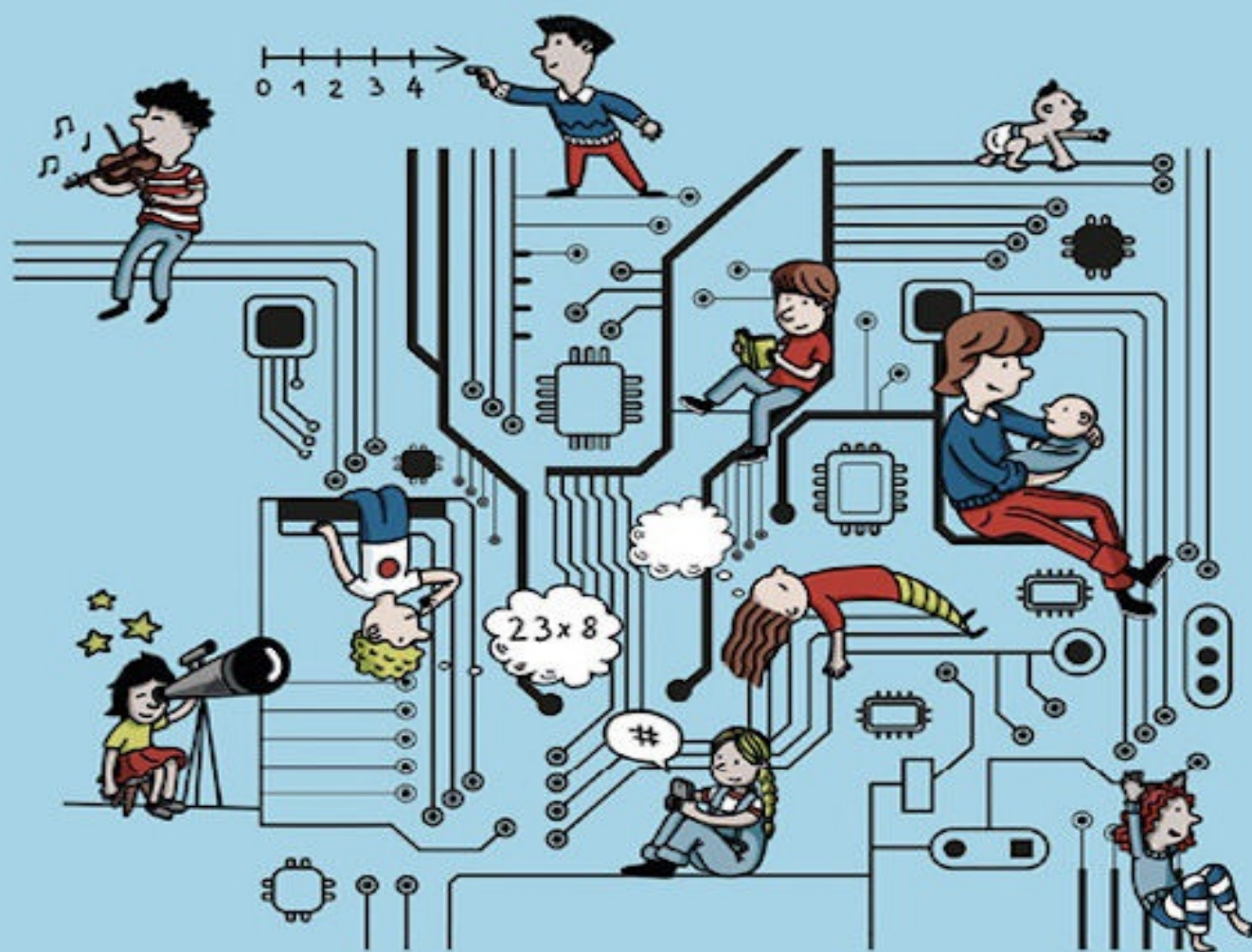


¿CÓMO APRENDEMOS?

Stanislas Dehaene

Los cuatro pilares con los que la educación puede potenciar los talentos de nuestro cerebro



Índice

Cubierta

Índice

Portada

Copyright

Este libro (y esta colección)

Dedicatoria

Epígrafe

Introducción a las ciencias del aprendizaje

¿Por qué el aprendizaje?

Homo docens

Aprender a aprender

El desafío de las máquinas

Parte I. ¿Qué es aprender?

1. Siete definiciones del aprendizaje

Aprender es ajustar los parámetros de un modelo mental

Aprender es aprovechar la explosión combinatoria

Aprender es minimizar los errores

Aprender es explorar el espacio de lo posible Aprender

es optimizar una función de recompensa

Aprender es acotar el espacio de investigación
Aprender es proyectar hipótesis a priori

2. Por qué nuestro cerebro aprende mejor que las máquinas actuales

Lo que aún le falta a la inteligencia artificial
Aprender es inferir la gramática de un dominio
Aprender es razonar como un buen científico

Parte II. Cómo aprende nuestro cerebro

3. El saber invisible: las sorprendentes intuiciones de los bebés

El concepto de objeto
El sentido del número
La intuición de las probabilidades

4. El nacimiento de un cerebro

Desde el comienzo, el cerebro del bebé está bien estructurado
Las autopistas del lenguaje
La autoorganización de la corteza
Los orígenes de la individualidad

5. Lo que adquirimos

¿Qué es la plasticidad cerebral?
El retrato de un recuerdo
Verdaderas sinapsis y falsos recuerdos
La nutrición, una pieza clave del aprendizaje
Posibilidades y límites de la plasticidad sináptica
¿Qué es un período sensible?
Una sinapsis debe estar abierta o cerrada Milagro en Bucarest

6. Reciclen su cerebro

La hipótesis del reciclaje neuronal
Las matemáticas reciclan los circuitos del número
La lectura recicla los circuitos de la visión y de la lengua hablada
Divisiones, ecuaciones y rostros
Los beneficios de un ambiente enriquecido

Parte III. Los cuatro pilares del aprendizaje

7. La atención

Alerta: el cerebro sabe cuándo prestar atención
Orientación: el cerebro sabe a qué prestar atención
Control ejecutivo: el cerebro sabe cómo procesar la información
Aprender a prestar atención
Presto atención si prestas atención
Enseñar es prestar atención a la atención del otro

8. El compromiso activo

Un organismo pasivo no aprende
Procesar en profundidad para aprender mejor
El fracaso de las pedagogías del descubrimiento
Sobre la curiosidad, y cómo despertarla
Saber qué y cuánto sabemos multiplica la curiosidad
Tres maneras de atentar contra la curiosidad en la escuela

9. El error es productivo y dar un buen feedback es garantía de mejores aprendizajes

La sorpresa, motor del aprendizaje
El cerebro está repleto de mensajes de error
Feedback no es sinónimo de castigo
La calificación, ese penoso sucedáneo del feedback
Evaluarse para aprender mejor
La regla de oro: planificar intervalos entre los aprendizajes

10. La consolidación

Liberar los recursos cerebrales
El sueño, un ingrediente clave
El cerebro dormido revive los episodios de la víspera
Sueño de un descubrimiento de verano
El sueño, la infancia y la escuela

Conclusión

Agradecimientos

Bibliografía

Créditos de material gráfico

Stanislas Dehaene

¿CÓMO APRENDEMOS?

Los cuatro pilares con los que la educación puede potenciar
los talentos de nuestro cerebro

Edición al cuidado de
Yamila Sevilla y Luciano Padilla López

Traducción de
Josefina D'Alessio

 **siglo veintiuno**
editores

Dehaene, Stanislas

¿Cómo aprendemos? / Stanislas Dehaene.- 1ª ed.- Ciudad Autónoma de Buenos Aires:
Siglo Veintiuno Editores Argentina, 2019.

Libro digital, EPUB.- (Ciencia que ladra... Serie Mayor // dirigida por Diego Golombek)

Archivo Digital: descarga

Traducción de Josefina D'Alessio // ISBN 978-987-629-974-9

1. Desarrollo mental. 2. Neurociencias. i. D'Alessio, Josefina, trad. II. Título.

CDD 612.825

Cet ouvrage, publié dans le cadre du Programme d'aide à la publication Victoria Ocampo, a bénéficié du soutien de l'Institut français d'Argentine. Esta obra, publicada en el marco del Programa Victoria Ocampo de Ayuda a la Publicación, cuenta con el apoyo del Institut Français d'Argentine.

Título original: *Apprendre! Les talents du cerveau, le défi des machines*

© 2019, Stanislas Dehaene

© 2019, Siglo Veintiuno Editores Argentina S.A.

www.sigloxxieditores.com.ar

Diseño de cubierta: Eugenia Lardiés

Ilustraciones de cubierta: Guido Ferro

Corrección: Mariana Gaitán y Héctor Di Gloria

Digitalización: Departamento de Producción Editorial de Siglo XXI Editores Argentina

Primera edición en formato digital: noviembre de 2019

Hecho el depósito que marca la ley 11.723

ISBN edición digital (ePub): 978-987-629-974-9

Este libro (y esta colección)

Adorable puente se ha creado entre los dos.

Gustavo Cerati, “Puente”

Un primer saber [...] necesario para la formación docente, desde una perspectiva progresista[:] *Enseñar no es transferir conocimiento, sino crear las posibilidades para su propia producción o construcción.*

Paulo Freire, *Pedagogía de la autonomía*

Cuanto más estudio el cerebro humano, más me impresiona.

Stanislas Dehaene, en este mismo libro

En muchas universidades del mundo existen facultades o escuelas de Ciencias de la Educación; por supuesto, siguen las líneas clásicas y las más renovadoras de los últimos siglos en cuanto a pedagogía y otras disciplinas sociales y humanas. Pero a veces da la sensación de que dejaron en suspenso algunas ciencias y que, pasados ya los tiempos de Piaget, hubo cierto divorcio con el trabajo de laboratorio. ¿Qué fue de los experimentos, las evidencias y los conocimientos que la psicología cognitiva, la computación y, muy especialmente, las neurociencias aportan para mejorar nuestras experiencias de aprendizaje y de enseñanza? Mientras vemos naufragar programas educativos, mientras nos quedamos con más errores o mitos que pruebas, nos llegan noticias de los enormes avances de los estudios acerca de nuestra conciencia, el procesamiento de la información en el cerebro o la plasticidad neuronal que deberíamos aprovechar cuanto antes en las aulas.

Al otro lado del río, el estudio del cerebro viene prometiendo una revolución en nuestro conocimiento de cómo y por qué hacemos lo que hacemos y hasta cómo mejorar nuestro desempeño en diversos órdenes de la vida. Así, aunque los frutos son muy recientes, la tentación de vincular la investigación con el mundo educativo siempre ha sido importante. Pero el

pasaje nunca es tan simple y la expectativa es tan grande que esas promesas se exponen al riesgo de resultar engañosas.

Lo cierto es que durante muchos años los grandes logros de los laboratorios neurocientíficos se quedaban allí... en el laboratorio y, aunque supiéramos cada vez más sobre la memoria, la motivación o el alerta, las consecuencias no se veían en las aulas. Quizá por esto mismo, en la década de 1990 –ayer nomás– apareció un trabajo de John Bruer llamado “Neurociencias y educación: un puente demasiado lejos”. La respuesta llegó ya avanzado este siglo, con investigaciones que respondían “es tiempo de construir el puente”, delineando cómo por fin la escuela podía considerarse un campo para aplicar los frutos de la cerebrología.

Uno de los constructores del puente es Stanislas Dehaene, sin duda uno de los más importantes neurocientíficos contemporáneos. Con un rigor y un carisma a toda prueba, nos convence de que si existe un destino para los humanos, es el de aprender, tanto con lo que traemos de fábrica como con ese acelerador de mentes que llamamos escuela. Pero allí, en esa escuela, debemos considerar también el funcionamiento de la memoria (necesaria aunque no goce de la mejor prensa), el rol de la atención, la importancia del sueño y hasta de una buena alimentación. Y, también, explorar ciertas patologías del desarrollo como ventanas abiertas que nos permiten contemplar y comprender las funciones cerebrales.

Si de aprendizaje se trata, no podemos dejar de lado a las máquinas, que prometen (o amenazan con) entender procesos cada vez más complejos e incluso enseñarse a sí mismas, configurando modelos del mundo que se acercan a la realidad y que algunos agitan como un fantasma. Sin embargo, el autor nos tranquiliza recordando que –al menos por ahora– detrás de toda gran máquina hay siempre un gran ser humano. Y que ese mismo ser humano procesa datos, aprende y resuelve problemas mil veces más rápido que cualquier inteligencia artificial que quiera hacerle sombra.

Y es que, en el fondo, ¿por qué aprendemos? ¿Tenemos un instinto de aprendizaje? Podemos considerar las investigaciones clásicas sobre el canto de los pájaros para proponer que sí, lo tenemos. Muchos pajaritos suelen aprender sus músicas de otros tutores a los que imitan, para luego agregar un toque personal que les permitirá desempeñarse mejor en la Ópera entre los árboles. Cual pajaritos, los bebés parecen venir de fábrica con ese instinto, lo que los lleva velozmente a hablar, cantar, comer caramelos o desarmar los juguetes. Las investigaciones de Dehaene y sus colegas demuestran

inequívocamente que el cerebro de los bebés ya cuenta con herramientas aritméticas, lingüísticas y con un GPS muy refinado: el bebé es, desde el comienzo, una máquina de aprender. Crecer es, quizá, exagerarse a uno mismo, poner en práctica ese plan innato que se va enriqueciendo a lo largo de la vida. Como en el Aleph de Borges, el cerebro en desarrollo puede ser “uno de los puntos del espacio que contienen todos los puntos” (algo que Dehaene nos aclara cuando encuentra en la teoría de Thomas Bayes la posibilidad de pensar al niño como a una suerte de estadístico).

Uno de los hallazgos prácticos de este libro es la propuesta de los cuatro pilares del aprendizaje, que permiten mejorar de verdad la educación. Ya los conocerán en detalle, pero vale la pena al menos enumerarlos para que esos principios virtuosos empiecen a abrirse camino en sus neuronas:

- la atención, ese mecanismo que nos permite darle importancia y amplificar ciertas señales e ignorar otras,
- el compromiso activo, o curiosidad, que nos obliga a tener cerebros exigentes y motivados en el aula,
- la detección y corrección de errores (el buen *feedback* que se aleja diametralmente del castigo frente al error) y
- la consolidación, esto es, la puesta en marcha de los diversos pasos en la formación de las memorias.

Con esos cuatro jinetes del aprendizaje, y desplegando la evidencia empírica que funda cada una de sus afirmaciones, Dehaene pone a la vista cuáles son las consecuencias prácticas de sus investigaciones.

Por si fuera poco, luego de este extraordinario paseo por los recovecos del cerebro que aprende, también conoceremos a otro Dehaene, el que se calza el traje de hacedor –no por nada es el presidente del primer Consejo Científico del Ministerio de Educación de Francia– y recuerda que la educación pública debe ser siempre una de las primeras prioridades del Estado. Así, en la conclusión nos regala trece recomendaciones para optimizar el potencial de los niños en el proceso de enseñanza y aprendizaje. Del laboratorio y la mente del autor al aula y a nuestras casas, sin escalas.

Adorable puente se ha creado entre las neurociencias y la educación. Stanislas Dehaene es ese puente. Este libro es ese puente. Podemos cruzar

tranquilos.

La Serie Mayor de Ciencia que ladra es, al igual que la Serie Clásica, una colección de divulgación científica escrita por científicos que creen que ya es hora de asomar la cabeza por fuera del laboratorio y contar las maravillas, grandezas y miserias de la profesión. Porque de eso se trata: de contar, de compartir un saber que, si sigue encerrado, puede volverse inútil.

Ciencia que ladra... no muerde, sólo da señales de que cabalga.

Diego Golombek

*Para Aurore, que acaba de nacer,
y para todas aquellas y todos aquellos
que fueron bebés alguna vez*

Comenzad, pues, por estudiar mejor a vuestros alumnos;
seguramente no los conocéis.

Jean-Jacques Rousseau, *Emilio o De la educación* (1762)

Cosa extraña y casi pasmosa: conocemos todos los recodos del cuerpo humano, ya catalogamos todos los animales del planeta, describimos y bautizamos cada brizna de hierba, y durante siglos dejamos las técnicas psicológicas libradas a su empirismo, como si fueran menos importantes que las utilizadas por quienes nos curan, nos crían y educan o cultivan los campos.

Jean Piaget, “La pedagogía moderna” (1949)

Si no sabemos cómo aprendemos, ¿cómo podríamos saber cómo enseñar?

Leo Rafael Reif, presidente del MIT (2017)

Introducción a las ciencias del aprendizaje

En septiembre de 2008, el encuentro con un niño fuera de lo común me forzó a revisar mis ideas sobre el aprendizaje. Estaba visitando uno de los hospitales de la Rede Sarah de Brasilia, esos centros de salud de arquitectura blanca que desarrolló João Filgueiras, inspirado en la estética de Oscar Niemeyer. Esa red de nueve unidades se especializa en la “rehabilitación” neurológica; desde hace unos diez años mi laboratorio (NeuroSpin, dependiente del Inserm)[\[1\]](#) sostiene proyectos en colaboración con ella. Su directora, la destacada psicóloga y neurocientífica Lúcia Braga, me propuso conocer a uno de los pacientes: Felipe, un niño de 7 años que había transcurrido la mitad de su vida en el hospital. Según me explicó Lúcia, a los 4 años este niño había recibido una bala perdida (por desgracia, en Brasil no es algo tan infrecuente). El proyectil le seccionó la médula espinal, de modo que lo dejó casi completamente paralizado en los cuatro miembros, es decir, cuadripléjico. La bala también arrasó con las áreas visuales de la corteza: Felipe quedó ciego. Para ayudarlo a respirar, se le hizo una traqueotomía en la base del cuello. Desde hace tres años, vive en una habitación del hospital, encerrado en su cuerpo inerte.

En el pasillo que me lleva a su habitación, me preparo mentalmente para enfrentarme a un niño con una gran discapacidad. Y me encuentro con... Felipe, un pequeño como todos los de 7 años, con el rostro lleno de vida, conversador y de una curiosidad inagotable. Habla a la perfección, con un vocabulario rico, y me pregunta con picardía sobre las palabras de mi lengua materna, el francés. Descubro que es un apasionado de los idiomas y que nunca pierde la ocasión de enriquecer su vocabulario trilingüe (portugués, inglés y español). Si bien es ciego y está inmovilizado en la cama, viaja con su imaginación y se distrae creando sus propios cuentos; el equipo del hospital lo alienta en la tarea. En pocos meses, Felipe aprendió a dictar sus historias a un asistente y luego a escribirlas con ayuda de un teclado conectado a una terminal informática y a una placa de sonido. Los pediatras y

los terapeutas del lenguaje de esa institución, deslumbrados, se turnan junto a la cama de Felipe para transformar esos relatos en verdaderos libros táctiles ilustrados con imágenes en relieve que él palpa con orgullo, con la poca sensibilidad de que dispone. Sus libros hablan de héroes y heroínas, de montañas y de lagos que jamás volverá a ver, pero con los que sueña como cualquier otro niño pequeño.

El encuentro con Felipe me conmocionó y al mismo tiempo me persuadió a optar por una exploración de lo que, sin lugar a dudas, es el mayor talento de nuestro cerebro: la capacidad de aprender. En efecto, este niño plantea a la vez una hermosa lección de esperanza y un desafío para la neurociencia. ¿Cómo puede ser que las facultades cognitivas resistan a una alteración tan grande del entorno? ¿Por qué Felipe y yo podemos compartir los mismos pensamientos, aunque tengamos experiencias sensoriales tan diferentes? ¿Cómo logran distintos cerebros humanos converger en los mismos conceptos, sin importar cómo ni cuándo los aprendan?

Muchos neurocientíficos son empiristas: consideran, como John Locke, que el cerebro obtiene sus conocimientos de su ambiente. Según ellos, la principal propiedad de los circuitos corticales es la plasticidad, la capacidad de adaptarse. En efecto, las células nerviosas ajustan permanentemente sus sinapsis en función de la información de entrada que reciben. Pero en este caso, dado el impedimento del ingreso de información visual y motriz, Felipe debería haberse convertido en un ser profundamente diferente. ¿Por obra de qué milagro logró desarrollar facultades cognitivas estrictamente normales?

El caso de Felipe está lejos de ser un hecho aislado: todos conocen las historias de Helen Keller o de Marie Heurtin, las dos fueron sordas y ciegas de nacimiento que, tras duros años de aislamiento social, aprendieron lengua de señas y lograron desarrollarse como pensadoras y escritoras brillantes.^[2] A lo largo de estas páginas, ustedes y yo tendremos otros encuentros que, según espero, cambiarán por completo sus ideas sobre el aprendizaje. Conocerán a Emmanuel Giroux, ciego desde los 11 años, eximio matemático especializado en geometría. Parafraseando al zorro de *El Principito* de Saint-Exupéry, Emmanuel afirma convencido: “En geometría, lo esencial es invisible a los ojos; solo se puede ver bien con la mente”. ¿Cómo llega este hombre ciego a pasearse ágilmente por los abstractos espacios de la geometría algebraica, a manipular planos, esferas y poliedros, sin haberlos visto siquiera una vez? Descubriremos que utiliza los mismos circuitos cerebrales que otros matemáticos, con la única salvedad de que su corteza

visual, lejos de permanecer inactiva, se recicló también para hacer matemáticas.

Además, les presentaré a Nico, un joven pintor que, durante una visita al museo Marmottan-Monet de París, logró hacer una excelente copia del famoso cuadro de Monet *Impresión, sol naciente* (figura 1). ¿Qué tiene esto de excepcional? Nada, excepto que su cerebro no posee más que un solo hemisferio, el izquierdo: ¡cuando Nico tenía 3 años le fue extirpada casi la totalidad del hemisferio derecho! Su cerebro, entonces, aprendió a alojar en un solo hemisferio todos sus talentos: el habla, la lectura y la escritura, el dibujo, la pintura, la informática e incluso la esgrima, deporte del que es campeón internacional en silla de ruedas. Por favor, olviden todo lo que crean saber acerca de los respectivos roles de los dos hemisferios, porque la vida de Nico prueba que es completamente posible convertirse en un artista sin ayuda del hemisferio derecho: la plasticidad cerebral parece obrar milagros.

En nuestra travesía visitaremos también los siniestros orfanatos de Bucarest donde se tenía a los niños en un estado de cuasiabandono desde su nacimiento. Al ampliar un poco nuestro rango de observación, notaremos que tiempo después, pese a todo, algunos de ellos, que fueron adoptados antes de cumplir 1 o 2 años, tuvieron una trayectoria escolar casi normal.

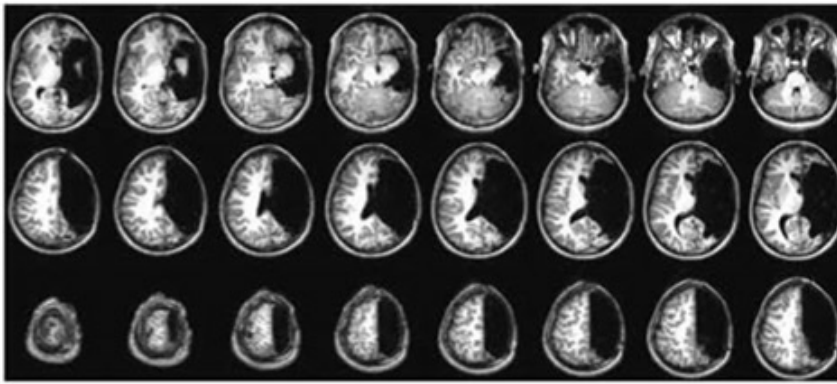


Figura 1. La plasticidad neuronal a veces logra compensar déficits impresionantes. Desde sus 3 años, el joven pintor Nico no posee más que un hemisferio de su cerebro, el izquierdo. Esto no le impidió volverse un artista consumado, capaz de pintar excelentes copias (abajo, su versión de *Impresión, sol naciente*, cuadro de Manet) y obras propias (arriba).

Todos estos ejemplos revelan la extraordinaria resiliencia del cerebro humano: ni siquiera un trauma grave como la ceguera, la pérdida de un hemisferio o el aislamiento social logra extinguir la chispa del aprendizaje. El lenguaje, la lectura, las matemáticas, la creación artística: todos estos talentos singulares de la especie humana, que ningún otro primate posee, resisten un daño masivo como la pérdida de un hemisferio, de la vista o de la motricidad. Aprender es un principio vital, y el cerebro humano tiene un enorme potencial para la plasticidad: para modificarse por sí solo y adaptarse. Pero en este itinerario descubriremos también contraejemplos trágicos, casos en los cuales el aprendizaje parece congelarse. Tomemos el ejemplo de la alexia pura, la imposibilidad de leer la mínima palabra. Investigué en persona qué les sucedía a muchos adultos, excelentes lectores, a quienes un minúsculo accidente cerebrovascular, limitado a una región muy pequeña del cerebro, volvió incapaces de descifrar incluso palabras tan simples como “pez” o “mar”. Recuerdo a una mujer brillante, trilingüe, lectora fiel del diario *Le Monde*, que se afligía porque, luego de su lesión cerebral, cada página del diario parecía escrita en hebreo. Su motivación para reaprender a leer estaba a la altura del desasosiego que había soportado. Sin embargo, dos años de esfuerzos no le permitieron superar el nivel de lectura de un niño de primer grado, que vocaliza letra por letra y tiene dificultades con cada palabra. ¿Por qué ya no podía aprender? ¿Y por qué algunos niños disléxicos, discalculicos o dispráxicos sienten la misma desesperanza radical al encarar la adquisición de la lectura, el cálculo o la escritura, mientras que otros transitan esos campos sin problema?

La plasticidad cerebral parece caprichosa: a veces se repone de déficits enormes y a veces deja con una discapacidad permanente a niños y adultos por demás motivados e inteligentes. ¿Depende de circuitos específicos? Y esos circuitos, ¿pierden su plasticidad a lo largo de los años? La plasticidad, ¿se puede reactivar? ¿Cuáles son las reglas que la gobiernan? ¿Cómo hace el cerebro de niñas y niños para ser tan eficaz desde el nacimiento y a lo largo de la infancia? ¿Qué algoritmos implantados por la evolución permiten que nuestros circuitos cerebrales elaboren una representación del mundo?

¿Comprender esos algoritmos nos garantizaría aprender mejor y más rápido?
¿Podríamos inspirarnos para construir máquinas más eficaces, inteligencias artificiales que nos imiten o incluso nos superen? Estas son algunas de las preguntas a las cuales este libro intenta dar respuesta, desde una perspectiva decididamente multidisciplinaria, valiéndose de los hallazgos recientes en los campos de las ciencias cognitivas y de las neurociencias, pero también de la inteligencia artificial y de la educación.

¿Por qué el aprendizaje?

Es más que lógico tomar como punto de partida la indagación de por qué debemos aprender. La existencia misma de la facultad del aprendizaje nos plantea una serie de preguntas. ¿No sería mejor que nuestros hijos supieran hablar y reflexionar desde el primer día, como Atenea, de quien cuenta la leyenda que salió del cráneo de Zeus provista de una armadura completa, casco y lanza, dando un grito de guerra? ¿Por qué no nacemos precableados, con un *software* programado de antemano y dotado de todos los conocimientos necesarios para nuestra supervivencia? En la lucha por la supervivencia que describe Charles Darwin, un animal que naciera maduro, con mayor conocimiento que los otros, ¿no debería al fin y al cabo ganar y propagar sus genes? Y entonces, ¿por qué la evolución habrá inventado el aprendizaje?

Mi respuesta es muy sencilla: el precableado completo del cerebro no es posible ni deseable. ¿De verdad es algo imposible? Sí, porque si nuestro ADN debiera especificar todos los detalles de nuestros conocimientos, simplemente no dispondría de la capacidad de almacenamiento necesaria. Nuestros 23 cromosomas incluyen 3.000.000.000 de pares de “letras” A, C, G, T: las moléculas adenina, citosina, guanina y timina. ¿Qué cantidad de información implica esto? La información se mide en bits: una decisión binaria, 0 o 1. Visto que cada una de las cuatro letras del genoma codifica 2 bits (podemos codificarlos como 00, 01, 10 y 11), esto da un total de 6.000.000.000 de bits. A primera vista, parece un número importante, pero atención: en las computadoras actuales contamos en bytes, que son secuencias de 8 bits. El genoma humano se reduce, entonces, a cerca de 750

megabytes: ¡el contenido de una pequeña memoria USB! Y este cálculo elemental ni siquiera contempla la gran cantidad de redundancias que tienen cabida en nuestro ADN.

A partir de esta modesta dote de informaciones heredadas de millones de años de evolución, nuestro genoma –inicialmente reducido a una sola célula, el óvulo fecundado– logra organizar todo el cuerpo, cada molécula de cada una de las células del hígado, los riñones, los músculos, y por supuesto, el cerebro: 86.000.000.000 de neuronas, billones de conexiones, sí, miles de miles de millones... ¿cómo podría definir las una por una? Si damos por sentado que cada conexión solo codifica 1 bit –lo cual es, por cierto, una subestimación–, la capacidad de nuestro cerebro está en el rango de los 100 terabytes (alrededor de 10^{14} bits), es decir, unas 15.000 veces más que la información contenida dentro del genoma humano. Nos vemos ante una paradoja: ¡el fabuloso palacio que es nuestro cerebro tiene capacidad para almacenar al menos quince mil veces más detalles que los planos del arquitecto que se usaron para construirlo! No veo más que una explicación: la estructura general del palacio se construye según las líneas rectoras del arquitecto (nuestro genoma), mientras que los detalles se dejan a cargo del contratista que los adapta al terreno (el entorno). Precablear un cerebro humano en todos sus detalles sería rigurosamente imposible; por ende, el aprendizaje debe prolongar la obra de los genes.

Este simple argumento contable, sin embargo, no es suficiente para explicar por qué el aprendizaje está universalmente extendido en el mundo animal. En efecto, hasta los organismos simples y desprovistos de corteza, como la lombriz, la mosca de la fruta o los pepinos de mar, aprenden una buena cantidad de sus comportamientos. Tomemos por ejemplo el *Caenorhabditis elegans*, el pequeño gusano del grupo que llamamos “nematodos”. En los últimos veinte años, este animalito de pocos milímetros se volvió una estrella de los laboratorios, en parte porque su arquitectura posee un increíble determinismo biológico y puede ser analizada hasta en sus menores detalles: la mayor parte de los individuos cuenta con un total exacto de 959 células, de las cuales 302 son neuronas. Todas sus conexiones son conocidas y reproducibles. Y sin embargo, aprende (Bessa y otros, 2013; Kano y otros, 2008; Rankin, 2004). En un comienzo, los investigadores lo consideraban una suerte de autómatas solo capaz de nadar hacia delante o hacia atrás, pero luego notaron que poseía al menos dos formas de aprendizaje: por habituación y por asociación. La habituación significa que el organismo se adapta a la

presencia repetida de un estímulo (por ejemplo, una molécula en el agua) y finalmente ya no responde a él. La asociación, por otro lado, consiste en describir y retener en la memoria qué elementos del ambiente predicen las fuentes de alimento o de peligro. Quedó comprobado que este gusano es un campeón de la asociación, capaz de recordar que en el pasado determinados gustos, olores o temperaturas estaban asociados al alimento (bacterias) o a moléculas repulsivas (el olor del ajo), y de utilizar esa información para elegir un camino óptimo a través de su ambiente.

Con tan pocas neuronas, el comportamiento de este gusano bien podría haber estado precableado por completo. Si no lo está, es porque adaptarse a las condiciones específicas en que vive resulta ventajoso para su supervivencia. Incluso dos organismos genéticamente idénticos no nacen siempre en el mismo ecosistema. En el caso del nematodo, la capacidad de adaptar rápidamente su comportamiento a la densidad, la química y la temperatura del lugar donde está le permite ser más eficiente. Por lo general, todos los animales deben adaptarse con rapidez a las condiciones imprevisibles de su existencia efectiva. Desde luego, la selección natural, el algoritmo increíblemente eficiente de Darwin, logra adaptar cada organismo a su nicho ecológico, pero lo hace con una lentitud desoladora: antes de que una mutación favorable pueda aumentar la supervivencia, hace falta que varias generaciones mueran en el intento. La facultad del aprendizaje, por su parte, actúa mucho más rápido: puede modificar el comportamiento en unos pocos minutos. Y es esto lo interesante del aprendizaje: la posibilidad de ajustarse, lo más rápido posible, a condiciones imprevisibles.

Por todos esos motivos, hubo una evolución en el aprendizaje. A lo largo del tiempo, los animales provistos de una capacidad siquiera rudimentaria de aprender tuvieron mayores oportunidades de supervivencia que aquellos que tenían conductas inamovibles. Además, eran más propensos a transmitir la información a la generación siguiente, que para entonces ya disponía de algoritmos de aprendizaje. Con esto, la selección natural propició el surgimiento del aprendizaje. De por sí, los algoritmos implican el descubrimiento de un buen recurso: es útil dejar que ciertos parámetros cambien enseguida para acomodarse mejor a las condiciones más variables de su ambiente.

Algunos factores de la física del mundo son estrictamente invariables: la gravitación es universal y la propagación de la luz o de los sonidos en el aire no cambian de un día para el otro, y he aquí por qué –¡afortunadamente!– no

tenemos necesidad de aprender a hacer crecer las orejas, los ojos, o los laberintos del sistema vestibular que miden la aceleración de nuestro cuerpo: todas estas propiedades se codifican genéticamente. En cambio, muchos otros parámetros como el espacio entre los ojos, el peso y la longitud de brazos y piernas o el tono de la voz, varían, y por ese motivo el cerebro debe aprenderlos. Nuestro cerebro es resultado de una solución de compromiso: mucho de innato (heredamos de la larga historia evolutiva gran parte de la circuitería responsable de codificar las grandes categorías intuitivas con las cuales subdividimos el mundo en imágenes, sonidos, movimientos, objetos, animales, personas, causas...), pero quizá todavía más de adquirido, gracias a ese sofisticado algoritmo que nos permite refinar esas competencias precoces en función de nuestra experiencia.

Homo docens

Con todos esos elementos a nuestro alcance, si hiciera falta resumir en una sola palabra el talento que caracteriza a nuestra especie, optaría por el verbo “aprender”. Más que ser integrantes de la especie *Homo sapiens*, formamos parte de *Homo docens*, la especie que se enseña a sí misma. Lo que sabemos del mundo, en su mayor parte, no es algo que se nos haya dado: lo aprendimos del ambiente o del entorno. Ningún otro animal pudo descubrir como nosotros los secretos del mundo natural. Gracias a la extraordinaria flexibilidad de sus aprendizajes, nuestra especie logró salir de su sabana natal para cruzar desiertos, montañas, océanos y, en apenas varios miles de años, conquistar las islas más remotas, las grutas más profundas, los hielos marinos más inaccesibles e inhóspitos, y hasta la luna. Desde la conquista del fuego y la fabricación de herramientas hasta la invención de la agricultura, la navegación (marina, aérea y extraplanetaria) o la fisión nuclear, la historia de la humanidad no es otra cosa que una reinención constante. La fuente secreta de todos estos logros es una sola: la extraordinaria facultad de nuestro cerebro de formular hipótesis y seleccionarlas para transformar algunas de ellas en conocimientos sólidos acerca del ambiente.

Nuestra especie hizo del aprendizaje su especialidad. En el cerebro, miles de millones de parámetros son libres de adaptarse al medio, la lengua, la

cultura, los padres, la alimentación... Esos parámetros son elegidos cuidadosamente: dentro del cerebro, la evolución definió, con precisión, qué circuitos están precableados y cuáles están abiertos al ambiente. En nuestra especie, la incidencia del aprendizaje es particularmente vasta, porque la infancia se prolonga muchos años. Gracias al lenguaje y a las matemáticas, nuestros dispositivos de aprendizaje tienen la posibilidad de transitar amplios espacios de hipótesis que se incrementan en una combinatoria potencialmente infinita, incluso si siempre se apoyan sobre bases fijas e invariables, heredadas de la evolución.

En fecha más reciente, la humanidad descubrió que esta notable capacidad de aprendizaje puede verse aún más fortalecida con ayuda de una institución: la escuela. La pedagogía activa es un privilegio de nuestra especie: ningún otro animal se toma el tiempo de enseñarles nuevos talentos a sus hijos, deliberadamente, prestando atención a sus dificultades y errores. La invención de la escuela, que sistematiza la instrucción informal presente en todas las sociedades humanas, supuso un incremento significativo en el potencial cerebral. Comprendimos que necesitábamos aprovechar esta pródiga plasticidad del cerebro del niño para inculcarle un máximo de informaciones y talentos. A lo largo de los años, las posibilidades de la escolarización no dejaron de ganar eficacia: comenzaron cada vez más temprano, desde el jardín de infantes, y se extendieron cada vez más. E incluso cada vez más mentes se benefician de una enseñanza superior en la universidad, auténtica sinfónica neuronal en que los circuitos cerebrales ponen a tono y potencian sus mejores talentos.

Hoy en día, la educación puede considerarse el principal acelerador de nuestro cerebro. Su lugar privilegiado, que recuerda por qué debe situarse entre los primeros puestos de las inversiones del Estado, se justifica fácilmente: sin ella, los circuitos corticales serían diamantes en bruto. La complejidad de las sociedades contemporáneas debe su existencia a las múltiples mejorías que la educación aportó a nuestra corteza: la lectura, la escritura, el cálculo, el álgebra, la música, las nociones de tiempo y espacio, el refinamiento de la memoria... ¿Sabían, por ejemplo, que la capacidad de memoria de corto plazo de un analfabeto, la cantidad de sílabas o de cifras que puede repetir, es casi una tercera parte de la de una persona escolarizada? ¿O que medidas tales como el coeficiente intelectual se incrementan varios puntos por cada año adicional de educación y alfabetización?

Aprender a aprender

La educación multiplica las ya considerables facultades del cerebro, pero ¿podría ser incluso mejor? En la escuela, la universidad o el trabajo, forzados a adaptarnos cada vez más rápido, hacemos malabares con nuestros algoritmos cerebrales de aprendizaje. Sin embargo, ese despliegue espectacular sucede de modo intuitivo, sin jamás haber aprendido a aprender. Nadie nos explicó las reglas que hacen que el cerebro memorice y comprenda o, por el contrario, olvide y se equivoque. Es una pena, porque los datos abundan. Un excelente sitio inglés, el de la Education Endowment Foundation (EEF), aporta largas listas de las más exitosas intervenciones pedagógicas.^[3] Y una de las más eficaces, según ellos, es la metacognición, vale decir, el hecho de conocer mejor el funcionamiento cognitivo. Saber aprender es uno de los factores más importantes del éxito escolar.

Por suerte, hoy en día sabemos mucho acerca de cómo funciona el aprendizaje. A lo largo de los últimos treinta años, la investigación en las fronteras de la ciencia de la computación, la neurobiología y la psicología cognitiva, permitió comprender los algoritmos que utiliza el cerebro, los circuitos involucrados, los factores que modulan su eficacia y los motivos de su tan excepcional eficiencia en los humanos. El funcionamiento de la memoria, el papel que desempeña la atención, la importancia del sueño son descubrimientos igualmente ricos en consecuencias para todos nosotros. Me ocuparé de cada una de estas cuestiones a lo largo de estas páginas. Por eso, espero que cuando cierren este libro sepan mucho más sobre sus propios procesos de aprendizaje. Me parece fundamental que cada niño, cada adulta, tenga plena conciencia del potencial de su propio cerebro y también, por supuesto, de sus límites. Al realizar una disección sistemática de los algoritmos mentales y los mecanismos cerebrales, las ciencias cognitivas contemporáneas recuperan y actualizan la célebre máxima de los Siete Sabios, inscripta en el Oráculo de Delfos: “Conócete a ti mismo”. Hoy en día, ya no es cuestión de practicar la introspección, sino de conocer mejor la refinada mecánica neuronal que da lugar a los pensamientos, para así dominarla mejor y ponerla al servicio de nuestras necesidades, metas y deseos.

La emergente ciencia del aprendizaje es de especial importancia para quienes hacen de la enseñanza su actividad profesional: docentes y educadores. Tengo la profunda convicción de que no podemos enseñar de

una manera conveniente sin poseer un modelo mental de lo que ocurre dentro de la cabeza del niño: cuáles son sus intuiciones, correctas o erróneas, cuáles son las etapas por las que debe pasar en su avance y qué factores lo ayudan a desarrollar sus capacidades.

Si bien las neurociencias cognitivas no tienen todas las respuestas, gracias a ellas en la actualidad sabemos que todos los niños comienzan la vida con una arquitectura cerebral similar: un cerebro de *Homo sapiens*, que difiere radicalmente de los de otros simios. Desde luego, no niego que los cerebros varían: tanto las peculiaridades de nuestros genomas como las excentricidades de nuestro desarrollo cerebral aseguran distintas fuerzas y velocidades de aprendizaje. Con todo, el bloque básico de circuitos es el mismo en cada cual (y otro tanto sucede con la organización de los algoritmos de aprendizaje). Entonces, hay principios fundamentales que cada modalidad de enseñanza, si pretende ser eficaz, debe respetar. En este libro, daremos numerosos ejemplos. Las habilidades que los niños muy pequeños tienen para el lenguaje, la aritmética, la lógica o la estimación de probabilidades demuestran la existencia de intuiciones precoces y abstractas sobre las cuales debe apoyarse la enseñanza. Todas ellas se potencian si se enfoca la atención, se adopta un compromiso activo, se reconocen y rectifican los errores (lo que se conoce como *feedback*) y se practica un ciclo de experimentación durante el día y de consolidación a la noche. Esos son para mí los cuatro pilares del aprendizaje, porque, como veremos, los encontramos desde que se echan los cimientos del edificio del algoritmo universal del aprendizaje humano, presente en todos los cerebros, tanto en la infancia como en la edad adulta.

Al mismo tiempo, nuestros cerebros presentan variaciones individuales, y en algunos casos extremos puede aparecer una patología. La realidad de las patologías del desarrollo como la dislexia, la discalculia, la dispraxia o los trastornos de atención ya está confirmada por completo, y hay estrategias para detectarlas y compensarlas. Uno de los objetivos de este libro es dar mayor difusión a estos conocimientos, en busca de que cada docente, así como cada familia, pueda derivar las consecuencias y adaptar su manera de enseñar. Por supuesto, hay variaciones muy grandes entre lo que los distintos niños saben, pero eso nunca significa que dejen de tener los mismos algoritmos de aprendizaje. Así, los recursos o “trucos del oficio pedagógico” que resultan más efectivos con todos los niños son aquellos que tienden a

serlo también con quienes tienen déficits de aprendizaje: solo hay que aplicarlos con mayor enfoque, paciencia, sistematicidad y tolerancia al error.

El último factor es decisivo. Si bien la detección del error y la consiguiente respuesta son indispensables, muchos niños pierden confianza, motivación y curiosidad porque en vez de una corrección reciben un castigo. Hay que prestar mucha atención para desvincular por completo error y castigo (y todavía queda mucho por decir acerca del triste papel que desempeñan las distintas instancias escolares en la perpetuación de esta confusión). Las emociones negativas aplastan el potencial de aprendizaje de nuestro cerebro, mientras que un entorno que haya desterrado el miedo y la amenaza puede reabrir las puertas de la plasticidad neuronal. No habrá un verdadero progreso en el campo de la educación si a la vez no se integran las facetas cognitiva y emocional del desarrollo del cerebro, dos ingredientes indispensables desde la perspectiva de la neurociencia cognitiva actual.

El desafío de las máquinas

Hoy en día, la inteligencia humana se enfrenta a un nuevo desafío: ya no es la única que sabe aprender. En todos los campos del saber existen algoritmos que desafían a nuestra especie, porque aprenden a reconocer los rostros o las voces, a transcribir el habla, a traducir las lenguas extranjeras, a controlar las máquinas, e incluso a jugar al ajedrez o al go (y muchas veces aprenden todo eso mejor que nosotros). Los algoritmos de *machine learning* nutren una industria multimillonaria que, cada vez más, se inspira en el cerebro humano. ¿Cómo funcionan estos algoritmos artificiales? ¿Sus principios pueden ayudarnos a comprender qué es el aprendizaje? ¿Lograron ya imitar el funcionamiento de nuestro cerebro o todavía les queda mucho por aprender?

Si bien los avances actuales de la informática son fascinantes, sus límites son claros. Los algoritmos convencionales de *deep learning* no hacen otra cosa que imitar una pequeña parte del funcionamiento del cerebro: la que corresponde a las primeras etapas del tratamiento sensorial, los famosos 200 o 300 milisegundos (ms) durante los cuales el cerebro opera de un modo no consciente. Eso no quiere decir que este tratamiento sea superficial: en una fracción de segundo, nuestro cerebro puede reconocer un rostro o una palabra

y, en este segundo caso, además asociarla a un contexto, comprenderla e integrarla a una pequeña frase... Sin embargo, este proceso todavía es estrictamente ascendente –*bottom-up* en inglés–, vale decir, sin una verdadera capacidad de reflexión. Solo en una segunda etapa, tanto más lenta, consciente y reflexiva, nuestro cerebro logra desplegar todas sus capacidades de razonamiento, de inferencia, de flexibilidad que las máquinas actuales todavía están lejos de igualar. Incluso las arquitecturas informáticas más avanzadas están muy por debajo de la capacidad de las crías humanas a la hora de construir modelos abstractos del mundo.

Aun dentro de su dominio preferencial, el reconocimiento veloz de las formas, los algoritmos actuales chocan con un segundo problema: son mucho menos eficaces que el cerebro humano. Y en el estadio que alcanzó en la actualidad, el *machine learning* consiste en poner a funcionar procesadores en millones, e incluso miles de millones, de pruebas de práctica. Esa modalidad pierde de vista la economía de los datos, ya que considera que *machine learning* es sinónimo de *big data*: sin una enorme cantidad de información, los algoritmos no consiguen extraer conocimientos abstractos generalizables a situaciones nuevas. En síntesis, no hacen el mejor uso de los datos.

Y en este virtual certamen, el bebé más pequeño se lleva las palmas sin mayores esfuerzos de su parte; no necesita más que una o dos repeticiones para aprender una palabra nueva. Su cerebro saca el mayor provecho de una porción extremadamente escasa de datos, capacidad que todavía resulta elusiva para los procesadores de última generación. Los algoritmos neuronales de aprendizaje suelen alcanzar un cómputo cercano al óptimo. A menudo, consiguen extraer la verdadera esencia aun de una observación ínfima. Si desean alcanzar el mismo desempeño, los investigadores en informática deben inspirarse en los numerosos trucos del aprendizaje que la evolución integró en nuestro cerebro: la atención, por ejemplo, que nos permite seleccionar y amplificar una información pertinente; o bien el sueño, un algoritmo mediante el cual el sistema nervioso central hace la síntesis de los aprendizajes del día. Comienzan a ver la luz máquinas provistas de estas propiedades, y su desempeño no deja de crecer. Sin duda, serán ellas las que, mañana, competirán con nuestra mente.

Una teoría emergente, llamada “teoría del cerebro estadístico”, explica que el cerebro humano todavía es superior a las máquinas porque actúa como un estadístico; constantemente atento a probabilidades e incertidumbres,

optimiza su capacidad de aprendizaje. Un teorema matemático lo verifica: solo la manipulación de probabilidades –es decir, de las incertidumbres sobre lo que aprendimos–, permite obtener el máximo provecho de cada información. Parece que a lo largo de su evolución nuestro cerebro descubrió este truco que consiste en tener un registro constante de la incertidumbre asociada a cada información, y actualizarlo durante cada aprendizaje.

Numerosos datos experimentales respaldan esta hipótesis. Hasta los bebés comprenden las probabilidades, y estas parecen profundamente inscriptas en los circuitos cerebrales. Cada niño actúa como un pequeño científico en ciernes: a sus espaldas, su cerebro formula hipótesis, verdaderas teorías científicas que cada experiencia pone a prueba. El razonamiento sobre las probabilidades –también inconsciente, pero con fuerte arraigo en la lógica de nuestros aprendizajes– permite rechazar gradualmente las hipótesis falsas y conservar solo las teorías que funcionan. A diferencia de otras especies animales, los seres humanos parecen poseer algoritmos muy particulares para formular teorías del mundo exterior. Solo el *Homo sapiens* logra generar de manera sistemática pensamientos simbólicos abstractos y actualizar su plausibilidad ante nuevas observaciones.

En la actualidad, nuevos algoritmos –llamados “bayesianos” en honor al reverendo Thomas Bayes (1701-1761), quien bosquejó esta teoría ya en el siglo XVIII– comienzan a formalizar e implementar esta visión del aprendizaje. Apuesto a que van a revolucionar el *machine learning*; de hecho, veremos que ya son capaces de obtener información abstracta con una eficacia próxima a la de un científico humano.

* * *

Ya podemos levar anclas y explorar juntos lo que hoy en día comprendemos del aprendizaje. Les propongo un viaje en tres etapas.

En la primera parte, titulada “¿Qué es aprender?”, analizaremos las teorías actuales del aprendizaje a la luz de su implementación concreta en las computadoras. Será el momento de formalizar lo que significa aprender. Y la idea es sencilla: aprender es configurar, en los circuitos de silicio o neurales, un modelo interno del mundo que nos rodea. Cuando uno camina por una ciudad desconocida, arma en su mente un mapa de su aspecto externo, un

modelo en miniatura de sus calles y pasajes. Lo mismo sucede cuando una niña empieza a aprender a andar en bicicleta y perfila en sus circuitos neurales una simulación inconsciente a propósito del modo en que las acciones ejercidas sobre los pedales y el manubrio afectan la estabilidad de la bicicleta misma. Y de un modo similar, un algoritmo informático de aprendizaje de reconocimiento facial adquiere plantillas de rasgos posibles respecto de ojos, narices, bocas y sus combinaciones.

Pero ¿cómo generamos un modelo mental adecuado? Como veremos, la mente de quien aprende puede compararse con una máquina gigantesca con millones de parámetros regulables, cuyos valores de ajuste definen en conjunto qué se aprende (por ejemplo, qué calles pueden llegar a figurar en el mapa mental del barrio). En el cerebro, los parámetros son sinapsis (las conexiones entre neuronas, que pueden variar en potencial); en la mayoría de las computadoras actuales, son los “pesos”, vale decir, las probabilidades ponderables de cada hipótesis factible. Así, tanto en los cerebros como en las máquinas, aprender requiere buscar la combinación óptima de parámetros que, una vez reunidos, definen cada detalle del modelo mental. En este sentido, el del aprendizaje es un problema de búsqueda a gran escala. Por eso, analizar cómo operan los algoritmos en las computadoras actuales puede ser de gran ayuda para entender cómo funciona el aprendizaje en el cerebro humano.

Mediante la comparación de los desempeños de los algoritmos informáticos con los de nuestro cerebro, *in silico versus in vivo*, comenzaremos a entrever cómo el aprendizaje, para ser óptimo, debe apoyarse sobre un uso razonado de las probabilidades y de las estadísticas. Por supuesto, los matemáticos y los especialistas en ciencia computacional no consiguieron (todavía) diseñar algoritmos tan poderosos como los del cerebro humano. Sin embargo, empiezan a destinar cada vez más atención a pensar cuál es el algoritmo de aprendizaje de eficiencia óptima para su uso en cualquier sistema. Entre los modelos del todo innato y el todo adquirido, emerge uno nuevo: el del cerebro bayesiano, un verdadero estadístico neuronal. Esta teoría postula una clara división del trabajo entre naturaleza y crianza. Nuestros genes, en el seno del cerebro en desarrollo, establecen vastos espacios de hipótesis *a priori*, así como los mecanismos que permiten adaptarlos a los *inputs* del mundo exterior, y el ambiente *selecciona*, entre estas hipótesis, las que mejor se corresponden con ese mundo. Así, el repertorio de hipótesis está

especificado genéticamente, mientras que su selección depende de la experiencia.

¿Esta teoría se corresponde verdaderamente con el funcionamiento del cerebro? ¿Cómo se implementa el aprendizaje en nuestros circuitos biológicos? ¿Qué se modifica en el cerebro cuando adquirimos nuevas habilidades? En la segunda parte, “Cómo aprende nuestro cerebro”, pondremos proa hacia la psicología y las neurociencias. El foco del catalejo quedará sobre la cuna del bebé humano, una auténtica máquina de aprender, muchas veces imitada pero jamás igualada. Los datos recientes demuestran que el niño es este estadístico en ciernes que predice la teoría bayesiana. Sus intuiciones fulgurantes en los ámbitos del lenguaje, de la geometría, de los números o de las estadísticas confirman que no existe algo como una pizarra en blanco, una *tabula rasa*. Desde el nacimiento, los circuitos neuronales del niño están bien organizados y proyectan hipótesis sobre el mundo exterior. Pero también poseen un considerable margen de plasticidad, que se traduce en una permanente ebullición de cambios a escala celular. Dentro de esta máquina estadística, lo innato y lo adquirido, lejos de oponerse, se combinan, dando como resultado un sistema estructurado pero plástico, capaz de autorrepararse en caso de una lesión cerebral tanto como de reciclar sus circuitos para aprender a leer o a hacer matemáticas.

En la tercera parte, “Los cuatro pilares del aprendizaje”, detallaré algunos de los trucos que hacen del cerebro el dispositivo de aprendizaje más eficaz que conocemos en la actualidad. Cuatro mecanismos esenciales modulan masivamente nuestra capacidad de aprender. En primer lugar, la atención: un conjunto de circuitos neuronales que seleccionan, amplifican y propagan las señales a las que damos importancia, y multiplican por cien o por mil su representación en la memoria. En segundo lugar, el compromiso activo: en los hechos, un organismo pasivo aprende poco y nada, porque el acto de aprender exige del cerebro la generación activa de hipótesis, con motivación y curiosidad. En tercer lugar, y como complemento natural del compromiso activo, la detección y corrección de errores, en un buen *feedback*: cada vez que nos sorprendemos porque el mundo contradice nuestras expectativas, las señales de error se propagan por todo el cerebro y se ocupan de corregir los modelos mentales, eliminar las hipótesis inadecuadas y estabilizar las más pertinentes. Por último, el cuarto factor es la consolidación: con el paso del tiempo, el cerebro compila lo adquirido y lo transfiere a la memoria de largo plazo, con el objetivo de liberar los recursos para otros aprendizajes. La

repetición desempeña un papel esencial en esta consolidación, y también el sueño, que, lejos de ser un período de inacción, constituye un momento privilegiado durante el cual el cerebro repite y recodifica las adquisiciones del día. Estos cuatro pilares del aprendizaje tienen validez universal. No importa si somos bebés, niños o adultos: los desplegamos a cualquier edad. Por este motivo debemos aprender a dominarlos; solo así podemos aprender a aprender.

En la conclusión, revisaré las consecuencias prácticas de los avances científicos. Cambiar las prácticas en la escuela, la familia o la oficina no es necesariamente tan complicado como pensamos. Existen algunas ideas muy sencillas acerca del juego, el placer, la curiosidad, la socialización, la concentración o incluso el sueño que pueden consolidar aún más lo que ya es el mayor talento de nuestro cerebro: la capacidad de aprender que ejercemos y ejercitamos constantemente.

[1] Instituto Nacional de Salud e Investigación Médica francés. [N. de E.]

[2] Recomiendo las películas *The Miracle Worker* (Arthur Penn, 1962) [que en castellano se conoce con distintos títulos: *El milagro de Ana Sullivan*, *Ana de los milagros*, *Un milagro para Helen* o *La maestra milagrosa*] y *Marie Heurtin* (Jean-Pierre Améris, 2014) [*El lenguaje del corazón* o bien *La historia de Marie Heurtin*]; así como los libros Arnould (1900) y Keller (1903).

[3] Invito a visitar ese sitio web, educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit.

Parte I

¿Qué es aprender?

La inteligencia puede considerarse como la capacidad de convertir información en bruto en conocimientos útiles y explotables.

**Demis Hassabis, fundador de la empresa de IA DeepMind
(2017)**

¿Qué es aprender? Este verbo posee la misma raíz latina que “aprehender”: tomar, atrapar, asir. Aprender, entonces, es asir con el pensamiento: llevarse una porción de realidad, un modelo de la estructura del mundo. Como bien dice Demis Hassabis (gerente general de la empresa inglesa DeepMind, filial de Google, y uno de los investigadores más activos en inteligencia artificial), aprender consiste en transformar la información que recibimos en un conjunto de conocimientos útiles y explotables. Gracias al aprendizaje, los datos en bruto que impactan nuestros sentidos se convierten en ideas abstractas, refinadas y lo suficientemente generales como para que podamos explotarlas en situaciones novedosas: en ciencias cognitivas, esos conjuntos de ideas reciben el nombre de “modelos internos”.

En las páginas siguientes, revisaremos lo que la inteligencia artificial y las ciencias cognitivas nos enseñaron acerca del funcionamiento de estos modelos internos, tanto en máquinas como en cerebros, y de la forma en que la representación de la información se transforma con el aprendizaje. Comenzaremos por examinar las redes de neuronas convencionales, modelos informáticos inspirados en el cerebro humano. ¿Cómo hacen para ajustar sus parámetros para modelar el mundo exterior? Veremos que, pese a su éxito, de momento no logran asir más que una fracción de las capacidades del cerebro humano. Como por encanto, al reseñar los distintos trucos de que se valieron los ingenieros para propiciar que poco a poco las máquinas aprendiesen, haremos aparecer una imagen más nítida de los fabulosos cálculos que los niños deben realizar mientras aprenden a ver, hablar o escribir. El lenguaje y las matemáticas requieren mucho más que una red de neuronas: necesitan una verdadera lengua interior, capaz de combinar los conceptos y de seleccionar entre estas combinaciones en función de su plausibilidad estadística. La perspectiva que emergerá de allí es la de un cerebro estadístico, que formula hipótesis como un científico y las adopta o rechaza en función de los datos

que recibe. Y en los hechos, como veremos, el cerebro del niño no cede la delantera: pese a sus logros, los algoritmos de aprendizaje actuales solo abarcan una fracción de las capacidades del cerebro humano. Al entender con exactitud dónde deja de funcionar la metáfora del *machine learning* y en qué instancia incluso un cerebro infantil supera a la computadora más potente, delinearemos con claridad qué significa “aprender”.

1. Siete definiciones del aprendizaje

Retomo y reitero la pregunta: ¿qué significa “aprender”? Como definición inicial y más amplia, sostengo que aprender es construir un modelo interno del mundo exterior. Incluso si no nos damos cuenta, nuestro cerebro es portador de miles de esos modelos internos (desde una perspectiva metafórica, equivalen a maquetas, modelos a escala, más o menos fieles a la realidad que representan). Tenemos todo en la cabeza: por ejemplo, un plano de nuestro barrio o un mapa mental de nuestra casa u oficina, podemos cerrar los ojos y verlos con el pensamiento. Por supuesto, nadie nació con este mapa mental, sino que cada cual tuvo que adquirirlo mediante el aprendizaje.

La riqueza de estas representaciones mentales –en su mayoría, inconscientes– supera la imaginación. Disponemos, por ejemplo, de un amplio modelo mental de la lengua castellana, que en este momento les permite comprender las palabras que leen y adivinar que “platsovski” no es una palabra de su idioma, mientras que “sextante” sí lo es y “blascón” podría serlo. Nuestro cerebro también alberga muchos modelos del cuerpo propio: se vale de ellos para codificar dónde están sus miembros, cómo dirigirlos, a qué velocidad moverlos, cómo mantener el equilibrio... Otros modelos mentales representan el conocimiento de los objetos y de nuestras interacciones con ellos: cómo sostener un lápiz, escribir o andar en bicicleta. Otros nos traen las mentes de los demás, un enorme catálogo mental de las personas que nos son próximas, de su aspecto, su voz, sus gustos y sus tics.

Estos modelos mentales pueden generar simulaciones hiperrealistas del universo que nos rodea. ¿Alguna vez notaron que el cerebro suele proyectar los más auténticos *reality shows* virtuales (valga la supuesta paradoja), en los cuales es posible caminar, moverse, bailar, visitar lugares nuevos, tener conversaciones brillantes o sentir emociones profundas? ¡Esos son sus sueños! Es fascinante tomar conciencia de que todos los pensamientos (a menudo muy complejos) que nos llegan durante los sueños son solo producto del libre funcionamiento de los modelos internos del mundo.

Pero también soñamos la realidad cuando estamos despiertos: nuestro cerebro proyecta permanentemente sobre el mundo exterior hipótesis, marcos de interpretación que le dan sentido al flujo de datos que nos llega por los sentidos. Por eso, sin que lo sepamos, cada imagen que aparece en nuestra retina es ambigua: cada vez que vemos un plato, por ejemplo, la imagen es compatible con una cantidad infinita de elipses. Si vemos el plato como redondo, incluso a pesar de que por los datos sensoriales en bruto nos presentan la figura de un óvalo, se debe a que nuestro cerebro aporta datos adicionales: aprendió que la silueta circular es la interpretación más pertinente. Entre bastidores, las áreas sensoriales constantemente computan a partir de probabilidades, y solo el modelo más probable logra acceder a la conciencia. En última instancia, lo que da sentido al flujo de datos que llega a nosotros desde las percepciones son las proyecciones obradas por el cerebro. Si no existiera un modelo interno, estos datos en bruto serían ininteligibles.

El aprendizaje permite que el cerebro atrape una porción de la realidad que antes le era ajena y la use para construir un nuevo modelo del mundo. Puede ser una porción de la realidad exterior, si es cuestión de aprender historia, botánica o el plano de una ciudad, pero también de la realidad interna, ya que buscamos aprender a coordinar los gestos y a concentrar los pensamientos con el objetivo de tocar el violín. En estos dos casos, nuestro cerebro *internaliza* un aspecto nuevo de la realidad: ajusta sus circuitos con intención de apropiarse de un campo que antes no dominaba. Desde luego, esos ajustes tienen que ser muy ingeniosos. La fuerza del aprendizaje reside en su capacidad de acomodarse al mundo externo y de corregir en caso de error; pero el cerebro de quien aprende ¿de qué modo “sabe” cómo actualizar su modelo interno cuando, digamos, se desorienta en su barrio, se cae de su bicicleta, pierde una partida de ajedrez o chapurrea la palabra “anfractuosa”?

Aprender es ajustar los parámetros de un modelo mental

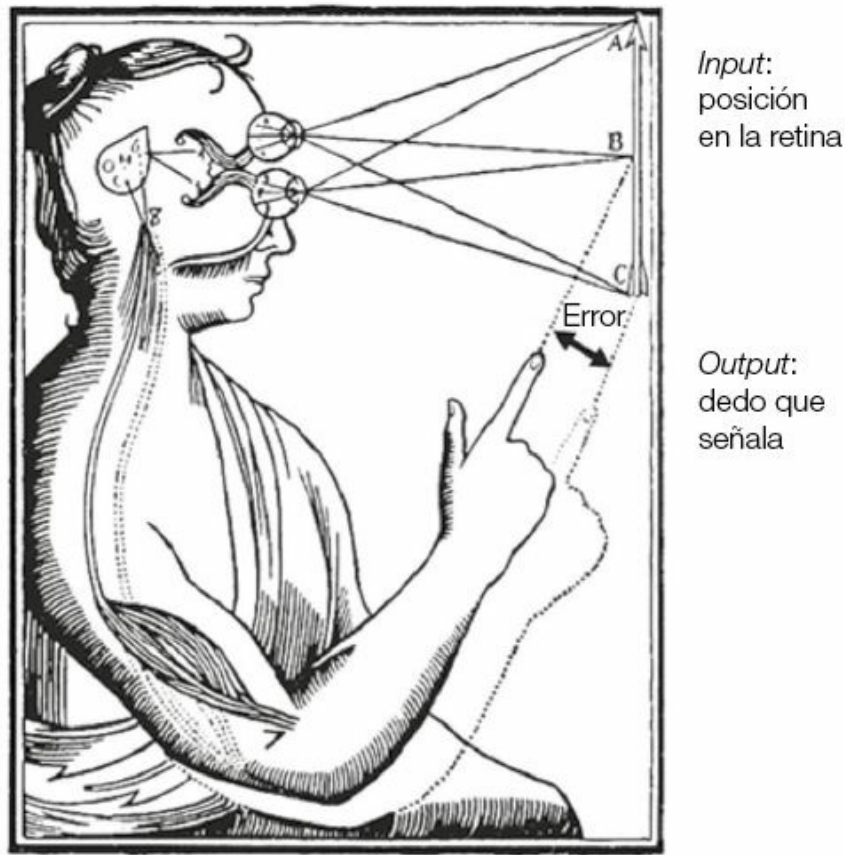
Ajustar un modelo mental a veces es muy sencillo. ¿Cómo hacemos, por ejemplo, para tender la mano y alcanzar un objeto que vemos? Ya en el siglo XVII, René Descartes lo había presentido: debemos aprender a transformar la información visual en órdenes musculares (figura 2). Pueden experimentarlo

ustedes mismos en pocos segundos: intenten tomar una varilla mientras llevan puestos los anteojos de otra persona (de ser posible, una muy miope). Mejor todavía, si pueden, consigan anteojos con lentes prismáticos o, sin más, prismas, que desplacen la visión una decena de grados hacia la izquierda, e intenten tomar un objeto. Verán que su primer intento es completamente fallido: a causa de los prismas, su mano aterriza muy a la derecha del palo que, sin embargo, ustedes ven. Poco a poco, adecuan sus movimientos, desplazándolos hacia la izquierda. Con un proceso de ensayo y error, sus movimientos se vuelven cada vez más exactos: su cerebro ha aprendido a compensar el desajuste de los ojos. Ahora quítense los anteojos y tomen la varilla: ¡los sorprenderá ver que su mano se dirige al lado equivocado, demasiado a la izquierda!

¿Qué ocurrió? Durante este breve aprendizaje, el cerebro ajustó su modelo interno de la visión. Un parámetro de este modelo, que corresponde al desfase entre la escena visual y la orientación del cuerpo, fue recalibrado. El cerebro se comportó como un tirador de élite, que primero realiza un disparo de prueba y luego ajusta la altura de su mira para así lograr mayor precisión. Este aprendizaje es muy rápido: bastan algunos ensayos para corregir el desfase entre el movimiento y la visión. Sin embargo, la nueva regulación no es compatible con la anterior; de aquí proviene el error sistemático que todos cometimos cuando nos quitamos los prismas y volvimos a tener una visión normal.

Es innegable que este aprendizaje es un poco particular, porque no requiere ajustar más que un solo parámetro: el ángulo de visión. Por este motivo es tan rápido. La mayor parte de los aprendizajes son tanto más elaborados y demandan el ajuste de varias decenas, centenas y hasta miles de millones de parámetros (cada una de las sinapsis que determinan la actividad de nuestros circuitos). Sin embargo, el principio siempre es el mismo: todo consiste en investigar, entre un sinnúmero de regulaciones posibles del modelo interno, aquellas que mejor se corresponden con el estado del mundo exterior.

Ajuste de un solo parámetro: el desfase entre la visión y la acción



Ajuste de millones de parámetros: las conexiones que permiten la visión

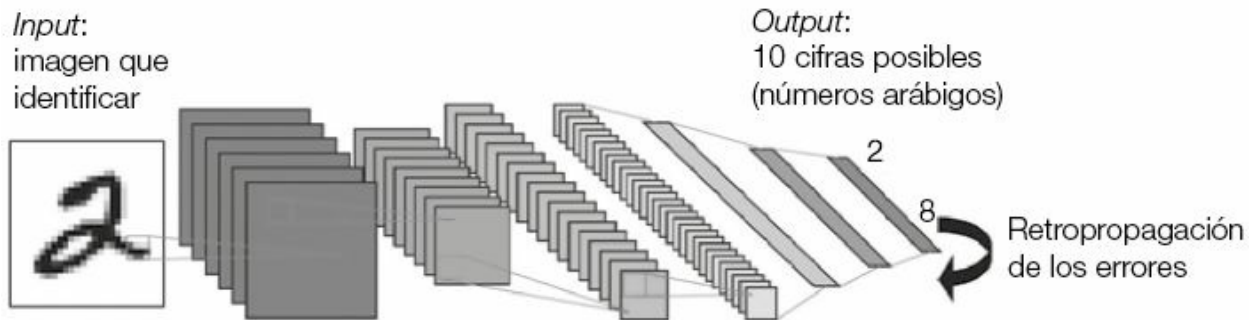


Figura 2. Aprender es ajustar los parámetros de un modelo del mundo. Aprender a señalar con el dedo, por ejemplo, consiste en ajustar el desfase entre la visión y la acción: cada error permite corregir la puntería. En la red neuronal artificial el principio es el mismo, pero los ajustes son infinitamente más numerosos. Reconocer qué número está presente en una imagen requiere ajustar millones de conexiones. También en este caso, todos los errores –

aquí, un incremento de la activación del número 8– permiten corregir su valor y, de este modo, mejorar el desempeño en el siguiente intento.

Consideremos ahora el aprendizaje de un idioma como el japonés. Cuando el cerebro de un bebé japonés aprende su lengua materna, ajusta gradualmente su modelo de la lengua a las características del idioma japonés. Intenten imaginar una máquina dotada de millones de regulaciones en todos los niveles. Algunas de estas regulaciones, en el nivel de la entrada auditiva, determinan el inventario de consonantes y de vocales que utiliza el japonés, y las reglas que permiten componerlas. El bebé que nace dentro de una familia japonesa debe descubrir qué categorías de sonidos son utilizadas en esa lengua, y dónde ubicar las fronteras entre ellas. Uno de los parámetros, por ejemplo, concierne a la distinción entre los sonidos /R/ y /L/: esta es crucial en castellano, pero no en japonés, que no hace diferencia alguna entre una “elección” y una “erección”... Cada bebé, entonces, debe fijar un conjunto de parámetros que, colectivamente, precisan qué categorías son pertinentes para su lengua materna.

Un procedimiento de aprendizaje similar se reproduce en todos los niveles: desde los patrones de sonido hasta el vocabulario, la gramática y el significado. El cerebro está organizado como una estructura de modelos de la realidad anidados –uno dentro del otro, como las muñecas rusas; y aprender significa utilizar la información que ingresa para fijar los parámetros en cada nivel de esa estructura jerárquica. Tomemos un ejemplo de un nivel superior: la adquisición de las reglas de la gramática. Otra diferencia entre el japonés y el castellano que el bebé debe aprender está relacionada con el orden de las palabras. En una oración bimembre canónica, con un sujeto, un verbo y un objeto directo, la lengua castellana sitúa primero el sujeto, luego el verbo y, por último, su objeto directo: “Juan come una manzana”. En japonés, en cambio, el orden más usual es sujeto, luego objeto, luego verbo: “Juan manzana come”. El orden se invierte también para las preposiciones (que, lógicamente, se llaman posposiciones), los posesivos y muchos otros grupos de palabras. Así, la oración “Mi tío visita museos en París” se convierte en una que puede parecernos un galimatías digno del venerable Yoda, de *La guerra de las galaxias*: “Tío mi París en museos visita”, lo que tiene mucho sentido para un hablante japonés.

Todas estas diferencias no son independientes unas de otras. Ciertos lingüistas piensan que se originan en un solo y mismo parámetro llamado

“posición del núcleo”: la palabra que otorga su categoría a un grupo –es decir, su núcleo– se sitúa siempre en posición inicial en español (*en* París, *mi* tío, *visita* museos) pero en último lugar en japonés (París *en*, tío *mi*, museos *visita*). Por cierto, este parámetro binario marca distinciones entre muchas lenguas, incluso sin vínculo histórico entre ellas (por ejemplo, el apache, que es una de las lenguas “atabascanas”, sigue las mismas reglas que el japonés). Así, con el fin de adaptarse al castellano o al japonés, es suficiente con que el niño ajuste el parámetro “posición del núcleo” en su modelo interno de la lengua.

Aprender es aprovechar la explosión combinatoria

¿Es verosímil que el aprendizaje de las lenguas se reduzca a la selección de algunos parámetros? Si eso nos parece difícil de creer, es porque no imaginamos la extraordinaria cantidad de posibilidades que se abren cuando se incrementa, siquiera un poco, la cantidad de parámetros ajustables. Esto se denomina “explosión combinatoria”: el aumento exponencial que se produce cuando se combina apenas un puñado de posibilidades. Supongamos que la gramática de las lenguas del mundo pudiera describirse con algo así como cincuenta parámetros binarios, según postulan algunos lingüistas. Eso da como resultado 2^{50} combinaciones, es decir, ¡más de mil billones, o un 1 seguido de quince ceros! Las reglas sintácticas de las tres mil lenguas del mundo caben con facilidad en este gigantesco espacio de las lenguas posibles. Sin embargo, en nuestro cerebro no hay cincuenta parámetros ajustables, sino una cantidad sorprendentemente mayor: 86.000.000.000 de neuronas, cada una de ellas provista de una decena de miles de contactos sinápticos con fuerzas que pueden variar. El espacio de representaciones que se abre es poco menos que infinito.

Las lenguas humanas aprovechan estas posibilidades de combinación en todos los niveles. Tomemos el ejemplo del léxico mental, es decir, el conjunto de las palabras que conocemos y cuyo modelo llevamos con nosotros. Cada cual aprendió, en su lengua materna, alrededor de 50.000 palabras con los significados más diversos. Es un número grande, cierto, pero parece escaso frente a las cantidades prodigiosas que ofrece la combinatoria.

Si consideramos que estas 50.000 palabras en promedio tienen 3 sílabas, cada una formada por alrededor de 2 fonemas, tomados entre los 24 fonemas del castellano, la codificación binaria de todas esas palabras requiere menos de 2.000.000 de elecciones binarias elementales (los bits, a los cuales se asigna el valor de 0 o 1). Digámoslo de otro modo: todo nuestro conocimiento del diccionario podría almacenarse en un pequeño archivo informático de 250 kilobytes (cada byte corresponde a 8 bits). A continuación, sería posible comprimir este léxico mental en un tamaño tanto menor si tuviéramos en cuenta las numerosas redundancias que rigen a las palabras. Si tomamos seis letras al azar para lograr una cadena como “xfrdrga”, estas no forman una palabra del castellano. Las palabras reales están compuestas por una pirámide de sílabas que se ensamblan de acuerdo con reglas estrictas. Y esto es así en todos los niveles: las frases o las oraciones son combinaciones regulares de palabras que, a su vez, son combinaciones regulares de sílabas que, a su vez, son combinaciones regulares de fonemas. En cada nivel, las combinaciones son simultáneamente amplias (porque elegimos entre varias decenas o centenas de elementos) y acotadas (porque solo determinadas combinaciones están permitidas). Aprender una lengua es descubrir los parámetros que rigen esas combinaciones en todos los niveles.

En conclusión, el cerebro humano segmenta el problema del aprendizaje mediante la construcción de un modelo jerárquico de múltiples niveles. Esto es más que obvio en el caso de la lengua –desde los sonidos elementales hasta la oración e incluso el discurso–, pero el mismo principio de análisis jerárquico se reproduce en todos los sistemas sensoriales. Determinadas áreas cerebrales captan las regularidades de bajo nivel: ven el mundo a través de una ventana temporal y espacial muy pequeña, y analizan las regularidades más nimias. Por ejemplo, en el área visual primaria, la primera región de la corteza en recibir los estímulos visuales, cada neurona no analiza más que una porción muy pequeña de la retina. Solo ve el mundo a través del ojo de una aguja y, como resultado, descubre las regularidades de muy bajo nivel, como la presencia de una línea oblicua en movimiento. Millones de neuronas hacen el mismo trabajo en diferentes puntos de la retina, y sus *outputs* se convierten en los *inputs* del nivel siguiente, que entonces detectará “regularidades de regularidades”, y así sucesivamente. En cada nivel, la escala aumenta: el cerebro busca las regularidades en rangos cada vez más amplios, tanto en el tiempo como en el espacio. En las sucesivas instancias de esta jerarquía emerge la capacidad de detectar objetos o conceptos cada vez

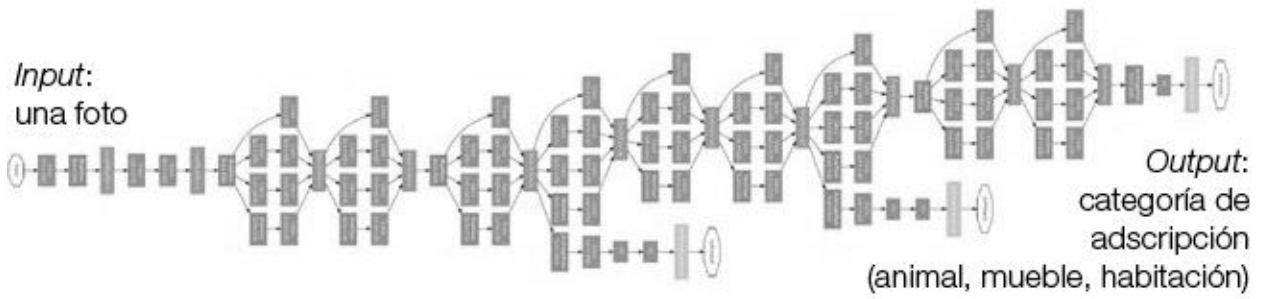
más complejos: una línea, un dedo, una mano, un brazo, un cuerpo humano... No, dos... Son dos personas que se miran cara a cara, es un apretón de manos... ¡Es el primer encuentro entre Charles Chaplin y Buster Keaton!

Aprender es minimizar los errores

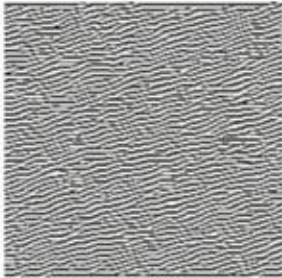
Los algoritmos informáticos que llamamos “redes de neuronas artificiales” se inspiran directamente en la organización jerárquica de la corteza. Al igual que ella, organizan una pirámide de capas sucesivas: cada una intenta descubrir regularidades más profundas que la capa previa. Debido a que estas capas consecutivas organizan la información entrante de forma cada vez más profunda, también se las llama “redes profundas”. Cada capa, por sí misma, solo puede descubrir una parte extremadamente simple de la realidad exterior (es del tipo de problemas que las matemáticas califican como “linealmente separables”; en este caso, cada neurona solo puede separar la información en dos categorías A y B trazando una estricta división entre ellas). En cambio, si se unen muchas de esas capas, se obtiene un dispositivo de aprendizaje sumamente robusto, capaz de descubrir estructuras complejas y de ajustarse a problemas muy diversos. Las redes de neuronas artificiales de última generación, que cuentan con el avance de los microchips, también son profundas, en el sentido de que incluyen decenas de capas sucesivas –cada vez más alejadas de la entrada sensorial y más astutas– capaces de identificar en sus *inputs* propiedades cada vez más abstractas.

Tomemos el ejemplo del algoritmo “LeNet”, creado por el pionero francés de las redes neuronales, Yann LeCun (figura 3; véase LeCun y otros, 1998). Desde los años noventa, esta red de neuronas alcanza desempeños notables en el reconocimiento de caracteres manuscritos. Durante años, el correo canadiense lo utilizó para el procesamiento automático de los códigos postales. ¿Cómo funciona?

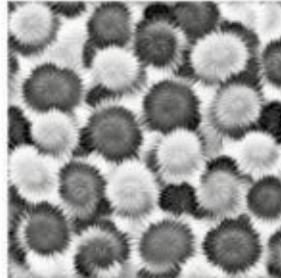
La red GoogLeNet



Patrones



Formas elementales



Partes de objetos



Objetos

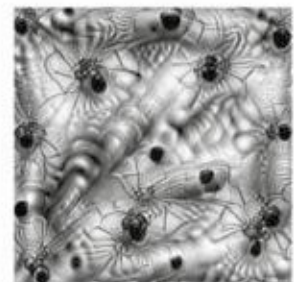
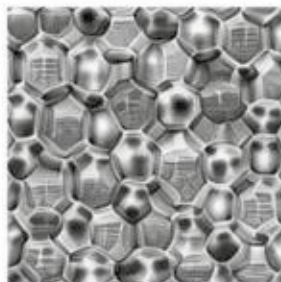
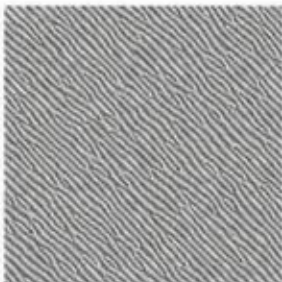
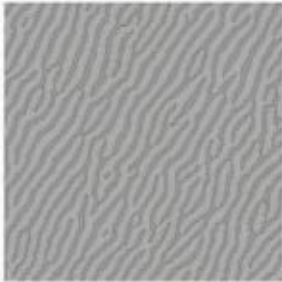


Figura 3. Aprender es construir una estructura de representaciones apropiadas al problema planteado. En la red GoogLeNet, que aprende a reconocer imágenes, millones de parámetros

se ajustan para que cada nivel de la estructura detecte determinado factor de la realidad. En el nivel más bajo, las neuronas artificiales son sensibles a los patrones y a las texturas. A medida que se asciende en la jerarquía, las neuronas responden a formas más complejas.

El algoritmo recibe como *input*, en forma de píxeles, la imagen de un carácter escrito y propone, como *output*, una interpretación tentativa: uno de los diez dígitos o de las veintiseis letras posibles. La red artificial posee una jerarquía de unidades de procesamiento que se parecen un poco a las neuronas y que forman capas sucesivas. Las primeras capas están conectadas directamente a la imagen: aplican filtros que reconocen fragmentos de rectas y curvas. Cuanto más se avanza en la jerarquía, mayores y más complejos se vuelven estos filtros. Las unidades más elevadas aprenden a reconocer porciones cada vez más amplias de la imagen: la curva de un 2, el remate de una O, o las líneas paralelas de una Z... hasta llegar, a la salida, a neuronas artificiales que responden a un carácter con independencia de su posición y de las peculiaridades de sus trazos. Todas estas propiedades no están impuestas por un programador: son resultado de millones de conexiones entre las unidades. Una vez ajustadas por un algoritmo automatizado, estas conexiones definen el filtro que cada neurona aplica a sus *inputs*: hacen que una neurona responda al dígito 2 y otra al 3.

¿Cómo se ajustan estos millones de conexiones? De la misma manera que en el caso de los anteojos con lentes prismáticos: en cada prueba, la red da una respuesta tentativa, constata que ha cometido un error e intenta ajustar sus parámetros para reducirlo en la prueba siguiente. Cada respuesta errónea provee información valiosa. Por su signo (como un gesto muy a la derecha o muy a la izquierda), el error indica lo que en verdad hacía falta realizar para tener éxito. Si nos remontamos a la fuente de ese error, la máquina es capaz de descubrir cómo debía (y debe) establecer los parámetros para evitar equivocarse.

Volvamos al ejemplo del deportista que ajusta la mira de su rifle. El procedimiento de aprendizaje es elemental. El tirador dispara y constata que ha enfocado unos centímetros (más exactamente, 5) a la derecha. Ahora tiene una información esencial, tanto sobre la amplitud (5 cm) como sobre el signo del error (demasiado a la derecha). Esta información le permite corregir el tiro. Si es un poco avisado, sabe en qué dirección debe realizar la corrección: si la bala se desvió hacia la derecha, hace falta mover la mira un poco a la izquierda. Incluso si no es tan astuto, puede hacer intentos al azar y

constatar que, si mueve la mira hacia un lado, el error aumenta, mientras que si la mueve hacia el otro, disminuye. Así, por prueba y error, el tirador puede determinar cómo reducir la magnitud del error. Al ajustar la mira para afinar su puntería, nuestro eximio deportista –tal como Monsieur Jourdain, el burgués gentilhomme de Molière, que al hablar “hacía” prosa sin saberlo– está aplicando un algoritmo de aprendizaje sin siquiera conocerlo. Está calculando implícitamente lo que los matemáticos llamamos la “derivada” o el “gradiente del sistema”, y hace lo que denominamos un “descenso del gradiente”: aprende a mover la mira de su rifle en la dirección más eficiente en busca de reducir la probabilidad de cometer un error.

La mayoría de las redes neuronales que se utilizan en la inteligencia artificial actual –y más allá de sus millones de *inputs*, *outputs* y parámetros ajustables– funciona de la misma forma que nuestro tirador: observa sus errores y los aprovecha para ajustar su estado interno en la dirección que considera mejor para disminuir el error. En muchos casos, ese aprendizaje está totalmente guiado: por un lado, le decimos con exactitud a la red qué respuesta debería haber activado en la salida (“es un 1, no un 7”); por otro lado, sabemos con precisión en qué dirección ajustar los parámetros si hay un error (un cálculo matemático permite saber exactamente qué conexiones ajustar cuando la red activa demasiado el *output* “7” en respuesta a una imagen del dígito “1”). En el lenguaje del *machine learning* esto es conocido como “aprendizaje supervisado” (porque alguien, a quien podríamos llamar el supervisor, conoce la respuesta correcta que debe dar el sistema) y “retropropagación de errores” (porque los errores se reenvían a la red con el objetivo de modificar los parámetros). El procedimiento es sencillo: intento dar una respuesta, me dicen lo que debería haber respondido, mido mi error y, en busca de reducirlo, corrijo todos mis parámetros. En cada etapa, doy apenas un pequeño paso, hago una pequeña corrección en la dirección correcta. Debido a este proceso el *machine learning* puede resultar increíblemente lento: aprender una actividad compleja, como jugar al Tetris, exige aplicar determinada receta miles, millones e incluso miles de millones de veces. En un espacio que abarca una multitud de parámetros ajustables, descubrir el ajuste óptimo de cada tornillo y de cada bulón puede insumir mucho tiempo.

Ya en la década de 1980 el funcionamiento de las primeras redes de neuronas artificiales se basaba sobre este principio de corrección gradual de los errores. Los progresos de la informática permitieron extender esta idea a

redes neuronales gigantescas, que incluyen centenas de millones de conexiones ajustables. Estas redes neuronales profundas están integradas por una sucesión de etapas que, a cada paso, se ajustan al problema planteado. A modo de ejemplo, la figura 3 muestra el sistema GoogLeNet, derivado de la arquitectura LeNet propuesta por Yann LeCun, que ganó una de las más importantes competencias internacionales en reconocimiento de imágenes. Expuesto a miles de millones de imágenes, este sistema aprendió a separarlas en casi mil categorías distintas: rostros, paisajes, barcos, autos, perros, insectos, flores, señales viales, etc. Cada nivel de su jerarquía se ajustó a una faceta útil de la realidad: las unidades de nivel bajo responden selectivamente a rasgos o a texturas, y a medida que se sube en la jerarquía, las neuronas responden más selectivamente a formas complejas: figuras geométricas (círculos, curvas, estrellas), partes de objetos (bolsillo de pantalón, asa de una taza, par de ojos...), e incluso objetos enteros (edificios, rostros, arañas...; Olah y otros, 2017).

En un intento por minimizar los errores, el algoritmo de descenso del gradiente descubrió que esas formas son las más útiles para la categorización de las imágenes. Pero si la misma red hubiera estado expuesta a textos o a partituras musicales, habría sido ajustada de un modo diferente y habría aprendido a reconocer palabras, notas musicales o cualquier otra forma recurrente en este nuevo entorno. La figura 4, por ejemplo, muestra cómo se autoorganiza una red de este tipo cuando se le pide que se especialice para el reconocimiento de miles de dígitos manuscritos (Guerguiev y otros, 2017). En el nivel más bajo, los datos están mezclados: existen formas muy parecidas, como un 3 y un 8, que pese a todo haría falta diferenciar, y, a la inversa, existen formas muy diferentes, como varias versiones del 8 con el bucle de arriba abierto o cerrado, que sin embargo habría que agrupar. En cada etapa, la red artificial de neuronas progresa en abstracción, hasta agrupar correctamente todos los ejemplares de un mismo dígito. Gracias al procedimiento de reducción de errores, se descubrió una jerarquía de índices que resuelve el reconocimiento de cifras manuscritas. Y ciertamente es muy notable que con solo corregir los errores propios se pueda descubrir una jerarquía completa de claves adecuadas para el problema en cuestión.

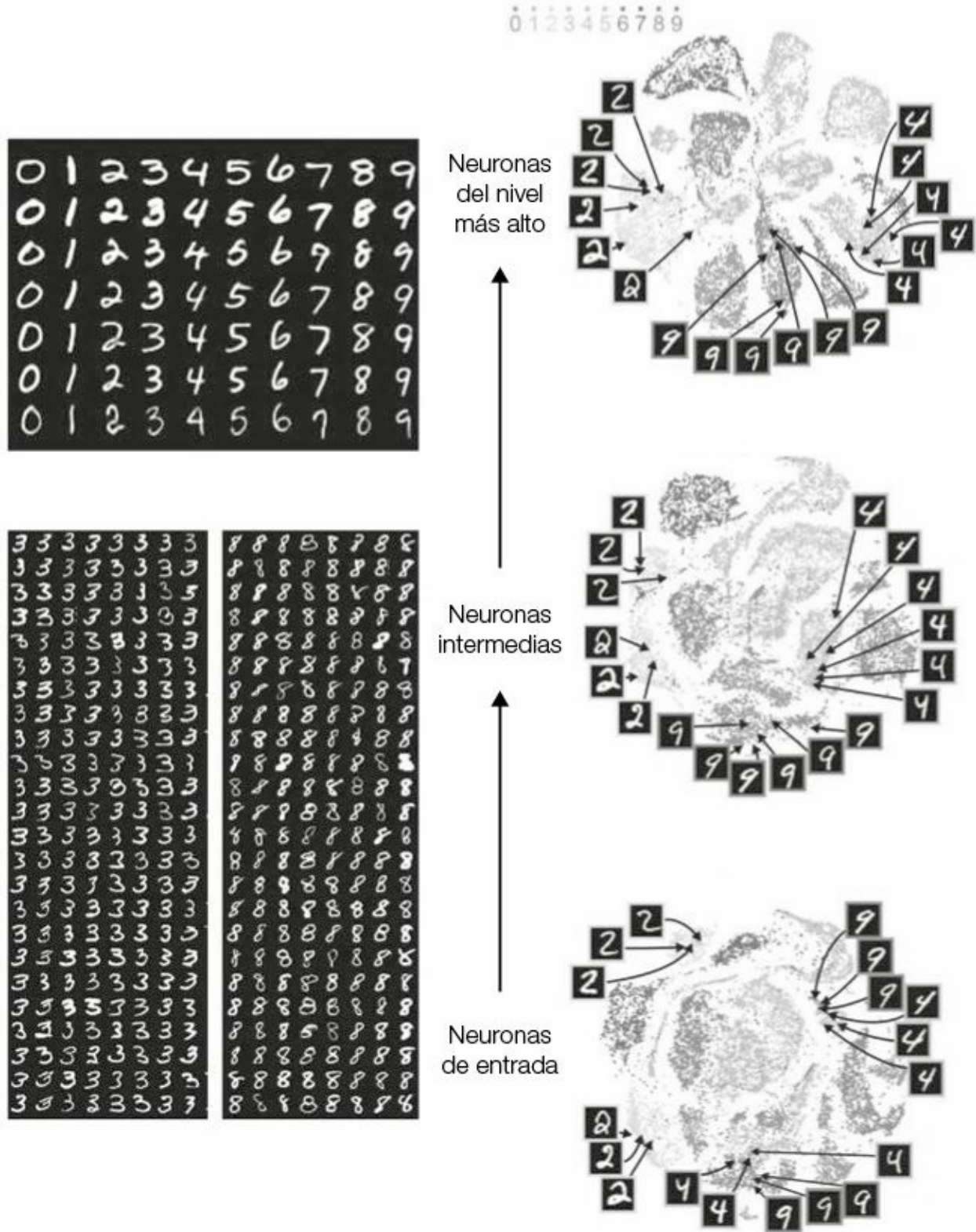


Figura 4. ¿Cómo aprende a categorizar los dígitos manuscritos una red neuronal profunda?

Este tema es difícil, porque un mismo dígito puede escribirse de cientos de maneras diferentes. En el nivel más bajo de la jerarquía neuronal (abajo a la derecha), todos los dígitos están mezclados, y aquellos que se parecen entre sí, como los 9 y los 4, se confunden. A medida que se asciende en la jerarquía, las neuronas logran agrupar todas las imágenes de un mismo dígito y separarlas con límites claros.

Hoy en día, el concepto de aprendizaje por retropropagación de los errores reside en la base de una gran cantidad de aplicaciones informáticas. Sobre él se asienta la capacidad de nuestros *smartphones* de reconocer su voz, o el novedoso talento de los autos inteligentes para distinguir a los peatones de las señales de tránsito; así, podemos colegir la posibilidad de que nuestro cerebro use una versión de este. Sin embargo, la retropropagación o *feedback* del error viene en varias presentaciones, pensadas para cada usuaria o usuario. El área del aprendizaje artificial progresó enormemente en treinta años, y los investigadores descubrieron gran cantidad de trucos que facilitan el aprendizaje. Sin darles un orden de prioridad, pasemos revista a algunos de ellos; veremos que nos dicen mucho sobre nosotros mismos y sobre el modo en que aprendemos.

Aprender es explorar el espacio de lo posible

Uno de los problemas con el procedimiento de corrección de errores que acabo de describir es que puede quedar atrapado en un conjunto de parámetros que no es el mejor. Imaginen una pelota de golf que rueda sobre el césped, siempre siguiendo la línea con la pendiente más pronunciada: es posible que quede bloqueada en una pequeña depresión del suelo, sin alcanzar necesariamente el punto más bajo de toda la superficie, el óptimo absoluto. Del mismo modo, a veces el algoritmo de descenso del gradiente se ve trabado en un punto del cual no puede salir. Eso es lo que llamamos un “mínimo local”, un pozo en el espacio de los parámetros, una trampa de la cual el algoritmo no logra salir porque le parece imposible hacerlo mejor. En ese momento, el aprendizaje se estanca, porque todos los cambios parecen contraproducentes: cada uno de ellos aumenta la tasa de error. El sistema estima haber aprendido todo lo que podía aprender. Ignora que en realidad existen, un poco más lejos en el espacio de los parámetros, otras

combinaciones mejores. El algoritmo de descenso del gradiente no los “ve”, porque se niega a subir la pendiente una vez más para volver a descender mejor del otro lado del hoyo. Como un miope, apenas ve a una corta distancia de su punto de partida y, por lo tanto, puede perderse configuraciones distantes pero mejores.

¿Este problema les parece demasiado abstracto? Piensen en una situación concreta: ustedes van a hacer las compras al mercado, donde pasarán algo de tiempo buscando los productos más baratos. Avanzan por un sector, pasan al primer vendedor, que les parece que está fuera de precio, evitan al segundo, que siempre es muy caro, y finalmente se detienen en el tercer puesto, que decididamente tiene mejores precios que los precedentes. Pero ¿quién les dice que en el pasillo de al lado, o tal vez incluso en el pueblo de al lado, los precios no serían todavía más interesantes? Detenerse en el mejor precio *local* no garantiza encontrar el mínimo *global*.

Como enfrentan a menudo esta dificultad, los informáticos emplean una serie de trucos. La mayor parte consiste en introducir una dosis de azar en la búsqueda de los mejores parámetros. La idea es simple: en vez de buscar solo en un sector del mercado, paseamos de modo aleatorio; en vez de dejar que la pelota de golf descienda suavemente por la pendiente, le damos algo de efecto, que reduce sus posibilidades de quedarse bloqueada en un hueco. En ciertas ocasiones, los algoritmos de búsqueda estocástica[4] operan con una configuración distante y parcialmente aleatoria, de modo que, si existe una solución mejor, tengan la posibilidad de encontrarla. En la práctica, podemos introducir una porción de esas fluctuaciones de varias maneras: configurar o actualizar los parámetros al azar, diversificar el orden de los ejemplos, agregar un poco de ruido a las entradas del sistema, o incluso utilizar solo una fracción aleatoria de las conexiones. Todas estas ideas mejoran el ímpetu del aprendizaje.

Algunos algoritmos de *machine learning* se inspiran también en el algoritmo darwiniano que rige la evolución de las especies: durante la optimización de los parámetros, incluyen mutaciones y cruces aleatorios de las soluciones descubiertas previamente. Tal como en biología, la tasa de estas mutaciones debe ser controlada con sumo cuidado para así explorar nuevas soluciones sin perder demasiado tiempo en intentos precarios y riesgosos.

Otro algoritmo se inspira en las fraguas, donde los obreros aprendieron a optimizar las propiedades del metal al “refundirlo”. Cuando se quiere forjar

una espada excepcionalmente dura, el método del refundido consiste en cocerla muchas veces, a temperaturas cada vez más bajas, para aumentar las posibilidades de que los átomos se dispongan en una configuración regular. En la actualidad este procedimiento se extrapoló a la informática: el algoritmo de recocido simulado introduce cambios aleatorios en los parámetros, con una temperatura virtual que decrece gradualmente. La posibilidad de un evento fortuito es elevada al principio, pero se reduce poco a poco hasta que el sistema alcanza una regulación óptima.

Los informáticos descubrieron que todos estos trucos son muy eficaces, por lo que no debería sorprendernos que algunos de ellos se hayan internalizado en nuestro cerebro a lo largo de la evolución. La exploración al azar, la curiosidad estocástica y la generación aleatoria de descargas neuronales desempeñan un papel primordial en el aprendizaje en el *Homo sapiens*. Tanto cuando tocamos rock, jugamos a piedra, papel o tijera, improvisamos sobre un *standard* de jazz o pensamos posibles soluciones para un problema matemático, el azar es un ingrediente esencial de la solución. Como veremos más adelante, mientras el niño está en modo “aprendizaje” –es decir, mientras juega–, explora decenas de posibilidades con una buena dosis de azar; y durante la noche, su cerebro continúa haciendo malabares con las ideas, hasta encontrar la combinación óptima. En la tercera parte de este libro, retomaré la cuestión del algoritmo semialeatorio que gobierna la extraordinaria curiosidad de los niños, así como la de los pocos adultos que supieron preservar su espíritu infantil, y comentaré lo que sabemos al respecto.

Aprender es optimizar una función de recompensa

¿Recuerdan la red de Yann LeCun, que reconocía la forma de los números arábigos? Para aprender, este tipo de inteligencia artificial necesita que le demos las respuestas correctas. Para cada imagen que ingresa, necesita saber cuál de las diez cifras le corresponde. La red solo logra corregirse calculando la diferencia entre la respuesta correcta y su propio intento. Este procedimiento es conocido como “aprendizaje supervisado”: un supervisor, por fuera del sistema, conoce la solución e intenta enseñársela a la máquina. Es eficaz, pero hay que reconocer que los aprendizajes de este tipo, en los

que conocemos la respuesta correcta con anterioridad, no suelen abundar. Cuando un niño aprende a caminar, nadie le dice exactamente qué músculos contraer; nos contentamos con alentarlos una y otra vez, hasta que deja de caerse. Los bebés solo aprenden sobre la base de una evaluación del resultado: “Me caí”, o bien, por el contrario, “¡Finalmente logré caminar de ese lado a este lado de la habitación!”.

La inteligencia artificial enfrenta el mismo problema del “aprendizaje no supervisado”. Cuando una máquina aprende a jugar a un videojuego, por ejemplo, lo único que sabe al principio es que debe intentar obtener el mejor puntaje. Nadie puede anunciarle por anticipado qué acciones específicas debe realizar para lograrlo. Pero ¿cómo consigue descubrir por sí sola el modo correcto?

En respuesta a este desafío, los investigadores inventaron el “aprendizaje por recompensa”, en el que no le damos al sistema todos los detalles de lo que debe hacer (porque no los conocemos), sino una “recompensa”, una evaluación en forma de puntaje cuantitativo (Mnih y otros, 2015; Sutton y Barto, 1998). A veces, incluso, el puntaje se da luego de un largo tiempo, mucho después de la acción evaluada. Sobre la base de este modelo, la empresa DeepMind creó una máquina que juega al ajedrez, a las damas y al go. Pero existe un enorme problema: la máquina recibe su puntaje, una sola señal –precisamente, la “recompensa”– que le indica si ganó o perdió, pero solo una vez concluida la partida. Mientras juega, el sistema no recibe devolución alguna, con la única excepción de que tiene relevancia el jaque mate. ¿Cómo es posible, entonces, que el sistema descubra qué hacer en cada momento dado? Y una vez que recibe el puntaje final, ¿cómo puede evaluar retrospectivamente las decisiones que tomó?

El truco que encontraron los informáticos consiste en aprender dos cosas al mismo tiempo: a (re)accionar y a autoevaluarse. La mitad del sistema, que llamamos “el crítico”, aprende a predecir el resultado final. A cada instante, esta red artificial de neuronas evalúa el estado del juego e intenta predecir la recompensa final: ¿estoy más cerca de ganar la partida o de perderla? ¿Mi equilibrio es estable o estoy a punto de caerme? Gracias a la crítica que se produce en esta mitad, el sistema puede evaluar sus actos constantemente, no solo al final de la partida. La otra mitad de la red, “el actor”, puede entonces utilizar esta evaluación para corregirse: “¡Atención! Será mejor que evite tal o tal otra acción, porque el crítico piensa que aumenta las posibilidades de perder”.

De intento en intento, el actor y el crítico progresan juntos: uno aprende a reaccionar oportunamente, enfocándose en las acciones más eficaces, mientras que el otro aprende a evaluar, de manera cada vez más refinada, las consecuencias de estos actos. Al final –como ese tan citado personaje de Vincent Cassel en la película *La Haine [El odio]* de Mathieu Kassovitz, el paciente que mientras cae desde lo más alto de un rascacielos no deja de repetir: “Por ahora, todo va bien”–, la red actor-crítico queda dotada de una notable previsión, la habilidad de anticipar, de entre océanos de partidas-que-todavía-no-están-perdidas, cuáles le prometen llevar las de ganar (y también cuáles van solo por la ruta perdedora, hacia un desastre previsible).

La combinación actor-crítico es una de las estrategias más eficaces de la inteligencia artificial contemporánea. Cuando se la adosa a una red jerárquica de neuronas, hace maravillas. Ya a finales de los años setenta, posibilitó que una red neuronal ganara el campeonato mundial de backgammon. En fecha más reciente, le permitió a DeepMind crear una red de neuronas multifunción, capaz de aprender a jugar a videojuegos de todo tipo, en una variedad que va de la galaxia Super Mario al Tetris.^[5] Simplemente le damos al sistema los píxeles de la imagen como *input*, las acciones posibles como *output* y el puntaje del juego como función de recompensa. La máquina aprende lo demás. Cuando juega al Tetris, descubre que la pantalla está tramada de formas geométricas, que la forma que cae es más importante que las otras (ya estables), que las acciones durante la caída pueden modificar la posición y la orientación de la figura en cuestión, y así sucesivamente, hasta converger en un jugador artificial de una eficacia formidable. Y cuando juega a alguna de las versiones de la “familia” Super Mario, el cambio de *input* y de recompensa le hace aprender otros factores: qué píxeles forman la imagen de Mario, cómo se desplaza, dónde están los enemigos, los muros, las puertas, las trampas, los puntos extra... y cómo reaccionar ante cada una de sus configuraciones. Al ajustar estos parámetros, es decir, los millones de conexiones que ligan las capas entre sí, la red misma logra adaptarse a todo tipo de juegos y aprende a reconocer las formas del Tetris, del Pac-Man o de Sonic.

¿Qué interés podría tener una empresa como Google por enseñarle a una máquina a jugar a videojuegos? Apenas dos años después de iniciada esa experiencia, los ingenieros de DeepMind le pedían a la misma red de neuronas que aplicara lo aprendido a la resolución de un problema económico de interés vital: ¿cómo optimizar la gestión del parque informático de Google

y sus distintos servidores? La red de neuronas era la misma; solo cambiaban el material sobre el cual se trabajaba (la cantidad de personas conectadas a cada servidor, día y hora, el estado del tiempo, eventos internacionales, solicitudes de búsquedas, etc.), la acción prevista (encender o apagar tal o cual servidor en los distintos continentes) y la función de recompensa (consumir menos energía). El resultado inmediato fue una disminución del consumo eléctrico. Google ahorró cerca de un 40% de energía y decenas de millones de dólares, y eso tiempo después de los miles de intentos de ingenieros especializados en ese mismo sentido. La inteligencia artificial alcanzó niveles de desempeño verdaderamente capaces de revolucionar la industria entera.

DeepMind no deja de cosechar triunfos. Como tal vez ya todos sepan, logró vencer dieciocho veces al campeón mundial Lee Sedol en el juego de go, que hasta hace muy poco se consideraba el Everest de la inteligencia artificial (Banino y otros, 2018; Silver y otros, 2016). Este juego se desarrolla sobre un damero grande (un *go-ban*) de 18 casilleros por lado, es decir, $18 \times 18 = 324$ lugares donde pueden disponerse las fichas blancas y negras. La cantidad de combinaciones es tan grande que resulta imposible explorar sistemáticamente las intrincadas ramificaciones de jugadas posibles que se ofrecen a cada jugador. Y sin embargo, el aprendizaje por recompensa permitió que el *software* AlphaGo reconociera las combinaciones favorables o desfavorables mejor que cualquier jugador humano. Uno de los numerosos trucos consistió en hacer jugar al sistema contra sí mismo, exactamente como un jugador de ajedrez se entrena jugando a la vez con las piezas blancas y las negras. La idea es sencilla: al final de cada partida, el programa que gana refuerza sus acciones, mientras que el perdedor las debilita, pero los dos aprendieron a evaluar mejor sus jugadas.

Nos reímos a carcajadas del barón de Münchhausen, quien, en sus legendarias *Aventuras*, llamativamente lograba salir de un pozo tirando de sus propios cabellos (otra versión, más divulgada en inglés, insiste en que las víctimas del tironeo fueron los cordones de sus botines, *bootstraps*, término que ahora nos resulta útil). Sea como fuere, en inteligencia artificial, la desaforada locura del buen Münchhausen dio nacimiento a una estrategia bastante astuta, el *bootstrapping*: poco a poco, y a partir de una arquitectura mínima, desprovista de conocimientos, una red de neuronas puede convertirse en campeona mundial tan solo jugando contra sí misma, sin perder pelos, mañas ni botines.[6]

La idea de aumentar la velocidad del aprendizaje al generar la colaboración entre las dos redes o, por el contrario, al hacerlas combatir entre ellas, no deja de redundar en avances muy importantes para la inteligencia artificial. Una de las ideas más recientes, llamada “aprendizaje adversarial” (esto es, contra un adversario; Goodfellow y otros, 2014), consiste en entrenar dos sistemas: uno que aprende a clasificar y atribuir correctamente, por ejemplo, los cuadros de Van Gogh en comparación con los de otros pintores, y otro cuyo único objetivo es hacer fracasar al primero, creando cuadros falsos. El primero obtiene la recompensa cuando, ya alcanzado el nivel de *connoisseur*, detecta los cuadros auténticos de Van Gogh, mientras que el segundo gana créditos cuando logra engañar al ojo aguzado del primero. El resultado es que el algoritmo hace emerger no una, sino dos inteligencias artificiales: una, especialista mundial en Van Gogh, apasionada por el detalle más nimio que permita autenticar un genuino cuadro del maestro –práctica que desde el siglo XIX es distintiva de los mejores críticos y peritos–; y otra, consumada farsante, con talento para producir falsificaciones y pastiches capaces de burlar a los más distinguidos tasadores y autenticadores. Este tipo de entrenamiento también podría compararse con la preparación para un debate presidencial: un candidato puede afilar su oratoria si contrata a un buen imitador de su contrincante, uno que se mimetice con sus tics y sus réplicas más efectivas.

¿Podría aplicarse este enfoque al funcionamiento del cerebro humano? Nuestros dos hemisferios y numerosos núcleos subcorticales también albergan a una vociferante asamblea de expertos que luchan, se coordinan o se evalúan entre sí. Algunas de las áreas cerebrales aprenden a simular lo que hacen las otras: nos permiten anticipar e imaginar el resultado de nuestras acciones, a veces con un grado de verdad digno de los más consumados actores. La memoria y la imaginación nos pueden hacer ver la ensenada donde nos bañamos el verano pasado o el picaporte que queremos accionar a oscuras. Otras áreas aprenden a criticar a las demás; todo el tiempo evalúan nuestras capacidades y saben predecir las recompensas o los castigos que podríamos obtener: nos impulsan a reaccionar o a quedarnos tranquilos. Veremos también que la metacognición, esta capacidad de conocernos, de autoevaluarnos, de simular mentalmente qué pasaría si reaccionáramos de tal o cual manera, tiene un papel fundamental en los aprendizajes humanos. La opinión que nos forjamos de nosotros mismos nos ayuda a progresar o, al contrario, nos encierra en el círculo vicioso del fracaso. Por lo tanto, no es

desacertado pensar el cerebro como una tumultuosa reunión de expertos que compiten o colaboran entre sí.

Aprender es acotar el espacio de investigación

La inteligencia artificial contemporánea se enfrenta a un problema importante. Cuanto más rico en parámetros es el modelo interno, más difícil se vuelve encontrar la forma de ajustarlo al medio. Y en las redes de neuronas actuales, el espacio de investigación es inmenso. Los informáticos, por lo tanto, tienen que lidiar con una colosal explosión combinatoria: en cada etapa, millones de opciones son factibles, y sus combinaciones son tan amplias que resulta imposible explorarlas todas. Como resultado, el aprendizaje a veces es excesivamente lento: hacen falta miles de millones de intentos para orientar al sistema en la dirección correcta dentro de esta inmensa comarca de posibilidades. Y si bien la magnitud de los datos es más que considerable, se vuelve exigua en relación con las gigantescas dimensiones de ese espacio. Estas secuencias poco simpáticas son conocidas como “maldición de la dimensionalidad” (o, en términos más neutros, “efecto Hughes”): el aprendizaje puede llegar a ser una tarea muy ardua cuando el tablero de opciones disponibles tiene millones de teclas.

La inmensa cantidad de parámetros que ponderan las redes de neuronas suele acarrear una segunda dificultad, que los anglosajones llaman *overfitting*, el sobreajuste. El sistema dispone de tantos grados de libertad que le resulta más fácil memorizar todos los detalles de cada ejemplo que dar con una propiedad general que permita explicarlos todos.

Como dijo John von Neumann, uno de los “padres fundadores” de la informática y la ciencia de la información, “con cuatro parámetros, puedo reproducir la forma de un elefante, y con cinco, puedo hacer que mueva la trompa”. Esto significa que disponer de demasiados parámetros libres puede equivaler a una maldición: es muy fácil sobreajustar cualquier información si se memoriza cada detalle; pero eso no significa que el sistema resultante logre percibir algo significativo. En efecto, pueden ajustarse los parámetros del perfil del paquidermo sin comprender en profundidad dato alguno sobre los elefantes como especie. Disponer de demasiados parámetros libres

perjudica la abstracción. El sistema aprende sin dificultad, pero es incapaz de aplicar lo aprendido a situaciones nuevas a las cuales no fue expuesto. Sin embargo, en esta facultad de generalizar reside la clave del aprendizaje. ¿Cuál sería la utilidad de una máquina que pudiera reconocer la imagen de un elefante exclusivamente en el caso de que ya la hubiese visto o que tan solo ganara las partidas de go que ya hubiese jugado? Desde luego, el verdadero objetivo es reconocer cualquier imagen y ganar frente a cualquier jugador, sin importar si las movidas son conocidas o novedosas.

Una vez más, los informáticos están investigando varias soluciones posibles a estos problemas. Una de las intervenciones más eficaces, que puede acelerar el aprendizaje y, al mismo tiempo, mejorar la generalización, consiste en simplificar el modelo: al reducir al mínimo la cantidad de parámetros que ajustar, se fuerza al sistema a encontrar soluciones generales. Fiel a este principio, Yann LeCun inventó lo que llamó “redes neuronales convolucionales”, un dispositivo de aprendizaje artificial que se volvió omnipresente en el campo del reconocimiento de imágenes (LeCun y otros, 1998; LeCun, Bengio y Hinton, 2015). La idea es sencilla: para reconocer los elementos de una imagen hace falta casi el mismo trabajo en todos lados. En una foto, por ejemplo, no hay reglas que estipulen en qué lugar pueden aparecer rostros. Para reconocerlos, es suficiente entonces aplicar en cada punto el mismo algoritmo (buscar un óvalo, un par de ojos, etc.). Así, no es necesario aprender un modelo diferente para cada punto de la retina: lo que aprendemos en un lugar puede reutilizarse en los restantes.

A lo largo del aprendizaje, las redes convolucionales de Yann LeCun transcriben en el conjunto de la red los aprendizajes que realizan en determinado entorno, y realizan esta operación en todos los niveles, a escalas cada vez mayores. Aprenden una cantidad acotada de parámetros: el sistema apenas tiene que sintonizar esporádicamente un solo filtro que se aplica en todos lados, en vez de un sinnúmero de conexiones diferentes para cada sector de la imagen. Este simple truco mejora a escala masiva los desempeños y, en especial, la generalización a nuevas imágenes. El algoritmo que se aplica a una imagen nueva beneficia una experiencia inmensa, la de todos los puntos de todas y cada una de las fotos vistas hasta entonces. Acelera también el aprendizaje, porque la máquina solo explora un subconjunto de modelos de la visión. Incluso antes de aprender, ya sabe algo importante sobre el mundo: que el mismo objeto puede aparecer en cualquier lugar de la imagen.

Este truco se generaliza a muchas situaciones similares. Reconocer el habla, por ejemplo, consiste en hacer abstracción de la voz del hablante. Esto se logra si se fuerza a una red de neuronas a utilizar las mismas conexiones en diversas bandas de frecuencias –lo que suele conocerse como “registros” o “timbres”–, sea la voz grave o aguda. Reducir el número de parámetros que aprender asegura una mayor velocidad y también una mejor generalización a voces nuevas: el beneficio es doble; y de ese modo nuestros *smartphones* responden a nuestras respectivas voces.

Aprender es proyectar hipótesis *a priori*

La estrategia de Yann LeCun es un buen ejemplo de una noción mucho más general: la explotación de los conocimientos innatos. Si las redes convolucionales aprenden mejor y más rápido que otros tipos de redes neuronales, es porque no aprenden todo. Incorporan en su propia arquitectura una hipótesis fuerte: lo que aprendo en un lugar puede generalizarse luego en los restantes.

El principal problema del reconocimiento de imágenes es la invariabilidad: debo reconocer un objeto, sean cuales fueren su posición y su tamaño, incluso si se mueve a la derecha o a la izquierda o si se aleja. Es un desafío, pero también es una restricción muy aguda: espero que los mismos indicios permitan reconocer un rostro en cualquier punto del espacio. Al repetir el mismo algoritmo en todos lados, las redes convolucionales sacan gran provecho de esta restricción, ya que la integran dentro de su propia estructura. De forma innata, antes de cualquier aprendizaje, el sistema ya “sabe” esta propiedad clave del mundo visual. No aprende la invariabilidad: la supone verdadera *a priori* y la utiliza para reducir el espacio de aprendizaje. ¡Cuánta sagacidad!

La moraleja de esta historia es que la naturaleza y el aprendizaje (lo que en inglés se conoce como *nature* y *nurture*) no deben oponerse: se hace y se nace. El aprendizaje solo, en ausencia de cualquier restricción innata, simplemente no existe. De una u otra manera, cualquier algoritmo de aprendizaje incluye un conjunto de hipótesis previas acerca del ámbito al cual se lo aplica. En vez de intentar aprender todo de cero, es mucho más eficaz

apoyarse sobre hipótesis por *default*, que delimiten claramente el campo de investigación e integren en su propia arquitectura las leyes que lo gobiernan. Cuantas más hipótesis innatas haya, más veloz será el aprendizaje (¡con la previsible condición de que estas hipótesis sean correctas!). Esto es verdadero y tiene alcance universal. Sería falso, por ejemplo, creer que el programa AlphaGo Zero, que de tanto jugar contra sí mismo se convirtió en campeón de go, partió de la nada: su representación inicial incluía, entre otros, el conocimiento de las simetrías del juego, lo que divide por ocho la cantidad de estados que estudiar.

Nuestro cerebro también está lleno de hipótesis de todo tipo. Pronto veremos que, desde el nacimiento, el cerebro de los bebés está muy organizado y ya conoce de antemano muchas cosas sobre el mundo exterior: de manera implícita, sabe que el mundo está hecho de objetos sólidos que se mueven solamente cuando se los toca y que en condiciones normales nunca se incrustan unos en otros; también sabe de otras entidades más extrañas, que se desplazan por sí solas en función de intenciones y de creencias, y que hablan (sí, se trata de las personas). Los bebés no tienen que aprender todo acerca del mundo, ya que su cerebro está lleno de hipótesis innatas. Así, solo deben adquirir los parámetros específicos que varían de manera impredecible (como la forma del rostro, el color de los ojos, el tono de la voz o los gustos personales de quienes los rodean).

Por supuesto, si desde el comienzo nuestro cerebro conoce cosas tan abstractas como la diferencia entre las personas y los objetos, eso quiere decir que no las aprendió durante las pocas semanas de su corta vida, sino a lo largo de las eras de su evolución. La selección darwiniana es, de hecho, un algoritmo de aprendizaje, un programa increíblemente poderoso que ha estado funcionando durante miles de millones de años al mismo tiempo sobre millones de millones de máquinas: cada ser vivo (Dennett, 1996). Somos los herederos de una sabiduría infinita: por obra del método de prueba y error darwiniano, nuestro genoma asimiló el saber de las generaciones ancestrales que nos precedieron. Estos conocimientos innatos son de una índole diferente a la de los hechos que aprendemos mediante la experiencia: son tanto más abstractos, porque confieren a nuestras redes de neuronas los sesgos que facilitan el aprendizaje de las leyes de la naturaleza.

En síntesis, durante el embarazo, nuestros genes dictan una arquitectura cerebral que, al acotar las dimensiones del espacio de investigación, facilita y acelera los aprendizajes posteriores. En la jerga de los informáticos, se dice

que los genes configuran los llamados “hiperparámetros” del cerebro: las variables de alto nivel que especifican la cantidad de capas, el tipo de neuronas, la forma general de sus interconexiones, si están duplicadas en algún punto de la retina (LeCun tomó como modelo la visión humana y sus sistemas de células), y así sucesivamente. En nuestro cerebro, estas variables forman parte del genoma humano, no necesitamos aprenderlas, ya que la especie misma las internalizó durante el transcurso de su evolución.

Ante los estímulos sensoriales, entonces, el cerebro nunca se propone la sumisión simple y pasiva. Muy por el contrario, desde el principio dispone de un conjunto de hipótesis abstractas heredadas de su evolución, que proyecta sobre el mundo exterior. No todos los científicos están de acuerdo con esta idea que considero crucial: la ingenua filosofía empirista, que sirve de base a una gran cantidad de redes de neuronas actuales, está equivocada. Lisa y llanamente, no es verdad que nazcamos con circuitos desorganizados por completo, vírgenes de cualquier forma de conocimiento, que luego reciben la impronta de su ambiente. Aprender, tanto en personas como en máquinas, siempre tiene como punto inicial un conjunto de hipótesis previas que se proyectan sobre la información recibida, y de las que el sistema selecciona aquellas que mejor se adaptan al entorno. Como dice Jean-Pierre Changeux en *El hombre neuronal* (1983), “aprender es eliminar”.

[4] Esto es, tal como lo define el propio Dehaene (2006), que contemplan fluctuaciones indiscriminadas. En su presencia, el cerebro –o, según señala el presente libro, la red neuronal– “debe comportarse como un estadístico que recopila múltiples muestras antes de llegar a una conclusión sólida”. [N. de E.]

[5] Sobre la red neuronal artificial que aprende a jugar en consolas Atari, véase Mnih y otros (2015).

[6] El término inglés *bootstrapping* significa sacarse a uno mismo de una situación utilizando los recursos disponibles. La expresión, se cuenta, surge de una de las hazañas del barón de Münchhausen, inmortalizado en la obra de Rudolf Erich Raspe, que, según la versión inglesa, consistió en salir de una ciénaga y elevarse al cielo tirando de los cordones de sus zapatos.

2. Por qué nuestro cerebro aprende mejor que las máquinas actuales

Los éxitos recientes de la inteligencia artificial pueden hacernos creer que por fin logramos comprender cómo imitar e incluso superar el aprendizaje y la inteligencia de la especie humana. De acuerdo con algunos autoproclamados profetas, las máquinas estarían a punto de superarnos. No existe afirmación más falsa que esta. De hecho, aunque la mayoría de los investigadores en ciencias cognitivas admira los progresos actuales de las redes neuronales artificiales, saben muy bien que estas máquinas todavía tienen grandes limitaciones. A decir verdad, la mayoría de las redes neuronales artificiales implementan apenas las operaciones que nuestro cerebro realiza de manera inconsciente, en dos décimas de segundo, cuando percibe una imagen: la reconoce, la categoriza y accede a su significado (Dehaene, Lau y Kouider, 2017). Sin embargo, el cerebro va mucho más lejos: es capaz de explorar la imagen de manera consciente, con mucha atención, punto por punto, durante varios segundos, y formular representaciones simbólicas y teorías explícitas del mundo que podemos compartir con otras personas por medio del lenguaje.

Las operaciones de este tipo (lentas, razonadas, simbólicas) aún son privativas de nuestra especie. Los algoritmos actuales de *machine learning* las plasman de manera muy precaria. Si bien hay un progreso constante en los ámbitos de la traducción automática o el razonamiento lógico, una crítica que solemos hacer a las redes neuronales artificiales es que intentan aprender todo al mismo nivel, como si, sin excepción, los problemas pudieran reducirse a una cuestión de clasificación automática. Para quien no posee más que un martillo, ¡todo se parece a un clavo! El cerebro, en cambio, es mucho más flexible. Logra jerarquizar la información con gran velocidad y, cuando es posible, extrae de ella principios generales, lógicos y explícitos.

Lo que aún le falta a la inteligencia artificial

Es interesante intentar una enumeración de lo que aún le falta a la inteligencia artificial, porque permite definir, con la mayor precisión posible, lo que hay de exclusivo en nuestra propia capacidad de aprendizaje. Aquí propongo una pequeña lista –por supuesto, incompleta– de las funciones que hasta un niño muy pequeño posee y que hacen fallar lastimosamente a la mayor parte de las redes actuales.

El aprendizaje de conceptos abstractos

La mayoría de las redes neuronales artificiales solo logra generar modelos adecuados del primerísimo período del tratamiento de la información: menos de un quinto de segundo, lapso en que las áreas visuales analizan una imagen. Estos algoritmos de aprendizaje profundo están lejos de ser tan profundos como afirman. En realidad, de acuerdo con uno de sus inventores, Yoshua Bengio, estas redes tienden a aprender regularidades estadísticas superficiales a partir de los datos, en vez de conceptos abstractos de alto nivel (Jo y Bengio, 2017). Para reconocer la presencia de un objeto, se apoyan en elementos anecdóticos de la imagen, como el color o la forma. Si estos detalles varían, su desempeño se desploma: las actuales redes neuronales convolucionales son absolutamente incapaces de reconocer qué constituye la esencia de un objeto y de concebir que una silla sigue siendo una silla ya sea que tenga cuatro patas o una, o bien esté hecha de vidrio, de metal doblado o de plástico inflable. Esta tendencia a no prestar atención más que a la superficie de las cosas causa que las redes puedan cometer errores muy importantes. Hay una amplia bibliografía sobre la forma de hacer que una red neuronal se equivoque: tomen una banana, modifíquenle algunos píxeles o péguenle una etiqueta muy particular, ¡y la red neuronal pensará que es una tostadora!

Es cierto que si se le muestra una imagen a una persona durante una fracción de segundo, muchas veces comete el mismo tipo de errores que la máquina, y puede confundir, por ejemplo, un perro con un gato (Elsayed y otros, 2018). Pero si le damos un poco más de tiempo, el cerebro ya no se

equivoca. A diferencia de la máquina, los humanos tenemos la capacidad de cuestionar nuestras creencias y volver a prestar atención a aquellos elementos de la imagen que no se corresponden con nuestra primera impresión. Este segundo análisis, consciente e inteligente, recurre a las capacidades generales de razonamiento y abstracción. Las redes neuronales artificiales descuidan un factor esencial: el aprendizaje humano no es solo un filtro de reconocimiento de formas, sino la formación de un modelo abstracto del mundo. Al aprender a leer, por ejemplo, adquirimos un concepto abstracto de cada letra del alfabeto, que nos permite reconocerla bajo cualquier tipo de ropaje o con cualquier adorno posible, así como generar nuevas variantes:



Douglas Hofstadter, informático e investigador en ciencias cognitivas, dijo una vez que el verdadero desafío para la inteligencia artificial consistía en reconocer la letra A... Una humorada, por supuesto, pero una humorada profunda. En efecto, incluso en esta situación trivial, los humanos desplegamos una habilidad incomparable para la abstracción. Esta hazaña está en la base de una ocurrencia divertida de la vida cotidiana: el captcha, esta pequeña cadena de letras que algunos sitios de internet nos piden que reconozcamos para probar que somos seres humanos y no máquinas.

Durante años, los captcha se les resistieron a las máquinas. Pero la informática evoluciona rápido: en 2017, un sistema artificial logró reconocer los captcha casi tan bien como un humano (George y otros, 2017). No sorprende que este algoritmo imite a nuestro cerebro en varios aspectos. Lo que hace es una verdadera proeza: extrae el esqueleto de cada letra, la esencia de la letra A, y utiliza todos los recursos del razonamiento estadístico para verificar en qué medida esta idea abstracta es compatible con la imagen actual. Sin embargo, este sofisticado algoritmo informático solo se aplica a los captcha. El cerebro, en cambio, utiliza esta facultad de abstracción para tratar todas las variables de la vida cotidiana.

La eficiencia del aprendizaje

Todo el mundo está de acuerdo en que las redes neuronales actuales aprenden

con gran lentitud: les hacen falta miles, millones, incluso miles de millones de datos para adquirir la intuición en determinado ámbito. Tenemos pruebas experimentales de esta lentitud. Por ejemplo, hacen falta más de novecientas horas de juego para que la red neuronal diseñada por DeepMind alcance un nivel razonable en una consola Atari, ¡mientras que un ser humano alcanza el mismo nivel en dos horas! (Lake y otros, 2017).

Otro ejemplo: el aprendizaje del lenguaje. El psicolingüista Emmanuel Dupoux estima que en la mayoría de las familias francesas, un niño oye entre quinientas y mil horas de habla por año, lo que le resulta suficiente para aprender el dialecto de Descartes, incluidas sus peculiaridades, estrictamente “idiomáticas”. Se trata por cierto de una sobrestimación, dado que entre los chimanes, una población indígena de la Amazonía boliviana, los niños no oyen más que sesenta horas de habla por año, lo que no les impide volverse excelentes hablantes de chimán. En comparación, los mejores sistemas informáticos actuales de Apple, Baidu o Google necesitan entre veinte y mil veces más datos para lograr una mínima competencia lingüística. En el ámbito del aprendizaje, la eficiencia del cerebro humano todavía es inigualable: las máquinas están ávidas de acumular datos, mientras que los humanos son eficientes en su utilización. En nuestra especie, el aprendizaje sabe extraer lo mejor del dato más pequeño.

El aprendizaje social

La especie humana es la única que voluntariamente comparte la información: por imitación o gracias al lenguaje, aprendemos muchísimo de otros seres humanos. Al menos por el momento, esta capacidad queda fuera del alcance de las redes neuronales. Entre ellas, el conocimiento está encriptado, diluido entre los valores de cientos de millones de pesos sinápticos. En esta forma oculta, implícita, es imposible extraerlo y compartirlo selectivamente con otros. En cambio, en nuestros cerebros, la información del nivel más alto, que accede a la conciencia, puede enunciarse de modo explícito a otros. El conocimiento consciente va a la par de la posibilidad de informarlo verbalmente: cada vez que entendemos algo con suficiente claridad, en nuestro lenguaje de pensamiento resuena una fórmula mental (y podemos

usar las palabras de nuestro idioma para informar al respecto). La extraordinaria eficacia con la cual logramos, en pocas palabras, compartir el conocimiento con otros (“Para llegar a la panadería, doble a la derecha en la callecita que está detrás de la iglesia”) todavía no tiene igual en el mundo animal ni en la informática.

El aprendizaje en un solo intento

Un caso extremo de esta eficacia se verifica cuando aprendemos algo nuevo en un solo intento. Si utilizo un nuevo verbo, por ejemplo, “vualar”, solo una vez, ya es suficiente para que ustedes lo conozcan. Por supuesto, algunas redes neuronales también son capaces de almacenar un episodio específico. Pero lo que las máquinas todavía no hacen bien, y el cerebro humano puede hacer de maravillas, es integrar esta nueva información dentro de una red de conocimientos. Ustedes logran instantáneamente no solo memorizar el verbo “vualar”, sino también conjugarlo, derivar formas nominales e insertar todos estos ítems en otras frases: “¿Vualamos este tango?”, “¿Si usted pudiera, vualaría?”, “La vualatud es un problema muy actual”, etc. Cuando yo les digo “Vengan a vualar un rato con nosotros”, ustedes no solo aprenden una palabra, sino que también la insertan en un vasto sistema de símbolos y de reglas: es el infinitivo de un verbo que sigue el paradigma de la primera conjugación regular (yo vualo, tú vualas, etc.). Aprender es lograr insertar los conocimientos nuevos dentro de una red existente.

La sistematicidad y el lenguaje del pensamiento

Las reglas de la gramática no son más que un ejemplo del particular talento del cerebro para descubrir las reglas generales que se esconden detrás de los casos específicos. En temas de matemáticas, lengua, ciencia, música, el cerebro humano logra derivar principios muy abstractos, reglas sistemáticas que puede volver a aplicar en contextos muy diversos. Tomemos el ejemplo de la aritmética: nuestra capacidad de sumar dos números es extremadamente

general, una vez que aprendemos el procedimiento con números bajos, podemos sistematizarlo a números altos. Además, a partir de ellos hacemos inferencias de una generalidad extraordinaria. Una gran cantidad de niños y niñas, a la edad de 5 o 6 años, descubre que a cada número n le sigue $n + 1$, y, por lo tanto, que la secuencia de los números enteros es infinita: no existe un número mayor en sentido absoluto. Todavía recuerdo con emoción el momento en que tomé conciencia de lo que en realidad era mi primer teorema matemático. ¡Qué extraordinario poder de abstracción! ¿Cómo es que nuestro cerebro, con su cantidad finita de neuronas, alcanza a conceptualizar el infinito?

Hoy en día, ninguna red neuronal sabe representar conocimientos tan sistemáticos como “a cada número le sigue otro”. Los valores de verdad absoluta no son su fuerte. La sistematicidad, esta capacidad de generalizar sobre la base de una regla simbólica más que sobre la de un parecido superficial, escapa todavía a los modelos actuales (Fodor y Pylyshyn, 1988; Fodor y McLaughlin, 1990). Paradójicamente, los autoproclamados algoritmos de aprendizaje profundo todavía son por completo incapaces de alcanzar semejante profundidad.

Además, nuestro cerebro parece tener una fluida habilidad para crear fórmulas en una suerte de idioma mental. Por ejemplo, puede expresar el concepto de conjunto infinito porque posee una lengua interna dotada de funciones abstractas como la negación o la cuantificación (infinito = *no* finito = más que *cualquier* número). El filósofo estadounidense Jerry Fodor (1975) teorizó sobre esta facultad al postular que el pensamiento humano consiste en símbolos que se combinan de acuerdo con las reglas de un “lenguaje del pensamiento” (véase también Amalric y otros, 2017). Su peculiaridad consiste en ser recursivo: cada objeto creado como nuevo (en este caso, el concepto de infinito) puede reutilizarse de inmediato en combinaciones inéditas, sin límites. ¿Cuántos infinitos existen? Esta es la absurda cuestión que indagó el matemático Georg Cantor, y que lo llevó a formular una de las teorías matemáticas más innovadoras del siglo XX. La capacidad de hacer “un uso infinito de medios finitos”, según la hermosa fórmula de Wilhelm von Humboldt (1767-1835), caracteriza al pensamiento humano.

Algunos modelos de ciencia computacional intentan alimentarse con la adquisición de reglas matemáticas en niñas y niños; pero para hacerlo tienen que incorporar una forma de aprendizaje muy diferente, una que involucra reglas y gramáticas y que además se las arregla para optar por la más simple

y probable entre ellas (Piantadosi, Tenenbaum y Goodman, 2012, 2016). Desde esta perspectiva, el aprendizaje se parece a la programación: consiste en seleccionar, de entre todas las fórmulas internas disponibles en el lenguaje del pensamiento, la más sencilla que se ajuste a los datos.

Hoy en día, las redes neuronales son incapaces de representar la gama de frases, fórmulas, reglas y teorías abstractas con las cuales el cerebro del *Homo sapiens* construye modelos del mundo. Sin dudas, esto no ocurre por casualidad: hay allí algo profundamente humano, que no encontramos en el cerebro de otras especies animales y que las neurociencias contemporáneas todavía no lograron resolver, una auténtica singularidad de nuestra especie. Entre los cerebros de los primates, el humano parece ser el único que dispone de un conjunto de símbolos que se combinan de acuerdo con una sintaxis compleja y arborescente.[7] Mi laboratorio, por ejemplo, demostró que el cerebro humano no puede evitar oír una serie de sonidos (*bip bip bip bup*) sin derivar de inmediato, a partir de ellos, una teoría de su estructura abstracta (tres sonidos idénticos y luego un sonido diferente). En la misma situación, los monos detectan una serie de tres sonidos, oyen que el último es diferente, pero no parecen integrar estos conocimientos parciales en una sola fórmula. [8] Llegamos a saber esto porque, al sondear la actividad cerebral, vemos que se activan distintos circuitos para el número y para la secuencia; pero nunca observamos el patrón de actividad integrada que encontramos en el área del lenguaje humano conocida como “área de Broca”.

Del mismo modo, hacen falta decenas de miles de intentos para que un mono aprenda a invertir el orden de una secuencia (pasar de ABCD a DCBA), mientras que cinco ensayos son suficientes para cualquier niño o niña de 4 años (Jiang y otros, 2018). Incluso un bebé de pocos meses de vida ya codifica el mundo externo mediante reglas abstractas y sistemáticas, una capacidad por completo ajena a las redes neuronales convencionales tanto como a otras especies de primates.

La composición de los conocimientos

Tan pronto como aprendí, digamos, a sumar dos cifras, esta destreza pasa a formar parte de mi repertorio de talentos: se pone inmediatamente a

disposición de mis demás objetivos. Puedo utilizarla en decenas de contextos distintos; entre otros, cuando tengo que pagar mi cena en el restaurante o revisar mis declaraciones de impuestos. Además, puedo combinarla con otras destrezas aprendidas. No tengo dificultad, por ejemplo, para seguir un algoritmo que requiere que tome un número, le sume 2 y decida si el resultado es mayor o menor que 5 (Sackur y Dehaene, 2009; Zylberberg y otros, 2011).

Es sorprendente notar que las redes neuronales actuales todavía no poseen esta flexibilidad. Lo que aprenden permanece confinado en conexiones ocultas e inaccesibles, lo que dificulta su reutilización en otras tareas más complejas. La capacidad de *componer* las destrezas que se aprendieron en un momento previo, es decir, de recombinarlas para resolver problemas nuevos, está fuera de su alcance. De momento, la inteligencia artificial no resuelve más que problemas extremadamente acotados: el programa AlphaGo, pese a su corona de campeón mundial en el juego de go, es un experto limitado, incapaz de generalizar sus habilidades a cualquier otro juego, siquiera un poco diferente (incluido el juego de go sobre un tablero de 15×15 líneas en vez de uno de 19×19). Para nuestro cerebro, en cambio, aprender es casi siempre volver explícitos los conocimientos de modo tal que podamos reutilizarlos y recombinarlos con otros.

Una vez más, nos vemos ante una característica singular del cerebro humano, ligada al lenguaje y difícil de reproducir en una máquina. René Descartes ya lo había constatado en 1637 en su célebre *Discurso del método*:

Si las hubiera [máquinas] que semejasen a nuestros cuerpos e imitasen nuestras acciones, cuanto fuere moralmente posible, siempre tendríamos dos medios muy ciertos para reconocer que no por eso son hombres verdaderos; y es el primero, que nunca podrían hacer uso de palabras ni otros signos, componiéndolos, como hacemos nosotros, para declarar nuestros pensamientos a los demás, pues si bien se puede concebir que una máquina esté de tal modo hecha, que profiera palabras, [...] sin embargo, no se concibe que ordene en varios modos las palabras para contestar al sentido de todo lo que en su presencia se diga, como pueden hacerlo aun los más estúpidos de entre los hombres; y es el segundo que, [aunque] hicieran varias cosas tan bien y acaso mejor que ninguno de nosotros, no dejarían de fallar en otras[; así,] se descubriría que no

obran por conocimiento, sino solo por la disposición de sus órganos, pues mientras que la razón es un instrumento universal, que puede servir en todas las coyunturas, esos órganos, en cambio, necesitan una particular disposición para cada acción particular.

La razón, instrumento universal... Las capacidades que menciona Descartes apuntan hacia un segundo sistema de aprendizaje, jerárquicamente superior al precedente, y basado sobre reglas y símbolos. En sus primeras etapas, nuestro sistema visual posee vagas similitudes con las redes neuronales actuales: aprende a filtrar las imágenes de entrada y a reconocer las combinaciones frecuentes, y eso es suficiente para que reconozca un rostro, una palabra o una configuración del juego de go. Pero a continuación, el estilo de procesamiento cambia de manera radical: el aprendizaje empieza a parecerse a un razonamiento, una inferencia lógica que intenta captar las reglas de un dominio. Crear máquinas que alcancen este segundo nivel de inteligencia es el gran desafío de la investigación contemporánea en inteligencia artificial. Analicemos dos elementos que definen lo que los humanos hacemos cuando aprendemos en este segundo nivel y que suponen desafíos para los algoritmos más usuales de *machine learning*).

Aprender es inferir la gramática de un dominio

Es una característica de la especie humana: permanentemente intentamos derivar de una situación específica reglas abstractas, conclusiones de alto nivel, que luego ponemos a prueba en nuevas observaciones. Jerarquizar los conocimientos de este modo, en un intento por formular leyes abstractas que den cuenta de nuestras observaciones, es un método de aprendizaje de una extraordinaria eficacia, ya que las leyes más abstractas son precisamente aquellas que se aplican a la mayor cantidad de observaciones. Encontrar la ley correcta o la regla lógica que explica todos los datos disponibles es la mejor forma de acelerar el aprendizaje, y el cerebro humano domina a la perfección este juego.

Veamos un ejemplo: imaginemos que les muestro una decena de cajas llenas de esferas de diferentes colores. Tomo una caja al azar de la que hasta

ahora no saqué nada, meto la mano y saco una esfera verde. ¿Pueden deducir algo respecto del contenido de esta caja? ¿De qué color será la siguiente esfera que saque?

La primera respuesta que sin duda viene a la cabeza es: “No hay modo de saberlo. Usted casi no dio información, ¿cómo podríamos tener idea del color de las otras esferas?”. Sí, pero... ahora imaginen que, previamente, yo extraje varias esferas de las otras cajas y que ustedes constataron la siguiente regla: en determinada caja, todas las esferas son siempre del mismo color. En ese caso, el problema se vuelve trivial. Cuando les muestro una nueva caja, les basta con tomar una sola esfera verde para deducir que las restantes serán de ese mismo color. Si cuento con esta regla general, se puede aprender en apenas un intento.

Este ejemplo demuestra que un conocimiento de alto nivel, que usualmente llamamos “nivel meta”, puede guiar un conjunto de observaciones de nivel más bajo. Una vez que se aprende la metarregla abstracta que dice que todas las esferas de una caja son del mismo color, el aprendizaje se acelera muchísimo. Por supuesto, la regla puede resultar falsa. Ustedes se verían muy sorprendidos (debería decir “metasorprendidos”) si la décima caja que exploraran contuviera esferas de todos los colores. En ese caso, deberían revisar su modelo mental y replantearse la hipótesis de que todas las cajas son idénticas en eso. Tal vez arriesgarían una hipótesis de un nivel todavía más alto, una metametahipótesis, de acuerdo con la cual las cajas son de dos tipos: unicolores o multicolores, en cuyo caso les harán falta por lo menos dos extracciones por caja antes de llegar a alguna conclusión. En cualquiera de los casos, formular una jerarquía de reglas abstractas les habrá hecho ganar un tiempo precioso.

Aprender, en este sentido, supone administrar una jerarquía interna de reglas e intentar inferir lo antes posible las reglas más generales, aquellas que resuman la mayor cantidad de observaciones. Desde la infancia, nuestro cerebro aplica este principio de jerarquía. Observen a un niño de 2 o 3 años que se pasea por un jardín y a quien sus padres le enseñaron una palabra nueva, por ejemplo, “mariposa”. A menudo, le basta oír la palabra una vez o dos... *et voilà!*: su significado está memorizado. Esta velocidad de aprendizaje es increíble. Supera todo lo que la inteligencia artificial de hoy en día logra imitar. ¿Por qué es difícil el problema? Porque cada enunciado de cada palabra no acota por completo su significado. La palabra “mariposa” se pronuncia cuando el niño está inmerso en una escena compleja, llena de

flores, árboles, juguetes, personas, todos candidatos potenciales para el significado de esa palabra. Ni siquiera hablo de significados menos evidentes: cada instante que vivimos está lleno de ruidos, olores, movimientos, acciones, pero también de ideas abstractas dotadas de sus propios nombres, que debemos aprender. Así, “mariposa” podría querer decir “color”, “cielo”, “moverse” o “simetría”. El problema se plantea de modo más desconcertante para las palabras abstractas. ¿Cómo aprende un niño las palabras “pensar”, “creer”, “no”, “libertad” o “muerte”? ¿Cómo comprende lo que quiere decir “yo”, si cada vez que oyó ese término sus interlocutores hablaban de... ellos mismos?!

La incorporación veloz de palabras abstractas es incompatible con los modelos simplistas del aprendizaje, como el condicionamiento clásico de Pávlov o el aprendizaje asociativo de Skinner. Las redes neuronales que intentan simplemente correlacionar entradas con salidas, imágenes con palabras, por lo general necesitan miles de intentos para comprender que “mariposa” designa ese objeto colorido que está allí, en una esquina de la imagen..., y este principio de correlación palabra-imagen nunca permitirá comprender términos desprovistos de referencia fija, como “nosotros”, “siempre” y “olor”.

El problema del aprendizaje de palabras plantea un inmenso desafío para las ciencias cognitivas. Sin embargo, sabemos que una parte de la solución reside en la capacidad de los niños de formular representaciones no lingüísticas, abstractas, lógicas. Incluso antes de aprender sus primeras palabras, el niño posee un tipo de lenguaje del pensamiento con el cual puede formular hipótesis muy abstractas, y ponerlas a prueba. Su cerebro no es una *tabula rasa*, y el conocimiento innato que proyecta sobre el mundo externo angosta el espacio abstracto en que aprende. Es más, si un niño aprende tan rápido el significado de las palabras, mucho más rápido que una red neuronal, es porque selecciona entre el conjunto de hipótesis posibles y se deja guiar por una panoplia de reglas de alto nivel: estas metarreglas aceleran enormemente el aprendizaje, tal como vimos en el problema de las cajas de esferas de colores.

Una de las reglas que facilitan la adquisición del vocabulario es la de privilegiar siempre la hipótesis más limitada y más simple que sea compatible con los datos. Por ejemplo, cuando un bebé escucha a su mamá decir “¡Ah, llegó el perro!”, nada excluye que la palabra “perro” haga referencia a *aquel* perro en concreto (Tobi), o, a la inversa, a cualquier mamífero, cuadrúpedo,

animal o ser vivo. ¿Cómo es que el niño descubre el verdadero significado de la palabra: solamente los perros, pero todos los perros? Los experimentos sugieren que razona como el buen lógico, que pone a prueba todas las hipótesis y opta por la más simple que explique la frase que oyó. Entonces, cuando los humanos pequeños oyen la palabra “Tobi” (o “Snoopy” o “Lassie”, según las épocas o las geografías), siempre la oyen en un contexto específico, referida a determinado animal, y el conjunto más pequeño compatible con los distintos elementos observables está acotado a ese y solo ese perro. La primera vez que oye la palabra “perro” en un contexto específico, puede creer temporariamente que esa palabra solo se refiere a determinado animal; pero cuando la oye por segunda vez, en un contexto diferente, puede deducir que se refiere más bien a una categoría completa (figura 5). Un modelo matemático de este proceso predice que son suficientes tres o cuatro ocurrencias para alcanzar el significado apropiado (Tenenbaum y otros, 2011; Xu y Tenenbaum, 2007). Es el tipo de inferencia que los niños hacen diez, cien, o mil veces más rápido que las redes neuronales actuales.

Otros trucos permiten que los niños aprendan a hablar en un tiempo récord, en comparación con los sistemas actuales de inteligencia artificial. Una de estas metarreglas expresa una perogrullada: por lo general, quien habla presta atención a aquello de lo que habla. Una vez que adquiere esta regla, el bebé puede acotar considerablemente su búsqueda del significado: no está obligado a poner en relación cada palabra con todos los objetos presentes en la escena visual, como haría una computadora, hasta obtener suficientes datos que le prueben que cada vez que oye hablar de “mariposa” el pequeño insecto colorido está presente. Le basta con seguir la mirada de sus padres o la dirección de sus dedos para deducir de qué hablan: es lo que llamamos “atención compartida”, un principio fundamental del aprendizaje del lenguaje.

Al respecto, ya contamos con experiencias muy refinadas: si ustedes quieren acompañarnos en esta aventura de investigación, pueden tener su propia vivencia: tomen a alguna criatura de 2 o 3 años (siempre pidiendo permiso, claro) y muéstrenle un juguete nuevo ante la mirada de una entusiasta persona adulta que diga: “¡Ah, un wug!”. No hace falta más de un intento para que el niño retenga que “wug” es el nombre de ese objeto.

Aquí hay tres “tufas”. ¿Podrían encontrar las otras?

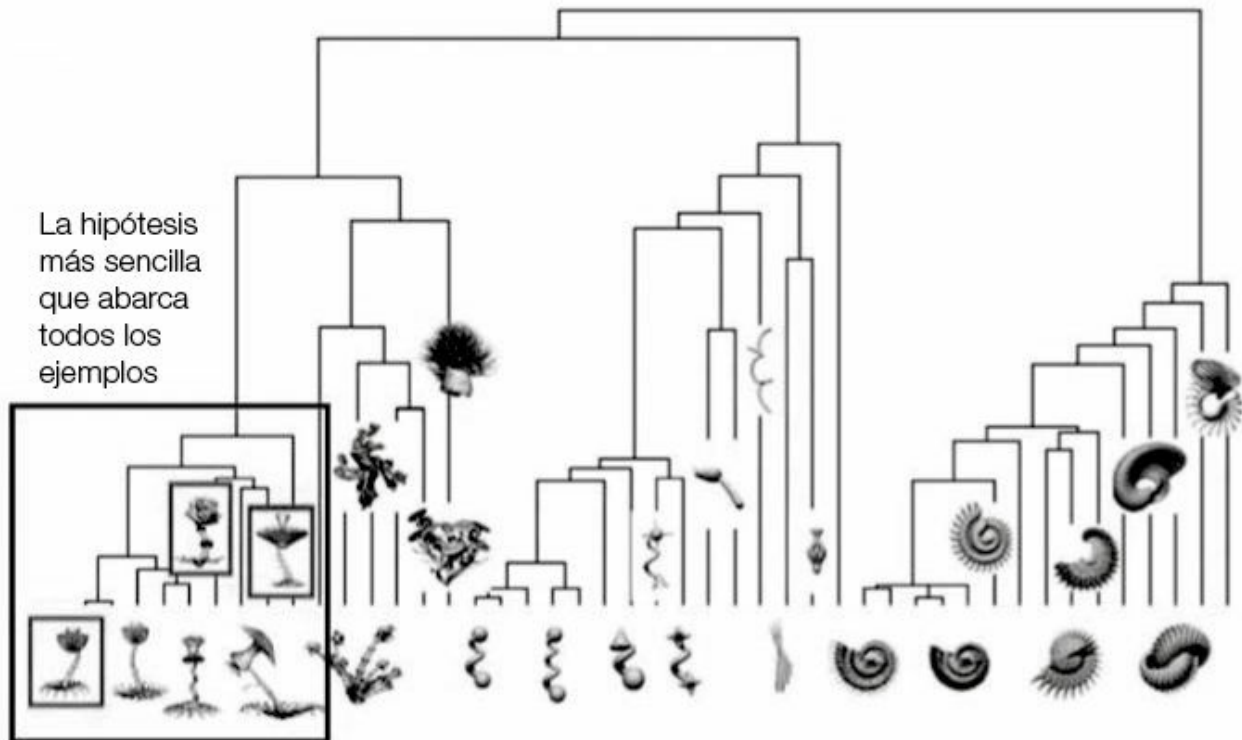


Figura 5. Aprender es seleccionar el modelo más sencillo que se ajuste a los datos.

Supongan que les muestro la tarjeta que se reproduce aquí arriba, y les digo que los tres objetos marcados con un contorno más oscuro son “tufas”. Con esta pequeña cantidad de datos, ¿cómo hacen para encontrar las otras tufas? Su cerebro construye un modelo de la manera en que se generaron esas formas, un árbol jerárquico de sus propiedades, y selecciona la rama más pequeña del árbol que sea compatible con todos los datos del problema.

Para proseguir la puesta en escena científica, compongan el mismo escenario y la misma situación, pero que el adulto, como en una obra de Jean Cocteau, se quede mudo: ahora deleguen en un altavoz –que habrán situado en el techo– la tarea de decir: “¡Ah, un wug!”. El niño no aprende cosa alguna, porque no logra comprender la intención del hablante (Baldwin y otros, 1996). Un bebé no retiene las palabras si no comprende la intención de quien habla. Esta capacidad, de una abstracción notoria, le permite aprender las palabras abstractas: para eso, debe literalmente ponerse en el lugar del locutor y comprender a qué pensamientos hacen referencia las palabras que pronuncia.

La niña o el niño de nuestra escena (al igual que sus colegas fuera de ella) utiliza todavía más metarreglas para aprender palabras. Por ejemplo, saca el mejor provecho del contexto gramatical: cuando oye: “Ey, miren la mariposa”, la presencia del determinante “la” le garantiza que esta palabra desconocida es un sustantivo. Esta es una metarregla que debe aprender: no nacemos con un conocimiento innato de todos los posibles determinantes de todas las lenguas. Sin embargo, la experiencia muestra que este aprendizaje es rápido: alrededor de los 12 meses, las criaturas ya registraron y retuvieron los principales determinantes y palabras funcionales de su lengua materna, y los utilizan para guiar los aprendizajes posteriores (Cyr y Shi, 2013; Shi y Lepage, 2008).

Si logran hacer esto, es indudablemente porque esas palabras son muy frecuentes y aparecen en numerosas frases, casi siempre ante un sustantivo o un sintagma nominal. El razonamiento puede parecer circular, pero en realidad no lo es: alrededor de los 6 meses, el bebé comienza por aprender sus primeros sustantivos, a menudo sustantivos ultrafamiliares como “taza” y “mesa”... Luego nota que estas palabras suelen estar precedidas por una palabra muy frecuente, el determinante “la”... de lo que deduce que, así escoltadas, todas esas palabras forman parte de la misma categoría, “sustantivo”, y que a menudo refieren a objetos de su entorno. Al fin y al cabo, esa metarregla le permite, cuando oye un enunciado como “la mariposa”, buscar en primer lugar un significado posible entre los objetos

que lo rodean, en vez de tratarlo como un verbo o un adjetivo. Luego, cada episodio de aprendizaje refuerza la regla, que a su vez facilita los aprendizajes posteriores, en un vasto movimiento que se acelera cada día. Los psicólogos del desarrollo afirman que niños y niñas confían en el *bootstrapping* sintáctico: el algoritmo de aprendizaje del lenguaje durante la infancia se las apaña para levantar vuelo por su cuenta,[9] gradualmente, al capitalizar una serie de pasos inferenciales. Pasos pequeños, pero sistemáticos.

Hay otra metarregla que el niño utiliza para acelerar el aprendizaje de palabras. Es el supuesto de exclusividad mutua, que podríamos sintetizar en esta fórmula: un solo nombre para cada cosa. En suma, esta regla dice que es poco probable que dos nombres diferentes designen un mismo concepto. Una palabra nueva, entonces, probablemente hará referencia a un concepto nuevo. Al tener presente este principio, cuando una niña oye una nueva palabra, puede acotar su búsqueda de significado a las cosas cuyo nombre todavía no conoce. Y, desde los 16 meses, lo aplica con bastante astucia.[10] Intenten realizar una experiencia. Tomen dos recipientes idénticos, uno de color azul, común y corriente, y el otro de un color poco habitual, por ejemplo, verde oliva; ahora díganle a un niño: “Dame el recipiente crapitado”. El niño les va a dar el recipiente que no es azul (palabra que ya conoce); parece dar por sentado que si ustedes hubiesen querido hablar del recipiente azul, habrían usado, precisamente, la palabra “azul”; *ergo...* ustedes deben estar refiriéndose al otro recipiente, menos conocido. Semanas más tarde, esa simple experiencia habrá sido suficiente para que recuerde que ese color extraño se llama “crapitado”.

Una vez más, notamos que la metarregla acelera enormemente el aprendizaje. Pero sin dudas hubo que aprenderla, o al menos consolidarla. En efecto, la observación demuestra que los niños de familia bilingüe la aplican tanto menos que los otros (Byers-Heinlein y Werker, 2009). Su experiencia bilingüe los hace darse cuenta de que papá y mamá utilizan palabras diferentes para decir lo mismo. Los niños monolingües, en cambio, se apoyan con más fuerza sobre la regla de exclusividad mutua. Han descubierto que, cuando ustedes usan una palabra nueva, es porque quieren que ellos aprendan un nuevo objeto o concepto. Si les dicen: “Dame el fax”, y a su alrededor no hay más que objetos familiares, van a buscar por todos lados, en un intento por encontrar ese misterioso objeto desconocido que les piden, no se les ocurrirá que podrían referirse a alguno de los objetos conocidos.

Todas estas metarreglas explican lo que llamamos “bendición de la abstracción”: las metarreglas más abstractas pueden ser la cosa más fácil de aprender, porque cada palabra que las criaturas oyen les aporta material de prueba al respecto. Entonces, una regla gramatical por el estilo de “las formas nominales suelen estar precedidas por un determinante” bien puede incorporarse tempranamente y guiar la subsiguiente adquisición de un amplio repertorio de sustantivos. Gracias a la bendición de la abstracción, hacia los 2 o 3 años, en este benemérito período bien llamado “explosión léxica”, el niño aprende entre diez y veinte palabras nuevas por día, sin dificultad alguna, sobre la base de tenues indicios que todavía hacen fracasar a los mejores algoritmos del planeta.

La capacidad de utilizar metarreglas parece demandar una buena dosis de inteligencia. Por lo tanto, ¿es exclusiva de la especie humana? No por completo. Existen otros animales también capaces de dar muestras de abstracción. Tomemos el caso de Rico, un perro pastor al que su dueña ha adiestrado para que le acerque los objetos más diversos (Kaminski, Call y Fischer, 2004). Es suficiente ordenarle “Rico, por favor: el dinosaurio” y el animal entra en la sala de juegos para volver, algunos segundos más tarde, con un dinosaurio de peluche en la boca. Los etólogos que lo investigaron demostraron que Rico conoce cerca de doscientas palabras. Pero lo más extraordinario es que también utiliza la regla de exclusividad mutua para aprender palabras nuevas. Si le dicen “Rico, el sikirid” (una palabra nueva), Rico se pone a buscar entre los objetos presentes, y solo regresa con un objeto nuevo, cuyo nombre todavía no conoce. Él también utiliza metarreglas como “un solo nombre para cada cosa”.

Matemáticos e informáticos ya comenzaron a diseñar algoritmos que permiten que las máquinas aprendan este tipo de jerarquía de reglas, de metarreglas e incluso de metametarreglas, hasta un nivel arbitrario. En estos algoritmos jerárquicos, cada episodio de aprendizaje restringe no solo los parámetros de nivel bajo, sino también los conocimientos de nivel más alto, los hiperparámetros abstractos que, a su vez, orientarán los aprendizajes posteriores. Si bien todavía no logran imitar la extraordinaria eficacia del aprendizaje de la lengua, a veces estos sistemas alcanzan notables desempeños. Por ejemplo, la figura 6 muestra uno de estos algoritmos recientes, una suerte de científico artificial que encuentra el mejor modelo del mundo exterior (Kemp y Tenenbaum, 2008). Este sistema dispone no solo de reglas abstractas, sino también de una verdadera gramática que le permite,

mediante la combinación de estas reglas elementales, engendrar una infinidad de estructuras de alto nivel. El sistema considera que una línea es un encadenamiento de puntos unidos el más próximo con el más próximo, por la aplicación de la regla “cada punto tiene dos vecinos, uno a la izquierda y uno a la derecha”, y así descubre por sí solo que esa es la mejor manera de organizar los entornos naturales (una línea que va de cero al infinito) o políticos (desde la extrema izquierda hasta la extrema derecha). Una variante de la misma gramática engendra un árbol binario en que cada nodo posee un padre y dos hijos. Esta estructura de árbol es la seleccionada automáticamente por el sistema cuando le pedimos que represente seres vivos. La máquina, como un verdadero Darwin artificial, redescubre de manera espontánea el árbol de las formas de la vida.

Otras combinaciones de reglas generan planos, cilindros y esferas y gracias a ellas el algoritmo va descubriendo de qué modo estructuras como esas se aproximan a la geografía local o global de nuestro planeta. Versiones más sofisticadas de ese mismo algoritmo son capaces de expresar ideas aún más abstractas. Por ejemplo, los informáticos estadounidenses Noah Goodman y Joshua Tenenbaum concibieron un sistema capaz de descubrir el principio de causalidad, la idea misma de que determinados acontecimientos causan otros (Goodman, Ullman y Tenenbaum, 2011; Tenenbaum y otros, 2011). Su formulación es abstrusa y matemática: “En un grafo direccional y acíclico que vincule diversas propiedades entre sí, existe un subconjunto de propiedades de las cuales dependen las demás”. No importa que esta expresión sea casi incomprensible, lo que importa es que da cuenta del tipo de fórmulas internas abstractas que esta gramática mental es capaz de expresar y de probar. El sistema pone a prueba miles de fórmulas como esta, y solo conserva aquellas que se ajustan a los datos. Como resultado, enseguida deduce el principio de causalidad (si en efecto algunas de las experiencias sensoriales que recibe son causas y otras, consecuencias).

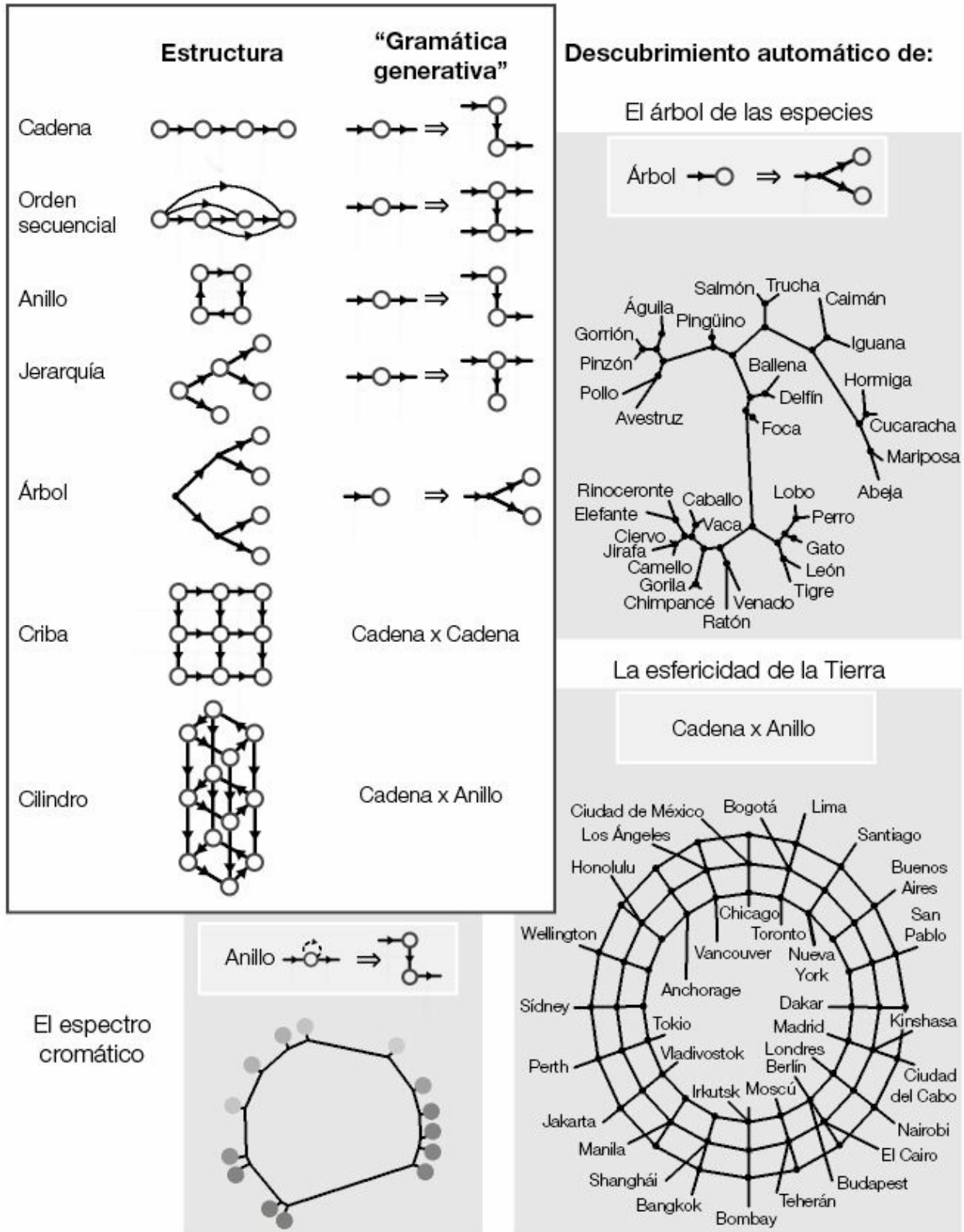


Figura 6. Aprender es inferir la gramática de un dominio. En el MIT, dos informáticos

inventaron un algoritmo que por sí solo descubre la estructura de un dominio científico. El sistema dispone de una gramática de reglas en que las combinaciones generan estructuras nuevas de los tipos más variados: líneas, planos, círculos, cilindros, árboles... Al seleccionar la estructura más apropiada para los datos, el algoritmo realiza descubrimientos que a los científicos les insumieron años de dedicación: el árbol de las especies animales (de Charles Darwin, en 1859), la esfericidad de la tierra (Parménides de Elea, hacia 600 a.C.), o incluso el espectro cromático (de Isaac Newton, en 1675).

Este es otro ejemplo de la bendición de la abstracción: el hecho de ser capaces de formular una regla de nivel alto permite acelerar enormemente el aprendizaje, porque la regla acota en gran medida el campo de posibilidades entre las cuales buscar. Gracias a esto, generaciones enteras de niñas y niños van en busca del Santo Grial de las Explicaciones; para conquistarlo, recurren a una inagotable reserva de preguntas y al omnipresente “¿por qué?”. Y así ponen a la especie en el camino de la búsqueda del conocimiento científico.

De acuerdo con este enfoque, aprender consiste en seleccionar, dentro de un gran conjunto de expresiones en el lenguaje del pensamiento, aquella que mejor se ajuste a los datos. Veremos que es un excelente modelo de lo que hace el niño: en su condición de científico en ciernes, formula teorías y las confronta con el mundo exterior. Esto supone que las representaciones mentales de los niños son tanto más estructuradas que las de las redes neuronales artificiales de la actualidad. Desde el nacimiento, el cerebro del niño dispone de dos ingredientes esenciales: por un lado, la maquinaria que permite engendrar fórmulas abstractas (un lenguaje del pensamiento, lenguaje que es combinatorio); y, por otro lado, la capacidad de seleccionar estas ideas con acierto, en función de su adecuación a los datos recibidos.

Ese es el nuevo enfoque acerca del cerebro: un inmenso modelo generativo, enorme arquitectura bien organizada, capaz de imaginar miles de reglas y estructuras hipotéticas, pero que poco a poco se atiene a solo aquellas que se condicen con la realidad.[\[11\]](#)

Aprender es razonar como un buen científico

¿Cómo selecciona el cerebro la mejor hipótesis? ¿Sobre la base de qué criterio acepta o rechaza un modelo del mundo exterior? Resulta que existe una estrategia ideal para hacerlo. Esta estrategia constituye el nodo de una de

las teorías más recientes y más productivas del aprendizaje: la hipótesis de que el cerebro se comporta como un científico en ciernes. Según dicho enfoque, aprender es razonar como un buen estadístico que elige, entre muchas hipótesis, aquella con mayor probabilidad de ser correcta, teniendo en cuenta los datos disponibles.

¿Cómo funcionan las teorías científicas? Cuando un científico formula una teoría, no se contenta con exhibir fórmulas matemáticas; también establece predicciones. La importancia de la teoría se juzga por la riqueza de las predicciones originales que se derivan de ella. La verificación o la refutación de estas predicciones acarrea la validación o la impugnación de una teoría. El investigador aplica una lógica simple: enuncia varias teorías, despliega la red de sus predicciones y elimina aquellas cuyas predicciones se ven invalidadas por la experiencia. Por supuesto, rara vez es suficiente un solo experimento: a menudo es necesario reiterarlo varias veces, en distintos laboratorios, para desentrañar qué es verdadero y qué falso. En palabras del filósofo de la ciencia, Karl Popper, la ignorancia retrocede constantemente a medida que una secuencia de conjeturas y refutaciones permite ir refinando la teoría.

El lento método de la ciencia se parece al modo en que aprendemos. En nuestras mentes, vemos disiparse la ignorancia, que retrocede a medida que el cerebro formula teorías más certeras del mundo exterior a partir de observaciones. Pero ¿esto es apenas una metáfora imprecisa? No: en efecto, es un enunciado bastante exacto acerca de lo que el cerebro debe de estar computando. Y hace ya unos treinta años o más que las hipótesis que describen al niño como un científico trajeron aparejada una serie de descubrimientos cruciales acerca del modo en que niñas y niños razonan y aprenden.

Durante mucho tiempo los matemáticos y los informáticos han teorizado sobre el razonamiento ideal en presencia de incertidumbre. Esta sofisticada teoría es llamada “bayesiana”, a partir del nombre de su primer descubridor, el reverendo Thomas Bayes, un pastor presbiteriano y matemático inglés, que llegó a ser miembro de la Royal Society británica. Pero tal vez deberíamos llamarla teoría “laplaciana”, dado que debemos su primera formalización completa al gran matemático francés Pierre-Simon, marqués de Laplace (1749-1827). Más allá de eso, hubo que esperar a los últimos veinte años para que este enfoque ganase prominencia y pasase a ser uno de los pilares de las ciencias cognitivas y el *machine learning*, porque de por sí garantiza la extracción de un máximo de información de cada dato. Aprender es ser capaz

de hacer la mayor cantidad posible de deducciones de cada observación, hasta de las más inciertas, y precisamente eso está respaldado por la regla de Bayes.

¿Qué descubrieron Bayes y Laplace? En pocas palabras: la forma correcta de hacer deducciones, razonando con las probabilidades para remontar cada observación, por tenue que sea, a su causa más plausible. Volvamos a los fundamentos de la lógica. Desde la Antigüedad, la humanidad ha comprendido cómo razonar con valores de verdad, “verdadero” o “falso”. Aristóteles introdujo las reglas de la deducción que conocemos como silogismos y que cada cual aplica de manera más o menos intuitiva. Por ejemplo, la regla llamada *modus tollens* (literalmente, “modo que niega”) dice que, si p implica q y se verifica que q es falsa, entonces p debe ser falsa. Esta es la regla que aplicó Sherlock Holmes en el famoso relato “Estrella de plata”, que presenta el episodio conocido como “El curioso incidente del perro a medianoche”:

—¿Existe algún otro detalle acerca del cual desearía usted llamar mi atención?

—Sí, acerca del curioso incidente del perro a medianoche.

—El perro no intervino para nada.

—Ese es precisamente el incidente curioso.

Efectivamente —razona Sherlock—, si el perro hubiera reparado en un desconocido, *entonces* habría ladrado. Como no lo hizo, eso quiere decir que el criminal era una persona familiar... un razonamiento que le permite al célebre detective acotar la investigación y desenmascarar al culpable.

¿Qué relación tiene esto, me dirán ustedes, con el aprendizaje? Bueno, aprender es también razonar como un detective: se trata siempre de volver a las causas ocultas de los fenómenos para deducir de ellas el modelo que las rige. Pero, en el mundo real, las observaciones pocas veces son verdaderas o falsas: son inciertas y probabilísticas. Y es exactamente allí donde se sitúan las contribuciones fundamentales del reverendo Bayes y del marqués de Laplace: la teoría bayesiana nos dice cómo razonar con las probabilidades, qué tipos de silogismos debemos aplicar cuando los datos no son perfectos, verdaderos o falsos, sino probabilísticos.

“La teoría de la probabilidad. La lógica de la ciencia”: esa es la traducción

posible del título de un libro fascinante del estadístico E. T. Jaynes (2003) sobre la teoría bayesiana, en cuyas páginas se demuestra que lo que llamamos “probabilidad” no es otra cosa que la expresión de nuestra incertidumbre. Con precisión matemática, la teoría explica las leyes según las cuales la incertidumbre debe evolucionar cuando realizamos una nueva observación. Es la extensión perfecta de la lógica entre la neblina de las probabilidades y las incertidumbres.

Tomemos un ejemplo, similar en espíritu al que el reverendo Bayes usó como base de su teoría en el siglo XVIII. Supongamos que observo a alguien lanzar una moneda, no trucada: tiene tantas posibilidades de caer sobre la cara como sobre la ceca (un 50% en cada caso). A partir de esta premisa, la teoría clásica de las probabilidades nos dice cómo calcular las probabilidades de observar determinado resultado (por ejemplo, cuál es la probabilidad de obtener tres caras y dos cecas). La teoría bayesiana nos permite realizar el camino inverso: de las observaciones a las causas. Nos dice, con exactitud matemática, cómo responder a preguntas del tipo “Si observo arrojar la moneda muchas veces, ¿debo revisar mi opinión sobre la moneda?”. La hipótesis *por default* es que la moneda no está trucada... pero si la veo caer veinte veces sobre la cara, debo revisar mi hipótesis: con seguridad la moneda sí está trucada. Mi hipótesis de partida se volvió improbable. ¿Pero en qué medida? La teoría de Bayes y Laplace explica precisamente cómo revisar nuestras creencias luego de cada observación. A cada hipótesis se le asigna un número que corresponde a un nivel de confianza. Luego de cada observación, ese número cambia en un valor que depende del grado de improbabilidad del resultado obtenido. Al igual que en las ciencias, cuanto más improbable es una observación experimental, más contradice las predicciones de nuestra teoría inicial, y con más confianza podemos rechazarla y buscar explicaciones alternativas.

La teoría bayesiana posee una eficacia notable: no es muy conocida por el gran público (entre otras cosas porque durante mucho tiempo fue un secreto militar muy bien guardado), pero tuvo una primera y contundente aplicación durante la Segunda Guerra Mundial, cuando el genial matemático y criptógrafo Alan Turing la utilizó para descifrar el código Enigma. Las distintas variedades de la máquina Enigma, que se utilizaban para encriptar todos los mensajes alemanes, estaban compuesta por numerosos engranajes, rotores y cables eléctricos, conectados de modo tal que sus combinaciones posibilitaban más de un millar de configuraciones diferentes que cambiaban

luego de cada letra. Cada mañana, el responsable del código ubicaba la máquina en la configuración prevista para ese día, tecleaba un texto y Enigma escupía una secuencia infernal de letras, que solo el poseedor de la clave podía decodificar. Para cualquier otra persona, el texto parecía perfectamente aleatorio. Pero aquí intervino la genialidad de Turing: descubrió que, si dos máquinas habían sido configuradas de la misma manera, eso inducía un ligero sesgo en la distribución de las letras; así, los dos mensajes tenían algunas posibilidades más de parecerse. Este sesgo era tan débil que ninguna letra aislada era prueba suficiente para sacar conclusiones certeras; pero al acumular esas improbabilidades, letra tras letra, Turing podía obtener mayor material probatorio de que la misma configuración había sido utilizada dos veces. Sobre esta base, y con la ayuda de lo que llamaron arbitrariamente “bomba” (un gran dispositivo electromecánico que prefiguró nuestras computadoras), Turing y su equipo lograron quebrar (“crackear”) el código Enigma, y así descifrar mensaje tras mensaje.

Una vez más, ¿cuál es la relevancia de estos pormenores a la hora de contemplar nuestro cerebro? El mismo tipo de razonamiento parece ocurrir dentro de nuestra corteza.[\[12\]](#) Según esta teoría, cada región del cerebro formula una o varias hipótesis y transmite sus predicciones a las otras regiones. De esta manera, cada módulo cerebral restringe las hipótesis del siguiente, mediante el intercambio de mensajes que son predicciones probabilísticas sobre el mundo exterior. Esas señales se llaman “descendentes” porque se inician en áreas cerebrales de alto nivel, como la corteza frontal, y poco a poco se desplazan hacia las áreas sensoriales de nivel más bajo, hasta, por ejemplo, la corteza visual primaria. La teoría dice que estas señales expresan la nube de hipótesis que nuestro cerebro juzga plausibles y desea poner a prueba.

En el nivel de las áreas sensoriales, estas hipótesis entran en contacto con los mensajes “ascendentes” provenientes del mundo exterior (de la retina, por ejemplo). En ese momento, el modelo se confronta con la realidad. El enfoque bayesiano supone que el cerebro debe calcular una señal de error: la diferencia entre lo que el modelo predice y lo que se observó. El algoritmo bayesiano luego indica cómo valerse de esta señal de error para ajustar el modelo interno del mundo. Si no hay error, eso significa que el modelo es correcto. Si no, la señal de error remonta la cadena de las áreas cerebrales involucradas y ajusta, paso a paso y en todas ellas, las hipótesis del modelo.

Con mayor o menor celeridad, el algoritmo converge en un modelo mental ajustado al mundo exterior.

De acuerdo con este enfoque acerca del cerebro, nuestro juicio de adultos se apoya sobre dos tipos de conocimientos: el saber innato de la especie (que los bayesianos llaman el *a priori*, la ponderación de los espacios de hipótesis que heredamos de la evolución) y la experiencia personal (la probabilidad *a posteriori*, la revisión de esas hipótesis, basada sobre todas las inferencias que pudimos hacer en el transcurso de nuestra vida). Esta división del trabajo disuelve el debate clásico entre lo innato y lo adquirido: la organización del cerebro nos aporta un potente equipo inicial y una máquina de aprendizaje igualmente poderosa. Cada conocimiento debe apoyarse sobre los dos componentes: un juego de hipótesis *a priori*, previo a cualquier interacción con el entorno, y la capacidad de organizarlas conforme a su plausibilidad *a posteriori*, una vez que disponemos de datos reales.

Podemos demostrar matemáticamente que el enfoque bayesiano es la mejor manera de aprender. Es la única forma de extraer la verdadera esencia de un episodio de aprendizaje, y de obtener el máximo provecho. Unos pocos bits de información, como las coincidencias sospechosas que Turing había detectado en el código Enigma, nos bastan para aprender. Si el sistema los trata como haría un buen estadístico, acumulando con paciencia las probabilidades, inevitablemente terminará por obtener una cantidad suficiente de datos para refutar determinadas teorías y validar otras.

¿Realmente el cerebro funciona así? ¿Es capaz de generar, desde el nacimiento, amplios espacios de hipótesis *a priori* entre las cuales aprende a elegir? ¿Procede por eliminación, formulando hipótesis sofisticadas y seleccionándolas en tanto y en cuanto los datos observables les sean favorables o contrarios? ¿El bebé se comporta, desde la cuna, como el más astuto de los estadísticos? ¿Es capaz de obtener el máximo de información posible de cada episodio de aprendizaje? Hagamos primer plano en los datos experimentales acerca del cerebro de los bebés.

[7] Sobre la recursividad como propiedad exclusivamente humana, véanse Dehaene, Meynel y otros (2015); Everaert y otros (2015); Hauser, Chomsky y Fitch (2002); Hauser y Watumull (2017).

[8] Sobre la singularidad humana para codificar una secuencia elemental de sonidos, véase Wang y otros (2015).

[9] Véase nota 6.

[10] Sobre el principio de exclusividad mutua en el aprendizaje de palabras, véanse Carey y Bartlett (1978); Clark (1988); Markman y Wachtel (1988); Markman, Wasow y Hansen (2003).

[11] Sobre el cerebro como un modelo generativo, véanse Lake, Salakhutdinov y Tenenbaum (2015); Lake y otros (2016).

[12] Acerca del modelo bayesiano del procesamiento de la información en la corteza, véase Friston (2005). Sobre datos empíricos en el paso jerárquico de mensajes de error probabilístico en la corteza, véanse también Chao y otros (2018); Wacogne y otros (2011).

Parte II

Cómo aprende nuestro cerebro

Esta facultad, existente en el alma de cada uno, y el órgano con que cada cual aprende, deben volverse, apartándose de lo que nace, con el alma entera, [...] hasta que se hallen en condiciones de afrontar la contemplación del ser.

Platón, *La República* (~380 a.C.)

El debate entre lo innato y lo adquirido se ha prolongado durante milenios. ¿El bebé se parece a una hoja en blanco, una *tabula rasa*, una botella vacía, que la experiencia debe llenar? Ya en el siglo IV a.C., en *La República*, Platón descartó de plano la idea de que nuestro cerebro viene al mundo desprovisto de cualquier forma de conocimiento y afirmó que, por el contrario, ya al nacer somos depositarios de dos enormes estructuras: “Esta facultad, existente en el alma de cada uno” (que en otras versiones se interpretó como el “poder del cerebro”) y “el órgano con que cada cual aprende”.

La convergencia es notable. Como acabamos de ver, a esa misma conclusión llega toda la investigación reciente en *machine learning*. El aprendizaje es infinitamente más eficaz si disponemos, por un lado, de un amplio espacio de hipótesis, un conjunto de modelos mentales con un sinnúmero de reglas entre las cuales elegir; y, por el otro, de algoritmos sofisticados que ajusten sus parámetros a partir de los datos recibidos del mundo exterior. Como decía uno de mis amigos, en el debate entre lo innato y lo adquirido, ¡hemos subestimado a los dos por igual! Aprender requiere una arquitectura bifronte: una inmensa cantidad de modelos generativos internos y algoritmos eficaces para ajustar estos modelos a la realidad.

Las redes neuronales artificiales lo hacen a su modo: encomiendan la representación de los modelos mentales a millones de conexiones ajustables. Sin embargo, pese a que estos sistemas procesan bastante bien el reconocimiento rápido e inconsciente de las imágenes o del habla, todavía no saben representar hipótesis más abstractas, como las reglas de la gramática o la lógica de las operaciones matemáticas.

El cerebro humano parece funcionar de modo diferente: nuestro saber se incrementa a partir de la gradual combinación de símbolos. Según la hipótesis del cerebro bayesiano, venimos al mundo con una enorme cantidad

de combinaciones factibles de pensamientos potenciales. Este lenguaje del pensamiento, dotado de hipótesis abstractas y de reglas gramaticales, preexiste a cualquier aprendizaje y constituye un inmenso espacio de hipótesis que se pondrán a prueba. Para hacerlo, el cerebro actúa como un científico: recolecta datos estadísticos y, entre un gran conjunto de modelos generativos posibles, selecciona aquel que mejor se ajusta al mundo exterior.

Esta perspectiva sobre el aprendizaje puede parecer contraintuitiva. Sugiere que en el cerebro de cada bebé ya anidan, en potencia, todas las lenguas del mundo, todos los objetos, todos los rostros, todas las herramientas que podrá encontrar, y también todas las palabras, todos los hechos, todos los acontecimientos que podrá recordar. La capacidad de combinación del cerebro debe ser tal que todos estos objetos del pensamiento ya estén a disposición –como veíamos, en potencia– con sus probabilidades *a priori* y con la capacidad de actualizarlas cuando la experiencia diga que deben ser revisadas. ¿Así aprende un bebé?

3. El saber invisible: las sorprendentes intuiciones de los bebés

¿Existe algo que parezca más indefenso que un recién nacido? ¿No parece completamente razonable pensar, como Locke, que su mente es una *tabula rasa* que espera recibir la huella del ambiente? Rousseau intentó resolver esta cuestión en su tratado *Emilio o De la educación* (1762): “Nacemos con capacidad para aprender, pero sin saber nada, sin conocer nada”. Cerca de dos siglos más tarde, la hipótesis fue retomada por Alan Turing, padre de la informática contemporánea: “Es verosímil que el cerebro del niño sea como una libreta que acabamos de comprar en la papelería. Un mecanismo muy pequeño y un montón de páginas en blanco”.

Hoy en día sabemos que esto es falso. Las apariencias engañan: pese a su inmadurez, el cerebro del recién nacido posee ya un amplio saber, heredado de su historia evolutiva. Sin embargo, la mayor parte de este conocimiento permanece invisible, ya que no se expresa en el comportamiento primitivo del bebé. Hicieron falta avances metodológicos importantes en el dominio de las ciencias cognitivas para poner en evidencia el extraordinario repertorio de capacidades con que cada bebé llega al mundo. Objetos, números, probabilidades, rostros, lenguaje... El alcance de los conocimientos innatos de los bebés está lejos de ser despreciable.

El concepto de objeto

Tenemos la intuición de que el mundo está hecho de objetos rígidos. En realidad, está constituido por átomos, pero en la dimensión en que vivimos, estos átomos se aglomeran y suelen comportarse como entidades coherentes que se mueven juntas y que a veces se entrecrocán sin perder su cohesión;

estos grandes haces de átomos forman lo que llamamos “objetos”. La existencia de los objetos es una propiedad fundamental de nuestro ambiente. ¿Tenemos que aprender esto? No. Millones de años de evolución parecen haber integrado este conocimiento en lo más profundo de nuestros cerebros. Desde que tiene unos pocos meses de vida, el bebé ya sabe que el mundo está formado por objetos, que se mueven como un bloque, que ocupan espacio, que no pueden desaparecer sin motivo, y que es imposible que dos de ellos ocupen un mismo lugar (Baillargeon y DeVos, 1991; Kellman y Spelke, 1983). En cierto sentido, su cerebro ya conoce las leyes de la física: espera que la trayectoria de un objeto, tanto en el espacio como en el tiempo, sea continua, sin desapariciones ni saltos repentinos.

¿Cómo lo sabemos? Porque los bebés manifiestan sorpresa en situaciones experimentales que violan las leyes de la física. Actualmente, en los laboratorios de ciencias cognitivas, los investigadores se transforman en ilusionistas (figura 7). Usan teatrinos especialmente creados para bebés y despliegan todos los trucos y las materializaciones de consumados prestidigitadores; hay objetos que aparecen, desaparecen, se multiplican, se disuelven... Cámaras ocultas monitorean los experimentos y los resultados son claros: hasta los bebés de unas pocas semanas son sensibles a la magia. Ya poseen intuiciones profundas acerca del mundo físico y, como todos nosotros, se quedan estupefactos cuando estas se refutan o quebrantan. Al hacer *zoom* en los ojos del niño, para así confirmar hacia dónde se orienta su mirada y durante cuánto tiempo, los investigadores logran medir el grado de sorpresa y deducir qué era lo que la criatura esperaba ver.

Les propongo que escondan un objeto detrás de un libro y luego, con un golpe seco, lo aplasten, como si ese objeto ya no existiera (en realidad, lo habrán hecho salir por una pequeña puerta trampa), y entonces lo comprobarán: ¡los bebés se quedan estupefactos! No logran comprender que algo sólido pueda evaporarse sin dejar huellas. También se sorprenden cuando un objeto desaparece detrás de una pantalla y reaparece detrás de otra, sin que lo hayan visto franquear el espacio vacío entre las dos. Del mismo modo, los desconcierta que un pequeño tren descienda por una pendiente y atraviese una pared rígida sin que esta lo detenga. Esperan que un objeto forme un todo coherente: si ven dos extremidades de una varilla moverse con coherencia de los dos lados de una pantalla, imaginan que pertenecen a una sola varilla, y se muestran muy sorprendidos cuando la pantalla se baja y deja ver dos varillas distintas (figura 8).

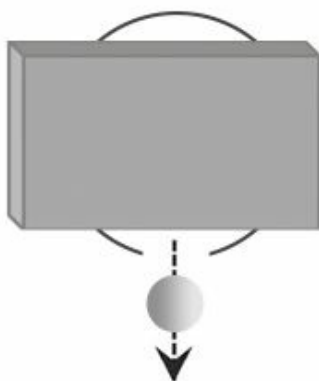


Figura 7. Lejos de ser pizarras vírgenes de conocimiento, los bebés poseen amplios saberes en cuestiones de física, aritmética, probabilidad o geometría. En el laboratorio, cuando se los somete a situaciones que violan las reglas de estos campos, la medida de su sorpresa revela la sofisticación de sus intuiciones acerca del mundo.

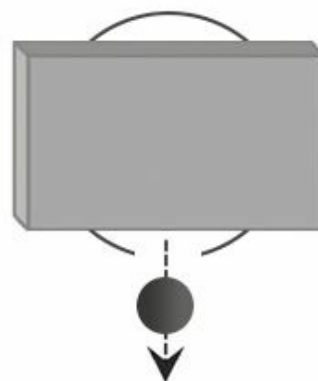
Intuición de los números y de las probabilidades:



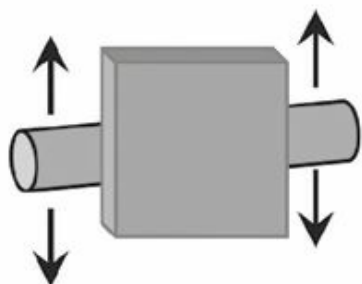
Sin sorpresa



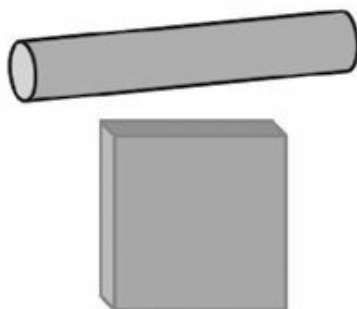
¡Sorpresa!



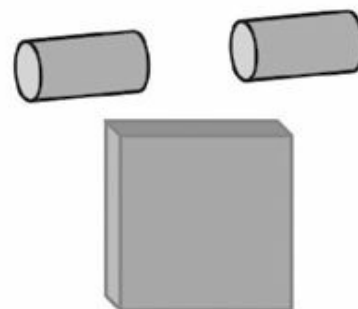
Intuición de los objetos:



Sin sorpresa

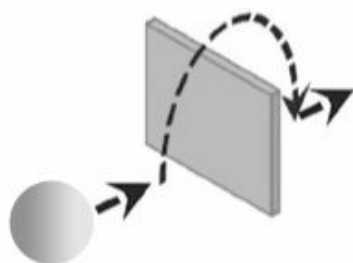


¡Sorpresa!



Intuición de la psicología:

Sin sorpresa



¡Sorpresa!

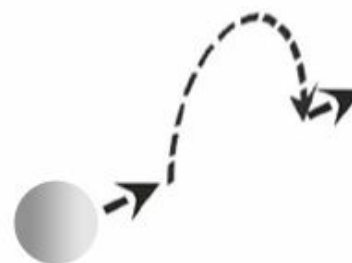
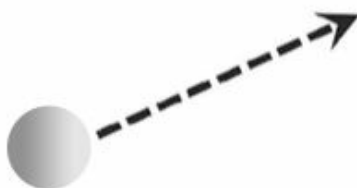


Figura 8. Para evaluar las intuiciones de los bebés, se determina si miran durante más tiempo

una escena sorprendente. Cuando en una caja hay mayoría de objetos blancos, el bebé se sorprende cuando ve salir de ella un objeto negro (intuición de los números y de las probabilidades). Si dos extremos de una varilla se mueven de forma coherente, el bebé se queda anonadado cuando se pone en evidencia que en realidad son dos objetos diferentes (intuición de los objetos). Y si una esfera salta de forma espontánea por encima de un muro, el bebé deduce que se trata de un ser animado dotado de una intención (en este caso, la de moverse hacia la derecha) y se queda atónito al ver que la esfera continúa saltando cuando el muro ya no está (intuición de la psicología: ¿para qué debería saltar un obstáculo inexistente?).

Esto demuestra que los bebés poseen un vasto saber acerca del mundo, aunque no conocen todo desde el principio, ni mucho menos. Hacen falta algunos meses para que los bebés comprendan cómo dos objetos se sostienen el uno al otro (Baillargeon, Needham y Devos, 1992; Hespos y Baillargeon, 2008). Al principio, ignoran que un objeto se cae cuando lo sueltan. Solo muy gradualmente toman conciencia de todos los factores que causan que un objeto se caiga o no se caiga. En un primer momento, se dan cuenta de que los objetos se caen cuando pierden su apoyo, pero creen que es suficiente cualquier tipo de contacto para mantenerlos quietos, por ejemplo, cuando ponemos un juguete al borde de una mesa. Poco a poco, descubren que el objeto no solo debe estar en contacto con la mesa, sino que es necesario que esté bien puesto ahí: ni debajo, ni al lado, sino sobre la mesa misma. Y todavía les harán falta algunos meses para descubrir que esta regla no es suficiente: lo que debe apoyarse arriba de la mesa es el centro de gravedad del objeto.

Piensen en esto la próxima vez que se desesperen al ver que su bebé deja caer, por décima vez, la cuchara mientras almuerza. ¡Solamente está experimentando! Como cualquier científico, necesita realizar una serie de experiencias para rechazar, una tras otra, todas las teorías incorrectas que postulan que

1. los objetos se sostienen por sí solos en el aire;
2. es suficiente que se toquen con otro para no caerse;
3. es suficiente que estén arriba de otro para no caerse;
4. la mayor parte del volumen del objeto debe estar encima de otro para no caerse; y así sucesivamente.

Esta actitud experimental no se extingue en la edad adulta. Nos encantan todos los dispositivos que parecen desobedecer las leyes habituales de la física (globos de helio, móviles en equilibrio, un móvil con un centro de gravedad periférico...) y disfrutamos de los espectáculos de ilusionismo en que los conejos desaparecen en un sombrero y las asistentes son cortadas en dos. Si estas situaciones nos divierten, es porque violan las intuiciones que poseemos desde el nacimiento y que hemos refinado durante el primer año de vida. Joshua Tenenbaum, especialista en inteligencia artificial del Massachusetts Institute of Technology (MIT), propuso la hipótesis de que los bebés tienen un motor físico en la cabeza, un modelo mental del comportamiento de los objetos que se asemeja al utilizado por los programas de videojuegos para simular mundos virtuales. Al hacer funcionar esas simulaciones dentro de su cabeza y compararlas con la realidad, los bebés descubren, muy temprano, lo que es físicamente posible o probable.

El sentido del número

Tomemos un segundo ejemplo: la aritmética. ¿Existe algo más evidente para nosotros que el hecho de que un bebé lo ignora todo acerca de las matemáticas? Y sin embargo, desde la década de 1980, existen experimentos que demuestran lo contrario (Izard, Dehaene-Lambertz y Dehaene, 2008; Izard y otros, 2009; Starkey y Cooper, 1980; Starkey, Spelke y Gelman, 1990).[\[13\]](#) En uno de ellos, se presentan reiteradas veces ante los bebés imágenes que muestran dos objetos. Después de un rato, se aburren... hasta que se les muestra una tarjeta con tres objetos: con la atención renovada, se quedan mirando esta nueva escena durante más tiempo, lo que pone de manifiesto que detectaron el cambio. Al manipular la índole, el tamaño y la densidad de los objetos, podemos probar que los niños perciben el número abstracto, el numeral cardinal del conjunto, y no tal o cual parámetro físico. La mejor prueba de que poseen este abstracto “sentido del número” es que generalizan de los sonidos a las imágenes: si escuchan *tu tu tu tu*, vale decir, cuatro sonidos, se interesan en una imagen que también tenga cuatro objetos, en lugar de prestar atención a una en que hay doce, y viceversa (Izard y otros,

2009). Los experimentos bien controlados, como este, abundan y demuestran que, desde el nacimiento, los bebés tienen la capacidad de reconocer de manera intuitiva un número aproximativo, sin contar, ya sea que la información provenga del oído o de la vista.

¿Los bebés también pueden calcular? Supongan ahora que un niño ve un primer objeto desaparecer detrás de una pantalla, y a este lo sigue un segundo objeto. A continuación, la pantalla se baja... ¡y resulta que hay solo un objeto! Los bebés expresan su sorpresa mediante una prolongada investigación de la escena inesperada (Koechlin, Dehaene y Mehler, 1997; Wynn, 1992). En cambio, si ven los dos objetos esperados, no le dedican más que un breve vistazo. Este comportamiento de “sorpresa cognitiva”, en reacción a la violación de un cálculo mental, muestra que, desde los pocos meses de vida, el niño comprende que $1 + 1$ debería ser 2. Lo que ocurre es que el bebé construye un modelo interno de la escena escondida, y sabe manipularlo agregándole o quitándole objetos. Y eso funciona para $1 + 1$ y $2 - 1$, pero también para $5 + 5$ y $10 - 5$... Siempre que el error es lo suficientemente grande, un bebé de 9 meses se sorprende si la pantalla presenta un cálculo incorrecto: infiere que $5 + 5$ no puede ser 5, y que $10 - 5$ no puede ser 10 (McCrink y Wynn, 2004).

¿Se trata en efecto de una capacidad innata? ¿Pueden ser suficientes los primeros meses de vida para que el niño aprenda el comportamiento de conjuntos de objetos? Si bien no cabe duda de que, con el transcurso de los meses, el niño refina la precisión con que percibe el número (Halberda y Feigenson, 2008; Piazza y otros, 2010, 2013), los datos también demuestran con plena claridad que el punto de partida no es una *tabula rasa*. Los recién nacidos perciben los números desde que tienen pocas horas de vida, y lo mismo ocurre con los monos, las palomas, los cuervos, las salamandras, los pollitos y hasta los peces. En el caso de los pollitos, los científicos pueden controlar todos los estímulos sensoriales desde el momento en que se rompe el huevo: el pequeño pollito no ha visto un solo objeto... y, sin embargo, reconoce los números (Rugani y otros, 2009, 2015).

Estos experimentos demuestran que la aritmética forma parte de las competencias innatas que la evolución otorga a numerosas especies, incluida la humana. Se detectó un circuito cerebral para los números en el mono y también en el cuervo. Sus cerebros albergan “neuronas numéricas” que se comportan de manera muy similar: son sensibles a cantidades específicas de objetos. Algunas neuronas prefieren ver un objeto, otras dos, tres, cinco o

incluso treinta, y estas células están presentes aun entre los animales que no han recibido entrenamiento específico (Ditz y Nieder, 2015; Viswanathan y Nieder, 2013). Mi laboratorio usó técnicas de imágenes del cerebro para poner en evidencia que en ubicaciones homólogas del cerebro humano los circuitos neuronales incluyen células similares especializadas en el cardinal de un conjunto concreto; además, en fecha reciente, gracias a los avances en las técnicas de registro, se tomaron imágenes directas de esas neuronas en el hipocampo humano (Piazza y otros, 2004; Kutter y otros, 2018).

De paso, estos resultados echan por tierra una de las principales teorías del desarrollo del niño, la del gran psicólogo suizo Jean Piaget, quien consideraba que los bebés ignoraban la “permanencia del objeto” –el hecho de que el objeto continúa existiendo cuando no lo vemos– y que esta capacidad recién aparecía hacia finales del primer año. Pensaba también que el concepto abstracto del número estaba fuera del alcance de los niños de escasos años de vida, y que lo aprendían lentamente por obra de una gradual abstracción de las medidas más concretas de tamaño, longitud o densidad. En realidad, es exactamente a la inversa. Los conceptos de objeto y de número son características fundamentales de nuestro pensamiento, forman parte del “núcleo de conocimientos” con que llegamos al mundo y que, por sus combinaciones, nos permite formular razonamientos más complejos (Spelke, 2003).

El sentido del número es solo un ejemplo de lo que llamo “saber invisible” de los niños, esas intuiciones precoces que poseen desde el nacimiento y que guían todos sus aprendizajes. Veamos otros ejemplos de las competencias que los investigadores demostraron que tienen los bebés de algunas pocas semanas de vida.

La intuición de las probabilidades

Del número a la probabilidad, no hay más que un paso... ¡que los investigadores dieron recientemente al preguntarse si los bebés de pocos meses podrían anticipar el resultado de un sorteo de lotería! En este experimento, se les presenta a los bebés una urna transparente llena de bolillas que se mueven al azar. La caja contiene cuatro bolillas: tres blancas y

una negra (figura 8). En la parte inferior, hay una salida. En un momento dado, un cartel esconde la urna... y de allí sale una bolilla, o negra o blanca. Frente a esto, el nivel de sorpresa del niño está directamente relacionado con el grado de improbabilidad de lo que ve: si lo que sale es una bolilla blanca, vale decir, el acontecimiento más probable (dado que en la caja hay una mayoría de bolillas de ese color), el bebé mira apenas un breve instante... En cambio, si se presenta el evento improbable, la salida de una bolilla negra, que solo tenía una posibilidad sobre cuatro de producirse, la mira durante mucho más tiempo. Controles posteriores confirman que, en sus pequeñas cabezas, las niñas y los niños ponen en juego una detallada simulación mental de la situación y de las probabilidades asociadas. Así, si se introduce un panel que bloquea las bolillas, o si se aleja más o menos las bolillas de la salida, o bien se hace variar el lapso previo a la aparición de los objetos, el niño integra todos estos parámetros a su cálculo mental de probabilidades. Como en el caso de los objetos que desaparecían detrás de una pantalla, la duración de su mirada atenta refleja la improbabilidad de la situación observada: los niños parecen calcular esta improbabilidad sobre la base de la cantidad de objetos involucrados.

Todas estas competencias superan a las redes neuronales actuales. En efecto, la reacción de sorpresa de los niños dista de ser una respuesta trivial: indica que el cerebro fue capaz de hacer un cálculo de probabilidades y llegó a la conclusión de que el acontecimiento observado tenía una ínfima posibilidad de producirse. Si las miradas de los bebés presentan reacciones de sorpresa muy elaboradas, se debe a que sus cerebros son capaces de realizar cálculos probabilísticos. De hecho, hoy en día, una de las teorías más en boga sobre el funcionamiento del cerebro lo considera un calculador probabilístico que manipula las distribuciones de probabilidad y las utiliza para anticipar acontecimientos futuros. Los experimentos revelan que los bebés están equipados con este sofisticado calculador.

Una serie de experiencias recientes demuestra que el bebé ya posee todos los mecanismos para realizar inferencias probabilísticas complejas. ¿Recuerdan la teoría matemática del reverendo Bayes, esta teoría de las probabilidades que permite retrotraer una observación a sus causas probables? Bueno, un bebé de apenas algunos meses parece seguir esta regla en su razonamiento (Xu y García, 2008). En efecto, no solamente sabe pasar de la caja de bolillas de colores a las probabilidades correspondientes (un

razonamiento hacia delante), como acabamos de ver, sino también de sus observaciones al contenido de la caja (un razonamiento en sentido inverso).

Un nuevo experimento: muéstrenle a un niño una caja opaca, sin que pueda ver su contenido. Después pídanle a una persona que, con los ojos tapados, extraiga al azar una serie de bolillas. Estas aparecen de a una, y se observa que la mayoría son rojas. ¿El bebé puede deducir que en caja hay una mayor cantidad de bolillas rojas? ¡Sí! Y si luego abrimos la caja y le mostramos que contiene una mayoría de bolillas verdes, se sorprende y se queda mirando durante mucho más tiempo que si la caja contuviera mayoría de rojas. Su lógica es impecable: si la caja está llena en mayor parte por bolillas verdes, ¿cómo se explica que la extracción aleatoria esté dominada por bolillas rojas?

Esto parece poco, pero involucra una extraordinaria capacidad de razonamiento implícito, inconsciente, pero muy sofisticado, que funciona en los dos sentidos: dada una muestra, los niños y las niñas pueden deducir las características del conjunto del cual proviene; o, al contrario, dadas las características del conjunto, pueden deducir a qué debería parecerse una muestra aleatoria.

Existe, desde el nacimiento, una lógica intuitiva. Numerosas variantes de esta experiencia demuestran en qué medida el niño se comporta como un científico en ciernes capaz de razonar como un buen estadístico, de eliminar las hipótesis menos probables y de investigar las causas ocultas de los fenómenos (Gopnik, 2000; Gopnik y otros, 2004). Por ejemplo, la psicóloga estadounidense Fay Xu ha demostrado que si un bebé de 11 meses ve a una persona que, con los ojos destapados, extrae una mayoría de bolillas rojas de una urna, y luego descubre que esa caja contiene en su mayoría bolillas amarillas, se sorprende, por supuesto, pero también hace una deducción adicional: ¡que esa persona prefiere las bolillas rojas! (Denison y Xu, 2010; Gweon, Tenenbaum y Schulz, 2010; Kushnir, Xu y Wellman, 2010). Y si ve que una serie no es aleatoria, sino que mantiene un patrón específico – digamos, la alternancia perfecta de una esfera amarilla, una roja, una amarilla, una roja, y así sucesivamente–, llega a la conclusión de que quien realiza la extracción es un humano, no una máquina (Ma y Xu, 2013).

La lógica y la probabilidad están estrechamente ligadas. Como bien decía Sherlock Holmes, “cuando se ha eliminado lo imposible, lo que queda, por improbable que parezca, debe ser la verdad”. En otras palabras, se puede pasar de la probabilidad a la certidumbre cuando, por obra del razonamiento, se llega a eliminar determinadas posibilidades. Si un bebé sabe lidiar con las

probabilidades, debe dominar en igual medida la lógica, porque el razonamiento lógico no es otra cosa que la restricción del razonamiento probabilístico a las probabilidades 0 y 1 (Cesana-Arlotti y otros, 2018). Exactamente esto demostró el filósofo y psicólogo del desarrollo Luca Bonatti poco tiempo atrás. Su experimento comienza así: esconde una flor y un dinosaurio detrás de una pantalla. Luego uno de esos objetos sale de la pantalla, pero no sabemos cuál, porque está parcialmente oculto en un recipiente, por lo que solo podemos ver su parte superior. Más tarde, el dinosaurio sale del otro lado de la pantalla, bien visible. En este momento, el niño puede efectuar un razonamiento lógico: “Lo que estaba dentro del recipiente era o la flor o el dinosaurio. Sin embargo, no es el dinosaurio, porque lo vi afuera. Entonces, es la flor”. Incluso los bebés de 10 meses parecen ser capaces de hacer este razonamiento: no se sorprenden si la flor sale del recipiente, sino solamente si aparece el dinosaurio. La mirada traduce la intensidad del razonamiento lógico: al igual que en los adultos, la pupila se dilata en el momento preciso en que la deducción se vuelve posible. Como un buen Sherlock Holmes, el bebé parece comenzar con varias hipótesis (es la flor, es el dinosaurio) y luego eliminar algunas de ellas (no puede ser el dinosaurio), en un movimiento que va desde las probabilidades hacia la certidumbre (debe ser la flor).

“La probabilidad es el lenguaje de la ciencia”, nos dice el físico estadístico Edwin Thompson Jaynes, uno de los grandes sistematizadores de, precisamente, la probabilidad bayesiana. Las experiencias más recientes demuestran que un bebé de unos pocos meses, incluso antes de pronunciar sus primeras palabras, ya domina este lenguaje no verbal, esta capacidad de manipular las probabilidades y de combinarlas para extraer de ellas conclusiones refinadas. No solo no tiene necesidad de aprenderlo, sino que, por el contrario, este conocimiento sirve de base para los demás aprendizajes de la vida: el bebé se apoya sobre esta habilidad porque la lógica y el sentido de las probabilidades le permiten derivar las consecuencias lógicas de las observaciones que realiza. Experimenta a cada instante, y así su cerebro de científico en ciernes acumula las conclusiones de sus investigaciones.

El conocimiento de los animales y las personas

Así como los bebés tienen una idea precisa del comportamiento de los objetos, también saben que existe otra categoría de entidades que se comportan de un modo muy diferente: los animales y las personas. Durante el primer año de vida, comprenden que estas cosas animadas tienen su idiosincrasia: son autónomas y se impulsan con su propio movimiento. Por eso, no necesitan esperar que otro objeto las toque para moverse, como sucede, por ejemplo, con una bola de billar. Su desplazamiento está motivado desde el interior, no es causado desde el exterior.

Por eso, los bebés no se sorprenden al ver a los animales desplazarse por sí mismos. De hecho, para ellos, cualquier objeto que se mueve por su cuenta, incluso si tiene la forma de un triángulo o de un cuadrado, queda etiquetado inmediatamente como “animal” y, a partir de entonces, todo cambia. Un niño pequeño sabe que los seres vivos no tienen que moverse necesariamente de acuerdo con las leyes de la física, sino que sus desplazamientos están regidos por sus intenciones y sus creencias.

Tomemos un ejemplo: si le mostramos a un bebé una esfera que se desliza en línea recta, salta por encima de un muro y luego continúa su movimiento hacia la derecha (figura 8), el niño de a poco se aburrirá de ella. ¿Simplemente se habrá habituado a esta extraña trayectoria? No, en realidad comprenderá mucho más: deducirá que se trata de un objeto animado dotado de una intención definida: ¡quiere moverse a la derecha! Más aún, podrá decir que el objeto está especialmente motivado, ya que no duda en saltar un muro para llegar allí.



Ahora, quitemos el muro. En este caso, el bebé no se sorprende de ver que la esfera cambia su movimiento y va hacia el objetivo en una trayectoria recta, sin saltar; esta es simplemente la mejor manera de alcanzar la meta. Por el contrario, el bebé abre los ojos de par en par si esta continúa saltando en el aire, sin motivo específico, porque ya no hay un muro. Cuando el obstáculo está presente, la trayectoria misma de la primera escena genera sorpresa en el bebé, porque no comprende qué extraña intención podría tener la esfera (Gergely, Bekkering y Kiraly, 2002; Gergely y Csibra, 2003; Warneken y Tomasello, 2006). Otros experimentos demuestran que el niño infiere las intenciones y las preferencias de las personas. En este caso específico, comprende que cuanto más alto sea el muro, mayor deberá ser la motivación de la persona para saltarlo. A partir de sus observaciones, los bebés pueden deducir no solamente los objetivos y las intenciones de las personas que los

rodean, sino también sus creencias, sus capacidades y sus preferencias (Liu y otros, 2017).

Esta noción que los niños tienen sobre los seres vivos no se detiene allí. Alrededor de los 10 meses, los bebés comienzan a atribuirles cualidades a las personas: si ven a alguien arrojar a un niño por una ventana, por ejemplo, deducen que tiene malas intenciones y se alejan de él. Desde luego, prefieren a una segunda persona que ayuda al niño a levantarse (Buon y otros, 2014). Mucho antes de poder pronunciar las palabras “malo” y “bueno”, ya formulan estas intuiciones en su lenguaje del pensamiento. Este juicio es refinado: los bebés de 9 meses pueden distinguir entre alguien que hace el mal de forma intencional y alguien que lo hace por accidente; o entre alguien que se niega a ayudar a una persona y alguien que quiere ayudar pero no tiene la posibilidad de hacerlo (Behne y otros, 2005). Como veremos más adelante, esta competencia social desempeña un papel fundamental en el aprendizaje. En efecto, un niño de 1 año comprende si alguien tiene la intención de enseñarle algo. Puede diferenciar entre una acción banal y una acción realizada con el objetivo de enseñar algo nuevo. Posee, como afirma el psicólogo húngaro György Gergely, un “sentido innato de la pedagogía”.

La percepción de rostros

Una de las manifestaciones más precoces de esta competencia social es la percepción de rostros. En el adulto, el indicio más nimio es suficiente para desencadenar el reconocimiento: una caricatura, un emoticón, una máscara... ¡Algunas personas hasta detectan la cara de Cristo en manchas de nieve o en los dibujos del viento en la arena de la playa! Sorprendentemente, esta hipersensibilidad está presente desde el nacimiento: un bebé de algunas horas de vida reacciona más rápido frente a un emoticón que frente a una imagen similar puesta al revés (incluso si el investigador se asegura de que el recién nacido jamás haya tenido la posibilidad de ver un rostro). Es más, ¡el reconocimiento de rostros parece comenzar *in utero*! Un equipo de investigación logró presentar un patrón de luces a fetos a través de las paredes del útero (Reid y otros, 2017). Para nuestra sorpresa, los resultados

demonstraron que tres puntos dispuestos con la forma de un rostro () atraen más al feto que tres puntos organizados en composición pirámidal ().

Muchos investigadores piensan que esta atracción casi magnética hacia los rostros desempeña un papel esencial en el desarrollo temprano del apego, sobre todo si se tiene en cuenta que uno de los síntomas precoces del autismo consiste en no mirar a las personas a los ojos. Un sesgo innato nos forzaría, llevando nuestra mirada hacia los rostros, a reconocerlos desde los pocos meses de vida, y, de hecho, desde esta edad, una región de la corteza visual del hemisferio derecho comienza a responder a los rostros más que a otras imágenes, por ejemplo, a paisajes (Adibpour, Dubois y Dehaene-Lambertz, 2018; Deen y otros, 2017; Livingstone y otros, 2017). La especialización para los rostros es uno de los mejores ejemplos de la colaboración armoniosa entre lo innato y lo adquirido. En este ámbito, el bebé presenta competencias estrictamente innatas (la ya mencionada atracción magnética ante los rostros, pero esta vez en representaciones gráficas que son homólogas), pero también un instinto extraordinario para incorporar las especificidades de la percepción correspondiente. Precisamente la combinación de estos dos factores permite, en poco menos de un año, superar la reacción absorta ante la mera presencia de dos ojos y una boca, y empezar a establecer una clara preferencia por los rostros humanos frente a los de otros primates como los gorilas o los chimpancés (Morton y Johnson, 1991).

El don de lenguas

La competencia social de los niños pequeños queda manifiesta no solo en la modalidad visual, sino también en la audición: el habla llega a ellos con tanta facilidad como la percepción de los rostros. Como señaló Steven Pinker, autor del exitoso libro *El instinto del lenguaje*, los humanos estamos innatamente preparados para el lenguaje, de modo tal que resulta tan inconcebible quitarnos la capacidad de aprenderlo y usarlo como pretender que no retiremos la mano ante una superficie caliente. Que esta afirmación no nos llame a engaño: por supuesto, los bebés no nacen con un despliegue completo de léxico y gramática, pero tienen una notable capacidad de

adquirirlos en un tiempo récord. Lo que está “cableado” en ellos no es tanto el lenguaje sino la habilidad para adquirirlo.

En la actualidad, contamos con mucha evidencia de esta observación. Ya desde el nacimiento, un bebé prefiere escuchar su lengua materna que una lengua extranjera (Mehler y otros, 1988), un hallazgo verdaderamente extraordinario, que implica que existe un aprendizaje del lenguaje *in utero*. En efecto, desde el tercer trimestre de embarazo, el feto ya es capaz de oír. La melodía del lenguaje, filtrada por la pared uterina, llega a los oídos del bebé, quien comienza a retenerla. “Tan pronto como llegó la voz de tu saludo a mis oídos, el niño saltó de alegría en mi vientre”, le dice Isabel, embarazada, a María, que la visita (Lucas 1, 44). El evangelista no se equivocó: en los últimos meses de embarazo, el cerebro en formación ya logra reconocer determinados patrones auditivos y melodías, incluso si todavía no puede ser consciente de eso.[\[14\]](#)

Por supuesto, es más fácil estudiar esta capacidad innata en bebés prematuros que en fetos. Una vez fuera del vientre, es posible equipar sus pequeñas cabezas con pequeños receptores de electroencefalografía y sensores de flujo sanguíneo cerebral, y observar qué sucede en sus cerebros. Con este método, mi esposa Ghislaine descubrió que hasta bebés prematuros, de 6 meses y medio, es decir, nacidos dos meses y medio antes de término, responden a la lengua hablada: sus cerebros ya reaccionan ante los cambios de las sílabas y de las voces (Mahmoudzadeh y otros, 2013).

Durante mucho tiempo pensamos que la adquisición del lenguaje no se desencadenaba hasta el primer o segundo año de vida. ¿Por qué? Porque, como su nombre lo indica, el infante recién nacido, *in-fans*, no habla y, por lo tanto, mantiene ocultos sus talentos. Sin embargo, en el plano de la comprensión del lenguaje, su cerebro es un verdadero genio de las estadísticas. Para demostrarlo, los científicos tuvieron que desplegar una panoplia de métodos originales, incluida la medición de la preferencia de los niños por los estímulos lingüísticos o no lingüísticos, sus respuestas al cambio, el registro de su actividad cerebral... Los resultados de esos estudios convergen y revelan lo mucho que los “infantes” ya saben acerca del lenguaje. Desde el nacimiento, el bebé puede reconocer la diferencia entre la mayor parte de las vocales y de las consonantes de todas las lenguas del mundo. Las percibe ya como categorías. Tomemos, por ejemplo, las sílabas /ba/, /da/ y /ga/: si bien el sonido de cada una de ellas varía continuamente en cuanto a sus propiedades acústicas, el cerebro del bebé las trata como

categorías distintas, separadas por límites precisos, al igual que lo hace un adulto.

Por influencia del ambiente lingüístico, estas competencias precoces, innatas, se especializarán a lo largo del primer año de vida. Muy rápidamente, el bebé repara en que determinados sonidos no se utilizan en su lengua materna: para los hablantes de español no existe la necesidad de diferenciar los sonidos [u] y [y], que son relevantes para el francés; en Japón, no hay necesidad de distinguir una “r” y una “l”. En unos pocos meses (seis para las vocales, doce para las consonantes), el cerebro del bebé hace una clasificación a partir de sus hipótesis iniciales y solo conserva los fonemas útiles para las lenguas que lo rodean.

Pero esto no es todo: muy pronto, el bebé comienza a aprender sus primeras palabras. ¿Cómo las identifica? En un principio, detecta las variaciones de prosodia, que hacen que la voz del locutor suba, baje, se detenga... y así marca las fronteras entre las palabras y las frases. Otro mecanismo consiste en identificar qué grupos de sonidos siguen a otros. En este campo, otra vez, el bebé se comporta como un estadístico en ciernes. Se da cuenta, por ejemplo, de que tras la sílaba /ba/ suele ir /bero/. Un rápido cálculo de probabilidades lo lleva a pensar que eso no puede ser resultado del azar: estas sílabas se siguen unas a otras con una probabilidad demasiado alta, deben formar una palabra, “babero”. Así es como esta palabra se suma al vocabulario del niño, y podrá, luego, ponerse en relación con un objeto o concepto precisos (Hay y otros, 2011; Saffran, Aslin y Newport, 1996). Desde los 6 meses, el niño ya conoce las palabras que se repiten con mayor frecuencia en su ambiente, como “bebé”, “papá”, “mamá”, “babero”, “pie”, “tomar”, “pañal”, etc. Estas palabras se graban en su memoria en grado tal que, durante la edad adulta, conservarán un estatus especial, y continuarán siendo procesadas con mayor eficacia que otras palabras comparables en significado, sonoridad y frecuencia, pero adquiridas en época más tardía.

Muy rápidamente, el mismo análisis estadístico le permite al bebé detectar que algunas palabras se repiten con más frecuencia que otras: son las pequeñas palabras gramaticales, como los artículos (el, la, los, las...) y los pronombres (yo, él, ella, nosotros...). Desde el final de su primer año, un bebé ya conoce un buen número de estas palabras, y se vale de ellas para detectar las otras. Si, por ejemplo, escucha a su mamá o a su papá decir: “Yo preparo unos postres”, sabe desligar las pequeñas palabras conocidas, “yo” y “unos”, y, por eliminación, descubrir que “preparo” y “postres” también son

palabras. Ya comprende que luego de “la” viene (a menudo) un sustantivo, y luego de “ella”, un verbo, de modo que alrededor de los 20 meses se verá desconcertado si le decimos frases falsas como “ella babero” o “los cucharos” (Bernal y otros, 2010).

Por supuesto, el análisis probabilístico no es del todo infalible. Así es como suele ocurrir, en cierta etapa, que los niños escuchan “un avión” e infieren la existencia de la palabra “navión”. A pesar de estas fallas, y por sucesivas aproximaciones, el niño logra superar el desempeño de cualquier algoritmo actual de inteligencia artificial. Para el momento en que celebra su primer cumpleaños, ya sentó las bases de las principales reglas de su lengua materna, y lo hizo en todos los niveles, desde los sonidos elementales (los fonemas) hasta la melodía (la prosodia), el vocabulario (el léxico mental) y las reglas de la gramática (la sintaxis).

Ninguna otra especie de primates presenta esta capacidad. El experimento se realizó muchas veces: varias parejas de científicos adoptaron a un bebé de chimpancé, lo trataron como a un miembro de su familia, hablándole en inglés o en lengua de señas, o incluso con símbolos visuales... para darse cuenta, algunos años más tarde, de que ninguno de estos animales dominaba un lenguaje digno de ese nombre: a lo sumo, conocían algunas decenas o centenas de palabras (Penn, Holyoak y Povinelli, 2008; Terrace y otros, 1979; Yang, 2013). Entonces, sin lugar a dudas el lingüista Noam Chomsky tiene razón al postular que nuestra especie posee lo que él denomina “dispositivo de adquisición del lenguaje”, un sistema especializado que se activa de forma automática en los primeros años de vida. Como ya decía Darwin en *El origen del hombre*, “ciertamente, el lenguaje no es un instinto en el sentido específico de la palabra, porque cada lenguaje debe ser aprendido”, pero es “una tendencia instintiva a adquirir un arte”. Lo que es innato en nosotros es el instinto de aprender una lengua, sea cual sea; se trata de un instinto tan irreprimible que el lenguaje aparece de forma espontánea, en el transcurso de algunas generaciones, en los humanos que carecen de él. En las comunidades de sordos, la lengua de señas aparece de forma espontánea y, desde la segunda generación, ya posee características lingüísticas universales (Senghas, Kita y Özyürek, 2004).

[13] Pueden encontrar una reseña detallada de estos hallazgos en la edición revisada de mi libro *El cerebro matemático* (Dehaene, 2010).

[14] Sobre este tema, invito a leer mi libro *La conciencia en el cerebro* (Dehaene, 2014).

4. El nacimiento de un cerebro

El niño nace con un cerebro inconcluso y no con un cerebro desocupado, como afirmaba el postulado de la pedagogía antigua.

Gaston Bachelard, *La filosofía del no* (1940)

La existencia de las capacidades precoces del bebé refuta la hipótesis de que su cerebro no es más que una pizarra virgen, una *tabula rasa*, una página en blanco que solo es capaz de absorber la huella de su ambiente. Esto desemboca en una sencilla predicción: si analizamos el cerebro del recién nacido, podremos observar, desde el nacimiento o incluso antes, estructuras cerebrales bien organizadas, que corresponden a cada uno de los grandes campos del saber.

Hasta hace una veintena de años, el cerebro del recién nacido era aún *terra incognita*. Las imágenes cerebrales acababan de inventarse, todavía no se habían aplicado al cerebro en desarrollo, y la perspectiva teórica predominante era la del empirismo, la idea de que el cerebro nace virgen de conocimiento y se somete a la huella de su ambiente. Solo con la llegada de refinados métodos de imágenes por resonancia magnética pudimos por fin visualizar la organización precoz del cerebro humano y descubrir que, conforme a lo que esperábamos, desde el nacimiento, casi todos los circuitos del cerebro adulto ya están presentes en el bebé.

Desde el comienzo, el cerebro del bebé está bien estructurado

Mi esposa Ghislaine Dehaene-Lambertz y yo, junto con nuestra colega neuróloga Lucie Hertz-Pannier, estuvimos entre los primeros en emplear la resonancia magnética funcional (fMRI, por sus iniciales en inglés) en bebés

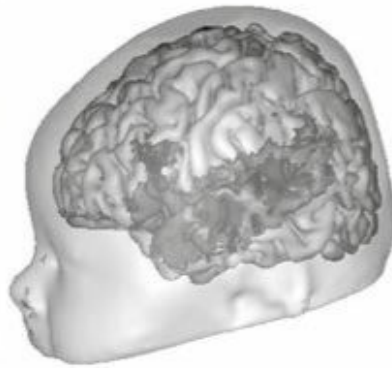
de 2 meses (Dehaene-Lambertz y otros, 2006; Dehaene-Lambertz, Dehaene y Hertz-Pannier, 2002). Por supuesto, nos basamos en la experiencia de los neuropediatras. Quince años antes, los médicos consideraban la MRI como una prueba anodina y sin dudar la utilizaban en pacientes de cualquier edad, incluso en bebés prematuros; sin embargo, esos estudios eran puramente anatómicos, solo apuntaban a detectar lesiones precoces. Hasta ese momento, nadie había utilizado la fMRI en bebés sin patologías, para ver si desde esta edad los circuitos del cerebro ya podían activarse de manera selectiva. Para lograrlo, tuvimos que resolver una serie de dificultades: proteger al niño del ruido de la máquina con un casco insonorizado, envolverlo dentro de una cuna con la forma del resonador para mantenerlo inmóvil, tranquilizarlo para que se acostumbrase a esta situación inusual, supervisar permanentemente al pie de la máquina para asegurarnos de que se sintiese bien...

Valió la pena, porque los resultados fueron espectaculares. Habíamos decidido estudiar el lenguaje, porque sabíamos que los bebés lo aprenden muy rápido, desde su primer año de vida. Y, efectivamente, observamos que a los 2 meses, cuando un bebé oye frases de su lengua materna, ya activa las mismas regiones del cerebro que un adulto (figura 9).

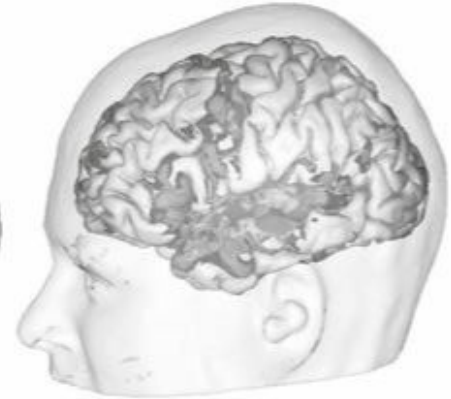
La primera región en activarse es la corteza auditiva primaria, el punto de entrada de toda la información auditiva a la corteza cerebral. Esta se enciende no bien comienza la frase. Si bien esto puede parecerles evidente, no era una certeza que ocurriera en las niñas y los niños de edad muy temprana. Hasta hace muy poco tiempo, algunos investigadores suponían que las áreas sensoriales del cerebro del bebé estaban tan desorganizadas que sus sentidos tendían a fusionarse. Según ese enfoque, durante varias semanas, el cerebro del bebé mezclaba la audición, la visión y el tacto, y le hacía falta tiempo para aprender a separar estas modalidades sensoriales (Elman y otros, 1996; Quartz y Sejnowski, 1997). Actualmente sabemos que esto es falso: desde el nacimiento, la audición activa las áreas auditivas, la visión activa las áreas visuales y el tacto las áreas táctiles, sin que haya necesidad de aprenderlo. Esta organización en territorios distintos para cada una de las modalidades sensoriales nos viene dada por los genes, sucede en todos los mamíferos y su origen se pierde en la noche de nuestra evolución (figura 10; Krubitzer, 2007).



Bebé de 2 meses



Adulto



Respuesta del cerebro del bebé a una sola frase

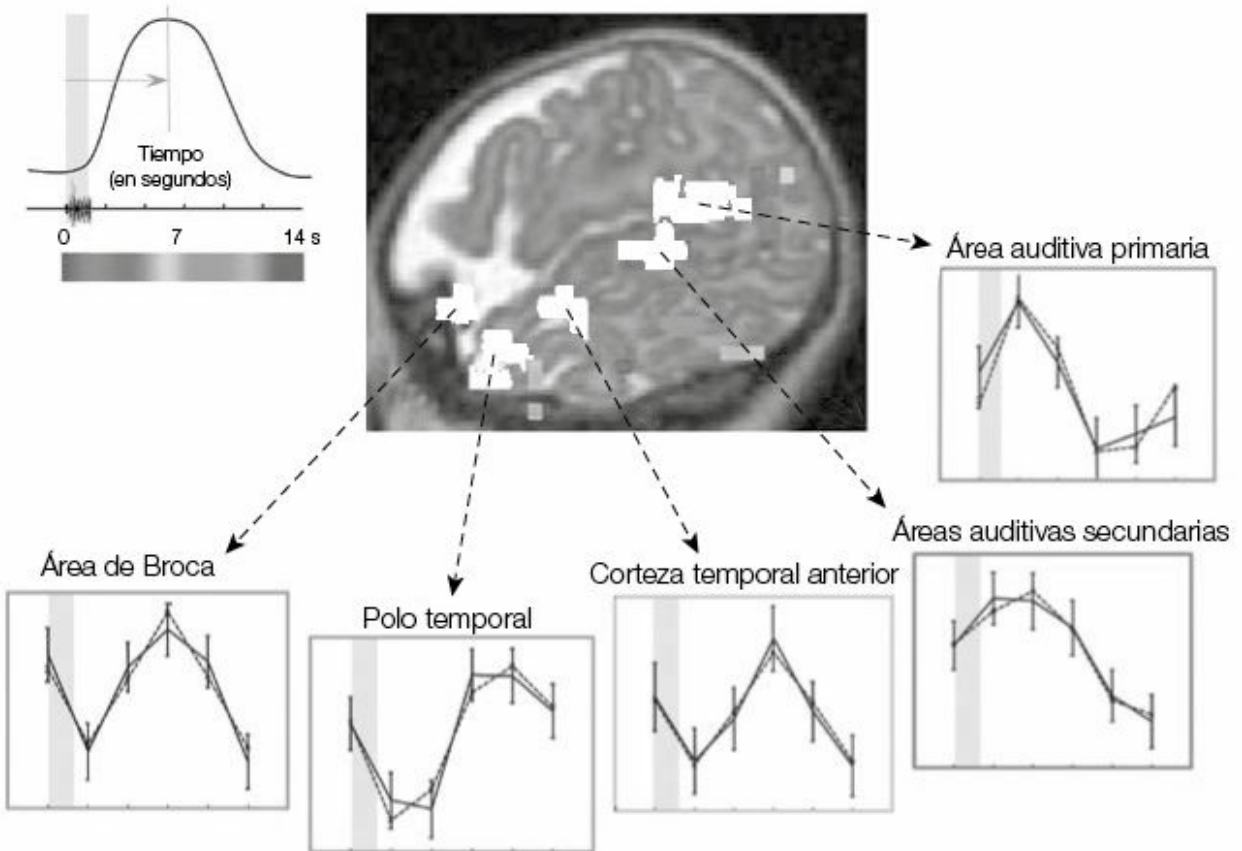


Figura 9. Desde el nacimiento, el cerebro humano aplica operaciones específicas en el campo del lenguaje. Cuando un bebé es sometido a una resonancia magnética (MRI), y se lo hace oír frases de su lengua materna, se constata que ya se activa una red muy específica de regiones cerebrales. La actividad comienza en el área auditiva primaria y poco a poco se extiende a las áreas temporales y frontales: exactamente como sucede en los hablantes adultos, en el mismo orden. Estos datos refutan la idea de un cerebro inicialmente desorganizado, simple *tabula rasa* sometida a la huella del entorno.

Pero volvamos al bebé que oye una frase. La actividad de su cerebro de inmediato se propaga tanto más allá de la región auditiva primaria. Una fracción de segundo después, otras áreas se encienden en un orden fijo: las regiones auditivas secundarias, vecinas de la corteza sensorial primaria; luego una sucesión de regiones que se alejan gradualmente hacia la parte delantera del lóbulo temporal, y, por último, el área de Broca, en la base del lóbulo frontal izquierdo, al mismo tiempo que el extremo del lóbulo temporal. Esta distribución sofisticada, lateralizada al hemisferio izquierdo, es llamativamente similar a la de un adulto. Desde los 2 meses, el bebé activa la misma jerarquía de áreas cerebrales fonológicas, léxicas, sintácticas y semánticas que un adulto. En este caso también, cuanto más se avanza en la jerarquía de la corteza, más lentas son las respuestas y en mayor medida integran información de niveles cada vez más elevados (figura 9; Lerner y otros, 2011; Pallier, Devauchelle y Dehaene, 2011).

Por supuesto, un bebé pequeño todavía no comprende las frases que le decimos, todavía no descubrió las palabras y las reglas de la gramática de su lengua materna. Sin embargo, en su cerebro, la información lingüística ya se procesa en circuitos altamente especializados, los mismos que en el adulto. No cabe duda: si un bebé aprende tan rápido a comprender y a hablar, mientras que los demás primates son incapaces de hacerlo, es porque su hemisferio izquierdo ya alberga una estructura de regiones especializadas en la detección de regularidades estadísticas en todos los niveles: sonido, palabra, frase y texto.

Las autopistas del lenguaje

Si efectivamente todas estas áreas cerebrales se activan en cadena, debe significar que están interconectadas. En el adulto, comenzamos a comprender

qué haces de axones conectan las regiones del lenguaje unas con otras. En especial, hay un gran cable, formado por millones de fibras nerviosas, que llamamos “fascículo arqueado”; vincula las áreas temporales y parietales con las áreas frontales y, sobre todo, con la famosa área de Broca. De por sí, nos sirve como indicador: en el 96% de los humanos es mucho más grueso en el hemisferio izquierdo, que se ocupa del lenguaje. Esta asimetría solo existe en la especie humana, no en otros simios, ni siquiera en el chimpancé.

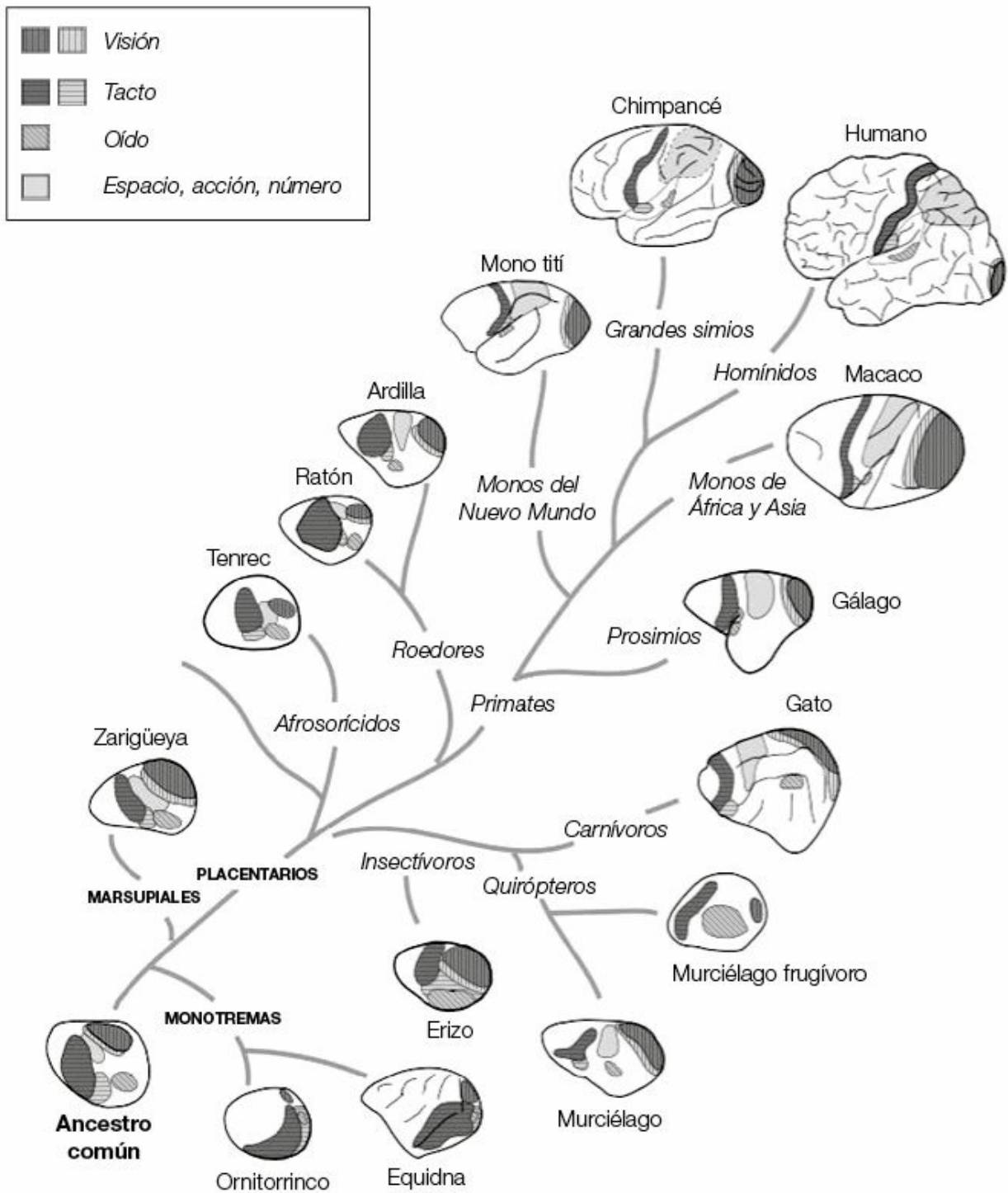
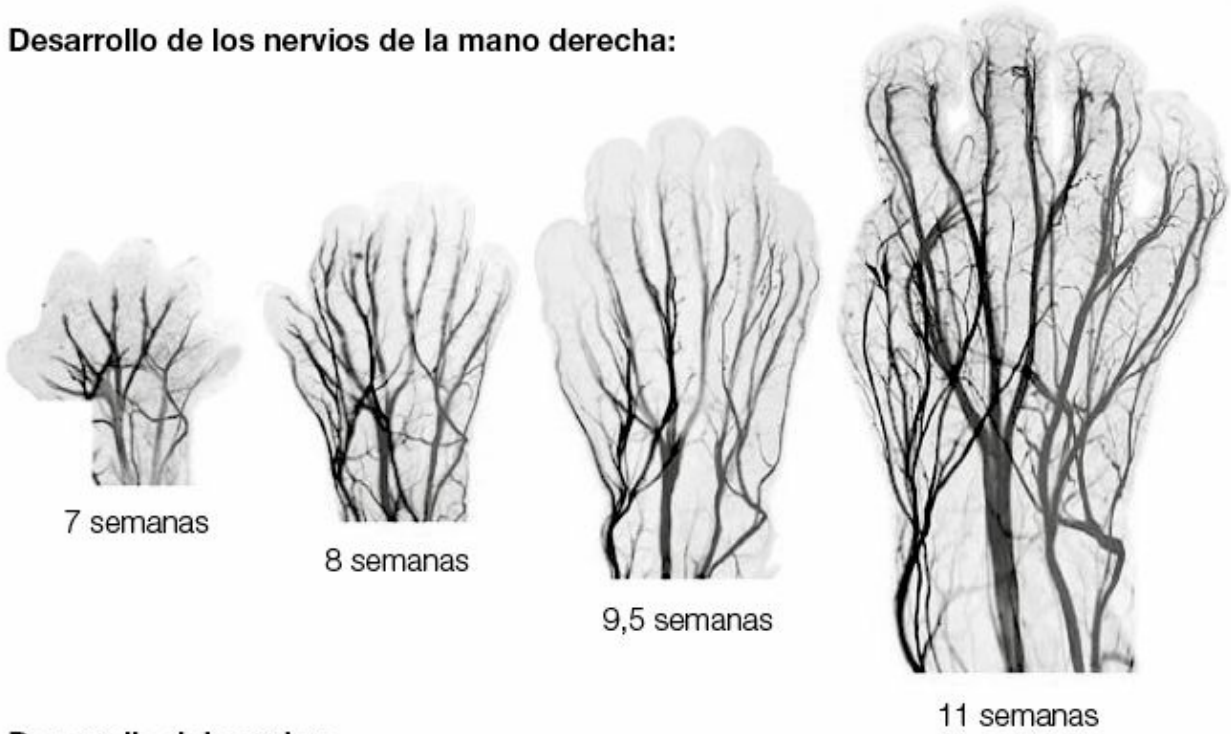


Figura 10. El cerebro humano hereda su arquitectura de la evolución de los mamíferos. Numerosas regiones especializadas (en este caso, las principales áreas sensoriales) comparten su mapa de base con otras especies. Se van posicionando durante la gestación, por acción de numerosos genes, y ya están activas en el tercer trimestre del embarazo. El cerebro de nuestra especie se caracteriza por tener áreas sensoriales proporcionalmente más pequeñas, mientras

que se observa una enorme expansión de las regiones cognitivas de la corteza parietal posterior, temporal y, sobre todo, prefrontal. Estas regiones albergan el lenguaje del pensamiento y nos permiten aumentar los conocimientos durante toda la vida.

Desarrollo de los nervios de la mano derecha:



Desarrollo del cerebro:

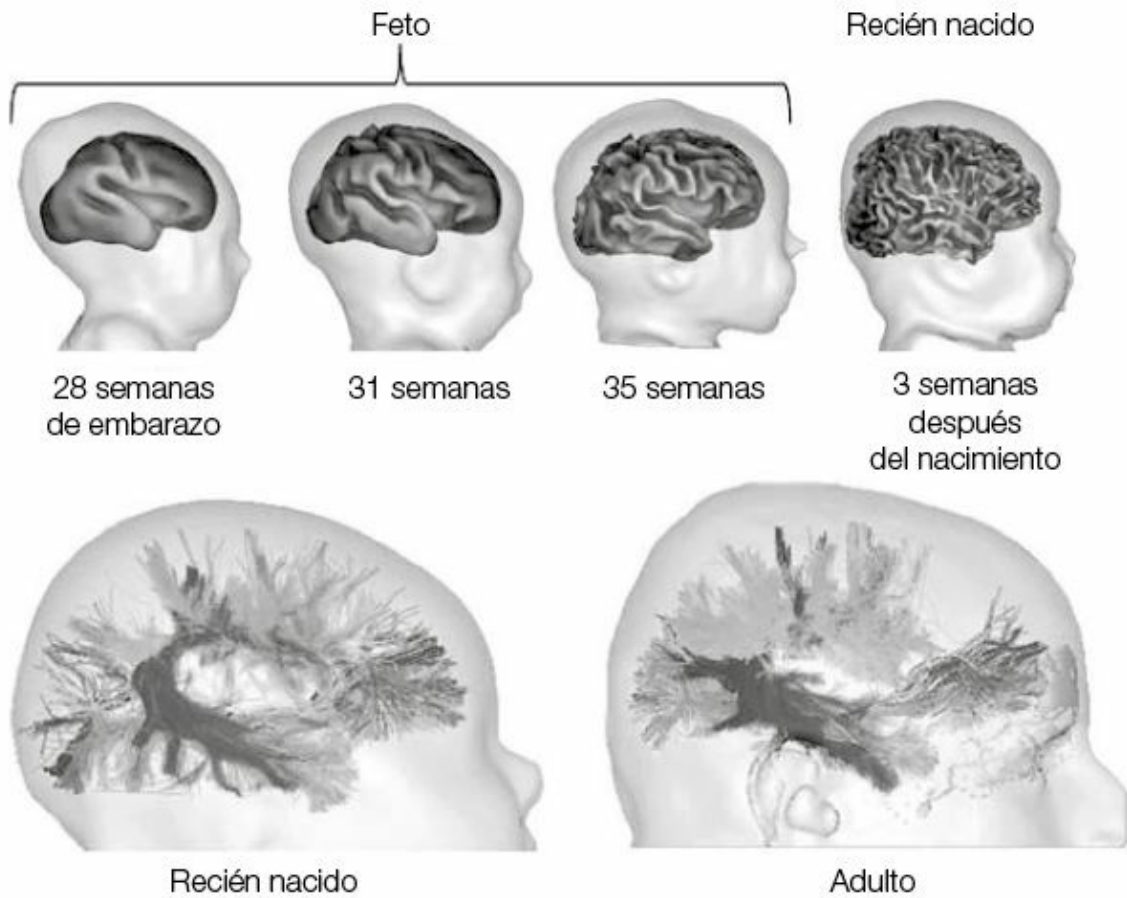


Figura 11. Durante las primeras semanas del embarazo, el organismo se autoorganiza sobre una base genética. No hace falta aprendizaje alguno para que los cinco dedos se formen y reciban una inervación específica. Del mismo modo, la arquitectura elemental del cerebro se pone en funcionamiento sin aprendizaje. Desde el nacimiento, la corteza ya está organizada, plegada y conectada de una forma específica que es igual para todos los seres humanos. Los detalles, sin embargo, están librados a la variación en función del entorno. Desde el tercer trimestre de gestación, el cerebro del feto comienza a adaptarse a la información que recibe del mundo exterior.

Una vez más, esta característica anatómica no es resultado del aprendizaje: está presente desde el vientre materno. De hecho, cuando se examinan las conexiones del cerebro de un recién nacido, se descubre que desde el nacimiento no solo está en su lugar el fascículo arqueado, sino también todos los grandes haces de conexiones de la corteza y de las regiones subcorticales (figura 11; Dehaene-Lambertz y Spelke, 2015; Dubois y otros, 2015).

Estas “autopistas del cerebro” se construyen durante el tercer trimestre del embarazo. Cada neurona excitatoria en formación envía su axón a explorar las regiones que la rodean, como un Marco Polo del cerebro, a veces hasta una distancia de varios centímetros. La exploración es guiada y canalizada por mensajes químicos, moléculas cuya concentración varía de una región a otra y que entonces reaccionan como paneles de señalización. La cabeza del axón olfatea, literalmente, este ambiente químico, de origen genético, y deduce qué dirección debe seguir. Con esto, y sin intervención del mundo exterior, se pone en funcionamiento una red de conexiones nerviosas cruzadas propia de la especie humana. Como veremos en un instante, esta red será refinada más tarde por el aprendizaje, pero el andamiaje inicial es innato, se construye *in utero*.

¿Deberíamos sorprendernos por esto? Hace apenas un par de décadas numerosos investigadores consideraban extremadamente improbable que el cerebro fuera otra cosa que una masa desorganizada de conexiones aleatorias (Quartz y Sejnowski, 1997). No podían imaginar que nuestro ADN, que no contiene más que una cantidad acotada de genes, pudiese albergar el plan detallado de circuitos tan especializados como los de la visión, el lenguaje o la motricidad. Pero ese era un razonamiento erróneo: el genoma humano abarca todos los detalles del plan corporal, sabe fabricar un corazón y sus cuatro ventrículos, dos ojos, veinticuatro vértebras, un oído interno y sus tres canales perfectamente perpendiculares; por añadidura, puede formar grupos de cinco dedos, con sus falanges y hasta su inervación; todo con precisión

extrema... Entonces, ¿por qué no podría crear un cerebro dotado de múltiples subregiones especializadas?

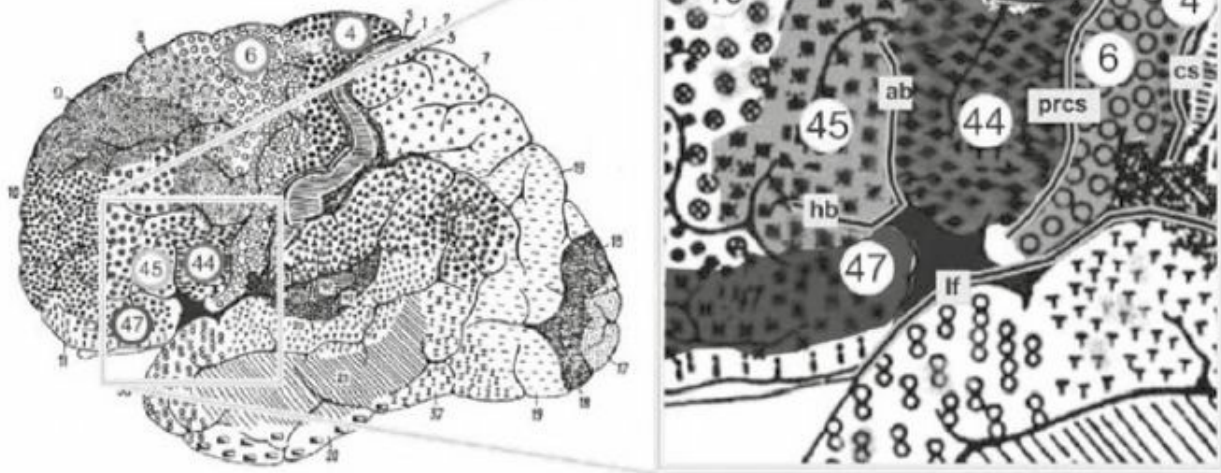
Las proezas de las imágenes cerebrales contemporáneas revelan que ya desde los dos primeros meses de embarazo, cuando los dedos de la mano no son más que yemas, están invadidos por tres nervios, el radial, el mediano y el cubital, que apuntan hacia sendos blancos específicos (figura 11; Belle y otros, 2017). Esa misma mecánica de alta precisión existe en el cerebro: así como la yema que será la mano se escinde en cinco dedos, la corteza se subdivide en varias decenas de regiones de enorme especialización, separadas por fronteras muy definidas (figura 12; Amunts y otros, 2010; Amunts y Zilles, 2015; Brodmann, 1909). Desde los primeros meses del embarazo, numerosos genes se expresan de manera selectiva en diferentes puntos de la corteza (Kwan y otros, 2012; Sun y otros, 2005). Cerca de las 28 semanas de gestación, el cerebro comienza a plegarse y se ven aparecer los principales surcos que caracterizan el cerebro humano. En un feto de 35 semanas, ya están bien formados los principales pliegues de la corteza y se ve asomar la asimetría característica de la región temporal, que alberga las áreas del lenguaje (Dubois y otros, 2009; Leroy y otros, 2015).

La autoorganización de la corteza

El pliegue de la corteza y el crecimiento de sus conexiones continúa durante todo el embarazo. Aparecen sucesivamente pliegues secundarios y terciarios, pliegues sobre pliegues; su epigénesis se vuelve cada vez más dependiente de la actividad del sistema nervioso. De acuerdo con la respuesta que el cerebro reciba de los órganos de los sentidos, determinados circuitos se estabilizarán, mientras que otros, ya inútiles, se degenerarán. Así, el pliegue de la corteza motora no es exactamente el mismo en zurdos y diestros, mientras que los zurdos contrariados, a quienes se forzó a escribir con la mano derecha, muestran una suerte de solución de compromiso: la forma de su corteza motora es típica de un zurdo, pero su tamaño presenta la asimetría típica de un diestro. Como concluyen Sun y sus colegas (2012), autores de un estudio comparativo, “la morfología de la corteza de un adulto refleja la acumulación de sesgos innatos y de experiencias precoces durante el desarrollo”.

Si los pliegues corticales emergen de manera espontánea en el cerebro del feto, es porque deben su formación a un proceso de autoorganización bioquímico que depende tanto de los genes como del entorno químico de las células, pero requiere muy poca información genética, y no necesita ningún tipo de aprendizaje (Lefèvre y Mangin, 2010).

Áreas corticales de Brodmann (1909)



Fronteras entre regiones definidas por cuatro moléculas diferentes

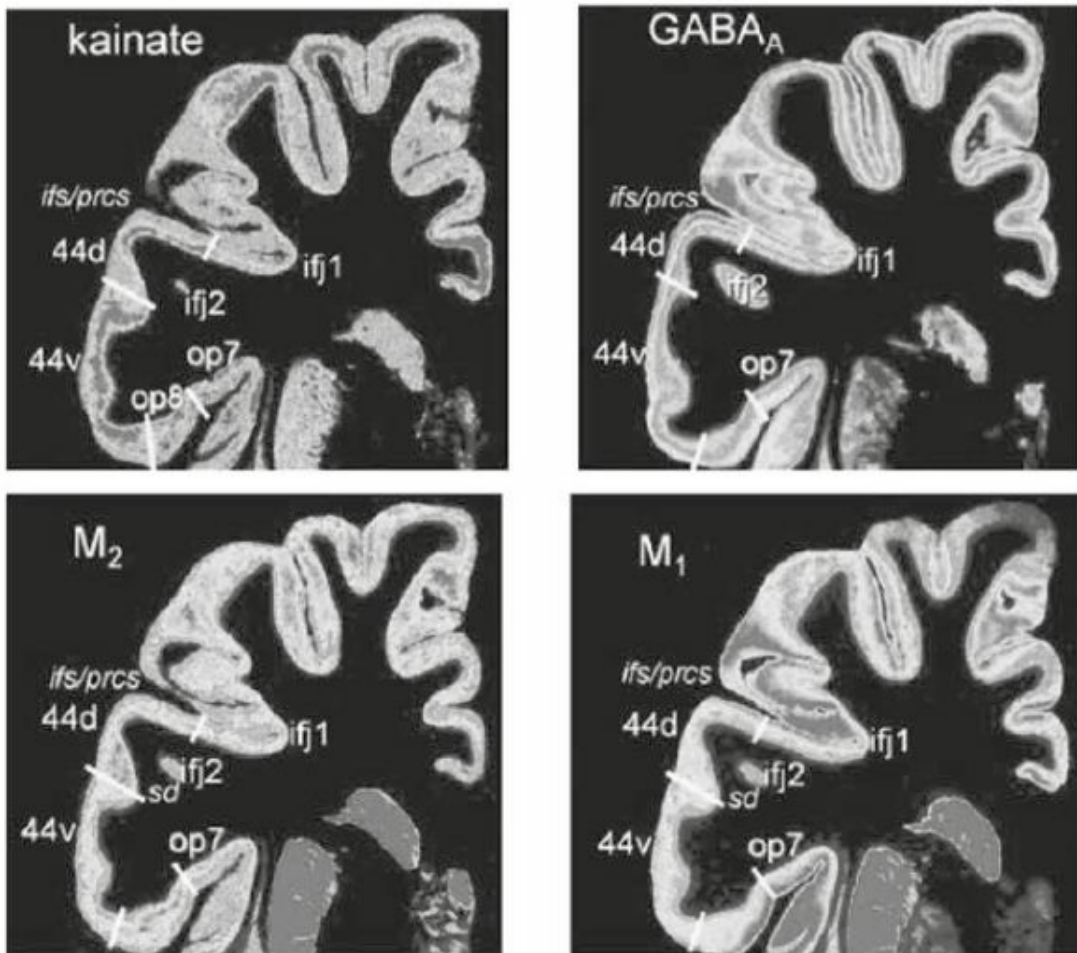


Figura 12. La corteza del cerebro humano está subdividida en regiones especializadas. Ya en

1909, el neurólogo alemán Korbinian Brodmann constató que el tamaño y la organización de las neuronas varían de una región a otra. De este modo, reconoce tres áreas (numeradas 44, 45 y 47) en la región de Broca, que está involucrada en el lenguaje. Estas distinciones actualmente se ven confirmadas y refinadas por las imágenes moleculares. La corteza está tapizada de distintas áreas, cuyas fronteras están marcadas por considerables variaciones de densidad de los receptores de neurotransmisores. Durante el embarazo, algunos genes se expresan selectivamente en las diferentes regiones de la corteza y contribuyen a subdividirla en módulos especializados.

Esa autoorganización no es tan paradójica como parece; la notamos dondequiera que miremos en la tierra. Podemos imaginar la corteza como una playa de arena donde se forman ondas y charcos o canales, a múltiples escalas, según las mareas que van y vienen. O como un desierto donde aparecen ondulaciones y dunas bajo la acción implacable del viento. De hecho, en todo tipo de sistemas biológicos o físicos surgen rayas, manchas y células hexagonales, también a diferentes escalas, desde las huellas dactilares hasta la piel de las cebras o las manchas del leopardo, las columnas de basalto en los volcanes o las nubes espaciadas con determinada regularidad en un cielo de verano.

El matemático Alan Turing fue el primero en explicar este fenómeno: es suficiente con que exista una amplificación local y una inhibición a distancia. Cuando sopla el viento en una playa, la arena comienza a acumularse y se inicia un proceso de autoamplificación: el pequeño bulto tiende a sumar cada vez más granos de arena; mientras tanto, delante de él, el viento forma un torbellino y se lleva parte de lo acumulado; en algunas horas, se ha formado una duna. Cuando hay agitación local e inhibición a distancia, se ve aparecer una región densa (la duna) rodeada por una región menos densa (el hueco), a la cual sigue otra duna, y así *ad infinitum*. Según las circunstancias específicas, de manera espontánea pueden surgir patrones de manchas, rayas o hexágonos.

Este mecanismo es omnipresente en la autoorganización del cerebro en desarrollo: la corteza está llena de columnas, de rayas o de fronteras que constituyen una cantidad de módulos especializados para procesar determinada información. En la corteza visual, por ejemplo, aparece una alternancia de franjas especializadas para procesar la información del ojo izquierdo y del ojo derecho: las “columnas de dominancia ocular”. Pero el mismo mecanismo se produce en un nivel más alto, no necesariamente en la superficie de la corteza, sino en un espacio más abstracto. Uno de los ejemplos más espectaculares, que le ha valió el Premio Nobel a sus

descubridores, es la existencia de “células de grilla” [*grid cells*], en las cuales las descargas forman hexágonos casi perfectos (figura 13).



Neurona de la corteza entorrinal

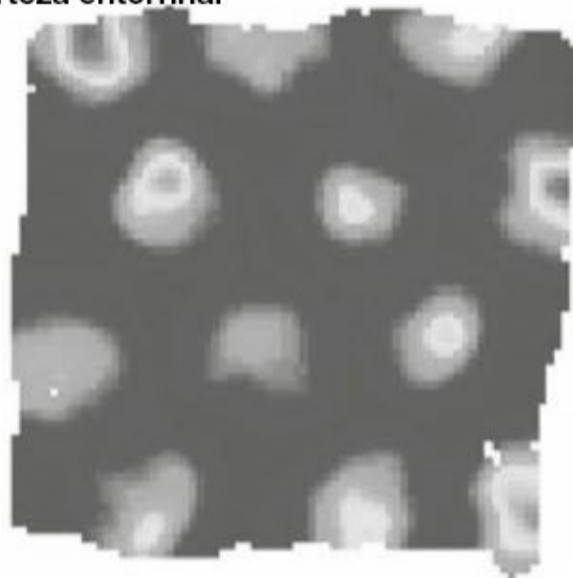
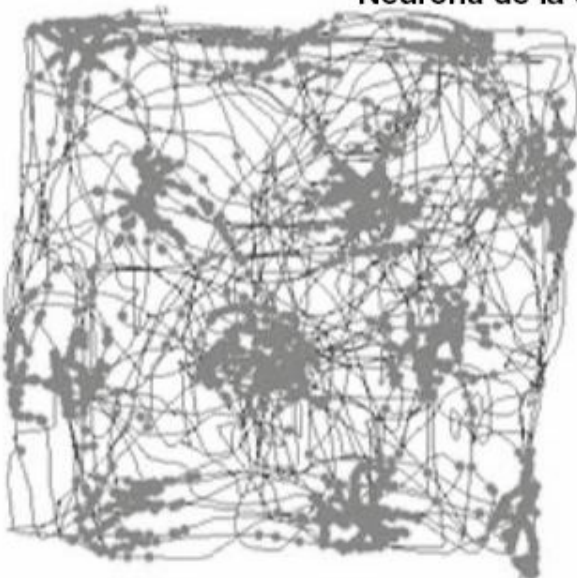


Figura 13. Cuando un sistema físico (puede ser lava o cera de abeja) se autoorganiza, no es extraño que se formen hexágonos. El sistema nervioso no es una excepción: en una región de la corteza entorrinal, verdadero GPS del cerebro, aparecen células de grilla que recubren el espacio con una suerte de adoquinado hexagonal. Cuando una rata explora una sala de 1 m², las descargas neuronales solo se producen cuando el animal está en la cima de una red de hexágonos. Este tipo de células de grilla aparece espontáneamente, un día después de que la rata comenzó a moverse: el sentido del espacio reposa sobre un GPS presente desde antes de nacer.

Las células de grilla son neuronas localizadas en una región específica del cerebro llamada “corteza entorrinal”. Sus notables propiedades geométricas fueron descubiertas al tomar registros del cerebro de un ratón mientras se desplazaba por una gran habitación (Banino y otros, 2018; Brun y otros, 2008; Fyhn y otros, 2004; Hafting y otros, 2005). Ya sabíamos que en una región vecina, el hipocampo, existen “células de lugar”, encargadas de “mapear” el espacio, es decir, que solo se activarían si el animal estaba en un sector muy específico del cuarto. El descubrimiento novedoso fue que las células de grilla no respondían solo a un lugar de la habitación, sino a una serie de posiciones diferentes. Además, estas localizaciones privilegiadas se distribuían regularmente formando una red de triángulos equiláteros que se agrupaban para formar hexágonos, ¡tal como las manchas sobre la piel de una jirafa o las columnas de basalto que aparecen en la roca volcánica!

Cuando el animal pasea sin rumbo fijo, cada célula de grilla le indica, con sus descargas, en qué lugar está con respecto a una red de triángulos que abarca el espacio completo. Con mucho acierto, el Comité Noruego del Nobel bautizó este sistema como “GPS cerebral”, ya que proporciona un sistema de coordenadas neuronales de una enorme fiabilidad que mapea el espacio externo y permite que el animal sepa en cada momento, incluso en la oscuridad, dónde está.

Pero ¿por qué las neuronas producen triángulos y hexágonos, en lugar de los rectángulos que estamos acostumbrados a ver en nuestros mapas? Desde los tiempos de Descartes, los matemáticos y los cartógrafos se basaron siempre sobre dos ejes perpendiculares (abscisas y ordenadas, x e y , longitud y latitud); entonces, ¿por qué el cerebro reacciona de otra manera? Sin duda, porque las neuronas se autoorganizan durante el desarrollo, y las formaciones en hexágono aparecen con toda la naturalidad del mundo en el momento en que un sistema físico se estabiliza (figura 13). De hecho, varios modelos matemáticos inspirados en las ideas de Turing dan cuenta de la aparición de

las células de grilla: esta organización en triángulos y en hexágonos es un atractor espontáneo de la dinámica de la corteza entorrinal (Kropff y Treves, 2008; Shipston-Sharman, Solanka y Nolan, 2016; Widloski y Fiete, 2014; Yoon y otros, 2013).

En la actualidad, esta teoría de la autoorganización de los mapas cerebrales comienza a ponerse a prueba. Una serie de experiencias extraordinarias demostró que, en efecto, el GPS cerebral surge de forma muy precoz en el transcurso del desarrollo. Dos grupos independientes de investigadores lograron la hazaña de implantar electrodos en ratones muy pequeños, recién nacidos, incluso antes de que comenzaran a caminar (Langston y otros, 2010; Wills y otros, 2010). También pudieron rastrear si en la corteza entorrinal ya estaban presentes las células de grilla y también las células de lugar (aquellas que responden a un solo lugar), e incluso las células de dirección de cabeza – células de un tercer tipo que todavía no mencioné; funcionan como la brújula de un barco: cada neurona se activa cuando el animal se orienta en determinada dirección, por ejemplo, noroeste o sur–. Este sistema es casi por completo innato: las células de dirección de cabeza están presentes desde el inicio de los registros, y las células de lugar y de grilla surgen apenas uno o dos días después de que el ratón comienza a desplazarse. Y esto no causa sorpresa: no bien empieza a moverse, es crucial que el pequeño animal sepa permanentemente dónde está para poder encontrar el hogar donde lo espera su madre. La evolución, entonces, encontró el modo de dotar a su cerebro de una brújula, un mapa y una memoria de los lugares.

¿Qué sucede con las humanas y los humanos? Por medios indirectos, sabemos que el cerebro del adulto también incluye un mapa neuronal con una simetría hexagonal, situado exactamente en el mismo lugar que en los ratones (Doeller, Barry y Burgess, 2010; Nau y otros, 2018). Sabemos también que en la infancia, incluso a edad muy temprana, ya poseemos un sentido del espacio. Hasta un niño ciego de corta edad logra orientarse dentro de una habitación: si lo llevamos del punto A al punto B y luego al punto C, sabe volver en línea recta desde C hacia A. Entonces, las criaturas poseen, tal como el ratón, un módulo mental para la navegación espacial (Landau, Gleitman y Spelke, 1981). Si todavía no logramos ver sin más este mapa en el cerebro del bebé, es porque no están superadas las extremas dificultades para obtener imágenes del cerebro en actividad a una edad tan temprana (intenten realizarle una resonancia –MRI– a un bebé mientras avanza en cuatro patas por el laboratorio...). Pero estamos casi seguros de que lo

lograremos cuando contemos con métodos de registro móvil de la actividad cerebral.

Podría multiplicar los ejemplos de otros módulos especializados en el cerebro del bebé. Por ejemplo, sabemos que, desde los pocos meses de vida (si no desde el nacimiento), una región de la corteza visual responde preferentemente a los rostros, más que a las imágenes de casas (Deen y otros, 2017; Livingstone y otros, 2017). Por cierto, esto es en parte resultado de un aprendizaje, pero estrechamente encauzado, guiado y acotado por la conectividad cerebral para aterrizar siempre en el mismo lugar, en un módulo específico de la corteza. También sabemos que la corteza parietal del bebé muy pequeño responde a la cantidad de objetos (Nieder y Dehaene, 2009), en la misma región que se activa cuando un adulto realiza cálculos o un mono memoriza una cantidad de objetos. El neurocientífico alemán Andreas Nieder consiguió la proeza de demostrar que en el mono esta región contiene neuronas sensibles al número de objetos: hay neuronas especializadas para un objeto, otras para dos objetos, tres objetos, y así sucesivamente... Y estas neuronas están presentes aunque el mono nunca haya recibido entrenamiento para realizar una tarea numérica. Se piensa, entonces, que surgen de modo innato. Mis colegas y yo propusimos un modelo matemático de gran precisión para la autoorganización de las neuronas detectoras del número, esta vez basado en la propagación de actividad de un tipo de onda a lo largo de la superficie de la corteza en desarrollo. Esta teoría puede explicar todos los detalles de las propiedades de las neuronas detectoras del número. Según el modelo, estas células terminan por formar una suerte de línea numérica, una cadena lineal en que los números 1, 2, 3, 4... ocupan posiciones sucesivas y que surge espontáneamente en el seno de una red de neuronas aleatorias (Hannagan y otros, 2017).

La idea de autoorganización difiere radicalmente del enfoque actual de la inteligencia artificial, basada en *big data*. A diferencia de las redes neuronales actuales, nuestro cerebro no parece necesitar millones de datos para organizarse. Por el contrario, sus núcleos de conocimiento se desarrollan espontáneamente, por estimulación interna. Son escasos los informáticos que, como Joshua Tenenbaum, profesor del MIT, contemplan la posibilidad de copiar esta autoorganización en una inteligencia artificial. Tenenbaum trabaja para simular un bebé virtual que llegaría al mundo con la capacidad de generar, por sí solo, millones de imágenes realistas. Estas le servirían de base para el aprendizaje, sin que exista necesidad de algún otro dato. Según este

enfoque radical, incluso antes del nacimiento, los circuitos neuronales se entrenarían con ayuda de una base de datos generada por ellos mismos (Lake y otros, 2016; de paso, señalemos que es otro ejemplo del *bootstrapping* del barón de Münchhausen). La mayor parte de su aprendizaje se produce de forma interna, sin interacción con el exterior, y luego solo resta realizar algunos ajustes en función de los datos suplementarios tomados del mundo real.

A partir de todas estas investigaciones se va bosquejando la perspectiva de la potencia de los genes y de la autoorganización en la puesta en marcha del cerebro humano. Desde el nacimiento, la corteza del bebé está plegada casi como la de un adulto, subdividida en áreas sensoriales y cognitivas especializadas que ya están interconectadas por haces de fibras nítidos y reproducibles. Consiste en un conjunto de módulos que, individualmente, proyectan sobre el mundo exterior un tipo específico de representación. La corteza entorrinal, con sus células de grilla, traza planos en dos dimensiones perfectos para codificar el espacio y navegar por él. Como veremos más adelante, regiones como la corteza parietal trazan líneas, excelentes para codificar las cantidades lineales como el número, el tamaño o el tiempo que pasa; mientras que la región de Broca proyecta árboles incrustados unos en otros, ideales para codificar las estructuras del lenguaje. Heredamos de nuestra evolución un conjunto de reglas fundamentales entre las cuales podremos, a continuación, seleccionar las que mejor representen las situaciones y los conceptos que deberemos aprender durante nuestra vida.

Los orígenes de la individualidad

Al afirmar la existencia de una naturaleza humana universal, un circuito cerebral innato establecido por los genes y la autoorganización, no pretendo negar las diferencias entre los individuos. Cuando los contemplamos exhaustivamente, todos y cada uno de nuestros cerebros tienen características únicas, y eso se verifica incluso desde el vientre materno. Por ejemplo, los pliegues corticales, al igual que las huellas digitales, se disponen antes del nacimiento y varían de formas distintivas. Del mismo modo, la fuerza y la densidad de las conexiones corticales de larga distancia, e incluso sus

trayectorias exactas, varían en gran medida y hacen que cada uno de nuestros “connectomas” sea único.

Sin embargo, es importante reconocer que esas variaciones se basan sobre un tema común. Como cualquier buen *standard* de jazz, el diseño del cerebro del *Homo sapiens* obedece a un esquema fijo, similar a la sucesión de acordes que los músicos de jazz memorizan cuando aprenden una canción. Solo en la parte superior de esta configuración humana universal los caprichos de los genomas y las peculiaridades de los embarazos agregan sus improvisaciones personales. Nuestra singularidad es real, pero no debe exagerarse: cada uno de nosotros no es más que una variación de la línea melódica del *Homo sapiens*. En cualquier individuo, blanco, negro, asiático o nativo sud-, centro- o norteamericano, en cualquier parte del planeta, la arquitectura del cerebro humano siempre es obvia. En ese sentido, la corteza de cualquier humano difiere de la de su pariente vivo más cercano, el chimpancé, tanto como cualquier improvisación sobre “My funny Valentine” se aparta, digamos, de una sobre el tema de “My romance”, canción compuesta por la misma dupla fundamental de los musicales y el jazz: Rodgers y Hart.

Debido a que todos compartimos la misma estructura cerebral básica que nos hace humanos, el mismo conocimiento central y también los mismos algoritmos de aprendizaje que nos permiten adquirir talentos adicionales, tenemos acceso a los mismos conceptos –ya sea en lectura, ciencias o matemáticas, sin importar que seamos ciegos, sordos o mudos–, dado que está presente el mismo potencial humano. Como ya en el siglo XIII observó el filósofo británico Roger Bacon, “el conocimiento de las cosas matemáticas es casi innato en nosotros. Esta es la más fácil de las ciencias, lo cual [resulta] obvio en el hecho de que ningún cerebro la rechaza; en el hecho de que los legos y los analfabetos saben contar y calcular”. Es evidente que lo mismo podría decirse del lenguaje: casi en ningún niño está ausente el impulso poderoso e innato para adquirir el lenguaje de su entorno, mientras que, como observamos antes, ningún chimpancé, incluso si es adoptado por una familia humana desde su nacimiento, logra murmurar unas pocas palabras ni componer algunos signos.

En resumen, las diferencias individuales son reales, pero casi siempre son de grado más que de tipo. Sin embargo, en los extremos de la distribución, las variaciones neurobiológicas terminan marcando una verdadera diferencia cognitiva. Cada vez más, los hallazgos de los nuevos estudios nos hacen pensar que los niños con trastornos del desarrollo están en un extremo de la

distribución normal de la organización del cerebro. Sus cerebros parecen haber tomado la curva equivocada en el camino del desarrollo que conduce desde la herencia genética hacia la migración neuronal y la autoorganización del circuito durante el embarazo.

La demostración científica es cada vez más sólida en el caso de la dislexia, un trastorno específico del desarrollo que afecta la capacidad de aprender a leer, mientras deja intacta la inteligencia y otras facultades. Si usted es disléxico, cualquiera de sus hermanos tiene un 50% de posibilidades de sufrir también dislexia, lo que indica la sólida determinación genética de este trastorno del desarrollo. Hasta ahora se detectaron al menos cuatro genes implicados en la dislexia y, curiosamente, los cuatro afectan la capacidad de las neuronas para migrar hacia sus ubicaciones finales en la corteza durante el embarazo (Galaburda y otros, 2006). La resonancia magnética (RMI) a personas afectadas por la dislexia también muestra anomalías profundas en las conexiones que alojan la lectura en el hemisferio izquierdo de los lectores normales. Resulta crucial que las anomalías se pueden encontrar desde el principio: en niños con una predisposición genética a la dislexia, a los 6 meses de edad, un déficit en la distinción de los fonemas del lenguaje hablado ya separa a aquellos que desarrollarán esa condición de aquellos que se convertirán en lectores normales (Lepanen y otros, 2002; Lyytinen y otros, 2004). De hecho, se sabe que los déficits fonológicos son un factor importante en la aparición de la dislexia, pero no la única causa: el circuito de la lectura es tan complicado que hay muchos lugares donde puede fallar. Ya contamos con descripciones de varios tipos de dislexia, incluidos los déficits de atención que hacen que los pequeños mezclen las letras en palabras afines (Friedmann, Kerbel y Shvimer, 2010) y los déficits visuales que causan las confusiones “en espejo” (McKloskey y Rapp, 2000). La dislexia parece estar en el extremo de un continuo de habilidades visuales, atencionales y fonológicas que adopta la forma de una curva de campana y va desde la completa normalidad hasta el déficit severo. Todos compartimos la misma composición de *Homo sapiens*, pero diferimos ligeramente en la dosis cuantitativa de nuestra herencia, tal vez debido a variaciones semialeatorias en el diseño temprano de los circuitos neuronales (Shaywitz y otros, 1992).

Podríamos contar casi la misma historia respecto de otros trastornos del desarrollo. La discalculia, por ejemplo, se vio relacionada con déficits tempranos de materia gris y blanca en los circuitos dorsal parietal y frontal (que en las niñas y los niños sin dificultades alojan el cálculo y las

matemáticas). Los niños prematuros, que pueden sufrir infartos periventriculares en la región parietal que alberga el sentido numérico, tienen mayores riesgos de discalculia. La desorganización neurológica temprana puede causar discalculia al impactar directamente en el conocimiento de los conjuntos y las cantidades, o al desconectarlo de otras áreas involucradas en la adquisición de las palabras numéricas y los símbolos de la aritmética. En cualquiera de esos casos, el resultado es una predisposición a las dificultades para adquirir las matemáticas durante la infancia. Es probable que esos niños requieran ayuda específica para fortalecer sus débiles intuiciones iniciales para las cantidades.

Muchas veces se tiende a exagerar las consecuencias deterministas de los descubrimientos científicos sobre las bases genéticas del aprendizaje. Al respecto, cabe señalar que ningún gen involucrado en la dislexia, la discalculia o cualquiera de los demás trastornos del desarrollo, como el autismo o la esquizofrenia, tiene un determinismo del 100%. En el mejor de los casos, inclinan drásticamente la balanza, pero el entorno también tiene una gran participación en la trayectoria del desarrollo que finalmente seguirá un *infans*. Mis colegas en educación especial son optimistas: con suficiente esfuerzo, ninguna dislexia o discalculia es tan grave como para quedar fuera del alcance de la rehabilitación. Ya es hora de que pasemos a este segundo jugador importante en el desarrollo del cerebro: la plasticidad cerebral.

5. Lo que adquirimos

En las páginas previas, insistí mucho en el aporte de la naturaleza a la construcción de nuestro cerebro: la interacción de los genes y la autoorganización. Pero, por supuesto, el aprendizaje es igual de importante. La organización innata del cerebro no permanece inalterada por mucho tiempo, sino que se modifica y se enriquece con el efecto de la experiencia, que constituye la otra cara de la moneda: ¿cómo modifica el aprendizaje los circuitos del cerebro del niño? Para comprenderlo, necesitamos volver atrás más de un siglo, hasta los descubrimientos fundacionales de Santiago Ramón y Cajal.

¿Qué es la plasticidad cerebral?

La labor de un pianista es inaccesible para el hombre sin estudios ya que la adquisición de nuevas habilidades requiere muchos años de práctica mental y física. Para entender plenamente este complejo fenómeno, se hace necesario admitir, además del refuerzo de vías orgánicas preestablecidas, la formación de vías nuevas por ramificación y crecimiento progresivo de la arborización dendrítica y axonal.

Santiago Ramón y Cajal, *Textura del sistema nervioso del hombre y los vertebrados* (1899-1904)

El gran anatomista español Santiago Ramón y Cajal, uno de los héroes de las neurociencias, fue el primero en mapear la microorganización del cerebro. Dibujante genial, produjo diseños a la vez realistas y simplificados que se cuentan entre las obras maestras de la ilustración científica, apenas provisto

de su microscopio. Pero, sobre todo, supo pasar de la observación a la interpretación y de la anatomía a la función con una seguridad de juicio impresionante. Si bien su microscopio le permitía ver solo la forma general de las neuronas y de sus circuitos, ese pionero supo llegar a deducciones atrevidas y a menudo exactas sobre el modo en que funcionaban.

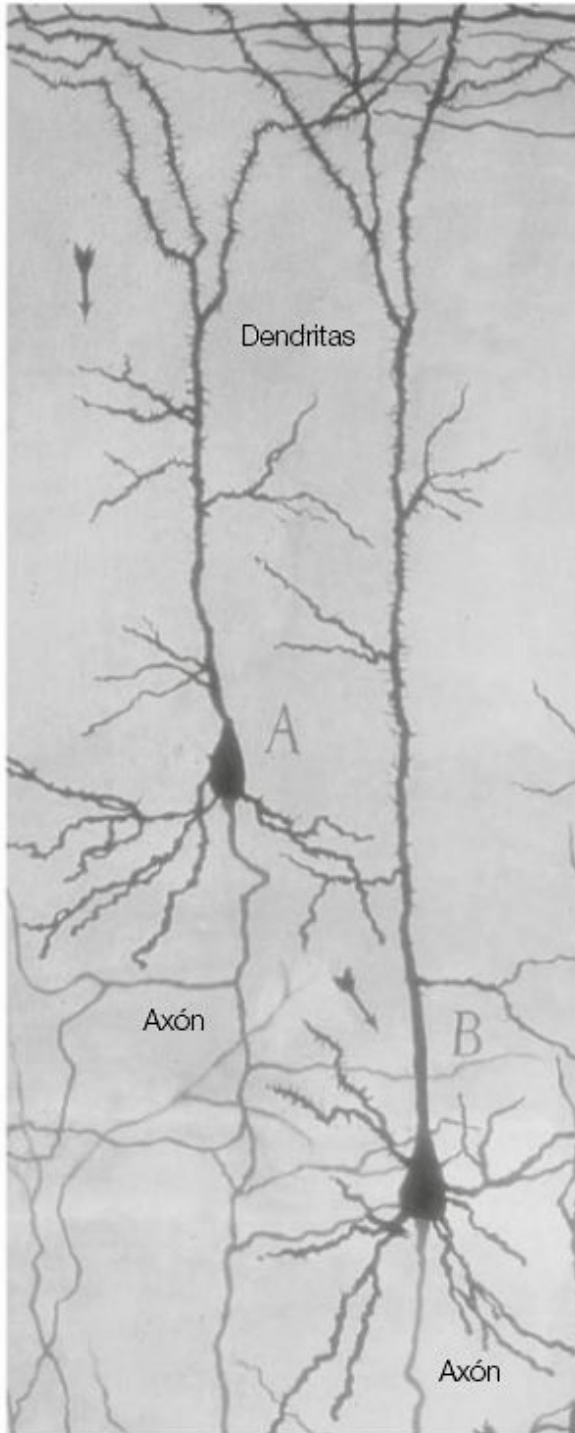
El mayor mérito de Cajal fue su comprensión de que el cerebro está formado por células nerviosas discretas (las neuronas), no por una red continua, un *reticulum*, como se creía antes de él. También notó que estas células tenían una forma muy particular: a diferencia de otras –como los glóbulos rojos, que son compactos y casi redondos–, las neuronas presentan ramificaciones de una complejidad increíble. Cada neurona despliega un inmenso árbol constituido por varios miles de ramas cada vez más pequeñas, las dendritas (en griego, *dendron* significa “árbol”). Las poblaciones de neuronas forman entonces un bosque inextricable de arborizaciones neuronales.

Esta complejidad no desanimó al neurobiólogo español. En un diagrama que se volvió famoso, a partir de un maravilloso dibujo de la anatomía del hipocampo, Cajal añadió algo muy sencillo y, sin embargo, iluminador y con un gran impacto teórico: ¡flechas que indicaban la dirección en que circulaba el impulso nervioso! Iban desde las dendritas hacia el cuerpo celular de la neurona y, por último, a lo largo del axón (figura 14). Esto suponía una especulación audaz, pero que resultó exacta. Cajal había comprendido que la forma de las neuronas se corresponde con su función: con ayuda de su árbol dendrítico, una neurona recolecta la información proveniente de otras células; todos estos mensajes convergen en el cuerpo celular, donde se los compila para enviarlos reunidos en uno solo, llamado “potencial de acción” o *spike*. El potencial de acción luego es transmitido por el axón, un tronco que sale del cuerpo celular y se extiende en dirección de otros miles de neuronas, en ocasiones a varios centímetros de distancia.

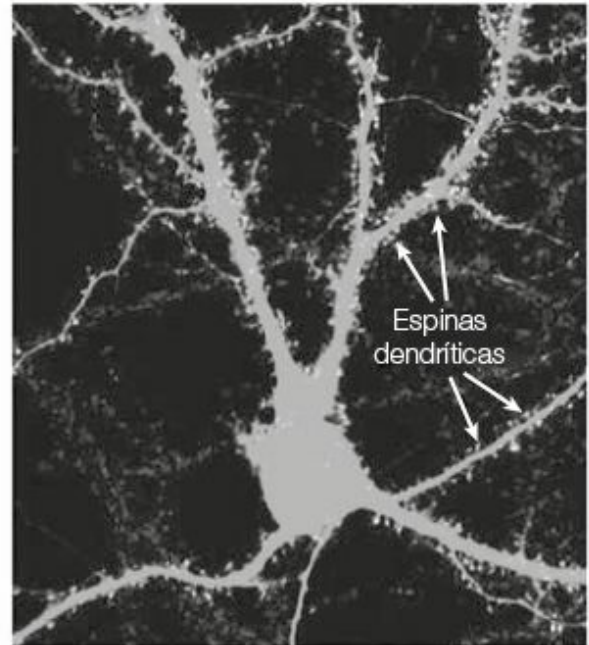
Cajal también llegó a una tercera hipótesis muy importante: que las neuronas se comunican entre sí por medio de las sinapsis. Fue el primero en comprender que las neuronas constituyen células distintas unas de otras; pero su microscopio también le reveló que estas células entran en contacto en algunos puntos, zonas de unión que en la actualidad llamamos “sinapsis” (nombre que debemos al gran fisiólogo británico Charles Sherrington). Cada sinapsis, entonces, es el punto de encuentro de dos neuronas o, más exactamente, del axón de una neurona y la dendrita de otra. La neurona que

llamamos “presináptica” envía su axón muy lejos, hasta encontrarse con una dendrita de una segunda neurona, la postsináptica, a la cual se conecta.

Microcircuitos



Neuronas



Sinapsis

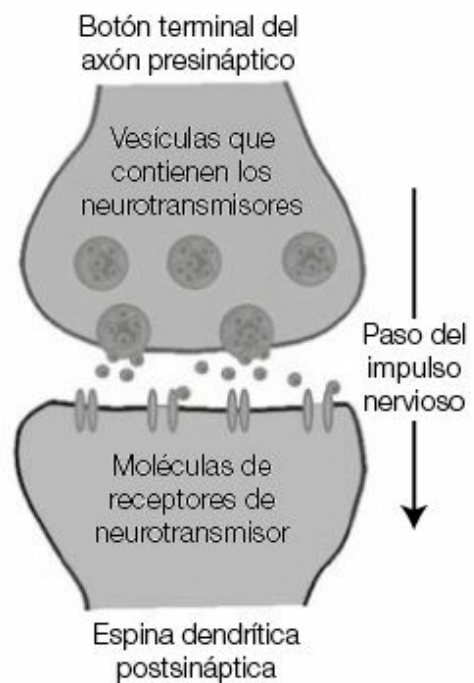


Figura 14. Microcircuitos, neuronas y sinapsis son el soporte material de la plasticidad del cerebro: se ajustan cada vez que aprendemos algo. El gran anatomista Ramón y Cajal trazó los circuitos de la corteza y comprendió que las neuronas son células individuales, cuyas arborescencias, las dendritas (arriba a la izquierda), registran la información proveniente de otras neuronas y envían mensajes a lo largo de su axón. Hoy en día, tenemos la posibilidad de observar las neuronas con una precisión pasmosa, hasta el detalle de las espinas dendríticas, los pequeños hongos que albergan las sinapsis, los puntos de conexión entre dos neuronas. Todos estos elementos cambian durante nuestro aprendizaje: la presencia, la cantidad y la fuerza de las sinapsis, el tamaño de las espinas dendríticas, la cantidad de ramas de las dendritas y de los axones (e incluso la aislación de los axones, gracias a una vaina de mielina, que determina la velocidad de transmisión del impulso nervioso).

¿Qué ocurre en una sinapsis? Otro premio Nobel, el neurofisiólogo Thomas Südhof, centró todas sus investigaciones en esta cuestión, y llegó a la conclusión de que las sinapsis son las unidades de cómputo del sistema nervioso, verdaderos nanoprocesadores. Recordemos que en nuestro cerebro hay alrededor de catorce mil millones de ellas: es una maquinaria de una complejidad sin igual. Todo lo que puedo hacer aquí es resumir los detalles más sencillos del proceso (figura 14). El mensaje que viaja en el axón es eléctrico, pero la mayoría de las sinapsis lo transforman en un mensaje químico. A escala de la sinapsis, el axón termina en un botón que contiene vesículas, algo así como almohadillas rellenas de una molécula que llamamos “neurotransmisor” (el glutamato, por ejemplo). Cuando la señal eléctrica llega hasta un botón terminal, estas almohadillas se abren y la molécula se difunde en el espacio sináptico entre las dos neuronas. A este mecanismo se debe el nombre “neurotransmisor”: transmite el mensaje de una neurona a la otra. En efecto, la molécula liberada se une a la membrana de la segunda neurona, en puntos específicos llamados “receptores”. El neurotransmisor obra sobre el receptor como una llave sobre una cerradura: literalmente, abre un canal en la membrana postsináptica. Allí se difunden los iones, átomos cargados positiva o negativamente, cuyo flujo genera una corriente eléctrica en el interior de la célula blanco. El ciclo está cerrado: el mensaje pasó de eléctrico a químico, de químico a eléctrico, y pudo cruzar el espacio entre las dos neuronas.

¿Qué relación tiene esto con el aprendizaje? Bueno, nuestras sinapsis se modifican permanentemente a lo largo de la vida, y estas modificaciones reflejan lo que aprendemos (Holtmaat y Caroni, 2016; Takeuchi, Duzskiewicz y Morris, 2014). Cada sinapsis es una pequeña usina química, y numerosos elementos de esta usina pueden cambiar durante el aprendizaje: la cantidad de vesículas, su tamaño, la cantidad de receptores, su eficacia, el

tamaño e incluso la forma de la sinapsis... Todos estos parámetros afectan la fuerza con que un mismo mensaje eléctrico, presináptico, se transmitirá a la segunda neurona, postsináptica, y por lo tanto proveen un espacio de almacenamiento útil para la información aprendida.

Además, estos cambios no se producen al azar: tienden a estabilizar la actividad de las neuronas al reforzar su capacidad de activarse unas a otras si ya lo han hecho en el pasado. La regla básica es tan elemental que ya fue hipotetizada en 1949 por el psicólogo Donald Hebb. La podemos resumir en una fórmula sencilla: *Neurons that fire together, wire together*, o, en castellano: “Coactivarse es conectarse”. Cuando dos neuronas se activan en el mismo momento, sus interconexiones se hacen más fuertes. Más exactamente, si la neurona emisora (es decir, la presináptica) descarga y, algunos milisegundos más tarde, la neurona receptora (postsináptica) entra también en actividad, entonces la sinapsis se afianza: en el futuro la transmisión entre estas dos neuronas será incluso más eficiente. Si en cambio la sinapsis no logra una comprensión adecuada, de modo que la neurona postsináptica no se activa, entonces la sinapsis se debilita.

Cada cambio sináptico que sigue la regla de Hebb refuerza la probabilidad de que se reproduzca la misma actividad. Gracias a la plasticidad sináptica, vastos tapices neuronales formados por millones de neuronas se suceden en un orden cada vez más exacto y reproducible. El ratón que recorre su laberinto por el camino óptimo, la violinista que con sus dedos hace brotar un sinfín de notas o la niña que recita una poesía despiertan una sinfonía neural en que cada movimiento, nota o palabra están registrados en varias centenas de millones de sinapsis.

Desde luego, el cerebro no almacena todos los eventos de nuestra vida. Solo los momentos que juzga más importantes se imprimen en las sinapsis. Por ende, la plasticidad sináptica es modulada por las grandes redes de neurotransmisores –sobre todo la acetilcolina, la dopamina y la serotonina–, que determinan qué episodios tienen la importancia suficiente para ser recordados. La dopamina, por ejemplo, es el neurotransmisor asociado a todas las recompensas: alimento, sexo, droga... y, por si se lo están preguntando, ¡sí, también al rock and roll! (Salimpoor y otros, 2013). Las descargas del circuito de la dopamina marcan todo lo que nos gusta, y le muestran al resto del cerebro que el episodio en curso es positivo, ya que está resultando mejor que lo previsto. Por otro lado, la acetilcolina suele involucrarse en todos los momentos importantes. Sus efectos son enormes. Si

ustedes recuerdan tan bien qué estaban haciendo el 11 de septiembre de 2001, cuando se enteraron del atentado contra el World Trade Center, es porque ese día un huracán de neurotransmisores desbarató sus circuitos cerebrales – especialmente en la amígdala, localización de las emociones fuertes– y sus sinapsis se alteraron en gran medida. La amígdala, grupo subcortical de neuronas que se disparan sobre todo por emociones fuertes, envía señales al hipocampo, que está muy cerca y almacena los episodios más relevantes de nuestra existencia. Así, las modificaciones sinápticas priorizan los hechos que los circuitos emocionales del cerebro consideran más significativos.

La capacidad de las sinapsis para modificarse en función de la actividad de sus neuronas pre- y postsinápticas se descubrió inicialmente en condiciones artificiales: los investigadores tuvieron que “aturdir” a las neuronas estimulándolas a un ritmo desenfrenado con una fuerte corriente eléctrica para poder observar que la fuerza de sus sinapsis cambiaba. Después de esta experiencia traumática, la sinapsis permaneció modificada durante varias horas, fenómeno llamado “potenciación a largo plazo”, que parecía ideal para lograr recuerdos muy perdurables (Bliss y Lømo, 1973; Lømo, 2018). Pero ¿era este realmente el mecanismo utilizado por el organismo en condiciones normales para almacenar información? Las primeras pruebas fueron aportadas por un molusco, la *Aplysia californica*), conocida también como liebre de mar o babosa marina, que tiene neuronas gigantes. En sus ganglios nerviosos –porque este invertebrado primitivo no tiene un cerebro propiamente dicho–, Eric Kandel, que obtendría el Premio Nobel, detectó una serie de modificaciones sinápticas y moleculares que se daban en cascada cuando se condicionaba al animal para que aprendiese a esperar la comida, un poco como el célebre perro de Pávlov (Pittenger y Kandel, 2003).

A medida que las técnicas de registro y visualización de las sinapsis progresaron, se acumularon pruebas de que su plasticidad es responsable del aprendizaje. Las modificaciones se producen precisamente en los circuitos que el animal utiliza con ese fin. Cuando un ratón aprende a evitar un lugar porque allí recibió una pequeña descarga eléctrica, se nota cómo cambian las sinapsis del hipocampo, región responsable del sentido del espacio y de la memoria episódica (Whitlock y otros, 2006): las conexiones entre el hipocampo y la amígdala registran el episodio traumático. En cambio, si el ratón se ve aterrorizado por un sonido, se modifican las sinapsis que ligan la amígdala con la corteza auditiva (Kim y Cho, 2017). Por lo demás, estos cambios no son simplemente concomitantes con el aprendizaje: parecen

desempeñar un papel causal en él. La prueba es que si en los minutos posteriores a un evento traumático se interfiere con los mecanismos moleculares que permiten a las sinapsis someterse a los cambios del aprendizaje, el animal no lo retiene en la memoria (Takeuchi, Duszkievicz y Morris, 2014).

El retrato de un recuerdo

¿Qué es, entonces, un recuerdo? ¿Y cuál es su base material en el cerebro? La mayoría de los investigadores coincide con la siguiente explicación, que distingue períodos de codificación y de recuerdo (Josselyn, Köhler y Frankland, 2015; Poo y otros, 2016).

Comencemos por la codificación. Sin excepción, nuestras percepciones, acciones o nuestros pensamientos dependen de que entre en actividad un subconjunto muy específico de neuronas (mientras que las otras permanecen inactivas, o incluso se inhiben). La identidad de las neuronas activas, distribuidas en numerosas regiones del cerebro, define el contenido del pensamiento. Por ejemplo, cuando veo a Ennio Morricone en la Arena de Verona, algunas neuronas (en la región temporal inferior) responden a su rostro, otras (en la región temporal superior) a su voz, otras (en la región del parahipocampo) a la distribución de los sectores del auditorio, y así sucesivamente: cada neurona por separado puede proveer cierta información, pero la memoria general siempre está codificada por varios grupos de neuronas interconectados. Si me encuentro con la señora a quien compro queso en el mercado, la activación de grupos de neuronas ligeramente diferentes me permite, en principio, no confundirla con el gran compositor y director de orquesta italiano.

Supongamos ahora que mis sistemas emocionales consideran que esta experiencia tiene la importancia suficiente para que mi memoria la almacene. ¿Cómo hace mi cerebro para registrarla? Para consolidar el evento, los grupos de neuronas que recientemente se activaron sufren cambios físicos importantes. Modifican la fuerza de sus interconexiones, lo que aumenta el apoyo del grupo y vuelve más probable que en el futuro estas neuronas se activen juntas. Algunas sinapsis se hacen físicamente más grandes e incluso

pueden llegar a duplicarse. A veces las neuronas desarrollan nuevas espinas dendríticas, botones axonales o dendritas. Todas estas modificaciones anatómicas implican la expresión de nuevos genes, en el transcurso de varias horas o incluso días, y forman la base concreta del aprendizaje: colectivamente, constituyen el sustrato de la memoria.

Una vez formada la memoria sináptica, las neuronas pueden descansar: cuando dejan de descargar, el recuerdo permanece latente, inconsciente pero muy real, inscripto en la anatomía misma de nuestros circuitos neuronales. Más tarde, gracias a estas conexiones, será suficiente un indicio externo (por ejemplo, una foto aérea de la Arena) para producir una cascada de actividad neuronal en el circuito original; poco a poco esto hará que el cerebro logre recrear un patrón de descargas similar a aquel que ocurrió cuando se creó el recuerdo y, por último, me permitirá traer al presente el rostro de Ennio Morricone. Según esta teoría, cada recuerdo es una reconstrucción; recordar es intentar reproducir en el cerebro las descargas neuronales asociadas a una experiencia pasada.

Así, el recuerdo no está localizado en una sola región del cerebro, sino distribuido en la mayoría (o acaso en cada uno) de los circuitos cerebrales, porque todos ellos son capaces de modificar su sinapsis en respuesta a una configuración frecuente de la actividad neuronal. Sin embargo, no todos los circuitos tienen la misma función. Pese a que la terminología todavía es confusa y continúa evolucionando, los investigadores distinguen por lo menos cuatro tipos de memoria.

- *Memoria de trabajo*: conserva una representación mental activa durante un lapso muy breve, digamos unos pocos segundos. Depende principalmente de la descarga vigorosa de numerosas neuronas en la corteza parietal y prefrontal, que a su vez estimulan a las neuronas de otras regiones más periféricas (Courtney y otros, 1997; Ester, Sprague y Serences, 2015; Goldman-Rakic, 1995; Kerkoerle, Self y Roelfsema, 2017; Vogel y Machizawa, 2004). Esta memoria nos permite retener un número de teléfono en la mente: durante el tiempo que nos insume marcarlo en nuestro *smartphone*, las neuronas se respaldan unas a otras y así conservan la información en estado activo. Este tipo de memoria se basa principalmente sobre la preservación de un patrón sostenido de actividad, aunque estudios recientes descubrieron que

también implica cambios sinápticos de corta duración, lo que permite que las neuronas pasen un breve período de inactividad y enseguida vuelvan a su estado activo (Mongillo, Barak y Tsodyks, 2008). En cualquier caso, la memoria de trabajo nunca dura más que unos pocos segundos: tan pronto como nos distraemos con algo más, el conjunto de neuronas activas se desvanece. Es el búfer a corto plazo del cerebro, que solo tiene en cuenta la información más reciente y actual.

- *Memoria episódica*: el hipocampo –estructura situada en la profundidad de los hemisferios cerebrales, debajo de la corteza– registra los episodios de nuestra vida cotidiana a medida que se desarrollan. Las neuronas del hipocampo parecen memorizar el contexto de todos los eventos: codifican dónde, cuándo, cómo y con quién ocurrieron. Almacenan el episodio en cuestión gracias a los cambios sinápticos, de modo que podamos recordarlo más tarde. El célebre paciente H. M., quien había sufrido la extirpación del hipocampo en una intervención quirúrgica bilateral destinada a controlar su epilepsia, ya no registraba nada: vivía en un eterno presente, incapaz de agregar el menor elemento nuevo a su biografía mental. Hoy en día, sabemos que el hipocampo interviene en todo tipo de aprendizaje rápido. Siempre que la información aprendida sea única, se trate de un acontecimiento particular o de un descubrimiento digno de interés, las neuronas del hipocampo le asignarán una secuencia de descargas específica (Genzel y otros, 2017; Lisman y otros, 2017; Schapiro y otros, 2016; Shohamy y Turk-Browne, 2013).
- *Memoria semántica*: los recuerdos no parecen quedarse en el hipocampo para siempre; durante la noche, el cerebro los repasa y los mueve hacia una nueva ubicación en la corteza. Allí, se transforman en conocimientos permanentes: el cerebro extrae la información que está presente en los episodios que vivimos, la generaliza y la integra en nuestra vasta biblioteca de conocimientos del mundo. Luego de unos días, recordamos el nombre del autor de la música de *La Misión* (sí, Ennio Morricone), aunque no tengamos el menor trazo del lugar o del momento en que lo oímos por primera vez: el recuerdo de la memoria episódica se tornó semántico. Lo que inicialmente era un episodio único se convirtió en conocimiento duradero y su código neural se desplazó desde el hipocampo hacia los circuitos relevantes

en la corteza (Kitamura y otros, 2017).

- *Memoria procedural*: cuando repetimos la misma actividad una y otra vez (atarnos los zapatos, recitar un poema, calcular, hacer malabares, tocar el violín, andar en bicicleta...), las neuronas de la corteza y de otros circuitos subcorticales terminan por modificarse para que la información circule mejor en el futuro. Más rápida, más reproducible, depurada de cualquier actividad parasitaria, la actividad neuronal se desarrolla con la precisión de un mecanismo de relojería. Se trata de la memoria procedural: el registro, inconsciente y compacto, de las configuraciones de actividades rutinarias, que realizamos a intervalos periódicos. En ella, el hipocampo ya no tiene función alguna, porque la repetición del gesto es lo que transfiere el recuerdo hacia un soporte implícito. A esto se debe que el paciente H. M., incluso desprovisto de cualquier recuerdo consciente, todavía pudiese aprender nuevos procedimientos. Los investigadores incluso le enseñaron a escribir a la inversa, observando su mano en un espejo. Dado que no recordaba la cantidad de veces que lo había practicado antes, H. M. se sorprendía al descubrirse tan bueno para lo que creía que era un “truco” nuevo.

Verdaderas sinapsis y falsos recuerdos

Hoy en día, el conocimiento de los investigadores acerca de los circuitos de la memoria es tal que ya no estamos muy lejos de la fantasía de Michel Gondry en la película *Eternal Sunshine of the Spotless Mind* [*Eterno resplandor de una mente sin recuerdos* u *¡Olvídate de mí!*]: borrar del tapiz de la memoria, de forma selectiva, los recuerdos que nos envenenan la vida –como el estrés postraumático de los soldados de la Guerra de Vietnam–, o bien, bordar en su lugar una composición ilusoria, un recuerdo falso.

El equipo de otro premio Nobel, el profesor Susumu Tonegawa, ya realizó esas dos manipulaciones en ratones. Primero les hizo vivir un pequeño trauma: ubicado en una habitación, el ratón recibe pequeñas descargas eléctricas. Es fácil asegurarse de que ese episodio está bien impreso en su memoria, ya que el ratón evita el lugar donde se produjo.

A partir de ese momento, podemos jugar con el recuerdo. En primer lugar,

tenemos la opción de registrarlo. Gracias a un microscopio sofisticado de dos fotones, logramos visualizar la respuesta de las neuronas y constatamos que, en el hipocampo, se activan neuronas diferentes para la habitación A, asociada a la descarga eléctrica, y para la habitación B, donde no pasó nada. Luego es la hora de manipularlo: mientras el animal está en la habitación A, se le envían descargas que activan, en su cerebro, la población de neuronas que codifican la habitación B. Cuando regresa a la habitación B, el ratón se alarma y se paraliza de miedo: es como si el mal recuerdo estuviera ahora asociado a la habitación B, donde, sin embargo, nunca ocurrió nada (Ramirez y otros, 2013).

El equipo de Susumu Tonegawa luego transformó el mal recuerdo en uno bueno: al reactivar las mismas neuronas de la habitación B mientras se exponía a los ratones a la presencia de compañeros del sexo opuesto –un buen momento garantizado–, los investigadores lograron borrar la asociación con la descarga eléctrica. Para entonces los ratones, lejos de evitar la habitación maldita, se ponían a explorarla frenéticamente como si buscaran allí a los compañeros eróticos que recordaban (Ramirez y otros, 2015).

Por ende, ¿podemos borrar un recuerdo traumático? Sí, es suficiente con volver a evocarlo mientras se debilitan las sinapsis unidas a él en la memoria. En los días siguientes, el ratón ya no muestra la menor reminiscencia del mal momento (Kim y Cho, 2017).

Siguiendo esa misma línea, el investigador francés Karim Benchenane logró la proeza de imprimir un nuevo recuerdo en el cerebro de un ratón durante el sueño. Para esto aprovechó que, mientras dormimos, el hipocampo reactiva una huella de los recuerdos de las horas previas, con especial énfasis en los lugares por los que pasamos (retomaremos este tema en el capítulo 10). Cada vez que el cerebro del ratón dormido evocaba determinado lugar de su entorno, el investigador le daba una pequeña inyección de dopamina, el neurotransmisor de la recompensa. Resultado: ¡tan pronto como se despertaba, el ratón se precipitaba hacia ese lugar! Durante el sueño, lo que inicialmente era un lugar neutral había adquirido un sentido muy especial en la memoria, tan adictivo como los dulces de nuestra abuela o el primer lugar en el que nos enamoramos.

En un intento por emular la vivencia de los humanos, algunas investigaciones con animales comenzaron a imitar los efectos en el cerebro de los aprendizajes escolares. ¿Qué sucede cuando un mono aprende las letras, los dígitos o el modo de uso de una herramienta? (Iriki, 2005;

Obayashi y otros, 2001; Srihasam y otros, 2012). El investigador japonés Atsushi Iriki mostró que un mono podía aprender a utilizar un rastrillo para acercar a su mano porciones de alimento demasiado alejadas para agarrarlas. Luego de varios miles de ensayos, el animal se volvía tan experto como un crupier: no necesitaba más que unas décimas de segundo para acercarse la comida tirando del rastrillo con un gesto vivaz. ¡El mono también lograba utilizar un primer rastrillo, de tamaño mediano, para trasladar otro, más largo, que le permitía finalmente alcanzar un alimento que estaba incluso a mayor distancia! Este tipo de aprendizaje desencadena una serie de cambios en el cerebro. El consumo de energía aumenta en un área muy específica de la corteza, la región parietal anterior, la misma que un ser humano utiliza para controlar los movimientos de la mano, escribir, atrapar un objeto o utilizar un martillo o una pinza. Allí se expresan nuevos genes, las sinapsis brotan, las arborescencias de dendritas y de axones se multiplican y, en el animal experto, todas estas ramificaciones suplementarias se traducen en un aumento del 23% del espesor de la corteza. También vemos alterarse los haces de conexiones: los axones provenientes de una región lejana, en la unión con la corteza temporal, se desarrollan hasta crecer varios milímetros e invadir una porción de la región parietal anterior que antes no estaba cableada por estas neuronas.

Estos ejemplos explican en qué medida los efectos de la plasticidad cerebral se extienden en el tiempo y en el espacio. Les propongo que repasemos juntos los puntos fundamentales. Un conjunto de neuronas, que codifica un evento o un concepto que deseamos memorizar, se activa en nuestro cerebro. ¿Cómo se registra este recuerdo? En el principio es la sinapsis: un microscópico punto de contacto entre dos neuronas que se refuerza cuando las neuronas que vincula se activan juntas en breve sucesión: “Coactivarse es conectarse”, como rezaba la famosa regla de Hebb. Una sinapsis que se afianza es como una fábrica que aumenta su productividad: recluta más neurotransmisores del lado presináptico y más moléculas de receptores del lado postsináptico. También, con el objetivo de alojarlos, aumenta su superficie.

Cuando una neurona aprende, su forma también cambia. Una estructura con forma de champiñón, que llamamos “espinas dendríticas”, se forma en el lugar de la dendrita que alberga la sinapsis. De ser necesario, emerge una segunda sinapsis que duplica la primera. Otras sinapsis distantes, pero que aterrizan

sobre la misma neurona, también se refuerzan (Fitzsimonds, Song y Poo, 1997).

De este modo, cuando el aprendizaje se extiende, la anatomía misma del cerebro termina por cambiar. Con los avances de la microscopía, y, muy especialmente, con la revolución aportada por los novedosos microscopios de dos fotones, que se valen del láser y de la física cuántica, puede verse, sin mediaciones, el crecimiento de los botones sinápticos y axonales en cada episodio de aprendizaje, al igual que se ve brotar un árbol en primavera. Acumulados, esos cambios pueden ser sustanciales, del rango de milímetros, y gracias a las imágenes de resonancia magnética comienzan a ser detectables en los humanos. Aprender a tocar música (Gaser y Schlaug, 2003; Oechslin, Gschwind y James, 2018; Schlaug y otros, 1995), a leer (Carreiras y otros, 2009; Thiebaut de Schotten y otros, 2014), a hacer malabares (Draganski y otros, 2004; Gerber y otros, 2014) o incluso a manejar un taxi en una gran ciudad (Maguire y otros, 2000, 2003) son experiencias que se traducen en aumentos detectables en el espesor de la corteza o de la fuerza de las conexiones que vinculan las regiones corticales: las autopistas del cerebro mejoran cuanto más las utilizamos.

Si bien la sinapsis es el epítome del aprendizaje, no es el único mecanismo que genera cambios en el cerebro. Cuando aprendemos, el *boom* de nuevas sinapsis fuerza a las neuronas a engendrar nuevas ramificaciones, tanto en los axones como en las dendritas. Lejos de la sinapsis, los axones útiles se recubren de una vaina de aislante, la mielina, que se parece al adhesivo que se pone alrededor de los cables eléctricos para aislarlos. Cuanto más se utiliza un axón, más vueltas tiene esta vaina. Así, lo aísla cada vez mejor y le permite retransmitir la información con una velocidad cada vez mayor.

Las neuronas no son los únicos jugadores en este juego. A medida que un aprendizaje progresa, también cambia el entorno de las neuronas, incluidas las células gliales, que las nutren y las cuidan, así como la red vascular de venas y arterias que les aportan oxígeno, glucosa y nutrientes. Así, todo el circuito neural y su estructura de soporte se transforman.

Algunos investigadores cuestionan el dogma que hace de la sinapsis el actor indispensable de cada aprendizaje. Según datos recientes, las células de Purkinje (neuronas especiales del cerebelo) pueden memorizar un intervalo de tiempo sin que este aprendizaje requiera la menor modificación de sus sinapsis; el mecanismo parece ser interno a la célula únicamente (Johansson y otros, 2014; Rasmussen, Jirenhed y Hesslow, 2008). Es posible que la

dimensión del tiempo, que es una especialidad del cerebelo, se almacene en la memoria usando un truco evolutivo diferente, que no dependa de la sinapsis. Cada neurona del cerebelo, por sí sola, sería capaz de almacenar varios intervalos de tiempo, quizá por medio de modificaciones químicas estables de su ADN.

Por cierto, aún no comprendemos cómo estos cambios inducidos por el aprendizaje, sean sinápticos o no, implementan los tipos de aprendizaje más elaborados de los cuales es capaz el cerebro humano, aquellos basados sobre el “lenguaje del pensamiento” y la recombinación rápida de conceptos existentes. Según vimos antes, las redes neuronales convencionales explican cómo millones de sinapsis, al modificarse, nos permiten aprender a reconocer una cifra, un objeto o un rostro. En cambio, no existe modelo alguno verdaderamente satisfactorio de las redes neuronales que son soporte de la adquisición del lenguaje o de las reglas matemáticas. Aun hoy, pasar del ámbito de la sinapsis a las reglas simbólicas que aprendemos en la clase de matemáticas es un desafío. Tengamos la mente abierta, porque estamos muy lejos de alcanzar una comprensión cabal del código celular por cuyo intermedio el cerebro logra guardar nuestros aprendizajes y recuerdos.

La nutrición, una pieza clave del aprendizaje

Como vimos, cuando aprendemos se producen cambios biológicos enormes: no solo se modifican las neuronas, con sus dendritas y axones, sino que también cambia su entorno de células gliales. Todas estas transformaciones requieren tiempo. Cada episodio de aprendizaje abre la compuerta a una cascada de modificaciones biológicas que dura varios días. Numerosos genes especializados en la plasticidad deben expresarse para que las células produzcan el andamiaje de proteínas y membranas que van a constituir las sinapsis, las dendritas y los axones. Este proceso absorbe mucha energía: el consumo del cerebro del niño pequeño representa hasta un 50% del balance energético del cuerpo. Glucosa, oxígeno, vitaminas, hierro, iodo, ácidos grasos, nutrientes de lo más diversos: todos estos ingredientes son esenciales en el cóctel de una infancia exitosa. El cerebro no se alimenta solo de estimulación intelectual; para hacer y deshacer algunos millones de sinapsis

por segundo, necesita una dieta equilibrada, una buena oxigenación y ejercicio físico (Prado y Dewey, 2014; Voss y otros, 2013).

Un triste episodio deja en claro la extrema sensibilidad del cerebro en desarrollo y el papel clave de una buena nutrición. En noviembre de 2003, un vergonzoso desastre golpeó a los niños de Israel (Fattal, Friedmann y Fattal-Valevski, 2011). De un día para el otro, los hospitales pediátricos de todo el país recibieron a decenas de bebés afectados por graves síntomas neurológicos: letargo, vómitos, trastornos oculares y de alerta que alcanzaban incluso el coma; para dos de ellos, significó la muerte. Desde luego, hubo una carrera contrarreloj: ¿cuál era el origen de esta enfermedad hasta entonces desconocida?

La investigación derivó en el campo de la nutrición: todos los bebés enfermos habían sido alimentados con la misma leche en polvo con base de soja. El análisis de esta leche de fórmula confirmó los peores temores: de acuerdo con la etiqueta, y conforme a las normas, debería haber incluido 385 mg de tiamina, más conocida como vitamina B1. En realidad, no contenía siquiera un dejo de este elemento. Al ser contactado, el fabricante admitió que había alterado la composición de su leche a principios de ese año: por razones económicas, había dejado de agregar tiamina, nutriente esencial del cerebro. Para empeorar las cosas, dado que el organismo no almacena esa sustancia, cualquier desequilibrio alimentario acarrea rápidamente una grave carencia.

Los neurólogos ya sabían que en los adultos el déficit de tiamina causa problemas neurológicos severos, un cuadro que llamamos “síndrome de Wernicke-Korsakoff” y que la mayoría de las veces encontramos en los casos graves de alcoholismo. En la fase aguda, esta deficiencia induce una enfermedad neurológica a veces mortal, la encefalopatía de Wernicke. Confusión mental, problemas del movimiento de los ojos, incapacidad de coordinación y déficit de alerta que podría llegar hasta el coma: estos síntomas se parecen mucho a los de los bebés de Israel.

La última prueba llegó con la intervención terapéutica. Tan pronto como se volvió a incorporar esta vitamina esencial a la alimentación de los lactantes, su estado mejoró en unos pocos días y pudieron volver a sus casas. Se estima entre seiscientos y mil la cantidad de bebés israelíes que, en el transcurso de sus primeros meses de vida, fueron privados de tiamina durante un lapso de dos a tres semanas. En ese momento, se pensó que la recuperación de una alimentación equilibrada los había salvado. Se ignoraba que, años más tarde,

presentarían serias dificultades en el lenguaje. La psicóloga israelí Naama Friedmann evaluó a sesenta de estos niños a la edad de 6 o 7 años. Pese al tiempo transcurrido, la mayoría todavía sufría déficits importantes de comprensión y de producción de lenguaje. Su gramática era anormal: tras leer o escuchar una oración, tenían problemas para decidir quién le había hecho qué a quién. En algunos casos les resultaba difícil siquiera nombrar una imagen, como la de una oveja. En cambio, su procesamiento conceptual parecía intacto: por ejemplo, eran capaces de asociar la imagen de un ovillo de lana con la de una oveja antes que con la de un león. En el resto de los parámetros, y muy especialmente el de la inteligencia (el famoso CI), no parecían estar afectados.

Esta triste historia –por fortuna, excepcional– esclarece los límites de la plasticidad cerebral, base del aprendizaje del lenguaje: si cualquier bebé es capaz de aprender todos los idiomas del mundo, desde los tonos del chino hasta los clics o chasquidos de las lenguas bantúes de Sudáfrica, es porque su cerebro cambia del modo adecuado en respuesta a la inmersión en una comunidad específica. Pero esta plasticidad no es infinita ni mágica, sino un proceso material estricto que requiere ciertas condiciones nutricionales y energéticas, de modo que apenas algunas semanas de privaciones pueden implicar un déficit permanente. Y debido a que la organización del cerebro es en gran medida modular, bien puede ocurrir que este déficit esté acotado a un campo específico, como la gramática o el vocabulario. Los estudios de casos pediátricos sobreabundan en ejemplos similares; el más común es el síndrome de la alcoholización fetal, causado por la exposición del feto al alcohol consumido por la madre: este poderoso teratogénico causa malformaciones embrionarias del cuerpo y del cerebro; como verdadero veneno para el desarrollo del sistema nervioso, debe evitarse a lo largo del embarazo. Para que se extiendan los árboles dendríticos, hay que aportar al cerebro todos los nutrientes que necesita.

Posibilidades y límites de la plasticidad sináptica

En casos de buena nutrición, ¿hasta dónde puede llegar la plasticidad cerebral? ¿Puede recablear nuestro cerebro? ¿La anatomía cerebral es pasible

de variaciones considerables en función de nuestras vivencias? La respuesta es no: la plasticidad es una variable de ajuste, ciertamente fundamental en el aprendizaje, pero limitada, y acotada por todo tipo de restricciones genéticas que nos hacen lo que somos: la intersección entre un genoma fijo y experiencias únicas.

Llegó el momento de que les comente un poco más acerca de Nico. Recuerden: es el artista que les presenté en la introducción y que pinta cuadros espléndidos con la sola ayuda de su hemisferio izquierdo (figura 1). En efecto, a los 3 años y 7 meses de edad, Nico perdió su hemisferio derecho. Se expuso a una intervención quirúrgica llamada “hemisferectomía”, que implica la ablación casi completa de un hemisferio; en su caso, con la intención de ponerle fin a una epilepsia devastadora.

Sin embargo, apoyado por su familia, por sus médicos y por un investigador argentino de la Facultad de Educación de Harvard, Antonio Battro, Nico logró seguir una escolaridad elemental en Buenos Aires, y luego fue a la secundaria en Madrid hasta los 18 años. Su lenguaje oral y escrito, su memoria y sus competencias espaciales son excelentes. Incluso obtuvo un título universitario en informática. Pero sobre todas las cosas, Nico presenta un talento notable para el dibujo.

¿Es este un buen ejemplo de la plasticidad cerebral en funcionamiento? Por supuesto que sí, dado que Nico domina numerosas funciones que en una persona normal suelen estar asociadas al hemisferio derecho: la capacidad de prestar atención a la globalidad de una figura, de copiar un dibujo disponiendo de modo conveniente los objetos en el espacio o hasta de comprender la ironía y los sobreentendidos de una conversación y adivinar los pensamientos de sus interlocutores. Pese a todo, si la misma lesión ocurriera en un cerebro adulto, estas funciones se perderían sin remedio.

Pero la plasticidad de Nico es limitada: quedó canalizada y confinada a ciertos circuitos que son los mismos que los de todos los niños. Cuando tomamos imágenes cerebrales mientras evaluábamos a Nico con una batería de pruebas, constatamos que había logrado situar todas sus habilidades aprendidas en el hemisferio izquierdo, pero sin que su organización habitual se viera alterada: ¡todas las funciones tradicionalmente lateralizadas a la derecha se habían localizado en lugares del hemisferio izquierdo simétricos de su posición habitual! Así, la región cortical que responde a los rostros y que en condiciones normales se aloja en el lóbulo temporal derecho, en el caso de Nico se situaba en el hemisferio izquierdo pero en un lugar

exactamente simétrico respecto de su localización habitual, y a menudo activado (con debilidad) por los rostros en todos los niños. Eso equivale a decir que su cerebro se había reorganizado, pero dentro de los límites de una organización preexistente, común a toda la especie. Los grandes haces de conexiones que, desde el nacimiento e incluso durante la gestación, ya pululan en el cerebro del bebé habían confinado sus aprendizajes dentro de los límites estrechos de un mapa cortical universal y compartido por todos los seres humanos.

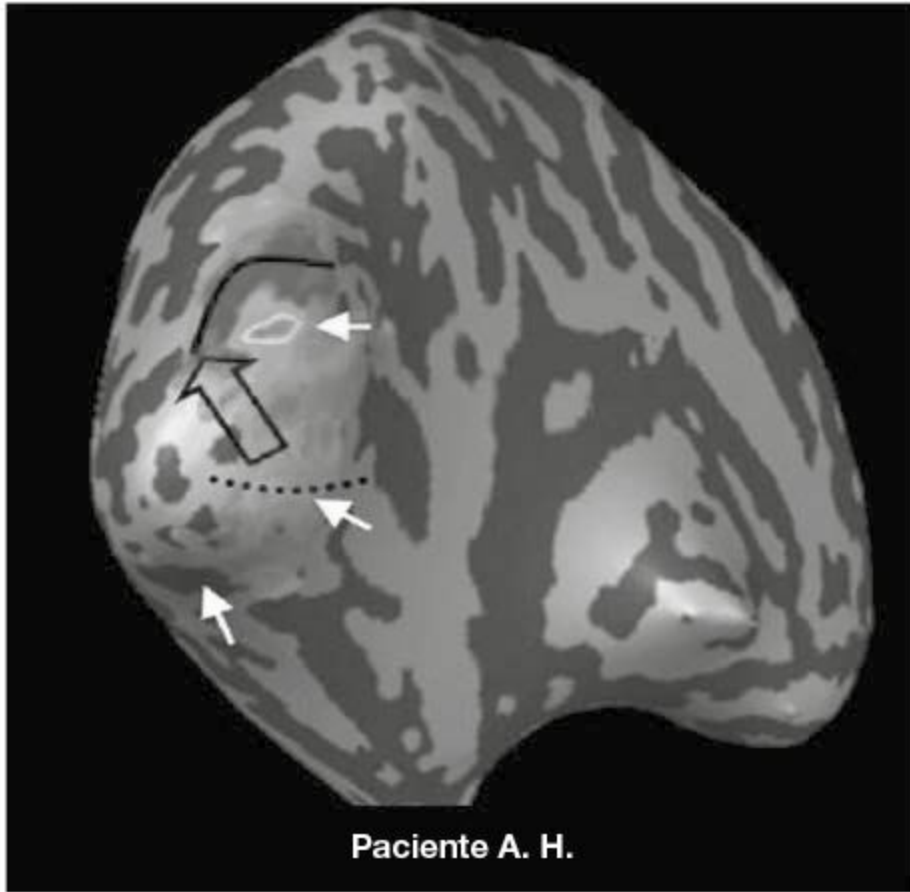
La potencia de la plasticidad, así como sus restricciones, jamás es tan evidente como cuando nos enfocamos en las capacidades visuales. No nos sorprende que Nico sea hemianópsico, lo que significa que su visión está cortada en dos: una mitad derecha con la cual ve sin problemas (en sus dos ojos) y una mitad izquierda con la cual es por completo ciego (también en sus dos ojos). En efecto, con el cruce de las vías visuales, los estímulos que provienen de la mitad izquierda de la retina, que usualmente aterrizan en el hemisferio derecho, caen en un vacío y no son procesados. Veinte años de vida no le permitieron compensar este fundamental problema de cableado: la plasticidad de las conexiones visuales, demasiado modesta, se fija a una gran velocidad durante la infancia temprana; en su caso, no pudo impedir esa ceguera para la mitad izquierda del mundo.

Ahora, examinemos a una pequeña de 10 años que conocemos solo por sus iniciales: A. H. (Muckli, Naumer y Singer, 2009). Esta niña, como Nico, no posee más que el hemisferio izquierdo, pero, a diferencia de él, sufrió una malformación embrionaria que hizo que el desarrollo de su hemisferio derecho se detuviera por completo antes de las 7 semanas de gestación. Por eso, todo su cerebro se desarrolló sin que el hemisferio derecho estuviera presente. ¿La plasticidad cambió todo su cerebro? No, pero le permitió ir un poco más lejos que Nico. En efecto, ella ve un poco del lado izquierdo del campo visual, aquel que debería proyectarse en el hemisferio derecho. Su visión está lejos de ser perfecta, pero distingue la luz y el movimiento en una región cercana al centro de la visión. Las imágenes cerebrales muestran que su cerebro se reorganizó parcialmente (figura 15). En el nivel del occipucio, en la corteza occipital que aloja la visión, notamos un mapa sin alteraciones en la parte derecha del mundo, pero también de zonas muy pequeñas que responden a la parte izquierda. Parecería haber axones provenientes de esta mitad de la retina que, en condiciones normales, debería haber sido ciega, y que se reorientaron hacia el otro lado del cerebro. Es un caso extremo de

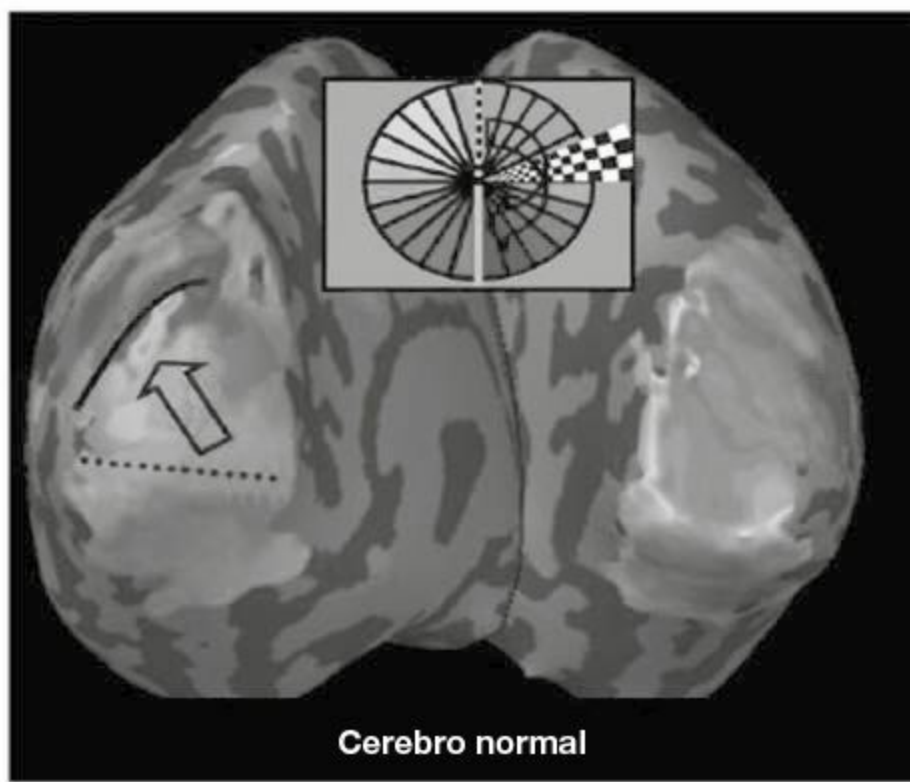
reorganización prenatal (y, aun así, solo parcial, insuficiente para restaurar una visión normal). En el sistema visual, las limitaciones genéticas dominan y la plasticidad actúa únicamente dentro de sus estrechos límites.

Los científicos quisieron saber hasta dónde podían llegar estos límites genéticos. Un experimento consiguió gran celebridad: el de Mriganka Sur, investigador del MIT, que logró transformar una corteza auditiva en corteza visual (Sur, Garraghty y Roe, 1988; Sur y Rubenstein, 2005). Para lograrlo, mediante una intervención quirúrgica practicada en fetos de hurón, cercenó los circuitos de entrada que, desde la cóclea y luego de varios pasos en el tronco cerebral, normalmente se dirigen hacia una región muy específica del tálamo y prosiguen hacia la corteza auditiva. Resultado indefectible: los hurones quedaron sordos. Sin embargo, se produjo una curiosa reorientación y las fibras visuales comenzaron a invadir el circuito auditivo desconectado, como si se ocuparan de reemplazar los circuitos auditivos faltantes. Así, un sector completo de la corteza que debería haber sido auditivo se comportaba ahora como un área visual. Notamos allí un mapa de neuronas sensibles a la luz y a la orientación de las barras, como en cualquier otra corteza visual. Las sinapsis se adaptaron a esta nueva configuración y empezaron a reflejar las correlaciones entre las neuronas que estaban destinadas a la audición, pero que, recicladas, se convirtieron en especialistas de la visión.

¿Estos datos deben invitarnos a la conclusión de que la plasticidad cerebral es “masiva” y que es la experiencia lo que “organiza la corteza”, como les gustaría a los más ardientes defensores de la metáfora de la *tabula rasa*? (Quartz y Sejnowski, 1997). Mriganka Sur jamás propondría esa tesis; por el contrario, insiste en que esa es una situación patológica, y que la reorganización está lejos de ser perfecta: en la corteza auditiva, los mapas visuales nunca llegan a estar tan bien diferenciados como deberían. La genética de la corteza visual la preparó para hacerse cargo de la visión. Durante el desarrollo normal, cada región cortical se especializa precozmente por influencia de muchos genes del desarrollo. Los axones encuentran su camino siguiendo pistas químicas predeterminadas que trazan protomapas. Solo al final del camino se someten a la creciente influencia de la actividad neuronal entrante y se adaptan a ella. El tapiz neuronal es fijo; apenas algunos detalles del bordado, pequeños aun que significativos, pueden cambiar.



Paciente A. H.



Cerebro normal

Figura 15. La plasticidad sináptica permite que el cerebro se reconfigure parcialmente cuando sufre alteraciones graves. La paciente A. H. (arriba) nació con un solo hemisferio cerebral: desde la séptima semana de embarazo, el hemisferio derecho de su cerebro dejó de desarrollarse. En un cerebro normal (abajo), las áreas visuales tempranas del hemisferio izquierdo solo representan la mitad derecha del mundo. Sin embargo, en el caso de esta paciente, regiones muy pequeñas se reorientaron y comenzaron a responder a la mitad izquierda del mundo (las señalamos con flechas blancas). Así, A. H. no es del todo ciega del lado izquierdo, como lo sería un adulto que perdiera su hemisferio derecho. Sin embargo, esta reorganización es modesta: la influencia de la genética predomina sobre la de la experiencia.

Además, es importante comprender que cuando las sinapsis se modifican, incluso influidas por la actividad de las neuronas, lo que deja una huella en el cerebro no es necesariamente el ambiente. Antes bien, el cerebro puede usar la plasticidad sináptica para autoorganizarse: primero genera patrones de actividad exclusivos desde dentro, en ausencia de cualquier *input* del entorno, y usa esos patrones de actividad, en combinación con la plasticidad sináptica, para conectar sus circuitos. Durante la gestación, incluso antes de recibir el menor *input* sensorial, el cerebro, los músculos y hasta la retina ya exhiben actividad espontánea (por eso el feto se mueve). Las neuronas son células excitables: pueden reaccionar con espontaneidad, y sus potenciales de acción se autoorganizan en ondas masivas que viajan por el tejido cerebral. Incluso en el útero, ondas aleatorias de espigas neuronales fluyen a través de las retinas del feto y, al llegar a la corteza, aunque no llevan ninguna información visual en el sentido estricto del término, ayudan a organizar los mapas visuales corticales (Goodman y Shatz, 1993; Shatz, 1996). Con esto, la plasticidad sináptica actúa inicialmente sin necesitar ninguna interacción con el mundo exterior. Recién durante el tercer trimestre de gestación, la frontera entre lo innato y lo adquirido se borra de a poco, a la par que el cerebro, ya bien formado, comienza a acomodarse al mundo interior y exterior.

Incluso en el momento del nacimiento, la corteza continúa transitada por descargas aleatorias que se desarrollan sin relación con los *inputs* sensoriales. Poco a poco, esta actividad endógena, preorganizada, evoluciona por influencia de los órganos de los sentidos. Este proceso es interpretable con precisión dentro del marco del modelo del cerebro bayesiano o estadístico (Berkes y otros, 2011; Orbán y otros, 2016). La actividad endógena inicial representa lo que los estadísticos llaman el *a priori*, las expectativas del cerebro, sus hipótesis previas a cualquier interacción con el entorno. Más tarde, estas hipótesis se ajustan gradualmente a los datos del ambiente, de

modo que a los pocos meses de vida la actividad espontánea empieza a condecirse cada vez más con las estadísticas del mundo real. El modelo interno del cerebro se refina a medida que compila estadísticas de sus *inputs* sensoriales. Al final hay una solución de compromiso, una selección del mejor modelo interno entre aquellos ofrecidos por la organización previa.

¿Qué es un período sensible?

Acabamos de ver que la plasticidad cerebral es a la vez inmensa y limitada. Todos los grandes haces de conexiones pueden y deben modificarse a medida que vivimos, maduramos y aprendemos, pero sin excepción están ya en su lugar desde el nacimiento y son esencialmente idénticos en todos los seres humanos. Nuestros aprendizajes son resultado de pequeños ajustes, sobre todo a escala de microcircuitos, a menudo en el rango de unos pocos milímetros. A medida que las neuronas crecen y sus ramas terminales desarrollan botones sinápticos en otras neuronas, los circuitos que forman permanecen firmemente arraigados dentro de los límites de su envoltura genética. En respuesta al entorno, las vías neuronales pueden cambiar su conectividad local, su fuerza y también su mielinización, al rodearse de una capa aislante de, precisamente, mielina que acelera sus mensajes y, por lo tanto, facilita la transmisión de información de una región a otra; sin embargo, no pueden reorientarse a voluntad.

Esta restricción espacial sobre la conectividad a larga distancia se suma a una restricción temporal: en numerosas regiones del cerebro, la plasticidad es máxima solo durante un lapso de tiempo limitado, que llamamos “período sensible”. Este período comienza en la infancia temprana, alcanza cotas máximas y luego concluye, a medida que crecemos. El proceso completo insume varios años y varía según las regiones del cerebro: las áreas sensoriales alcanzan su cota máxima de plasticidad en torno a la edad de 1 o 2 años, mientras que las regiones de mayor jerarquía como la corteza prefrontal alcanzan su máximo tanto más tarde en la infancia o incluso en la adolescencia. Sin embargo, lo cierto es que, con la edad, la plasticidad disminuye y el aprendizaje, aunque no es imposible, se vuelve cada vez más difícil (Werker y Hensch, 2014).

Si a menudo afirmo que el bebé es una auténtica máquina de aprender, es porque en los primeros años de vida su cerebro es el lugar de una verdadera efervescencia de plasticidad sináptica. Las dendritas de sus neuronas piramidales se multiplican a gran velocidad. En el nacimiento, la corteza de los bebés se parece a un bosque tras el paso de un huracán, cubierto solo por algunos troncos de árboles desnudos. Los primeros seis meses de vida son una eclosión primaveral para el cerebro del recién nacido: se multiplican las conexiones y las ramificaciones hasta formar una fronda inextricable (figura 16; Conel, 1939; Courchesne y otros, 2007).

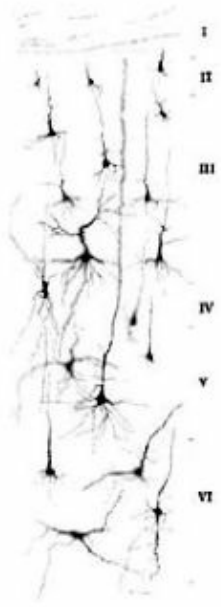
El incremento de la complejidad de los árboles neuronales podría hacernos pensar que el entorno deja su huella en el cerebro y lo obliga a crecer a la par de su almacenamiento de datos. La realidad es mucho más compleja: en el cerebro inmaduro, no hay una relación de proporcionalidad directa entre la cantidad de aprendizaje que se produce y las sinapsis que surgen, sino que estas siempre se generan en demasía, para luego, de acuerdo con su utilidad para el organismo, ser conservadas o eliminadas gracias a la acción del ambiente. Durante la primera infancia, la densidad de sinapsis aumenta hasta alcanzar el doble de la de un adulto, y solo entonces disminuye con lentitud. En cada región de la corteza, a incesantes ondas de sobreproducción les sigue la retracción selectiva de las sinapsis inútiles o, por el contrario, la multiplicación de las sinapsis y de las ramas dendríticas y axonales que se comprobaron útiles. Piensen en esto cuando miren a un niño o una niña: cada segundo que pasa, en su cerebro se están creando y eliminando muchos millones de sinapsis. Esta efervescencia explica, en gran parte, la existencia de períodos sensibles. En la primera infancia, el conjunto de la base dendrítica y sináptica todavía resulta muy maleable; a mayor maduración del cerebro, más se confina el aprendizaje a modificaciones marginales.

Con todo, es notable que las oleadas de sobreproducción sináptica no se producen en todas las regiones en el mismo momento (Rakic y otros, 1986). La corteza visual primaria, como otras regiones sensoriales, madura tanto más rápido que las áreas corticales de nivel más alto.

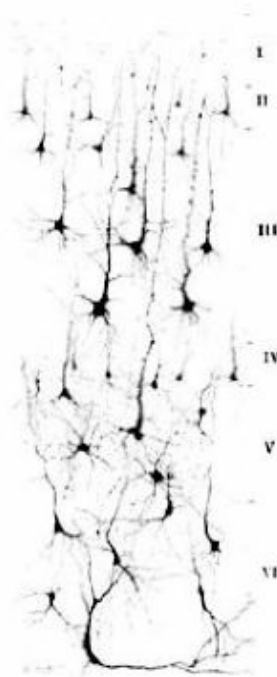
El principio de organización parece ser estabilizar enseguida aquello que constituye el *input* directo. Por eso, congela la organización cortical en las áreas sensoriales tempranas, mientras que deja las áreas de alto nivel abiertas para cambiar durante un lapso mucho más extenso. Así, las regiones más elevadas en la jerarquía, como la corteza prefrontal, son las últimas en estabilizarse: continúan modificándose durante la adolescencia e incluso

después de ella. En la especie humana, el máximo de sobreproducción sináptica se termina cerca de los 2 años de edad para la corteza visual, de los 3 o 4 años para la corteza auditiva y entre los 5 y los 10 años para la corteza prefrontal (Huttenlocher y Dabholkar, 1997). La mielinización sigue la misma progresión (Dubois y otros, 2007, 2015; Flechsig, 1876).

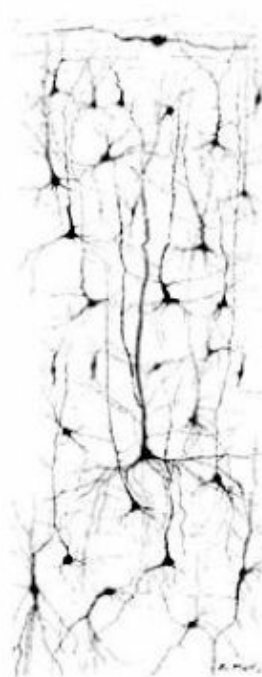
Recién nacido



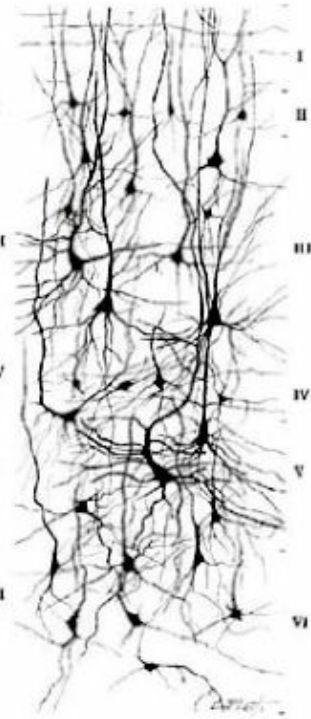
1 mes



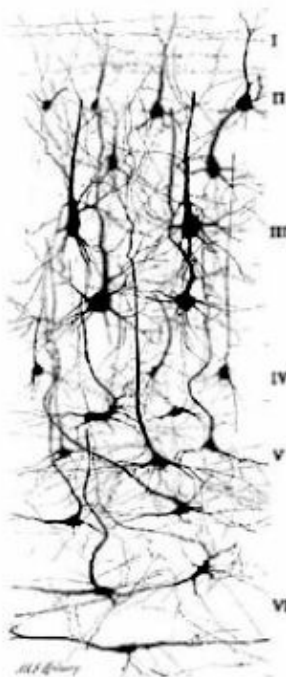
3 meses



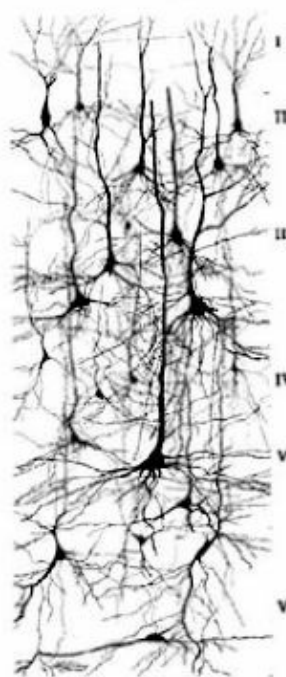
6 meses



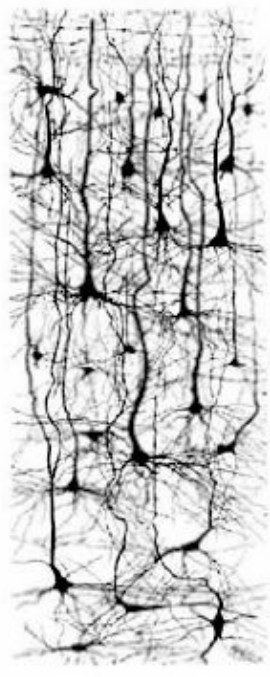
1 año



2 años



4 años



6 años

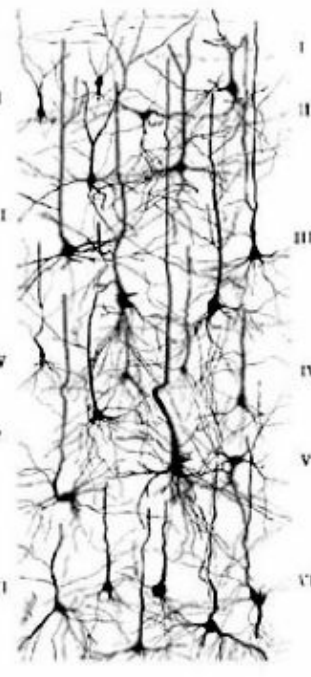


Figura 16. Durante los primeros años de vida, las arborescencias de las neuronas crecen y ganan exuberancia hasta formar un embrollo inextricable. En el cerebro de un niño de 2 años, la cantidad de sinapsis es casi el doble que en un adulto. Pero los árboles dendríticos no solo crecen, sino que también son podados: por influencia de la actividad neuronal, las sinapsis útiles se conservan y se multiplican, mientras que las otras son suprimidas.

En los primeros meses de vida, las áreas sensoriales son las principales beneficiarias de esto: su procesamiento de la información se acelera y la velocidad de transmisión de una información desde la retina hasta las áreas visuales pasa de un cuarto de segundo a una décima de segundo en algunas semanas (Adibpour, Dubois y Dehaene-Lambertz, 2018; Dehaene-Lambertz y Spelke, 2015). Esta aislación alcanza los haces que conectan la corteza frontal (sede del pensamiento abstracto, la atención y la planificación) con mucha más lentitud. Durante años, el niño es un ser híbrido: sus circuitos sensoriales y motores se encuentran bastante maduros, mientras que sus áreas de mayor nivel continúan operando de manera mucho más lenta, con una dotación de circuitos aún desmielinizados. Como resultado, durante el primer año de vida, necesita hasta tres o cuatro veces más tiempo que un adulto para tomar conciencia de una información elemental, como la presencia de un rostro (Kouider y otros, 2013).

Estas sucesivas oleadas de sobreproducción sináptica y de mielinización modulan la plasticidad cerebral. Así, una serie de períodos sensibles se abren y se cierran en momentos diferentes de acuerdo con las distintas regiones cerebrales que deben desarrollarse. Las áreas visuales y auditivas están entre las más propensas a perder su capacidad de aprender. El ejemplo mejor estudiado, tanto en la especie humana como en animales, es el de la visión binocular (Epelbaum y otros, 1993; Fawcett, Wang y Birch, 2005; Hensch, 2005). Para percibir la tercera dimensión, el sistema visual fusiona la información proveniente de los dos ojos. Esta “fusión binocular” exige que ambos envíen a la corteza *inputs* de alta calidad durante un período bastante breve: algunos meses en el gato, algunos años en el hombre. Si, durante este período, un ojo permanece cerrado, envía una imagen borrosa o se sufre una difracción por culpa de un estrabismo pronunciado, el circuito cortical de la fusión de los dos ojos no se pone en marcha, y esta pérdida es permanente: el estrabismo del niño debe ser corregido en los primeros años de vida, idealmente antes de los 3 años, o esto implicará un déficit definitivo a escala de la corteza.

Otro período sensible permite que dominemos los sonidos de nuestra lengua

materna. Los bebés son campeones del aprendizaje de las lenguas: cuando nacen, son capaces de escuchar todos los fonemas de todas las lenguas, y donde sea que nazcan, cualquiera sea su carga genética, es suficiente su inmersión en un baño de lenguaje (sea de una, dos o tres lenguas) para que en algunos años se conviertan en especialistas de la fonología de su o sus idiomas maternos. Como adultos, ya perdimos esta notable capacidad de aprendizaje: como vimos antes, un japonés puede pasar su vida entera en Francia sin llegar jamás a distinguir los sonidos /R/ y /L/. Es posible que toda su vida confunda “glacial” y “glaciar”, “promo” y “plomo”, “elección” y “erección”... Queridos lectores, no experimenten ningún sentimiento de superioridad porque, nacidos y criados como hablantes de castellano, ustedes nunca podrán distinguir ni las /T/ dentales y retroflejas que cualquier hablante del hindi oye como algo evidente, ni las vocales breves y largas del finlandés o el japonés, ni los cuatro o cinco tonos del chino mandarín.

La investigación demuestra que perdemos estas capacidades cerca del final del primer año de vida (Dehaene-Lambertz y Spelke, 2015; Maye, Werker y Gerken, 2002; Peña, Werker y Dehaene-Lambertz, 2012; Werker y Tees, 1984). Cuando somos bebés, compilamos de forma inconsciente estadísticas sobre lo que escuchamos, y nuestro cerebro se ajusta a la distribución de los fonemas que se emplean en el entorno. Alrededor de los 12 meses, algo se solidifica en el cerebro y perdemos esta capacidad de aprender. Salvo en casos extraordinarios, ya jamás podremos hacernos pasar por hablantes nativos de japonés, finlandés o hindi: nuestra fonología está (casi) cincelada en piedra. Hace falta un esfuerzo inmenso para que un adulto recupere la capacidad de discriminar sonidos de una lengua extranjera. Solo a fuerza de una reeducación intensa y focalizada, que comienza, por ejemplo, por amplificar las diferencias entre /R/ y /L/ para volverlas audibles, y luego las va reduciendo, un adulto japonés logra recuperar parcialmente la discriminación de estos dos sonidos (McCandliss y otros, 2002).

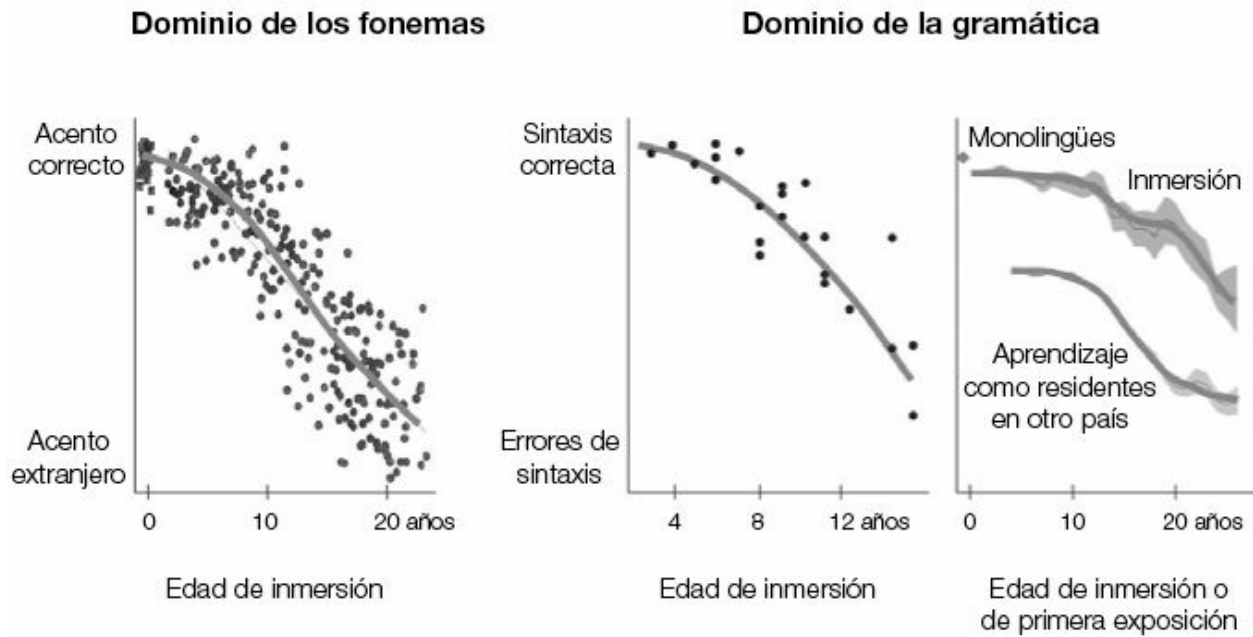
Esto motiva que hablemos de un período sensible en lugar de un período crítico: la capacidad de aprendizaje se reduce, pero nunca llega a cero. En la edad adulta, la capacidad residual de adquirir fonemas extranjeros registra variaciones significativas entre las personas. Para la mayoría de nosotros, intentar hablar correctamente un idioma extranjero en la edad adulta supone un esfuerzo insondable. Sin embargo, algunas personas mantienen la capacidad de aprender la fonología de lenguas extranjeras, y esta competencia se puede predecir parcialmente por el tamaño, la forma y el

número de conexiones de su corteza auditiva (Golestani y otros, 2007). Estos afortunados cerebros en apariencia estabilizaron un conjunto de conexiones más flexible, pero queda claro que son la excepción más que la regla.

El dominio de la fonología de una lengua extranjera es una de las primeras competencias en decaer con la edad: al respecto, desde los años más tiernos, un niño ya es mucho menos capaz que un bebé de unos pocos meses. La capacidad de aprender la gramática de una lengua extranjera permanece disponible un poco más de tiempo, pero de todas maneras termina por venirse abajo en las cercanías de la pubertad. Lo sabemos por el estudio de los migrantes: un niño o una niña que llega a un país extranjero puede descollar en su nueva lengua, pero siempre presenta un leve acento extranjero y ocasionales errores sintácticos que delatan su origen, y esta diferencia se acrecienta muchísimo en quienes migran durante la adolescencia o la adultez (figura 17; Flege, Munro y MacKay, 1995; Hartshorne, Tenenbaum y Pinker, 2018; Johnson y Newport, 1989; Weber-Fox y Neville, 1996).

Un estudio reciente recopiló datos de millones de estudiantes de una segunda lengua en internet y los utilizó para modelar la curva promedio de aprendizaje de idiomas. Los resultados sugieren que disminuye con lentitud durante la infancia, pero mengua bruscamente alrededor de los 17 años (Hartshorne y otros, 2018). Debido a que lleva tiempo aprender, los investigadores recomiendan comenzar antes de los 10 años. Además, enfatizan el valor adicional de una estadía en el país de interés, porque nada supera la interacción social: el éxito es tanto mayor si necesitamos hablar un idioma extranjero para conseguir comida o para subir a un tren que si simplemente lo aprendemos en el aula o gracias a series de televisión. Renueva su validez la máxima: cuanto más temprano, mejor; la plasticidad cerebral para el aprendizaje de gramática parece sufrir una reducción drástica al final de la pubertad (aunque esta caída no puede atribuirse exclusivamente a una pérdida de plasticidad cerebral; es probable que desempeñen un papel importante otros factores relacionados con la motivación y la socialización).

Dominio final de una segunda lengua



Conservación de la primera lengua en los niños adoptados

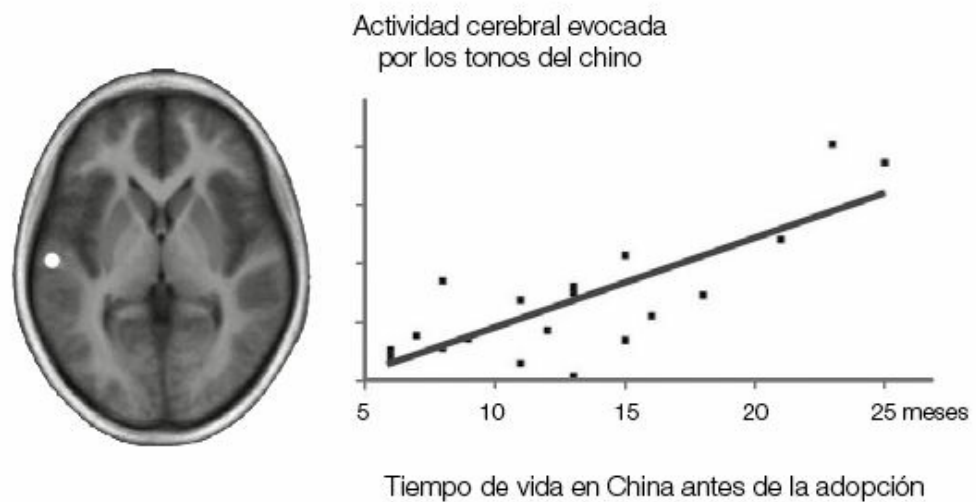


Figura 17. El aprendizaje de una segunda lengua o de lengua extranjera revela la

disminución de la plasticidad cerebral con la edad. Cuanto más tarde se aprende una lengua, menos posible es llegar a practicarla correctamente, sin acento extranjero o errores gramaticales (arriba). A la inversa, en las adopciones internacionales se observa con nitidez que cuanto más tiempo pasa un niño en su país de origen antes de ser adoptado, más conserva en su cerebro una huella de su primera lengua (abajo), que permanece parcialmente impresa en circuitos neuronales inconscientes.

De todos modos, se trata de la segunda lengua. Si el aprendizaje decae con relativa lentitud en el término de una decena de años, sin duda es porque se apoya, al menos en parte, sobre un cerebro ya modelado por la primera lengua. ¿Qué ocurriría si un niño estuviera privado de cualquier exposición al lenguaje durante los primeros años de su vida? La leyenda cuenta que la pregunta fue planteada por primera vez por el faraón Psamético I. Para responderla, le habría confiado dos niños a un pastor, con la estricta prohibición de hablar con ellos; pese a todo, parece que los niños hablaron... ¡en frigio! De acuerdo con la leyenda, este “experimento” habría sido repetido por el emperador Federico II Hohenstaufen en el siglo XIII; por Jacobo IV, rey de Escocia, en el siglo XV, e incluso por Yalaluddin Muhammad Akbar, el emperador mogol, en el siglo XVI, y los niños, privados del habla, ¡se habrían muerto!

Por desgracia, no es necesario difundir tales fábulas, dado que la experiencia ocurre a menudo en todos los países del mundo: cada día nacen niños sordos y, si no los ayudamos, permanecen prisioneros dentro de su burbuja de silencio. Hoy en día sabemos que desde el primer año de vida es indispensable darles una lengua: lo más natural es alguna de las lenguas de señas (los niños que las hablan se desarrollan de forma completamente normal), o bien una lengua hablada, cuando estos niños pueden beneficiarse de un implante coclear que restaure en parte su audición. También en esos casos, la investigación muestra que es necesario reaccionar muy rápido (Friedmann y Rusou, 2015): los niños implantados luego de los 8 meses de edad, presentan déficits permanentes en el dominio de la sintaxis y nunca logran comprender con fluidez frases en las cuales los elementos estén desplazados, fenómeno conocido como “movimiento sintáctico”. En la frase “¿Me mostrarías la niña a la que peina la abuela?”, no es evidente comprender que el sintagma nominal “la niña”, aunque esté antes del verbo “peinar”, no es su sujeto, sino su objeto. Los niños sordos, cuando son implantados de forma muy tardía, son incapaces de comprender frases como

esta y de elegir entre una imagen en que la abuela peina a la niña y otra en la que la niña peina a la abuela.

Así, la infancia temprana es una etapa clave para el desarrollo del movimiento sintáctico: en ausencia de interacciones lingüísticas, la plasticidad cerebral para este aspecto de la sintaxis se cierra cerca del final del primer año de vida. Recuerden a los niños de Israel: en 2003, unas semanas de privación de tiamina en sus primeros meses de vida fueron suficientes para hacerles perder definitivamente el sentido de la sintaxis. Estos resultados convergen con otros estudios llevados a cabo en niños salvajes, como el célebre Víctor de Aveyron, o maltratados, como la pequeña niña estadounidense llamada Genie y criada (o, más bien, degradada) en un armario durante más de trece años, casi sin que se le hablara. Al volver al mundo luego de tantos años, Víctor y Genie comenzaron a hablar y a adquirir vocabulario, pero años más tarde su gramática todavía era rudimentaria.

La adquisición del lenguaje provee entonces un excelente ejemplo de períodos sensibles, tanto para la fonología como para la gramática. También es una buena explicación de la modularidad del cerebro: mientras la gramática y los sonidos del lenguaje se fijan, otras funciones, como la capacidad de aprender palabras nuevas y su significado, permanecen abiertas durante toda la vida. Esta plasticidad residual es precisamente la que nos permite aprender, a cualquier edad, qué es un fax, un iPad o un meme, qué significa el verbo “googlear” o incluso neologismos humorísticos como el adjetivo “deprimencial”, creado por Julio Cortázar, o “satisfahaciente”, en la más epigramática Alejandra Pizarnik. Afortunadamente, en relación con la adquisición de vocabulario, nuestro cerebro adulto continúa gozando de un nivel de plasticidad similar al de los niños, aunque se desconoce aún la razón biológica por la cual estos circuitos léxicos no están expuestos a un período sensible.

Una sinapsis debe estar abierta o cerrada

¿Por qué se cierra la plasticidad sináptica? ¿Qué mecanismos biológicos la bloquean? El origen de la apertura y el cierre de los períodos sensibles es un gran tema de investigación de las neurociencias contemporáneas (Caroni,

Donato y Muller, 2012; Friedmann y Rusou, 2015; Werker y Hensch, 2014). El cierre del período sensible parece estar ligado al equilibrio entre la excitación y la inhibición. En el niño, las neuronas excitatorias son eficaces en muy poco tiempo, mientras que las neuronas inhibitorias se desarrollan de manera más gradual. Algunas, aquellas que contienen la proteína parvalbúmina, terminan por rodearse de una matriz dura, una suerte de red rígida que impide que las sinapsis se muevan y crezcan. Si se liberara a las neuronas de este lastre, por ejemplo, mediante la aplicación de un agente farmacológico como la fluoxetina (más conocida con el nombre comercial Prozac), la plasticidad sináptica podría regresar. Esta posibilidad es una inmensa fuente de optimismo para el tratamiento de los accidentes cerebrovasculares, casos en que haría falta que el paciente volviera a aprender a utilizar regiones preservadas.

También entran en juego otros factores. Existe, por ejemplo, una proteína llamada Lynx 1: cuando está presente en una neurona, inhibe los efectos masivos de la acetilcolina, que en condiciones normales señala los momentos de interés y modula masivamente la plasticidad. Cuando hay circuitos adultos invadidos por Lynx 1, se puede intentar restaurar la plasticidad, por ejemplo, aplicando una corriente que despolariza las neuronas y las acerca a su umbral de activación (Krause y otros, 2017). Esta incipiente terapia aporta también una esperanza para quienes padecen una depresión profunda, casos en los cuales en ocasiones es suficiente la aplicación de una pequeña corriente eléctrica transcraneal para encaminarlos hacia el alivio.

Podríamos preguntarnos por qué el sistema nervioso se obstina en bloquear de este modo su propia plasticidad. Sin duda, en determinado momento comienza a ser ventajoso dejar de modificar la estructura del cerebro. La simulación de las redes neuronales deja en evidencia que, con mucha rapidez, las neuronas de nivel bajo adquieren campos receptores simples y reproducibles, como detectores de contornos en la corteza visual. No existe interés en continuar actualizándolas, visto que este tipo de detector ya es casi óptimo: se ahorra así el costo energético ligado al crecimiento de botones sinápticos y axonales. Además, cambiar la organización de las áreas sensoriales primarias, la base sobre la cual descansa la visión, conlleva el riesgo de que el andamiaje entero de las áreas de nivel superior tambalee y pierda sus aprendizajes. Entonces, después de determinado momento, vale más dejarlo intocado; no cabe duda de que por eso la evolución construyó

diversos mecanismos de cierre del período sensible y los desencadenó antes en las áreas sensoriales que en las regiones corticales de nivel más alto.

La faceta positiva es que, como los circuitos se fijan, guardamos durante toda nuestra vida una huella sináptica inconsciente de algunos aprendizajes precoces. Incluso si estos se convierten en obsoletos, por ejemplo, porque volvemos a aprender algo nuevamente, los circuitos cerebrales conservan una huella latente de nuestros comienzos. Un ejemplo notable es el caso de los niños adoptados que aprenden una segunda lengua materna. Expuesto a la problemática de numerosos niños que viven en las calles, son víctimas de la guerra, de la separación o de la miseria, Corea es uno de los países que recurrió masivamente a la adopción internacional. Desde 1958, y durante un período de cuarenta años, cerca de 180.000 niños y niñas fueron adoptados. De estos, 130.000 fueron adoptados fuera de su país y más de 10.000 llegaron a Francia. Christophe Pallier y yo hicimos una evaluación de veinte de ellos durante su edad adulta. Habían llegado a Francia entre los 5 y 9 años, y casi no tenían recuerdos conscientes de su país natal, excepto por algunos olores. En una primera aproximación, su cerebro se comportaba como el de un niño nacido en Francia (Pallier y otros, 2003):[\[15\]](#) las áreas del lenguaje, en el hemisferio izquierdo, respondían con fuerza a las frases francesas, y ya no a las frases coreanas (en cualquiera de los casos, no más que a otro idioma desconocido, como el japonés). En los niveles léxico y sintáctico, entonces, la nueva lengua había suplantado a la antigua.

Y sin embargo... los trabajos de otro equipo de investigadores demostraron que un grupo de niños y niñas chinos adoptados en Canadá conservaban, en lo más profundo de su corteza, una huella latente de los tonos de su lengua de origen (Pierce y otros, 2014). Mientras un canadiense de nacimiento los procesa en el hemisferio derecho como una vaga melodía, ellos los tratan como sonidos del lenguaje dentro de una región fonológica del hemisferio izquierdo llamada *planum* temporal (figura 17). Este circuito se graba con la primera lengua desde el primer año de vida y, como resulta evidente, más adelante no se revierte.

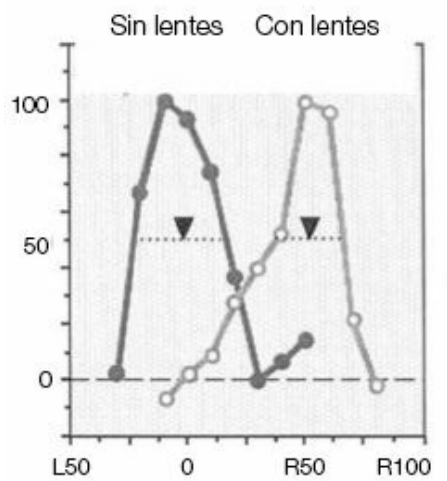
No es el único ejemplo posible. Como ya expliqué, el estrabismo infantil, si no se lo corrige, modifica definitivamente los circuitos visuales. El etólogo y neurofisiólogo Eric Knudsen estudió un modelo animal de este período sensible. Crio lechuzas haciéndoles llevar lentes prismáticos que desplazaban el conjunto del campo visual 20° hacia la derecha. Con estas lechuzas equipadas de vistosos y coquetos anteojos, realizó los estudios más sagaces

de los mecanismos neuronales del período sensible (Knudsen y Knudsen, 1990; Knudsen, Zheng y DeBello, 2000). Solo las lechuzas que habían llevado lentes durante la infancia lograban ajustarse a este estímulo sensorial inusual: sus respuestas auditivas se desplazaban para alinearse sobre la retina de modo de poder cazar respaldadas por el oído y, en simultáneo, la visión nocturna. Las lechuzas mayores, aunque llevaran los lentes durante semanas, no se acomodaban tan bien, o, incluso, no lo hacían en modo alguno. Pero, por encima de todo, los animales entrenados durante la infancia conservaban, para el resto de sus vidas, una huella neuronal permanente de su experiencia precoz. Antes del aprendizaje, se observaba un doble circuito: algunos axones de las neuronas auditivas, en el colículo inferior, habían conservado su posición normal, mientras que otras se habían reorientado en busca de alinearse en el mapa visual (figura 18). Cuando se le quitaba el lente, el animal reaprendía a orientarse correctamente; pero no bien se devolvían las lentes a su posición, se desplazaba 20°: con la misma efectividad de un hablante bilingüe las lechuzas lograban hacer el pasaje de una modalidad a la otra. Su cerebro había registrado, de una vez y para siempre, dos juegos de parámetros que le permitían cambiar de configuración sin sangre, sudor ni lágrimas, exactamente como los niños chinos adoptados en Canadá conservan una huella cerebral de la sonoridad de su lengua de origen.

También en nuestra especie los aprendizajes precoces dejan una huella permanente, ya sea que se trate de la práctica del piano, de la visión binocular, o incluso del léxico mental. Como adultos, siempre reconocemos con mayor velocidad las palabras que escuchamos desde la infancia, como “babero”, “papá”, “pañal”, todas aquellas que la plasticidad sináptica precoz ha grabado definitivamente en nuestra memoria (Ellis, Ralph y Matthew, 2000; Gerhand y Barry, 1999; Morrison y Ellis, 1995). La corteza juvenil registra casi todo sin esfuerzo, y resguarda este saber en la memoria permanente de la geometría de sus axones y de sus dendritas.



Respuesta de las neuronas



Retraso interaural, en microsegundos

Anatomía de las conexiones

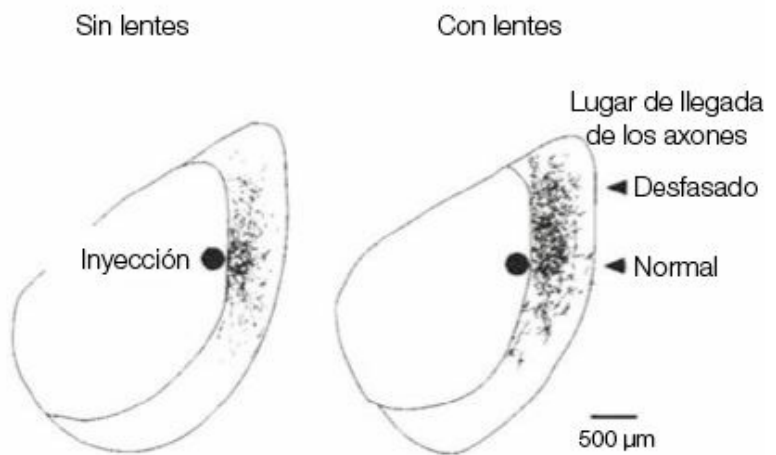


Figura 18. Las experiencias tempranas pueden imprimirse profundamente en los circuitos neuronales. Una lechuza logra acostumbrarse a llevar puestos lentes prismáticos que desfasan

la visión, siempre que esta experiencia tenga lugar durante la infancia. Sus neuronas auditivas, que localizan los objetos gracias al minúsculo desfase entre los sonidos que llegan a los oídos derecho e izquierdo, se ajustan con el objetivo de confluir con la visión. Esto se traduce en un desplazamiento de los axones en el rango de una décima de milímetro. Noten que el circuito preserva las dos opciones: desfasado o normal.

Milagro en Bucarest

La intensa plasticidad cerebral durante los primeros años de vida les da la razón a todos los progresistas: invertir en la primera infancia, y sobre todo en educación primaria e inicial, es una prioridad, porque durante este momento de gracia el cerebro del niño se transforma con mayor facilidad y de manera más completa. Más tarde, con el cierre progresivo de los períodos sensibles, el aprendizaje se volverá más complicado; pero no olvidemos que también gracias a esta cristalización el cerebro conservará, a lo largo de su vida, la huella de las modificaciones sinápticas que se realizaron durante la infancia.

Por fortuna, el cierre de los períodos sensibles no es implacable como la caída de una guillotina: la plasticidad no desaparece de un día para el otro. Además, depende mucho de las áreas cerebrales. En las áreas sensoriales, la plasticidad se cierra luego de unos pocos años, pero en otras como la fonología y la sintaxis de una segunda lengua se extiende por lo menos hasta la pubertad. En algunas áreas –por ejemplo, el vocabulario o los conocimientos semánticos–, la capacidad de aprender parece no disminuir nunca.

Si bien hace falta apurarse para aprender, sería falso por completo llegar a la conclusión de que todo se juega antes de los 3 años. El cerebro conserva su resiliencia durante mucho tiempo. Pasado el bendito período de la primera infancia, la plasticidad disminuye sin por eso desaparecer. Se atenúa con el transcurso del tiempo, inicialmente en las áreas sensoriales periféricas, pero en regiones como las áreas corticales de mayor jerarquía conserva toda su vida el potencial de adaptación. Por eso, una intervención pedagógica a veces hace milagros, sobre todo cuando obra de manera rápida e intensiva. Quizá no restaurará todas las sutilezas del movimiento sintáctico o de la percepción de los tonos del chino, pero logrará transformar al niño en riesgo en un adulto pleno y responsable.

Los huérfanos de Bucarest son un ejemplo desgarrador de la notable

capacidad de recuperación del cerebro en desarrollo. En diciembre de 1989, Rumania se rebeló repentinamente contra el régimen comunista. En menos de una semana, el pueblo sublevado echó al dictador Nicolae Ceausescu y a su esposa, que fueron fusilados el día de Navidad. Desde el exterior, alarmados, descubrimos las espantosas condiciones de vida de los habitantes de esta pequeña porción del mundo.



Habilidades sociales

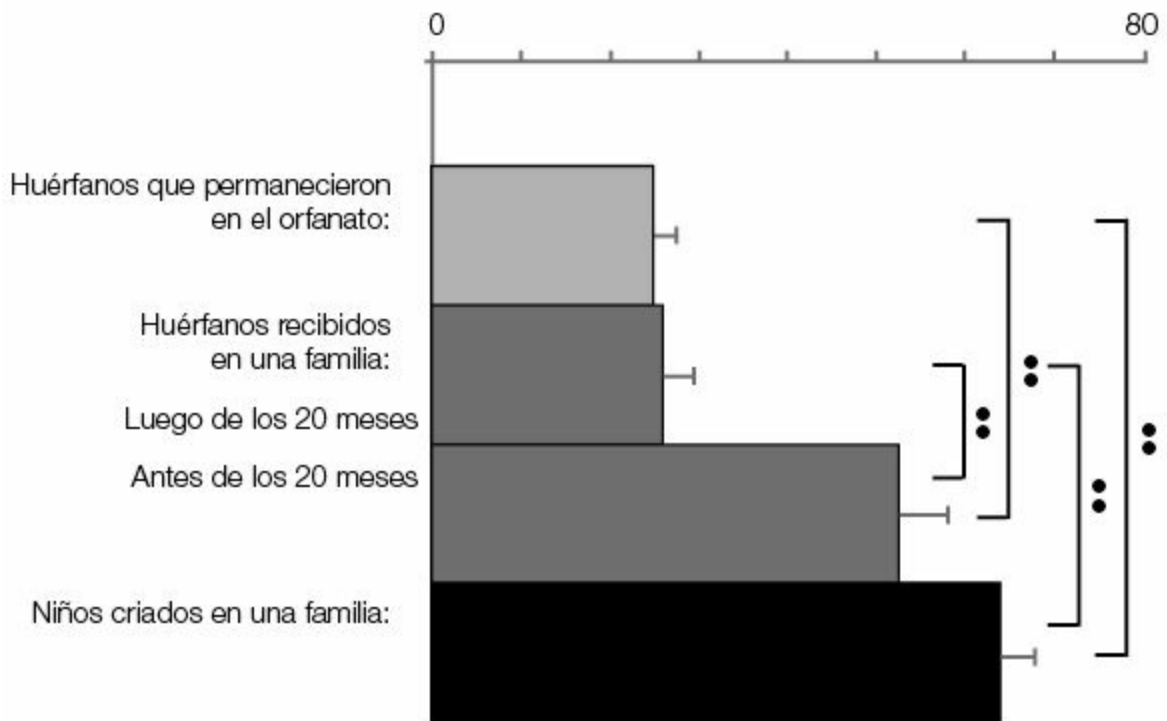


Figura 19. Los traumas de la infancia dejan marcas en el cerebro, pero una intervención precoz puede minimizar esas huellas. En los orfanatos de Rumania bajo el régimen del dictador Ceausescu, los niños eran maltratados y privados de interacción con los adultos. A los 8 años, estos huérfanos presentaban enormes déficits en las habilidades sociales, incluso si habían sido recibidos por familias adoptivas desde sus 20 meses. Sin embargo, los que habían sido adoptados antes de esa edad lograban reponerse y mostraban avances considerables.

Una de las imágenes más insoportables fue la de estos niños pequeños, con cuerpos demacrados y ojos mortecinos, abandonados en cerca de seiscientos orfanatos, verdaderos asilos donde se hacinaban alrededor de 15.000 niños y niñas, casi por completo librados a su (mala) suerte (figura 19). Convencido de que la fuerza de un país reside en su juventud, el régimen de Ceausescu había promulgado una delirante política natalista: impuestos para todos los solteros y las parejas sin hijos, prohibición de la anticoncepción y del aborto, pena de muerte para los abortistas... todo esto en busca de nacimientos por millares. Las parejas que no podían asegurar la subsistencia de sus hijos debían confiarlos a los entes estatales, lo que explica la existencia de estas centenas de orfanatos que no alcanzaban a asegurar la higiene, la alimentación, la calefacción ni el mínimo de calor humano y de estimulación cognitiva indispensables para el pleno desarrollo de los niños. Esta política desastrosa produjo cientos de niños desatendidos que presentaban enormes déficits cognitivos y emocionales de todo tipo.

Luego de la apertura del país, numerosas ONG se interesaron por esta catástrofe. En ese contexto nació un proyecto de investigación muy particular, el Bucharest Early Intervention Project (Almas y otros, 2012; Berens y Nelson, 2015; Nelson y otros, 2007; Sheridan y otros, 2012; Windsor y otros, 2013). Bajo la égida del secretariado de Estado rumano para la protección de la infancia, un investigador de Harvard, Charles Nelson, decidió hacer un relevamiento científico de las consecuencias de la vida en estos orfanatos, y de la posibilidad de salvar a estos niños ubicándolos en familias adoptivas. Como en Rumania no había programas de adopción dignos de ese nombre, creó su propio sistema y logró involucrar a 56 familias voluntarias, interesadas en recibir a uno o dos huérfanos en su hogar. Pero eso equivalía a una gota de agua frente a los sombríos abismos de los orfanatos rumanos: solo 68 niños pudieron beneficiarse del programa. Un artículo de *Science* describe en detalle el momento dramático en que se reunió a 136 niños, identificados del 1 al 136, y se extrajeron números de un gran sombrero: 68 de ellos permanecerían en el orfanato, mientras que la otra

mitad finalmente sabría qué era formar parte de una familia. Esta manera de proceder puede parecer chocante, pero ¿cómo hacerlo mejor? Dado que los recursos humanos eran limitados, el sorteo era sin duda la solución más equitativa. Por cierto, el equipo continuó recaudando fondos para sacar cada vez más niños de su miseria, y un segundo artículo de *Science* juzgó irreprochables las condiciones éticas del estudio inicial (Millum y Emanuel, 2007).

El sorteo permitió plantear una pregunta rigurosa: frente a las mismas condiciones de origen, ¿la ubicación precoz en una familia adoptiva permitió a los niños restablecerse? La respuesta fue positiva, pero con gran dependencia de la edad: solo los niños que fueron ubicados en una familia antes de los 20 meses de edad mejoraron mucho más que aquellos que permanecieron en el orfanato.

Decenas de estudios documentaron los efectos dramáticos del aislamiento afectivo y social en el desarrollo cerebral, y el de Bucarest no fue la excepción: en comparación con los niños nacidos en una familia promedio, todos los huérfanos presentaban graves déficits de las funciones cognitivas e incluso de los factores fundamentales del funcionamiento cerebral, como el metabolismo de la glucosa o el volumen total de materia gris. Luego de la ubicación en una familia, sin embargo, parte de estos índices evidenció grandes progresos. Seis años más tarde, hacia los 8 años de edad, los niños y las niñas que habían sido adoptados antes de alcanzar los 20 meses claramente habían progresado en comparación con el grupo control y no presentaban diferencias con los niños que se habían criado en una familia desde el nacimiento. La intensidad de las ondas alfa en sus electroencefalogramas, que es un marcador de la atención, había recuperado su nivel normal, y lo mismo había sucedido con las habilidades sociales (figura 19). El vocabulario también había progresado considerablemente.

Esos avances espectaculares no deben hacer olvidar que persistía un déficit de materia gris ni que los niños adoptados luego de los 20 meses de edad continuaban estancados en varios problemas. Nada reemplaza verdaderamente veinte meses de amor perdido, y estos niños llevarán siempre, en su cerebro, los estigmas de las graves privaciones que sufrieron. Pero el estudio de los huérfanos de Bucarest, como el de los niños adoptados de Corea, destaca hasta qué punto la plasticidad del cerebro facilita la resiliencia y la reversibilidad de los traumatismos precoces: mientras sean

tratados, y cuanto antes mejor, numerosos déficits están lejos de ser irreversibles.

[15] Se obtuvieron resultados similares en el campo del reconocimiento de rostros: adoptados antes de los 9 años, los niños coreanos reaprenden el prototipo del rostro occidental y pierden la habitual preferencia por los rostros de su propia raza (Sangrigoli y otros, 2005).

6. Reciclen su cerebro

Resumamos lo visto hasta aquí. Todos los bebés llegan al mundo con un rico bagaje, un conjunto de hipótesis universales. Al nacer, sus circuitos cerebrales están bien organizados y esto les confiere intuiciones fuertes en todo tipo de campos: objetos, personas, tiempo, espacio, números... Sus habilidades estadísticas son notables: ya actúan como científicos en ciernes y su sofisticada capacidad de aprender permite que paulatinamente converjan en los modelos más apropiados del mundo.

Al nacer, todos los grandes haces de fibras del cerebro ya están en su lugar. Sin embargo, la plasticidad cerebral puede reorganizar sus conexiones terminales. Millones de sinapsis se modifican cada vez que adquirimos nuevos conocimientos. Enriquecer el entorno de los niños, por ejemplo enviándolos a la escuela, significa modificar profundamente su cerebro y dotarlo de habilidades que mantendrán durante su vida entera. Sin embargo, esta plasticidad no es ilimitada. Está acotada en el espacio, en el rango de unos pocos milímetros, así como en el tiempo, dado que muchos circuitos comienzan a cerrarse después de unos meses o años.

En este capítulo, analizo el papel que desempeña la educación formal en el desarrollo temprano del cerebro. De por sí, la educación nos presenta una paradoja: ¿por qué el *Homo sapiens* puede tomar una tiza o un teclado y comenzar a escribir o hacer cálculos? ¿Cómo es que la especie humana puede expandir sus capacidades en direcciones novedosas, que no desempeñaron papel alguno en su evolución genética? Que los monos logren aprender a leer o calcular nunca debe dejar de sorprendernos. Lo expresó muy bien un personaje de Vladimir Nabokov:

Estamos absurdamente acostumbrados al milagro de unos pocos signos escritos capaces de contener una imagería inmortal, evoluciones del pensamiento, nuevos mundos con personas vivientes que hablan, lloran, se ríen. [...] ¿Y si un día nos

despertáramos, todos nosotros, y descubriéramos que somos absolutamente incapaces de leer?

Durante mucho tiempo, en zonas urbanas o rurales de Portugal y Brasil y hasta en la Amazonía, estudié el cerebro de adultos analfabetos, personas que no tuvieron la posibilidad de seguir estudios formales, porque su familia no podía solventarlo o tan solo porque no había ninguna escuela cerca. Las investigaciones demuestran que sus habilidades son, en algunos aspectos, profundamente diferentes (Kolinsky y otros, 1987, 2011; Szwed y otros, 2012): las letras escapan a su conocimiento, pero también muestran dificultades para reconocer formas, para distinguir entre dos imágenes en espejo (Kolinsky y otros, 1987, 2011; Pegado, Nakamura y otros, 2014), para prestar atención a una parte de un rostro (Ventura, 2013) o para memorizar y diferenciar las palabras habladas (Castro-Caldas y otros, 1998; Morais, 2017; Morais y otros, 1986; Morais y Kolinsky, 2005). Con cierta inocencia, Platón creía que el aprendizaje de la lectura, al permitirnos descansar en una memoria externa, la del objeto que le sirve de soporte, iba a arruinar nuestra memoria interna. No podría existir algo menos cierto que esa afirmación. El mito del bardo o del juglar que, pese a ser iletrado, poseería sin esfuerzo una inmensa memoria, es tan solo eso: un mito. Todos deberíamos ejercitar nuestra memoria, y concurrir a la escuela y aprender a leer ayuda enormemente.

El impacto de la educación es todavía más sorprendente en el ámbito de las matemáticas (Dehaene y otros, 2006, 2008; Piazza y otros, 2013; Pica y otros, 2004). Un aborigen de la Amazonía que jamás fue a la escuela ni aprendió a contar dispone, por cierto, de un conjunto universal de intuiciones aritméticas y geométricas, el mismo con el que cuentan los niños muy pequeños: distingue las formas geométricas básicas, comprende la organización del espacio, sabe navegar en línea recta, percibe la diferencia entre cantidades como 2 y 4 (o 30 y 50), y sabe que podemos ordenarlas de izquierda a derecha. Todas estas competencias innatas las heredamos de nuestra evolución y las compartimos con otros animales tan diversos como el cuervo, el mono macaco o el pollito recién nacido. Sin embargo, la educación nos permite superar ampliamente estas habilidades iniciales. Gracias a ella comprendemos, por ejemplo, que existe la misma distancia entre dos números sucesivos. Sin educación, nos habríamos quedado con la idea inocente, infantil, de que 1 es mucho más diferente de 2 de lo que 8 lo es de

9. La práctica de contar y la aritmética exacta, al mostrarnos que a cada número n le sigue un $n + 1$, nos hacen comprender que todos esos números son equidistantes y forman una escala lineal, la recta numérica; por el contrario, los animales, los niños muy pequeños y los analfabetos consideran que esta recta se comprime, porque los números grandes están más cerca unos de otros (Dehaene, 2003; Dehaene y otros, 2008; Siegler y Opfer, 2003).

Si solo nos valiéramos del sentido aproximativo del número, como los otros animales, no seríamos capaces de distinguir entre 11 y 12. La refinada precisión de nuestro sentido numérico es algo que debemos a la educación, y sobre ella descansa la construcción de las matemáticas. Sin educación, no se nos ocurriría diferenciar un número par de un número impar, para no hablar de los conceptos de número primo u objetos tan descabellados como $\sqrt{2}$, π , e , $+\infty\dots$

La hipótesis del reciclaje neuronal

¿Cómo logra la educación revolucionar nuestras representaciones mentales para convertirnos en primates lectores de Proust, Péric, Grothendieck o Einstein? Como ya vimos, todo lo que aprendemos se debe a la modificación de los circuitos cerebrales preestablecidos, en gran medida ya organizados al nacer, pero capaces de cambiar en la escala de algunos milímetros. Toda la diversidad de las culturas humanas se inscribe en ese recinto de restricciones que nuestra naturaleza impone.

Para resolver esta paradoja, propuse la hipótesis del reciclaje neuronal (Dehaene, 2005, 2014; Dehaene y Cohen, 2007). La idea es simple: pese a que la plasticidad sináptica otorga una gran maleabilidad, sobre todo en la especie humana, cuya infancia llega a durar unos quince años, nuestros circuitos cerebrales están sometidos a fuertes restricciones anatómicas, heredadas de la evolución. Por ende, cada nuevo objeto cultural que inventamos, como el alfabeto o los números arábigos, debe encontrar su “nicho” neuronal: un conjunto de circuitos cuya función originaria sea similar, como para adaptarse al nuevo invento, pero también lo bastante flexible como para que pueda reconvertirse a este uso hasta entonces inédito.

Cualquier adquisición cultural novedosa solo será posible en la medida en que la sustente una arquitectura neuronal preexistente, que esta innovación misma recicla. La educación debe hacer su trabajo respetando los límites de los circuitos neuronales, aprovechando su diversidad así como el extendido período de plasticidad que caracteriza a la especie humana.

De acuerdo con esta hipótesis, educarse entraña reciclar los circuitos cerebrales existentes. A lo largo de los milenios aprendimos a crear cosas nuevas a partir de lo viejo. Todo lo que aprendemos en la escuela reorienta un circuito neuronal preexistente en una nueva dirección. Para leer o para calcular, los niños y las niñas tienen el respaldo de estructuras que evolucionaron para otro uso, pero que, gracias a su margen de plasticidad, logran reorganizarse y especializarse para cumplir esta nueva función cultural.

Usamos esta expresión extraña, “reciclaje neuronal”, porque se trata de un concepto que une dos ideas que evocan lo que se produce en nuestro cerebro, vale decir, la reconversión para una nueva tarea y la reutilización de un material dotado de características propias:

- Para una persona, reciclarse es recibir una formación complementaria para adaptarse a la evolución de la profesión o reconvertirse a una nueva actividad. Exactamente lo mismo se produce cuando aprendemos a leer o a calcular. Cada aprendizaje cultural otorga a nuestra corteza nuevas competencias, que escapan a la esfera normal de competencias del cerebro de los primates. La educación recicla la corteza, al igual que un trabajador se recicla cuando encara un nuevo empleo.
- Para un material, reciclarse implica ingresar en un nuevo ciclo de producción, con el fin de lograr un nuevo objeto. Pero esta reutilización tiene sus límites: ¡es imposible construir una bicicleta con hojas de árbol! Cada material posee cualidades intrínsecas que lo vuelven más o menos apropiado para otros usos. Del mismo modo, desde el nacimiento cada región de la corteza –por sus propiedades moleculares, sus circuitos locales y sus conexiones de larga distancia– está dotada de características propias. El aprendizaje debe lidiar con estas restricciones materiales.

Al utilizar la expresión “reciclaje neuronal”, quise distinguir el aprendizaje rápido de una función cultural novedosa de las demás situaciones en las cuales la biología, durante un largo proceso evolutivo, elabora algo nuevo a partir de algo viejo. En efecto, en el proceso darwiniano de evolución por selección natural, la transformación de viejos materiales para propósitos innovadores es algo común: la recombinación genética moderniza órganos antiguos para construir elegantes máquinas de vanguardia. ¿Las plumas de los pájaros?: antiguos reguladores térmicos convertidos en pestañas aerodinámicas. ¿Las patas de los reptiles y de los mamíferos?: aletas antediluvianas. Como sostiene el premio Nobel de Fisiología François Jacob, en el taller de la evolución, los pulmones se vuelven un órgano de flotación, una articulación secundaria de los reptiles se separa del maxilar inferior y se transforma en huesecillos del oído interno, y hasta el gesto de contracción de la boca de un carnívoro hambriento se convierte en la delicada sonrisa de la Gioconda.

El cerebro no es una excepción. Los circuitos del lenguaje, por ejemplo, pueden haber aparecido durante la hominización mediante la duplicación y posterior reutilización de mapas corticales arcaicos (Chakraborty y Jarvis, 2015; Fukuchi-Shimogori y Grove, 2001). Pero este tipo de modificaciones lentas, de causa genética, no forman parte de mi definición de reciclaje neuronal. El término apropiado para describir las evoluciones biológicas es “exaptación”, un neologismo acuñado por Stephen Jay Gould, evolucionista de Harvard, a partir de la palabra “adaptación”. “Exaptar” consiste en dar a un mecanismo antiguo una nueva utilidad durante la evolución (así, se vuelve apto para otras tareas). Debido a que se basa sobre la propagación de genes en una población, a escala de especie, la exaptación ocurre a lo largo de decenas de miles de años. El reciclaje neuronal, por su parte, actúa en un tiempo más bien corto, en cuestión de días o, a lo sumo, pocos años. Reciclar un circuito cerebral consiste en reorientar parcialmente su función en poco tiempo, sin que esto implique una modificación genética, mediante el aprendizaje y la educación (y nada más).

Formulé la hipótesis del reciclaje neuronal con el objetivo de explicar el talento singular de la especie humana para salir de su nicho ecológico. Los humanos son únicos en sus capacidades para adquirir nuevas destrezas como leer, escribir, contar, calcular, cantar, vestirse, cabalgar o conducir un automóvil. La plasticidad cerebral extendida, combinada con nuevos

algoritmos de aprendizaje simbólico, nos dio una notable facultad de adaptación, y nuestras sociedades han descubierto medios para ampliar aún más esas habilidades al administrar a los niños, día tras día, el poderoso régimen de la escuela.

Por supuesto, enfatizar la singularidad de la especie humana no significa negar que el reciclaje neuronal también existe en otros animales, aunque a menor escala. Las tecnologías recientes permitieron poner a prueba la hipótesis del reciclaje en monos macacos a una escala sin precedentes. En efecto, se pudo registrar la actividad de las mismas cien neuronas durante varias semanas, mientras los monos adquirían una nueva habilidad. Estos experimentos respondieron a una pregunta sencilla pero profunda de la teoría: ¿puede el aprendizaje generar un cambio radical en el código neural de determinado circuito cerebral o, como predice el concepto de reciclaje, únicamente lo reutiliza?

En esta experiencia muy reciente, gracias a una interfaz cerebro-máquina, los investigadores le pidieron a un mono que aprendiera a controlar su propio cerebro. Le enseñaron que para que el cursor se moviese hacia la derecha debía activar esas diez neuronas; y para que el cursor fuera arriba tenía que activar otras diez, y así sucesivamente (Galgali y Mante, 2018; Golub y otros, 2018; Sadtler y otros, 2014). Y este procedimiento funcionó: en pocas semanas, el animal aprendió a modificar la actividad de una decena de neuronas elegidas arbitrariamente para así mover un cursor de acuerdo con sus deseos. Sin embargo –y allí está la clave– solo lograba hacerlo si las descargas neuronales que le pedían que produjera no se separaban demasiado de las que su corteza ya producía de forma espontánea antes del entrenamiento. En otras palabras: lo que el mono aprendía ya debía pertenecer al repertorio de la corteza que le pedían que reentrenase.

Para apreciar lo que demostraron los investigadores, es importante tomar conciencia de que la dinámica de la corteza está limitada. El cerebro no explora todas las configuraciones de actividad a las cuales podría acceder. En teoría, en un conjunto de cien neuronas, la actividad podría abarcar un espacio de cien dimensiones, lo que implica una gigantesca cantidad de estados (si consideramos que cada neurona podría estar activada o desactivada, este número excede el resultado de 2^{100} , que es más de un 1 seguido por treinta ceros). Pero, en realidad, no es el caso: las neuronas verdaderas solo visitan una fracción de este gran universo, por lo general restringida a una decena de dimensiones. Si tenemos presente esta idea, nos

resulta muy sencillo comprender la limitación sobre el aprendizaje: si lo que le pedimos a la corteza “se sostiene” en este espacio preexistente, el mono aprende sin dificultad; si le exigimos una configuración de actividad que supere estos límites, no logra aprenderla. El comportamiento que el animal aprende puede ser radicalmente nuevo en la evolución (¡un primate controla un cursor en un monitor!); sin embargo, los estados neuronales que lo implementan deben encajar en el espacio de los patrones de actividad cortical disponibles. Este resultado valida directamente una predicción clave de la hipótesis del reciclaje neuronal: la adquisición de una habilidad nueva no requiere una reescritura radical de los circuitos corticales, como si fueran una pizarra en blanco, sino tan solo una reutilización de su organización anterior.

Poco a poco resulta más claro que cada región del cerebro impone su propio conjunto de restricciones. En un área de la corteza parietal, por ejemplo, algunas neuronas se limitan a una sola dimensión, una línea recta (Chafee, 2013; Fitzgerald y otros, 2013). Dado que codifican todos los datos en un eje que va de poco a mucho, estas neuronas son ideales para representar cantidades y sus tamaños relativos. Esta dinámica neuronal puede resultar extraordinariamente limitada, pero lo que parece una desventaja podría ser beneficioso cuando se trata de expresar parámetros como el tamaño, el número, la superficie o cualquier otro que deba ordenarse de menor a mayor. En cierto modo, esta corteza está precableada para codificar las cantidades, y la experiencia demuestra que su intervención es sistemática desde el momento en que aprendemos a manipular cantidades en un eje lineal, ya sean los números o cualquier otro elemento conceptualizable mediante magnitudes, como el estatus social (quién está debajo de quién en determinada escala; Chiao, 2010).

Veamos otro ejemplo: una región de la corteza temporal, la corteza entorrinal, contiene las famosas células de grilla que trazan el espacio, que ya mencioné. En esta región, el código neural es bidimensional: a pesar de que sean millones, las neuronas no pueden evitar permanecer confinadas a un plano, vale decir, un espacio de dos dimensiones (Yoon y otros, 2013). Evidentemente, esto es perfecto para formar un mapa del ambiente, como si se lo viera desde lo alto (en efecto, esta región representa el GPS mental con el cual una rata se orienta en el espacio). Sin embargo, investigaciones recientes demostraron que esta misma región, y algunas otras, se encienden desde el momento en que debemos aprender a representar cualquier tipo de datos en dos dimensiones, incluso si los datos no son espaciales

(Constantinescu, O'Reilly y Behrens, 2016). Puede tratarse, por ejemplo, de pájaros en los cuales el largo del cuello constituya la primera dimensión y el de las patas la segunda. En un experimento, una vez que los participantes humanos aprendieron a representar este “espacio de aves” inusual, usaron su corteza entorrinal (y algunas otras áreas) para navegar mentalmente.

La lista podría continuar: la corteza visual representa conjuntos de trazos y de formas, la región de Broca codifica los árboles sintácticos (Musso y otros, 2003; Nelson y otros, 2017; Pallier, Devauchelle y Dehaene, 2011), etc. Cada región posee una dinámica propia que casi no cambia. Cada una de ellas proyecta sobre el mundo su espacio de hipótesis: una busca disponer los datos sobre una línea, otras fijarlos sobre un plano o sobre un árbol... estos espacios preceden al aprendizaje y, en cierto modo, lo tornan posible. Por supuesto, somos capaces de aprender hechos novedosos, pero para lograrlo hace falta que encuentren su nicho neuronal, un espacio de representación adaptado a su organización natural.

Veamos ahora cómo se expresa esta idea en el dominio de los aprendizajes escolares más fundamentales: la aritmética y la lectura.

Las matemáticas reciclan los circuitos del número

Tomemos primero el ejemplo de las matemáticas. Como expliqué en mi libro *El cerebro matemático* (Dehaene, 2010), tenemos muchas pruebas de que la educación matemática (como tantos otros aspectos del aprendizaje) no se imprime en el cerebro como sobre una tablilla de cera. Por el contrario, se apoya sobre una representación innata y preexistente de cantidades numéricas, que luego extiende y refina.

Tanto en humanos como en monos, los lóbulos parietal y prefrontal ya contienen un circuito neural que representa los números de manera aproximada. Antes de cualquier educación formal, este circuito ya incluye neuronas sensibles a la cantidad aproximada de objetos en un conjunto concreto (Viswanathan y Nieder, 2013), una recta numérica mental espontánea. ¿Qué hace el aprendizaje? En animales entrenados para comparar cantidades, se incrementan las neuronas detectoras de números en el lóbulo frontal (Viswanathan y Nieder, 2015). Lo más importante es que, cuando

aprenden a confiar en los símbolos de los dígitos arábigos, más que en la mera percepción aproximativa del conjunto, una fracción de estas neuronas se vuelve selectiva para los números arábigos (Diester y Nieder, 2007). Esta transformación (parcial) de un circuito para incorporar la invención cultural de los símbolos numéricos es un buen ejemplo de reciclaje neuronal.

Cuando aprendemos a realizar operaciones aritméticas básicas (como la suma y la resta), los humanos continuamos reciclando esa región, pero también los circuitos de una zona contigua, el lóbulo parietal posterior, que se usa para desplazar el foco de la atención. Parece que reutilizamos esa habilidad para movernos en el espacio numérico: la suma activa los mismos circuitos que desplaza la atención hacia la derecha, en la dirección de los números mayores, mientras que la operación de la resta recurre a circuitos que llevan la atención a la izquierda (Knops y otros, 2009; Knops, Viarouge y Dehaene, 2009). Todos poseemos una suerte de recta numérica en nuestra cabeza, un mapa mental del eje numérico en el cual hemos aprendido a movernos con precisión cuando realizamos cálculos.

Por complejos que sean, todos los objetos matemáticos, desde los *topoi* de Grothendieck hasta los múltiples complejos o espacios funcionales, encuentran sus raíces últimas en la recombinación de los circuitos neuronales elementales presentes durante la infancia.

Recientemente, mi equipo de investigación fue mucho más lejos: junto con Marie Amalric, joven matemática convertida en científica cognitiva, nos preguntamos si estos mismos circuitos continúan siendo utilizados para reflexionar sobre conceptos más abstractos de las matemáticas (Amalric y Dehaene, 2016, 2017). Entonces, reclutamos a quince matemáticos profesionales y escaneamos sus cerebros con resonancia magnética funcional mientras les presentábamos expresiones matemáticas abstrusas, que solo ellos podían entender, incluidas fórmulas como

$$\int_S \nabla \times F \cdot dS$$

o incluso proposiciones como

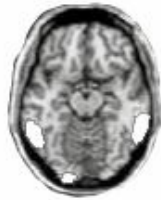
toda matriz cuadrada es equivalente a una matriz de permutación.

Corteza temporal inferior

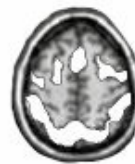
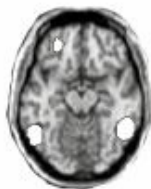
Surco intraparietal



Reflexión matemática en los matemáticos profesionales



Representación de los números



Cálculo mental



Intersección = red universal de las matemáticas

Figura 20. La educación consiste en reciclar circuitos cerebrales antiguos para reorientarlos hacia actividades nuevas. Desde la primera infancia, cada individuo posee circuitos de representación de los números y del cálculo mental. Los matemáticos profesionales utilizan

estos mismos circuitos cuando reflexionan acerca de los conceptos más complejos de su disciplina. Dichas redes neuronales responden inicialmente a los conjuntos concretos de objetos, pero logran reciclarse para procesar conceptos más abstractos.

Como habíamos predicho, estos objetos matemáticos de alto nivel continúan activando la misma red cerebral que, desde la más tierna infancia, se activa cuando un bebé ve uno, dos o tres objetos (Izard, Dehaene-Lambertz y Dehaene, 2008) o cuando un niño aprende a contar (figura 20; Cantlon y otros, 2006; Cantlon y Li, 2013). Todos los objetos matemáticos, por complejos que sean, desde los *topoi* de Grothendieck hasta las integrales dobles o los espacios funcionales, tienen sus raíces últimas en la recombinación de circuitos neuronales elementales presentes en la infancia. Todos nosotros, en cualquier etapa de la construcción cultural de las matemáticas, desde los estudiantes de primaria hasta los ganadores de la Medalla Fields, refinamos continuamente el código neuronal de ese circuito cerebral específico.

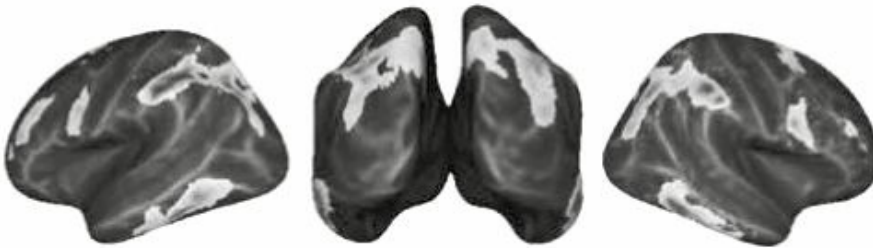
Lo cierto es que la organización de ese circuito está bajo fuertes restricciones hereditarias, las de la composición genética universal que nos hace humanos. Si bien el aprendizaje le permite acomodar muchos conceptos nuevos, su arquitectura general sigue siendo la misma en todos nosotros, independientemente de la experiencia. Mis colegas y yo obtuvimos sólidas pruebas de ello cuando estudiamos la organización cerebral de los matemáticos cuya experiencia sensorial, desde la infancia, había tenido diferencias radicales: matemáticos ciegos (Amalric, Denghien y Dehaene, 2017). Por sorprendente que parezca, no es raro que alguien ciego se convierta en un excelente matemático. Quizá el caso más conocido sea Nicholas Saunderson (1682-1739), quien perdió la visión a los 8 años, y era tan brillante que terminó ocupando la Cátedra Lucasiana de Matemáticas en la Universidad de Cambridge luego de su maestro y amigo Isaac Newton.

Saunderson ya no está entre nosotros, disponible para someterse a un escaneo cerebral, pero Marie Amalric y yo logramos contactar a tres matemáticos contemporáneos ciegos, todos profesores universitarios en Francia. Uno de ellos, Emmanuel Giroux, verdadero gigante de las matemáticas, es director de un laboratorio integrado por sesenta personas en la École Normale Supérieure de Lyon. Ciego desde los 11 años, es famoso sobre todo por su espléndida demostración de un importante teorema de geometría de contacto. La existencia misma de estos matemáticos ciegos

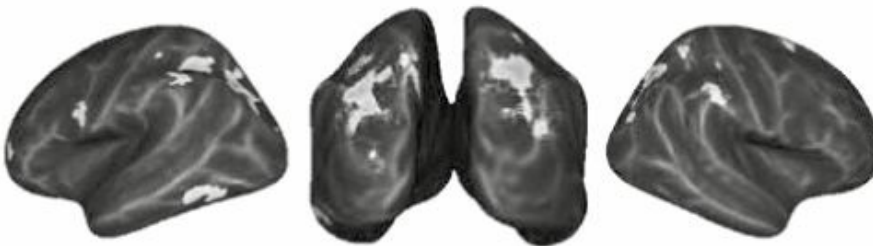
refuta la visión empirista de Alan Turing (el cerebro como una libreta en blanco que se empieza a llenar por influencia de los sentidos).



15 matemáticos no privados del sentido de la vista:

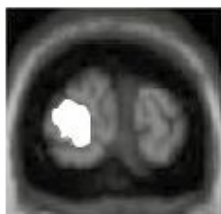


3 matemáticos ciegos:



Activación suplementaria de la corteza visual en los ciegos:

Experiencia 1



Experiencia 2

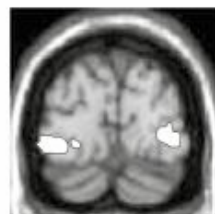


Figura 21. El aprendizaje de las matemáticas se apoya en conocimientos universales, en gran parte independientes de la experiencia sensorial. En los matemáticos ciegos, durante la reflexión matemática se activan las mismas regiones de la corteza parietal, temporal y frontal que en sus colegas con visión “normal”. Solo hay una diferencia: los ciegos también logran reciclar su corteza visual para dedicarla a las matemáticas.

En efecto, ¿cómo podría una persona ciega, con una experiencia tan diferente y limitada, derivar los mismos conceptos que los matemáticos con visión “normal”, a menos que contara ya con los circuitos capaces de generarlos? Al respecto, Emmanuel Giroux parafrasea *El Principito*: “En geometría, lo esencial es invisible a los ojos; solo se puede ver bien con la mente”. En matemáticas, las experiencias sensoriales no importan mucho, son las ideas y los conceptos los que hacen el trabajo pesado.

Si los estímulos externos determinaran la organización de la corteza, un matemático ciego que aprendió todo gracias al sentido del tacto debería activar áreas cerebrales muy diferentes a las de otro sin esa condición a la hora de realizar operaciones matemáticas. Por el contrario, la teoría del reciclaje neuronal predice que los circuitos de los matemáticos son fijos. Y este último resultado es el que observamos cuando estudiamos a los tres matemáticos ciegos. Como esperábamos, cuando “ven” lo que significa un teorema de matemática, lo hacen gracias a los mismos circuitos de los lóbulos parietales y frontales que utiliza un matemático sin impedimentos visuales (figura 21). En nada incide la experiencia sensorial: este circuito es el único capaz de reciclarse para hacer operaciones matemáticas.

La única diferencia es que cuando nuestros tres matemáticos ciegos piensan acerca de su dominio predilecto utilizan también las áreas visuales de la corteza. Debemos esta hipótesis a una intuición de Cédric Villani (otro matemático genial, medalla Fields). Cuando discutimos acerca de esta experiencia, me dijo en broma: “Emmanuel Giroux es un gran matemático, pero también tiene mucha suerte: ¡como es ciego, puede dedicar aún más corteza a las matemáticas!”. Tenía razón, y advertía otro magnífico ejemplo del reciclaje. En las personas ciegas, la corteza occipital, normalmente dedicada a la visión, no permanece inactiva: se dedica a nuevas funciones, en especial al cálculo mental y a las matemáticas (Amalric y otros, 2017; Kanjlia y otros, 2016). En los ciegos de nacimiento, la reorganización es todavía más extrema, dado que observamos en su corteza visual respuestas verdaderamente inesperadas a la gramática de la lengua hablada, similares a

las del área de Broca (Amedi y otros, 2003; Bedny y otros, 2011; Lane y otros, 2015; Sabbah y otros, 2016).

La presencia de respuestas arbitrarias en la corteza visual de las personas ciegas todavía es objeto de debate teórico (Bedny, 2017; Hannagan y otros, 2015): ¿consiste en un verdadero reciclaje, o bien es una prueba extrema de plasticidad que conduce a la reorganización total de la corteza? En mi opinión, la balanza se inclina en favor de la hipótesis del reciclaje neuronal, porque la organización preexistente de esta región no desaparece por completo, como ocurriría en una pizarra que se expusiera a los borrones de un entorno muy diferente. En efecto, la corteza visual de las personas ciegas conserva la organización de sus conexiones y de sus mapas neuronales (Bock y otros, 2015), siempre reorientados hacia otros dominios. Como esta corteza es muy amplia, encontramos numerosas regiones “visuales” que responden no solo a las matemáticas y el lenguaje, sino también a las letras y los nombres (presentados en braille), a los objetos, los lugares y los animales (Abboud y otros, 2015; Amedi y otros, 2003; Bedny y otros, 2011; Mahon y otros, 2009; Reich y otros, 2011; Striem-Amit y Amedi, 2014; Strnad y otros, 2013). En la mayoría de las ocasiones, estas categorías se encuentran en lugares similares de la corteza, sin importar si vemos o somos ciegos. La “caja de letras” del cerebro, por ejemplo, está situada casi en el mismo lugar en un lector con visión “normal” y en un lector ciego que aprendió braille: la función de esta región parece estar determinada en gran medida por sus conexiones con las áreas del lenguaje, y lo mismo ocurre, sin duda, con otras propiedades innatas (Bouhali y otros, 2014; Hannagan y otros, 2015; Saygin y otros, 2012, 2013, 2016).

Para regresar a las matemáticas, la hipótesis del reciclaje neuronal no se basa solo en la idéntica localización en el cerebro de los conceptos elementales ($1 + 1 = 2$) y de las ideas matemáticas más avanzadas ($e^{-i\pi} + 1 = 0$). Otros descubrimientos, puramente psicológicos, indican que las matemáticas que aprendemos en la escuela dependen del reciclaje de circuitos antiguos consagrados a las cantidades aproximativas.

Piensen en el número 5. En este mismo momento, su cerebro reactiva una representación de las cantidades, la misma que compartimos con otros primates. Ahora, intenten decidir si 5 es mayor o menor que 6. Los experimentos muestran que el razonamiento es mucho más lento cuando los números están próximos, como 5 y 6, que cuando están distantes, como 5 y 9. Este efecto de distancia (Dehaene, 2007b; Dehaene, Dupoux y Mehler, 1990;

Moyer y Landauer, 1967) es una de las huellas de la representación antigua de los números que hemos reciclado cuando aprendimos a contar y a calcular. La mente intenta concentrarse en los símbolos, pero no puede evitar activar las cantidades correspondientes, y estas se superponen mucho más cuando los números son contiguos. Incluso para decidir si dos cifras como 8 y 9 son diferentes una de otra, lo que debería ser inmediato, todavía influye la distancia que las separa, y exactamente lo mismo ocurre con los monos que aprendieron a reconocer los símbolos de los números arábigos (Dehaene y Akhavein, 1995; Diester y Nieder, 2010).

Podría abundar en ejemplos que ponen en evidencia la existencia del eje numérico mental que utilizamos cuando comparamos cantidades y cuando hacemos cálculos. El tiempo que nos toma restar dos números varía en proporción directa con el tamaño del número restado (Groen y Parkman, 1972; Pinheiro-Chagas y otros, 2017). Todo sucede como si la mente tuviera que desplazarse para llegar a los números grandes: cuanto más lejos nos vamos, más tiempo le dedicamos. Del mismo modo, cuando pensamos en un precio, no podemos evitar atribuirle mayor grado de imprecisión cuanto mayor es el número: los números más grandes son también los más borrosos (Dehaene y Marques, 2002; Marques y Dehaene, 2004). Por este motivo, en contra de cualquier racionalidad, cuando negociamos, no tenemos problema en ceder algunos miles en el precio de un departamento y, el mismo día, regateamos unos centavos en la cuenta de la panadera: el grado de imprecisión tolerable es proporcional al número representado, tanto para los seres humanos como para el mono macaco.

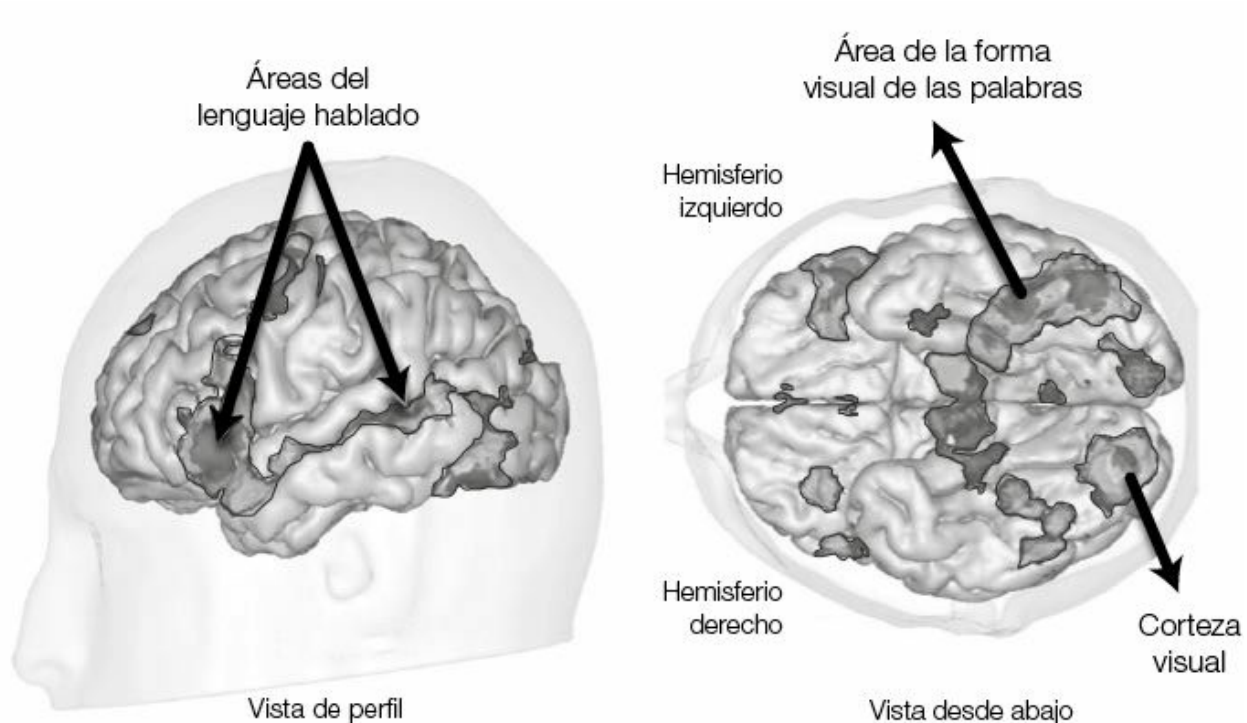
Y la lista podría continuar: paridad, números negativos, fracciones... todos esos conceptos se sustentan en la representación inicial de las cantidades.[\[16\]](#) A diferencia de una computadora digital, somos incapaces de manipular símbolos en un plano abstracto. Siempre los anclamos en lo concreto de las cantidades. La persistencia de efectos analógicos reales en un cerebro educado deja entrever las antiguas ataduras de nuestro concepto del número.

El número aproximativo es uno de los viejos pilares sobre los cuales se apoya la construcción de las matemáticas. Sin embargo, la educación también lo enriquece. Cuando aprendemos a contar y a calcular, esta aculturación matemática introduce símbolos precisos allí donde la evolución se contentó, durante millones de años, con cantidades borrosas. Es un poderoso factor de cambio: todos nuestros circuitos aritméticos sufren modificaciones mínimas con el fin de permitir la manipulación de los símbolos. De todos modos, este

sentido innato de las cantidades, incluso si, en tono jocoso, lo llamé “giba de las matemáticas”,[\[17\]](#) no es, por cierto, su única base. Como ya vimos, heredamos de nuestra evolución un sentido del espacio, con sus propios circuitos neuronales y sus células de lugar, de grilla y de dirección. Poseemos también un sentido de las formas, que permite que un niño pequeño reconozca un rectángulo o un triángulo. De un modo que todavía no se comprende del todo, bajo la influencia de símbolos como las palabras y los números, todos estos conceptos se reciclan cuando aprendemos matemáticas. El cerebro humano los recombina, en un lenguaje del pensamiento, para formar nuevos conceptos (Amalric y otros, 2017; Piantadosi y otros, 2012, 2016). Los bloques de construcción básicos que heredamos de la evolución son como las palabras de una lengua nueva, productiva, en la que los matemáticos escriben cada día nuevas páginas.

La lectura recicla los circuitos de la visión y de la lengua hablada

¿Qué sucede con el aprendizaje de la lectura? En mi segundo libro, *El cerebro lector* (Dehaene, 2007a), describí en detalle los circuitos que nos permiten aprender a leer. Es otra prueba del reciclaje neuronal: para leer, reutilizamos un amplio conjunto de áreas cerebrales que inicialmente están dedicadas a la visión y al procesamiento de la lengua hablada. Cuando aprendemos a leer, reorientamos estos circuitos para que las regiones visuales reconozcan las cadenas de letras y las envíen hacia las áreas del lenguaje. En un buen lector, el resultado es que las palabras que se leen son tratadas exactamente como palabras que se oyen: la alfabetización crea una nueva puerta de entrada visual hacia los circuitos del lenguaje.



Actividad cerebral en respuesta a las frases escritas



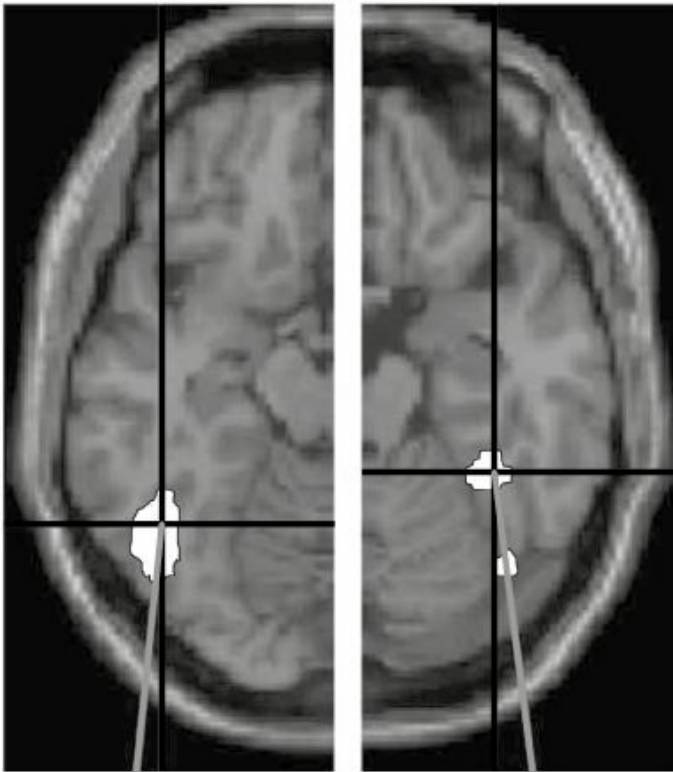
Figura 22. El aprendizaje de la lectura recicla una red de áreas cerebrales involucradas en la visión y en el procesamiento del lenguaje hablado. En todas las regiones que aquí están contorneadas, la actividad cerebral evocada por frases escritas aumenta a medida que se incrementa el puntaje de los sujetos en tareas de lectura, desde los analfabetos completos

hasta los lectores expertos. Se observa que la lectura tiene un doble efecto: por un lado, hace que se especialicen las áreas visuales, sobre todo una región del hemisferio izquierdo llamada “área de la forma visual de las palabras”; y por otro, activa las áreas del lenguaje hablado a partir de la visión.

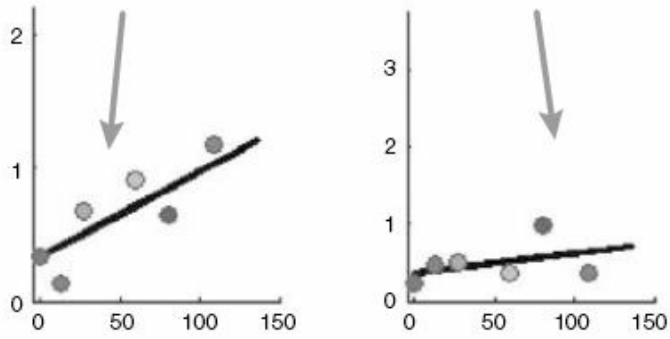
Mucho antes de aprender a leer, cada niño dispone de un sistema visual sofisticado que le permite reconocer y nombrar los objetos, los animales y las personas. Reconoce una imagen (sin que importen su tamaño, su posición ni su orientación en tres dimensiones) y sabe asociarla a un nombre. La lectura recicla una parte de este circuito: la educación en la lectura desarrolla una región muy especializada de la corteza visual, que junto con mi colega Laurent Cohen denominamos “área de la forma visual de las palabras”, que se convierte en la verdadera “caja de letras” de nuestro cerebro y nos permite reconocer las palabras, cualquiera sea su tamaño, su posición, su CaJa (MAYÚSCULA o minúscula) o su *fuentes* (Dehaene y otros, 2001, 2004). Gracias a sus conexiones muy directas en dirección hacia las áreas del lenguaje (Bouhali y otros, 2014; Saygin y otros, 2016), esta región de la corteza aprende a traducir rápidamente la información visual en sonidos y en significados.

¿Qué ocurriría si analizáramos el cerebro de una persona analfabeta a medida que aprende a leer? Si la teoría es correcta, deberíamos ver sin ambages que la corteza visual se reorganiza. La teoría del reciclaje neuronal sostiene que la lectura invade una región de la corteza normalmente destinada a actividades cercanas, y la reorienta hacia esta nueva tarea. La predicción es inmediata: este aprendizaje debería entrar en competencia con las otras funciones preexistentes de la corteza. ¿Podría ocurrir que perdiéramos determinadas funciones heredadas de la evolución a medida que aprendemos a leer? ¿O, por lo menos, que las reorganizáramos en gran medida?

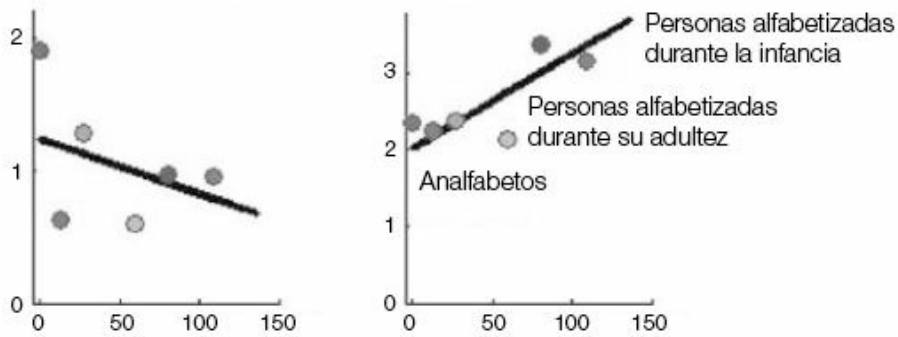
Junto con mis colegas, tuvimos ocasión de evaluar en una serie de experimentos esta predicción que va a contrapelo de la intuición. Comenzamos por estudiar el cerebro de adultos analfabetos, en Portugal y en Brasil, comparándolo con el de personas de los mismos pueblos pero que habían tenido la oportunidad de aprender a leer en la escuela, ya fuera en la niñez o en la adultez (Dehaene y otros, 2010; Dehaene, Cohen y otros, 2015; Pegado, Comerlato y otros, 2014). Los resultados nos permitieron construir un mapa completo de todas las regiones del cerebro donde la respuesta aumenta con el aprendizaje (figura 22).



Respuesta a las cadenas de letras



Respuesta a los rostros



Puntaje en lectura: Cantidad de palabras leídas por minuto

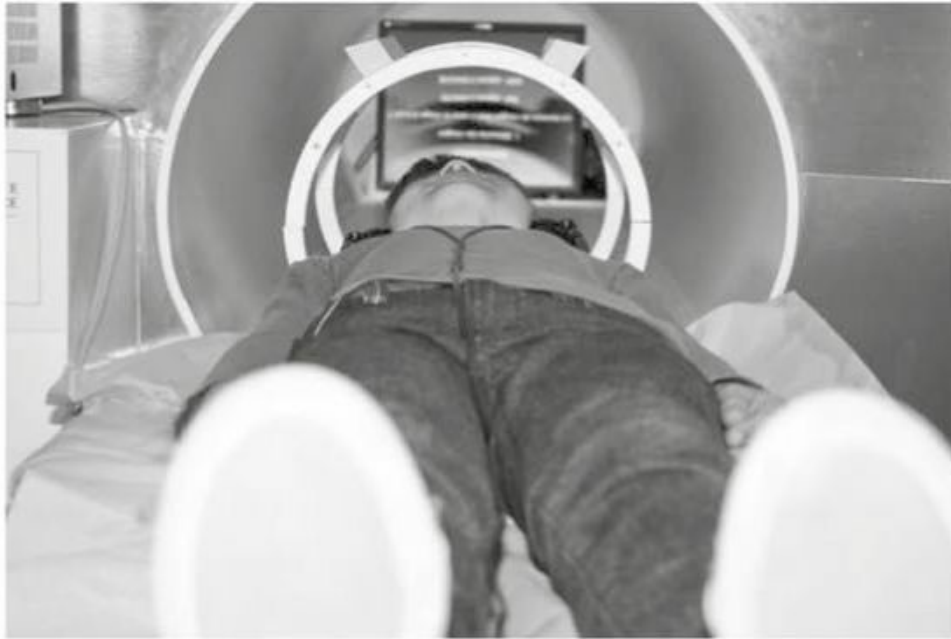
Figura 23. De acuerdo con la hipótesis del reciclaje neuronal, el aprendizaje de la lectura entra en competencia con funciones que previamente cumplía la corteza visual (en este caso, el reconocimiento de rostros). Conforme mejora el puntaje en tareas de lectura, aumenta la activación evocada por las palabras escritas en el hemisferio izquierdo y la activación evocada por los rostros se desplaza desde el hemisferio izquierdo hacia el hemisferio derecho.

Es un resultado fácil de comprender. Primero, muéstrenle una frase, palabra por palabra, a una persona analfabeta y verán que su cerebro no responde para nada. Es evidente: no sabe leerlas, por lo que la actividad se detiene muy pronto, en el nivel de las áreas visuales. Ahora, preséntele la misma frase escrita a una persona que aprendió a leer: un circuito cortical mucho mayor se enciende, de forma directamente proporcional al puntaje que obtuvo en la lectura. Los cambios se extienden desde las áreas visuales primarias, especializadas en el reconocimiento de los caracteres (Chang y otros, 2015; Dehaene y otros, 2010; Szwed y otros, 2014), hasta las regiones del lenguaje asociadas con la comprensión de las frases, pasando por la famosa “caja de letras del cerebro”, que se encuentra en la corteza t́mporo-occipital del hemisferio izquierdo. Cuanto mejor sabemos leer, más se activan estas regiones ante las palabras escritas y más se fortalecen sus conexiones: automatizar la lectura es hacer más fluida la relación directa entre las letras y los sonidos del lenguaje.

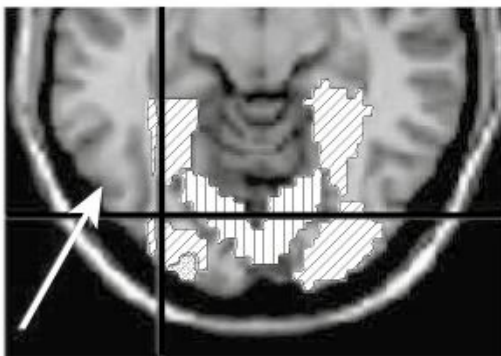
Pero también cabe plantear la pregunta en sentido inverso: ¿existen regiones que se activen más en los malos lectores y en las que la actividad disminuya a medida que se aprende a leer? Hemos descubierto que sí: las personas analfabetas presentan mayor activación para el procesamiento de rostros. Cuanto mejor sabemos leer, más disminuye esta actividad en el hemisferio izquierdo, en el lugar exacto en el que se instalan las palabras escritas. Todo parece indicar que hay que hacer lugar en la corteza para las letras: la lectura interfiere en una función mucho más antigua, la de reconocer rostros. Pero, por supuesto, esta función no es desechada fuera de la corteza. Más bien, pudimos observar que, con la alfabetización, la respuesta a los rostros se incrementa del lado del hemisferio derecho. Al verse desplazados del hemisferio izquierdo, que es la localización del lenguaje en la mayoría de las personas, los rostros se refugian en el otro lado (figura 23; Dehaene y otros, 2010; Pegado, Comerlato y otros, 2014).

En primer lugar, hicimos esta observación en adultos alfabetizados y

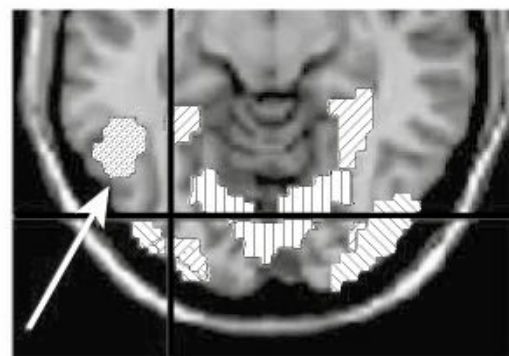
analfabetos, y luego la replicamos en niños que estaban aprendiendo a leer (Dehaene-Lambertz, Monzalvo y Dehaene, 2018; Dundas, Plaut y Behrmann, 2013; Li y otros, 2013; Monzalvo y otros, 2012). Tan pronto el niño comienza a leer, el área de la forma visual de las palabras se desarrolla en el hemisferio izquierdo y su simétrica contraparte, en el hemisferio derecho, refuerza su respuesta a los rostros (figura 24).



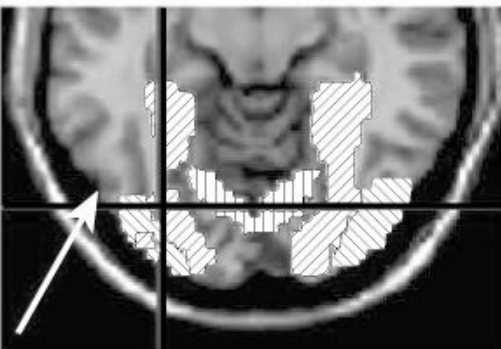
6 años, no lector



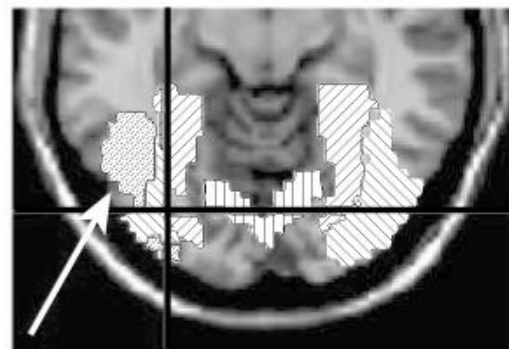
6 años, lector



9 años, disléxico



9 años, lector



Respuesta a las

palabras

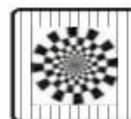
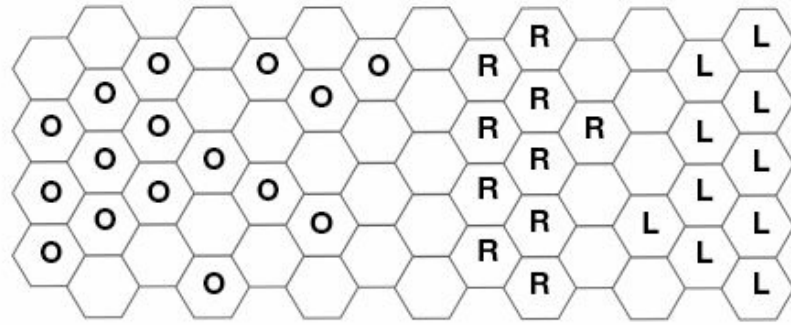


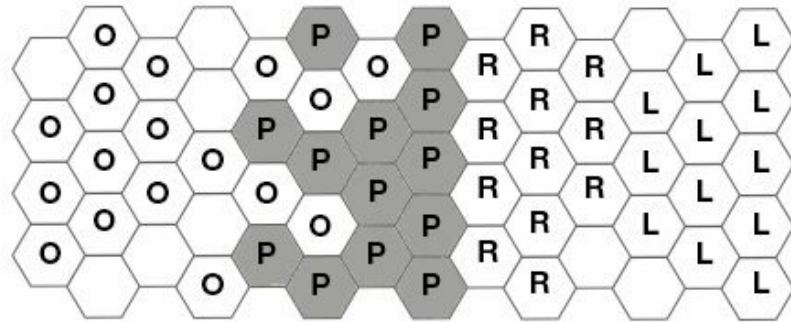
Figura 24. La resonancia magnética (MRI) permite seguir el desarrollo de la lectura en el niño. Desde el momento en que un niño aprende a leer, se ve aparecer una región que responde a las letras y a las cadenas de letras. La lectura recicla una parte del mosaico de regiones que en todos los primates sirven para reconocer los rostros, los objetos y los lugares.

Antes de la escuela primaria



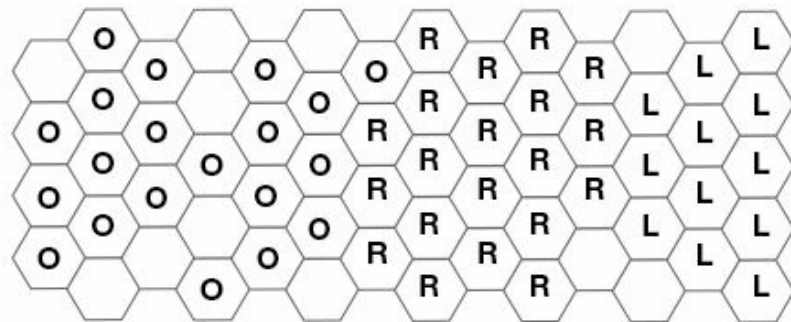
Respuestas a los **O**jetos - **R**ostros - **L**ugares

Después de la escolarización: cerebro alfabetizado



Respuestas a las **P**alabras

Infantes sin escolarización: cerebro analfabeto



Expansión de los **R**ostros

Figura 25. El aprendizaje es más fácil durante la infancia, cuando la corteza todavía es maleable. Antes de que un niño pequeño vaya a la escuela, determinadas regiones visuales de su cerebro ya se especializaron para el reconocimiento de objetos, de rostros o de lugares, pero todavía quedan grandes franjas poco o nada especializadas (simbolizadas por los hexágonos vacíos). El aprendizaje de la lectura (**P**) invade estas regiones inestables e impide el desarrollo de otras categorías de objetos. Si el niño no aprende a leer, esta región se utiliza para el reconocimiento de los rostros y los objetos, y pierde gradualmente su capacidad de

aprender las letras.

El efecto es de tal magnitud que, a una edad dada, con solo examinar la actividad evocada por los rostros, un algoritmo logra decidir si un niño ya aprendió a leer o no. Y si el niño sufre de dislexia, estas regiones no se desarrollan con normalidad, ni a la izquierda, en el área de la forma visual de las palabras, ni a la derecha, en la región del giro fusiforme especializada para el reconocimiento de los rostros (Monzalvo y otros, 2012). La actividad reducida de la corteza t́mporo-occipital izquierda frente a palabras escritas es un marcador universal de las dificultades de la lectura en todas las lenguas en que se la evaluó (Rueckl y otros, 2015).

Recientemente, obtuvimos la autorización para realizar una investigación audaz. Queríamos ver cómo emergen los circuitos de lectura en niños individuales, y para lograrlo, convocamos a los mismos niños cada dos meses a nuestro centro de neuroimágenes, con el fin de escanearlos al final del jardín de infantes y a lo largo de todo el primer año de escuela. Los resultados estuvieron a la altura de nuestras expectativas. Al principio, no había nada para señalar: cuando el niño, todavía en jardín de infantes, no había aprendido aún a leer, su corteza respondía a los objetos, a los rostros, a las casas, pero no a las letras. Luego de dos meses de escuela, en cambio, vimos aparecer una respuesta específica a las palabras escritas, en la misma posición que en el adulto: en la corteza t́mporo-occipital izquierda. Muy lentamente, la representación de los rostros se modificó: a medida que el niño se alfabetizaba, la actividad del reconocimiento de los rostros aumentaba en el hemisferio derecho, en proporción directa con el incremento de las habilidades en lectura. Una vez más, en sintonía con la hipótesis del reciclaje neuronal, pudimos ver cómo la adquisición de la lectura entraba en competencia con la función anterior de la corteza, el reconocimiento visual de rostros.

Al realizar este trabajo, nos dimos cuenta de que esta competencia podía explicarse de dos formas diferentes. La primera posibilidad consiste en el *knock-out*: desde el nacimiento, los rostros comienzan a instalarse en la corteza visual del hemisferio izquierdo, pero la lectura los desaloja de allí, como de un puñetazo. La segunda posibilidad es el bloqueo: la corteza se desarrolla lentamente y poco a poco se especializa para reaccionar a los rostros, los objetos o los lugares. Cuando las letras penetran en este terreno en

desarrollo, ocupan una parte del territorio y obturan la expansión de las otras categorías visuales.

¿La alfabetización hace que la corteza quede noqueada o bloqueada? Nuestros experimentos se inclinan por la segunda posibilidad: aprender a leer bloquea la expansión de los rostros en el hemisferio izquierdo (figura 25). Gracias a los registros de MRI que realizamos cada dos meses, vimos crecer la especialización de la corteza visual con mucha lentitud. A esta edad, cerca de los 6 o 7 años, todavía está lejos de completarse. Existen, por supuesto, territorios especializados para los rostros, los objetos o los lugares, pero también muchas regiones corticales que todavía no se especializaron para ninguna de estas categorías. Y las letras, precisamente, invaden algunas de estas regiones no especializadas y las reciclan. Pero, en contra de lo que pensábamos al comienzo de la investigación, no aplastan el área de los rostros, sino que se instalan justo al costado, en una región vecina, de modo similar a como un agresivo supermercado se instala justo al lado de un pequeño comercio. La expansión de uno bloquea la del otro (Dehaene-Lambertz, Monzalvo y Dehaene, 2018).

En síntesis, por supuesto, no es azaroso que aprendamos a leer en los primeros años de la infancia. El aprendizaje de la lectura saca partido de este período bendito en que la corteza visual todavía es lábil. Su arquitectura está lejos de carecer de restricciones, pero ha evolucionado para aprender a reconocer todo tipo de imágenes y adaptarse a su forma. Cuando la exponemos a miles de palabras escritas, se adapta y se recicla para esta nueva actividad.

Con la edad, muy paulatinamente, un número creciente de columnas corticales se fijan y dejan de modificarse. Entonces, se vuelve cada vez más difícil automatizar el reconocimiento de las letras y sus combinaciones. Junto con mis colegas pudimos estudiar a dos personas que intentaron aprender a leer en la adultez: una de ellas nunca había tenido la oportunidad de ir a la escuela y la otra había sufrido un pequeño accidente cerebrovascular justo en el área de la forma visual de las palabras, que la había vuelto aléxica, incapaz de leer una sola palabra. Las evaluamos y las escaneamos periódicamente durante dos años (Braga y otros, 2017; Cohen y otros, 2016). Sus progresos eran de una lentitud increíble. La primera terminó desarrollando una región especializada para las letras, pero sin que por eso se vieran afectados los rostros: lo que ya estaba impreso no se movía más. La segunda no logró jamás hacer emerger una nueva “caja de letras” en su corteza visual. Su

lectura, aunque mejoraba, todavía era lenta y se asemejaba al descifrado laborioso de un principiante.

Divisiones, ecuaciones y rostros

La conclusión es simple: para reciclar en profundidad la corteza visual y convertirnos en excelentes lectores, es necesario apoyarse sobre el período de plasticidad máxima que ofrece la infancia. Nuestras investigaciones ofrecen muchos más ejemplos. Tomemos el aprendizaje de la lectura musical: un músico que aprendió a leer partituras a una edad temprana prácticamente duplica la superficie de su corteza visual dedicada a los pentagramas, en comparación con alguien que nunca ha aprendido música. Este crecimiento masivo y precoz ocupa lugar en la superficie de la corteza, y parece desalojar de allí el área de la forma visual de las palabras: en los músicos, la región cortical que responde a las letras, la “caja de letras del cerebro”, se desplaza casi un centímetro en comparación con su posición normal en quienes no se dedican a la música (Mongelli y otros, 2017).

Otro ejemplo: el reconocimiento de las ecuaciones matemáticas. Un matemático consagrado debe saber analizar, a primerísima vista, expresiones tan abscónditas como

$$\begin{aligned} \pi &= 3,141592\dots \\ \phi &= 1,61803394\dots \\ f(x) &= \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right) \end{aligned}$$

o

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

con tanta facilidad como nosotros leemos una novela. Esa es una pericia visual de alto nivel. Una vez asistí a una conferencia en que el matemático Alain Connes (otro ganador de la Medalla Fields) exhibió una ecuación que tenía veinticinco líneas de una extraordinaria densidad. Según explicó, allí figuraba la expresión matemática de todos los efectos físicos de todas las partículas elementales conocidas. Un segundo matemático levantó el dedo y dijo: “¿No hay un error en la línea 13?”. “No”, respondió Connes de inmediato sin perder la compostura, “porque el término compensatorio correspondiente está justo en la línea 14”.

¿Cómo se refleja en el cerebro de los matemáticos tamaña solvencia para las fórmulas? Las imágenes cerebrales muestran que estos objetos matemáticos invaden buena parte de las regiones occipitales laterales de ambos hemisferios; después del entrenamiento matemático, estas regiones terminan por responder a las expresiones algebraicas mucho más que en quienes carecen de esa práctica. Y, una vez más, somos testigos de una competencia que decae y se contrae, como esa piel de asno (o más precisamente, de zapa) que tanto dio que hablar a partir de la novela de Balzac: en este caso, notamos que las áreas de la corteza sensibles a los rostros se desvanecen en ambos hemisferios (Amalric y Dehaene, 2016). En otras palabras, mientras la alfabetización se limita a quitar los rostros del hemisferio izquierdo y los obliga a moverse hacia el hemisferio derecho, la práctica intensa con números y ecuaciones interfiere con su representación en ambos lados, lo que lleva a una contracción global de los circuitos de reconocimiento facial.

Resulta tentador trazar un vínculo entre este hallazgo y el famoso mito del matemático excéntrico, incapaz de interesarse por otra cosa que por las ecuaciones y que no reconoce ya a su vecina, su perro o su reflejo en el espejo. Las anécdotas sobre los matemáticos abundan, y dieron lugar a chistes como el siguiente: “¿Sabes cómo se reconoce a un matemático extrovertido? Es el que mira *tus* zapatos”.

En realidad, todavía no sabemos si la reducción de la respuesta cortical para los rostros en los matemáticos tiene una relación directa con su supuesta falta de competencia social (mito, antes que realidad, dado que muchos matemáticos se sienten maravillosamente a gusto en sociedad). La causalidad todavía debe establecerse: ¿pasar la vida entre fórmulas matemáticas reduce la respuesta a los rostros? O bien, por el contrario: ¿los matemáticos se sumergen en los cálculos porque les resultan más sencillos que las interacciones sociales? En cualquiera de los casos, la competencia cortical está demostrada y la representación de los rostros muestra una notable sensibilidad a los aprendizajes escolares, en grado tal que puede proveer un marcador confiable de la educación que un niño recibió en matemáticas, en música o en lectura. Es otra prueba del reciclaje neuronal.

Los beneficios de un ambiente enriquecido

La principal conclusión es que, en las antípodas de la caricatura que sugiere que todo es innato o que todo es adquirido, el cerebro del niño es simultáneamente estructurado y plástico. Desde el nacimiento, está dotado de una inmensidad de circuitos especializados que resultaron modelados por los genes y seleccionados por decenas de millones de años de evolución. Esta autoorganización le confiere intuiciones profundas en los grandes ámbitos del saber: un sentido de la física que gobierna los objetos y su comportamiento, un don para la navegación espacial, intuiciones sobre los números, la probabilidad y las matemáticas, un apego a los otros seres humanos, e incluso, en nuestra especie, un auténtico despliegue de genialidad para las lenguas. ¡Las metáforas de la *tabula rasa* o la pizarra en blanco no podrían ser más erróneas! Y, sin embargo, la evolución también dejó la puerta abierta a muchas oportunidades de aprendizaje. No todo está predeterminado en el cerebro del niño. Todo lo contrario: el detalle de los circuitos neuronales, en la escala de algunos milímetros, exhibe una amplia apertura a las interacciones con el mundo exterior.

Durante los primeros años de la vida, los genes suscitan una exuberante superproducción de circuitos neuronales: dos veces más sinapsis que las necesarias. Gracias a algún mecanismo que todavía no comprendemos bien, esta sobreabundante combinatoria abre una inmensa variedad de modelos mentales del mundo, un verdadero lenguaje interior. Gracias a ella, el cerebro del niño está en plena ebullición de posibilidades y explora un conjunto de hipótesis tanto más vasto que el del adulto. Cada bebé está abierto a todas las lenguas, todas las ortografías, todas las matemáticas posibles (por supuesto, dentro de los límites que impone la genética de la especie).

Luego intervienen poderosos algoritmos de aprendizaje para seleccionar las sinapsis y los circuitos más útiles. En algunos años, el cerebro se especializa y se fija. Las regiones sensoriales son las primeras en fijarse: desde el primer año de vida, las áreas auditivas comienzan a converger hacia las vocales y las consonantes de la lengua materna. En algunos años, a medida que se cierran sucesivamente los diferentes períodos sensibles de la plasticidad cerebral, nos convertimos, para siempre, en nativos de una lengua, de una ortografía, de una cultura específica. Y si se nos priva de la estimulación en determinado ámbito, porque somos huérfanos en Bucarest o bien analfabetos en los suburbios de Brasilia, quedamos expuestos al riesgo de perder la fluidez mental en este campo del saber —y quizá perderla para siempre—. La intervención siempre es deseable, porque durante toda su vida el cerebro

conserva parte de su plasticidad, sobre todo en las regiones cerebrales de nivel más alto, como las de la corteza prefrontal. Sin embargo, quedó demostrado que las intervenciones más precoces también son las más eficaces: tanto para hacerle llevar lentes prismáticos a un búho, como para enseñarle una segunda lengua a un niño, para ajustarse a la sordera, a la ceguera o a la pérdida de un hemisferio cerebral completo, cuanto más temprano se comience siempre será mejor.

Las escuelas son instituciones concebidas para aprovechar al máximo la plasticidad. La educación depende en gran medida de la espectacular flexibilidad del cerebro del niño para reciclar algunos de sus circuitos y reorientarlos hacia nuevas actividades como la lectura o las matemáticas. Numerosos experimentos demuestran que, cuando interviene de forma precoz, la escolarización transforma la vida: luego de seguirlos durante décadas, se observó que los niños provenientes de contextos desfavorecidos que participaron en intervenciones precoces mostraban mejores indicadores en muy diferentes ámbitos de la vida, como la salud, los ingresos promedio o la exposición a infracciones penales.[\[18\]](#)

Pero la escuela no es una panacea ni una píldora sintética que todo lo puede. Las familias y los cuidadores también tienen el deber de estimular el cerebro del niño para enriquecer su entorno tanto como sea posible. Cada lactante es un físico en ciernes que adora experimentar con la gravedad y la caída de los objetos. Entonces, tenemos que dejarlo improvisar, construir, fracasar y recomenzar, en vez de inmovilizarlo en una silla durante horas. Cada niño es un matemático incipiente que adora contar, medir, trazar rectas y círculos, ensamblar las formas, siempre y cuando le demos reglas, brújulas, papeles, además de acertijos matemáticos atractivos... Cada bebé es un lingüista genial: a partir de los 18 meses, aprende un mínimo de 10 a 20 palabras por día, pero solo si le hablamos... La familia y los amigos deben alimentar este apetito por el conocimiento y nutrirlo con frases bien formadas, sin dudar en utilizar un léxico rico. El vocabulario que dominará a los 3 o 4 años dependerá de manera directa de la cantidad de discurso que le hayamos dirigido, y no simplemente de la exposición a las palabras del entorno (Shneidman y otros, 2013; Shneidman y Goldin-Meadow, 2012).

Los resultados de las investigaciones son convergentes: enriquecer el ambiente de los niños muy pequeños es construir el futuro de su cerebro. Cuando alcance la edad de 4 años, la niña a quien le leen cuentos cada noche mostrará en las neuroimágenes una activación superior que la de otras

criaturas en las regiones cerebrales para el lenguaje hablado, y más tarde precisamente estos circuitos corticales fortalecidos le permitirán entender textos o formular pensamientos complejos (Hutton y otros, 2015, 2017; véase también Romeo y otros, 2018). Del mismo modo, el niño que tiene la suerte de nacer en una familia bilingüe, y a quien sus padres le hacen el maravilloso regalo de hablarle cada cual en su lengua materna, sin esfuerzo adquiere dos léxicos, dos gramáticas y dos culturas (Bialystok y otros, 2009; Costa y Sebastián-Gallés, 2014; Li, Legault y Litcofsky, 2014). Conservará toda su vida una mejor capacidad de análisis del lenguaje y de aprendizaje de una tercera o cuarta lengua. Y cuando se acerque a la vejez, es probable que su cerebro resista durante más tiempo a los estragos de la enfermedad de Alzheimer.

En efecto, confrontar a los niños y las niñas desde la edad más temprana a entornos estimulantes les permite conservar sinapsis más cuantiosas, dendritas más vastas, circuitos más flexibles y más redundantes (Donato, Rompani y Caroni, 2013; Knudsen, Zheng y DeBello, 2000; Van Praag, Kempermann y Gage, 2000; Voss y otros, 2013; Zhu y otros, 2014), como la lechuga que había aprendido muy temprano a llevar lentes prismáticos, y por eso conservó toda su vida axones más diversificados y una mayor capacidad de pasar de un comportamiento al otro. Diversifiquemos el bagaje de aprendizajes tempranos de nuestros niños: el florecimiento de sus cerebros depende, en parte, de la riqueza de la estimulación que reciben de su entorno.

[16] Pueden consultarse Dehaene, Bossini y Giraux (1993) sobre la representación mental de la paridad; Blair y otros (2012), Fischer (2003) y Gullick y Wolford (2013) sobre la representación de los números negativos; Jacob y Nieder (2009) y Siegler, Thompson y Schneider (2011) sobre la representación de las fracciones.

[17] Alude a la ya mencionada *bosse des maths*, que también puede traducirse como “el don de las matemáticas”. Esa expresión dio título a Dehaene (2010). [N. de T.]

[18] Hay numeroso material probatorio de los efectos en el largo plazo de una educación temprana de calidad. Véanse, por ejemplo, el programa Abecedarian (Campbell y otros, 2012, 2014; Martin, Ramey y Ramey, 1990), el programa Perry de jardín de infantes (Heckman y otros, 2010; Schweinhart, 1993) y el programa de Jamaica (Gertler y otros, 2014; Grantham-McGregor y otros, 1991; Walker, Chang y otros, 2005).

Parte III

Los cuatro pilares del aprendizaje

Acabamos de ver que todos los aprendizajes dependen de que se modifiquen los circuitos del cerebro durante un período sensible en que todavía son flexibles y poseen un enorme margen de plasticidad, pese a estar parcialmente especificados por la evolución. Sin embargo, la existencia de la plasticidad sináptica no es suficiente para explicar el extraordinario éxito de nuestra especie. En efecto, dicha plasticidad está presente en todo el mundo animal: incluso la mosca, el nematodo o la babosa de mar poseen sinapsis modificables. Si nos hemos convertido en *Homo docens*, si el aprendizaje se convirtió en nuestro nicho ecológico y en la principal causa de nuestros logros planetarios, es porque el cerebro dispone de una variedad de trucos adicionales.

A lo largo de su evolución, el cerebro de los seres humanos consiguió y pulió cuatro funciones de gran importancia que maximizan la velocidad con la cual extraemos información del entorno. Las llamo “los cuatro pilares del aprendizaje”, porque son indispensables para todas las construcciones mentales que realizamos: si solo uno de estos pilares falta o es inestable, todo el edificio tambalea. A la inversa, cada vez que necesitamos aprender, y aprender rápido, podemos confiar en ellos para optimizar los esfuerzos. Estos pilares son:

- La atención, que amplifica la información sobre la que nos concentramos.
- El compromiso activo, un algoritmo que llamamos también “curiosidad”, y que incita al cerebro a evaluar constantemente nuevas hipótesis.
- La revisión o *feedback* a partir del error, que compara las predicciones con la realidad y corrige los modelos que elaboramos acerca del

mundo.

- La consolidación, que automatiza y vuelve fluido lo que aprendimos, especialmente durante el sueño.

Lejos de ser exclusivas de los humanos, estas funciones se comparten con muchas otras especies animales. Sin embargo, gracias a nuestro cerebro social y a las habilidades lingüísticas, las aprovechamos de manera más eficaz que cualquier otro animal, especialmente en nuestras familias, escuelas y universidades.

Prestar atención, involucrarse, poner a prueba lo adquirido y saber consolidarlo son los secretos de un aprendizaje exitoso. Y estos componentes fundamentales de la arquitectura cerebral se implementan con igual eficiencia en la familia y en la escuela. El maestro o la profesora que logre movilizar estas cuatro funciones en cada uno de sus alumnos sin duda maximizará la velocidad y la eficacia con que aprenderá su clase. Todos deberíamos, entonces, aprender a dominarlas, y con ese objetivo, necesitamos comprender bien cómo funcionan y para qué sirven.

7. La atención

Imaginen que llegan al aeropuerto justo a tiempo para tomar un avión. Todo lo que hacen revela la mayor concentración de la atención: con la mente alerta, buscan el anuncio de las partidas y, sin dejarse distraer por la marea de pasajeros, pasan rápidamente por las filas hasta notar cuál corresponde a su vuelo. Si bien se abren camino por entre una sucesión de estridentes publicidades, casi no las ven y avanzan en línea recta hacia el mostrador de facturación. De pronto, se dan vuelta, porque en la multitud un amigo acaba de pronunciar su nombre de pila. Este mensaje, que el cerebro evaluó como prioritario, se apodera de la atención e invade la conciencia... por lo tanto, les hace olvidar el número de ventanilla.

En unos minutos, el cerebro pasó por la mayoría de los estados clave de la atención: vigilancia y estado de alerta, selección y distracción, orientación y filtrado. En las ciencias cognitivas, llamamos “atención” al conjunto de mecanismos mediante los cuales el cerebro selecciona una información, la amplifica, la canaliza y la profundiza. Son mecanismos antiguos en la evolución: el perro que orienta sus orejas y el ratón que se queda inmóvil cuando oye un crujido despliegan circuitos atencionales muy similares a los de los seres humanos (Wang y Krauzlis, 2018).

¿Por qué la evolución incorporó mecanismos atencionales en tantas especies animales? Porque estos circuitos resuelven un problema muy específico: la saturación de información. A cada instante, el cerebro recibe un bombardeo de estímulos: los sentidos de la vista, el oído, el olfato, el tacto transmiten millones de bits de información por segundo. Al comienzo, todos estos mensajes son procesados en paralelo por neuronas distintas; pero enseguida se vuelve imposible tratarlos a todos con la misma profundidad: los recursos del cerebro no son suficientes. Por este motivo, una pirámide de mecanismos atencionales, organizados como un filtro gigantesco, realiza una selección. En cada etapa, nuestro cerebro decide qué importancia es conveniente otorgar a

tal o cual estímulo, y asigna recursos solo a la información que considera prioritaria.

Esa selección es fundamental para el aprendizaje. En ausencia de atención, descubrir regularidades en un cúmulo de datos se parece a buscar la legendaria aguja en un pajar. Este es uno de los motivos de la lentitud de las redes neuronales artificiales convencionales: pierden un tiempo considerable en analizar todas las combinaciones posibles de los datos que se les proporcionan, en lugar de clasificar la información y enfocarse solo en los bits relevantes. Recién en 2014 dos investigadores, el canadiense Yoshua Bengio y el coreano Kyunghyun Cho, expusieron cómo se puede integrar la atención a algunas redes neuronales artificiales (Bahdanau, Cho y Bengio, 2014; Cho, Courville y Bengio, 2015). Su primer modelo aprendía a traducir frases de un idioma a otro. Al dotarlo de una función atencional, lograron que el sistema aprendiera mejor y más rápido, porque en cada paso de la traducción se concentraba solo en las palabras pertinentes de la frase original.

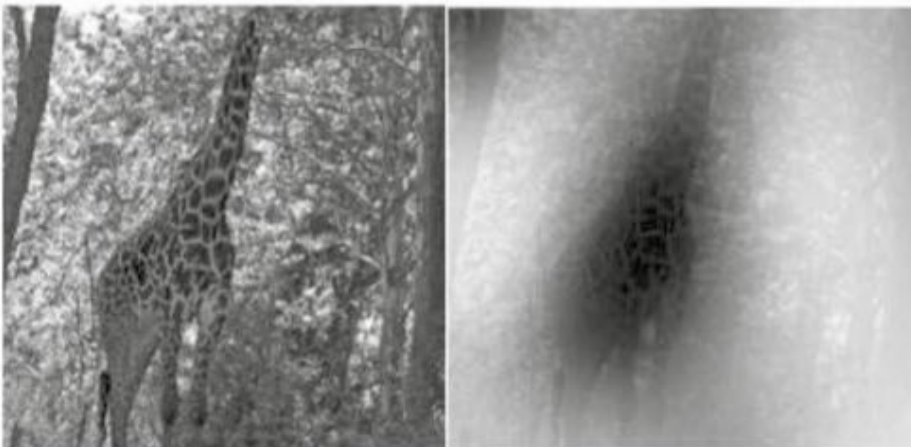
En muy poco tiempo, entre los estudiosos de la inteligencia artificial la idea de aprender a prestar atención se esparció y expandió como un incendio en un bosque. Los sistemas artificiales actuales logran describir una imagen (“una mujer que lanza un *frisbee* en un parque”) porque canalizan la información prestando atención a cada parte pertinente: en el momento de describir el color del disco, la red se concentra sobre los píxeles de la cabeza y aparta temporariamente todos cuantos corresponden a las personas o al parque, para regresar a ellos más tarde (figura 26; Xu y otros, 2015). Un ingeniero en inteligencia artificial sensato ya no conecta todas las entradas con todas las salidas. Sabe que el aprendizaje será más rápido si en lugar de una red simple, en que cada píxel de entrada tiene la posibilidad de predecir cualquier palabra en la salida, creamos una arquitectura organizada donde el aprendizaje se divide en dos módulos: uno que aprende a prestar atención y otro que aprende a nombrar los datos filtrados por el primero.



Una mujer que lanza un *frisbee* en un parque.



Una *niña pequeña* sentada en un sillón con un oso de peluche.



Una *jirafa* de pie en un bosque, con *árboles* en segundo plano.

Figura 26. La atención es el primer pilar del aprendizaje, un componente tan crucial que está integrado a la mayor parte de las redes neuronales artificiales contemporáneas. Aquí, una red aprende a encontrar las palabras para describir una imagen. El aprendizaje se basa sobre un mecanismo de atención selectiva que pone el foco en determinadas zonas de la imagen (que tienden al blanco en las imágenes de la derecha). La atención selecciona una zona y concentra allí, en un instante dado, todos los recursos.

La atención es indispensable, pero trae aparejado un problema: si está mal orientada, el aprendizaje puede atascarse (Ahissar y Hochstein, 1993). Si no se presta atención al *frisbee* que la mujer está lanzando, el procesamiento continúa como si esa parte de la imagen no existiera (y el disco mismo queda borrado). Sus datos se descartan desde el comienzo y permanecen confinados a las áreas sensoriales más tempranas. Los objetos desatendidos solo causan una activación modesta que induce poco o ningún aprendizaje (Seitz y otros, 2005; Watanabe, Nanez y Sasaki, 2001). Esto nada tiene que ver con la extraordinaria amplificación que se produce en nuestro cerebro cuando prestamos atención a un objeto y tomamos conciencia de él: las neuronas sensoriales que lo codifican se amplifican y prolongan masivamente, y sus mensajes se propagan hasta la corteza prefrontal, donde poblaciones enteras de neuronas se encienden y permanecen activas durante mucho tiempo, mucho más allá del acontecimiento originario, por ejemplo, la duración de la exposición de la imagen (Dehaene y Changeux, 2011; Van Vugt y otros, 2018). Un aumento tan fuerte de la activación neuronal es exactamente lo que necesitan las sinapsis para modificar su fuerza, lo que los neurocientíficos llaman “potenciación a largo plazo”.

Cuando un alumno presta atención consciente a, por ejemplo, una palabra en un idioma extranjero que su docente acaba de presentar, permite que esa palabra se propague profundamente en sus circuitos corticales, hasta la corteza prefrontal. Así, esa palabra tiene muchas más posibilidades de ser recordada. En gran medida, las palabras inconscientes o desatendidas permanecen confinadas a los circuitos sensoriales del cerebro y nunca consiguen alcanzar las representaciones léxicas y conceptuales más profundas que sustentan la comprensión y la memoria semántica.

Por ese motivo, cada estudiante debe aprender a prestar atención ¡y también los maestros deben prestar más atención a la atención! Si los estudiantes no atienden a la información correcta, es muy poco probable que aprendan algo. El mayor talento de un maestro consiste en canalizar y cautivar

constantemente la atención de los niños para guiarla hacia el objeto o el nivel adecuado.

La atención desempeña un papel tan crucial en la selección de información relevante que está presente en muchos circuitos diferentes en el cerebro. El psicólogo estadounidense Michael Posner diferencia al menos tres sistemas de atención principales:

1. El alerta, que indica *cuándo* prestar atención y adapta nuestro nivel de vigilancia.
2. La orientación de la atención, que muestra *a qué* prestar atención y amplifica cada objeto de interés.
3. El control ejecutivo, que decide *cómo* procesar la información a la que atendemos: selecciona los procesos que son apropiados para determinada tarea y controla su ejecución.

Cada uno de estos sistemas modula masivamente la actividad cerebral y puede entonces facilitar el aprendizaje, pero también orientarlo en la dirección equivocada. Los exploraremos uno por uno.

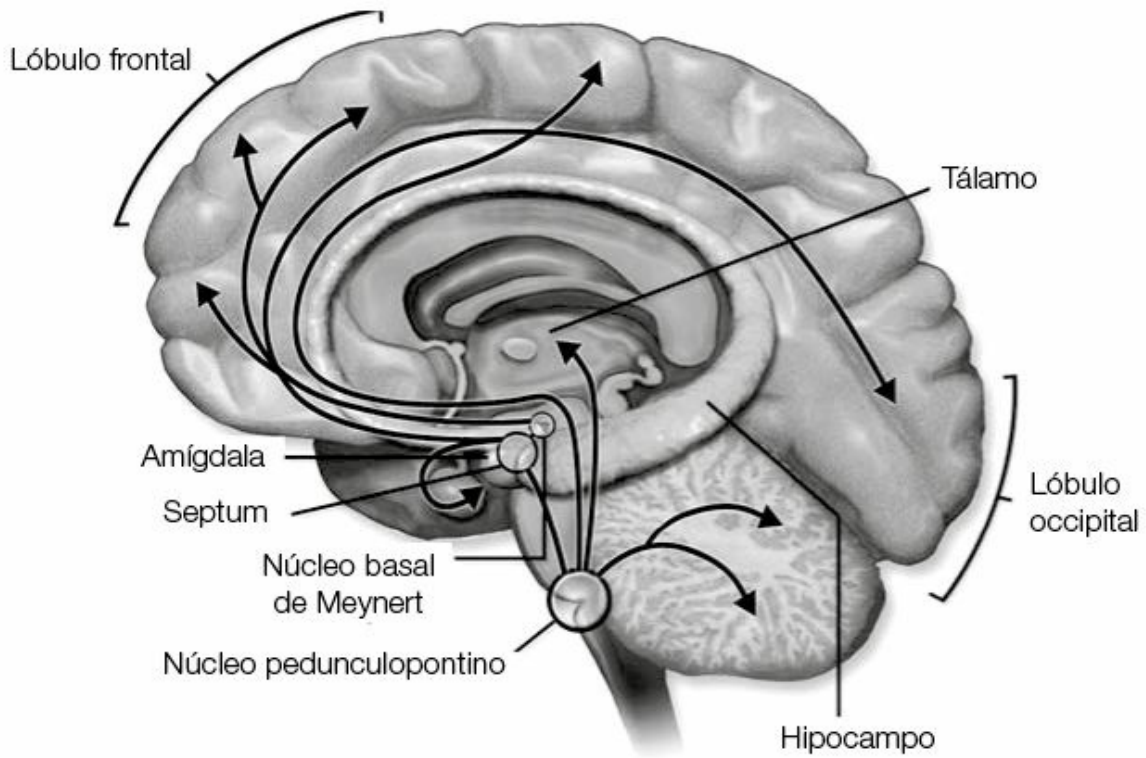
Alerta: el cerebro sabe cuándo prestar atención

El primer sistema atencional, y tal vez el más antiguo en la evolución, nos dice *cuándo* prestar atención. Envía señales de advertencia que movilizan todo el organismo en caso de que las circunstancias lo exijan. Si se acerca un predador o nos invade una emoción fuerte, una serie de núcleos subcorticales decide de inmediato aumentar la vigilia y la vigilancia de la corteza. Ordena la liberación masiva y difusa de neuromoduladores como la serotonina, la acetilcolina y la dopamina (figura 27). Con axones de largo alcance, dotados de numerosas ramas difusas como portavoces, estos mensajes de alerta alcanzan casi la totalidad de la corteza. Allí van a modular la actividad cortical y el aprendizaje, en grado tal que los investigadores hablan de una señal *now print*, “imprimir ahora”, como si estos sistemas directamente le

ordenaran a la corteza que guarde en la memoria el contenido actual de la actividad neuronal.

Un conjunto de experiencias realizadas con animales muestra que, en efecto, la descarga de los sistemas de alerta puede provocar alteraciones radicales en la plasticidad de los mapas corticales (figura 27). El neurofisiólogo estadounidense Michael Merzenich realizó numerosas experiencias en las que se confunde al animal estimulando eléctricamente sus circuitos subcorticales de la dopamina o de la acetilcolina. El resultado fue un cambio masivo en los mapas corticales. Todas las neuronas que se activaron en ese momento, incluso si no tenían importancia objetiva, estaban sujetas a una amplificación intensa.

Circuitos cerebrales de la acetilcolina



Modulación del aprendizaje auditivo

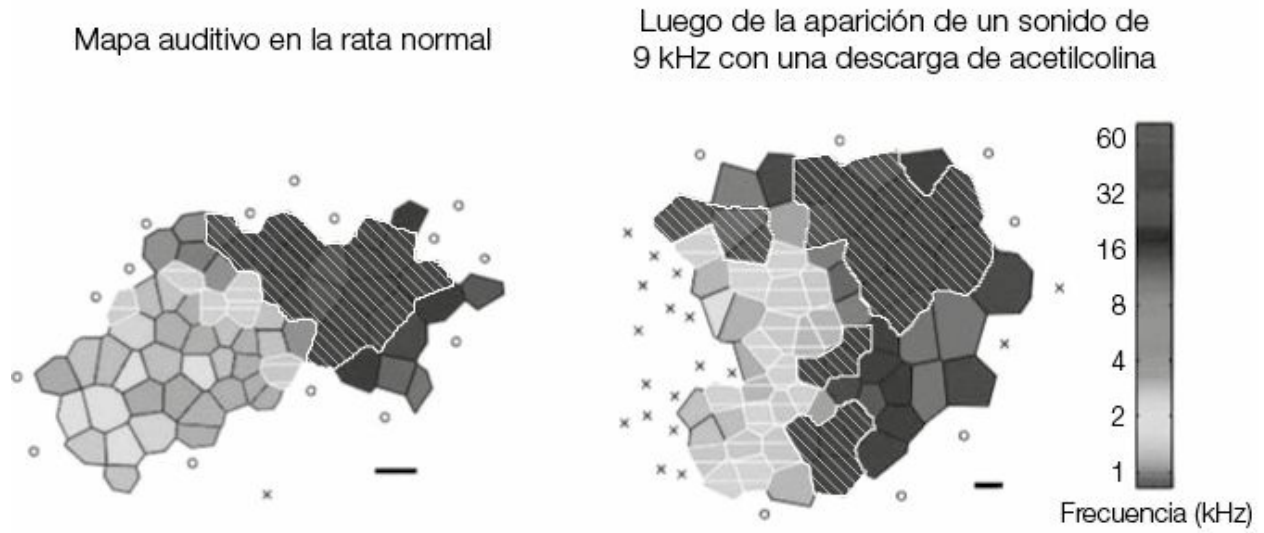


Figura 27. Las señales de alerta desencadenan el aprendizaje. Algunos neuromoduladores

como la serotonina, la acetilcolina y la dopamina nos dicen *cuándo* prestar atención y parecen forzar el cerebro a aprender. En esta experiencia, un sonido banal, de una frecuencia de 9 kHz, se asocia a la estimulación del núcleo basal de Meynert (arriba), que libera una descarga de acetilcolina en la corteza. Luego de algunos días de seguir este régimen, toda la corteza auditiva se ve invadida por esta frecuencia y sus vecinas (abajo, regiones marcadas con rayas).

De este modo, cuando un sonido –por ejemplo, un la sostenido o la# o A#, según la notación que prefieran– se asocia sistemáticamente a una descarga de dopamina o de acetilcolina, todo el mapa auditivo se modifica y termina invadido por esa nota. El animal se torna cada vez más hábil para discriminarla, pero pierde la capacidad de representar otras frecuencias (Bao, Chan y Merzenich, 2001; Froemke, Merzenich y Schreiner, 2007; Kilgard y Merzenich, 1998).

Es notable que esta plasticidad cortical inducida por el sistema de alerta tenga lugar incluso en el animal adulto. El análisis de los circuitos implicados revela que los neuromoduladores como la serotonina y la acetilcolina, en especial por obra del receptor nicotínico (sensible a la nicotina, otro actor importante en la vigilia y el estado de alerta), modulan la actividad de las neuronas inhibitoras de la corteza, y afectan así el equilibrio entre la excitación y la inhibición (Werker y Hensch, 2014). Recuerden que la inhibición interviene en el cierre de los períodos sensibles para la plasticidad sináptica. Al estar desinhibidos por las señales de alerta, los circuitos corticales parecen recuperar su plasticidad juvenil: el período sensible se reabre para las señales etiquetadas como críticas.

¿Qué sucede en el *Homo sapiens*? Es tentador pensar que una reorganización similar de los mapas corticales se produce cada vez que, llevado por la pasión, un compositor o un matemático se sumerge en su ámbito predilecto, sobre todo si esa pasión comienza a una edad temprana. Un Mozart o un Ramanujan tal vez estén tan exaltados, electrizados, por el fervor con que su cerebro se ve literalmente invadido de modelos mentales de música o de matemáticas. Con todo, tiendo a pensar que esto puede aplicarse no solo a los genios, sino a cualquier persona capaz de apasionarse lo suficiente por su trabajo, ya sea un alfarero o una científica que diseña transbordadores espaciales. Al permitir que los mapas corticales se remodelen a escala masiva, la pasión genera talento.

Pese a que no todo el mundo es Mozart, los mismos circuitos cerebrales del alerta y de la motivación están presentes en todos los cerebros. ¿Qué

circunstancias de la vida cotidiana lograrían movilizarlos? ¿Hace falta esperar una situación traumática o una emoción fuerte? Tal vez no. Algunos estudios sugieren que los videojuegos, sobre todo los de acción, en los que se juega a la vida y la muerte, entrañan un medio impetuoso de involucrar estos mecanismos atencionales. Al movilizar los sistemas del alerta y de la recompensa, modulan los aprendizajes con enorme eficacia. El circuito de la dopamina, por ejemplo, se enciende cuando jugamos a un juego de acción (Koepp y otros, 1998). La neurocientífica Daphne Bavelier demostró que esto se traduce en un aprendizaje rápido (Bavelier y otros, 2011; Cardoso-Leite y Bavelier, 2014; Green y Bavelier, 2003). Los juegos de acción más violentos parecen tener los efectos más intensos: diez horas de juego son suficientes para mejorar la detección visual, refinar la estimación rápida de la cantidad de objetos presentes en la pantalla y fortalecer la capacidad de concentrarse en un objetivo sin dejarse distraer por otros. Un jugador de videojuegos toma decisiones ultrarrápidas sin comprometer su rendimiento.

Los padres, las madres y los docentes se quejan constantemente de que los niños de hoy en día, criados entre computadoras, tablets, consolas y otras pantallas, cambian de una actividad a otra en permanente estado de distracción, pero esto no es cierto. Lejos de disminuir la capacidad de concentración, los videojuegos también pueden aumentarla. ¿Sabremos, en el futuro, volver a movilizar la plasticidad sináptica en todos los niños? Dado que los videojuegos son un poderoso estimulante, con el transcurso de los años mi laboratorio desarrolló una serie de juegos para tablets basados en los principios de las ciencias cognitivas.[\[19\]](#)

Pero los videojuegos también tienen su costado oscuro: implican riesgos ya muy conocidos, como el aislamiento social, la pérdida de tiempo y la adicción. Por suerte, existen también muchas otras formas de desbloquear los efectos del sistema de alerta sin dejar de aprovechar el sentido social del cerebro. Los profesores que cautivan a sus alumnos, los libros que dejan absortos a sus lectores, las películas o las obras de teatro que transportan a los espectadores y los sumergen en experiencias de la vida real son, sin duda, señales de alerta igualmente poderosas que estimulan la plasticidad cerebral.

Orientación: el cerebro sabe a qué prestar atención

El segundo sistema atencional del cerebro decide *a qué* le prestamos atención. Como un poderoso reflector enfocado en el mundo exterior, este sistema de orientación selecciona, entre los millones de estímulos que nos bombardean, aquello a qué adjudicarle recursos mentales, porque evaluamos que es algo urgente, peligroso, atractivo... o simplemente relevante para nuestros objetivos actuales.

El gigante de la psicología estadounidense, William James, fue quien mejor definió esta función de la atención; en sus *Principios de psicología* (1890) sostuvo:

Millones de elementos del mundo exterior se presentan a mis sentidos, pero nunca ingresan en mi experiencia consciente. ¿Por qué? Porque no tienen interés alguno para mí. Lo que percibo es aquello a lo que acepto prestar atención. Solo los elementos que detecto afectan mi pensamiento; si no existe selección, la percepción es un caos total. Solo el interés le da acento y énfasis, luz y sombra, un segundo plano, un primer plano, en una palabra, una perspectiva inteligible. La atención varía de una criatura a la otra, pero sin ella la conciencia de cada criatura sería un desorden brumoso y sin discernimiento, que es prácticamente imposible concebir.

La atención selectiva obra en todos los ámbitos sensoriales, incluso en los más abstractos. Por ejemplo, podemos orientar la atención a los sonidos que nos rodean: los perros mueven sus orejas, pero en nosotros, integrantes de la especie humana, lo que se mueve es solo un puntero interno que sintoniza con lo que decidimos enfocar. En una fiesta ruidosa, podemos seleccionar una entre diez conversaciones, sobre la base de la voz y el significado que nos interesan. En la visión, la orientación de la atención suele ser más obvia: por lo general, movemos la cabeza y los ojos hacia lo que nos atrae. Al desplazar la mirada, llevamos el objeto de interés al centro de la fóvea, que es un área de gran sensibilidad en el centro de la retina. Incluso sin mover los ojos, podemos prestar atención a cualquier lugar o cualquier objeto, donde sea que esté, y amplificar sus características (Posner, 1994). O hasta podemos seleccionar uno entre varios dibujos superpuestos, al igual que una entre varias conversaciones simultáneas. Y nada impide prestar atención al color de

una pintura, la forma de una curva, la velocidad de un corredor, el estilo de un escritor o la técnica de un pintor. Cualquier representación en nuestro cerebro puede convertirse en el foco de atención.

En todos los casos, el efecto es el mismo: la orientación de la atención amplifica aquello que está en su foco. Las descargas de las neuronas que codifican la información a la cual se presta atención se incrementan, mientras se aplaca el ruidoso parloteo de las otras. El impacto es doble: la atención hace que estas neuronas se vuelvan más sensibles a la información que juzgamos pertinente, pero, por sobre todas las cosas, aumenta su influencia en el resto del cerebro. Los circuitos posteriores se hacen eco del estímulo que recibimos con el ojo, el oído o la mente. Por último, grandes extensiones de la corteza se reorientan para codificar cualquier información que esté en el centro de la atención (Çukur y otros, 2013; Desimone y Duncan, 1995; Kastner y Ungerleider, 2000). La atención actúa como un amplificador y un filtro selectivo.

El ensayista Alain sostuvo: “El arte de prestar atención, que es el gran arte, supone el arte de no prestar atención, que es el verdadero arte”. En efecto, prestar atención implica también elegir qué elementos se pasarán por alto. Para que un objeto quede dentro del cono de luz proyectada por ese reflector que es la atención, millares de otros deben permanecer en la sombra. Orientar la atención supone elegir, filtrar. Esta atención selectiva amplifica la señal seleccionada, pero también reduce drásticamente aquellas que se consideran irrelevantes. La designación técnica para este mecanismo es “competencia sesgada”: en cualquier momento, muchos *inputs* sensoriales compiten por los mismos recursos, y la atención otorga una orientación a esta competencia al fortalecer la representación del elemento seleccionado mientras suprime los demás. Para entonces, la metáfora del reflector alcanza sus límites: en busca de iluminar mejor una región de la corteza, el foco atencional también reduce la iluminación de las restantes. El mecanismo se vale de ondas que interfieren la actividad eléctrica: para suprimir un área, la inunda con ondas lentas en el rango de frecuencia “alfa” (entre 8 y 12 Hz), que parece utilizarse específicamente para inhibir un circuito y evitar que desarrolle una actividad neuronal coherente.

Por lo tanto, prestar atención consiste en suprimir la información no deseada; al hacerlo, el cerebro se expone al riesgo de quedar ciego a lo que elige no ver. ¿Ciego, de verdad? Sí, de verdad. El término es completamente apropiado: una experiencia célebre, la del “gorila invisible”, es un

maravilloso ejemplo de la ceguera que causa la falta de atención (Mack y Rock, 1998; Simons y Chabris, 1999). En este clásico experimento, se pide a los participantes que miren una breve película en la que dos equipos de basquetbolistas, vestidos uno de blanco y el otro de negro, se hacen pases. La propuesta consiste en contar el número de pases del equipo blanco. Nada más fácil que esto –piensan los voluntarios– y, de hecho, treinta segundos más tarde contestan de modo triunfal, con el número correcto. Pero hete aquí que el experimentador hace una extraña pregunta: “¿Viste al gorila?”. ¿El gorila? ¿Qué gorila? Al rebobinar la cinta –sí, todavía son los años del VHS– oh, sorpresa: aparece alguien que siempre estuvo ahí. Es un actor disfrazado de gorila que ingresa en la escena, la cruza y, cual émulo de King Kong, se golpea el pecho. Resulta imposible perderselo, y hasta se puede probar que los ojos estuvieron posados en él. Si no lo vieron, fue porque, al estar concentrados en los jugadores del equipo blanco, estaban inhibiendo a los personajes vestidos de negro... ¡gorila incluido! Su espacio de trabajo mental, ocupado con la tarea de contar los pases, fue incapaz de tomar conciencia de este cuadrúmano incongruente. El experimento del gorila invisible constituye un descubrimiento fundamental de las ciencias cognitivas, muchas veces recreado en diferentes condiciones: el simple hecho de focalizar la atención en un objeto del pensamiento nos vuelve ciegos a otros estímulos. Por ejemplo, cuando se le pide a alguien que juzgue si el tono de un sonido es alto o bajo, puede quedar “ciego” a otro estímulo (por ejemplo, una palabra escrita que aparece en la siguiente fracción de segundo). Los psicólogos conocen este fenómeno como “parpadeo atencional”: si bien los párpados permanecen completamente abiertos, la mente “parpadea” durante un lapso breve.

En experimentos como ese, en realidad sufrimos dos ilusiones distintas. La primera consiste en que no vemos la palabra o el gorila, lo cual es bastante malo (otras investigaciones demuestran que la falta de atención puede llevarnos a pasar por alto una luz roja o atropellar a un peatón, ¡lo que nos recuerda que nunca debemos usar el teléfono cuando estamos al volante!). Pero la segunda ilusión es aún peor: no nos damos cuenta de nuestra propia inconciencia; por lo tanto, ¡estamos absolutamente convencidos de que hemos visto todo lo que había que ver! La mayoría de las personas que pasan por el experimento del gorila invisible por primera vez no pueden creer su propia ceguera. Piensan que se les jugó una mala pasada; por ejemplo, que les cambiaron la película. Dicen: “¡Si hubiera un gorila en el corto, lo habría

visto!”. Por desdicha, esto es falso: nuestra atención es extremadamente limitada y, a pesar de toda la buena voluntad, cuando los pensamientos se concentran en un objeto, otros, por destacados, divertidos o importantes que parezcan, pueden ser eludidos sin más y volverse invisibles. Los límites intrínsecos de nuestra conciencia nos llevan a sobreestimar lo que somos capaces de percibir.

La experiencia del gorila realmente merece ser conocida por todos; en especial, por las familias y los docentes. Cuando enseñamos, tendemos a olvidar qué significa no saber. Pensamos que todo el mundo puede ver aquello que vemos. Eso causa que a menudo nos cueste entender por qué, aun con las mejores intenciones, un niño no puede ver, en el sentido más literal del término, lo que estamos tratando de enseñarle. Sin embargo, la experiencia deja un claro mensaje: para ver, hay que prestar atención. Si los estudiantes están distraídos o no comprenden a qué deben prestar atención, permanecen por completo ajenos al mensaje de su maestro; simplemente no lo ven; y lo que no se ve no puede aprenderse (Leong y otros, 2017).

Contemplemos, por ejemplo, una experiencia reciente del psicólogo estadounidense Bruce McCandliss sobre la importancia de la atención en el aprendizaje de la lectura (Yoncheva y otros, 2010). ¿Es mejor prestar atención a cada una de las letras o a la forma global de las palabras? Para saberlo, los investigadores le enseñaron a un grupo de adultos un sistema de escritura no habitual, formado por elegantes curvas. Primero, los sujetos recibían un entrenamiento con dieciséis palabras, y luego se les escaneaba el cerebro mientras intentaban leer esas mismas dieciséis palabras y otras dieciséis palabras nuevas. Sin embargo, sin que lo supieran, también se manipulaba su atención (figura 28). A la mitad de ellos se les decía que hacía falta prestar atención a todas las curvas, porque cada una de ellas, como en un carácter chino, correspondía a una palabra. Al otro grupo se le explicaba que, en realidad, los conjuntos de curvas estaban formados por letras superpuestas –las palabras eran de tres letras– y que debían prestar atención a cada una de ellas. Por lo tanto, el primer grupo prestó atención en el nivel de la palabra completa, mientras que el segundo atendió a las correspondencias locales de letras y sonidos que realmente se habían utilizado para escribir las palabras.

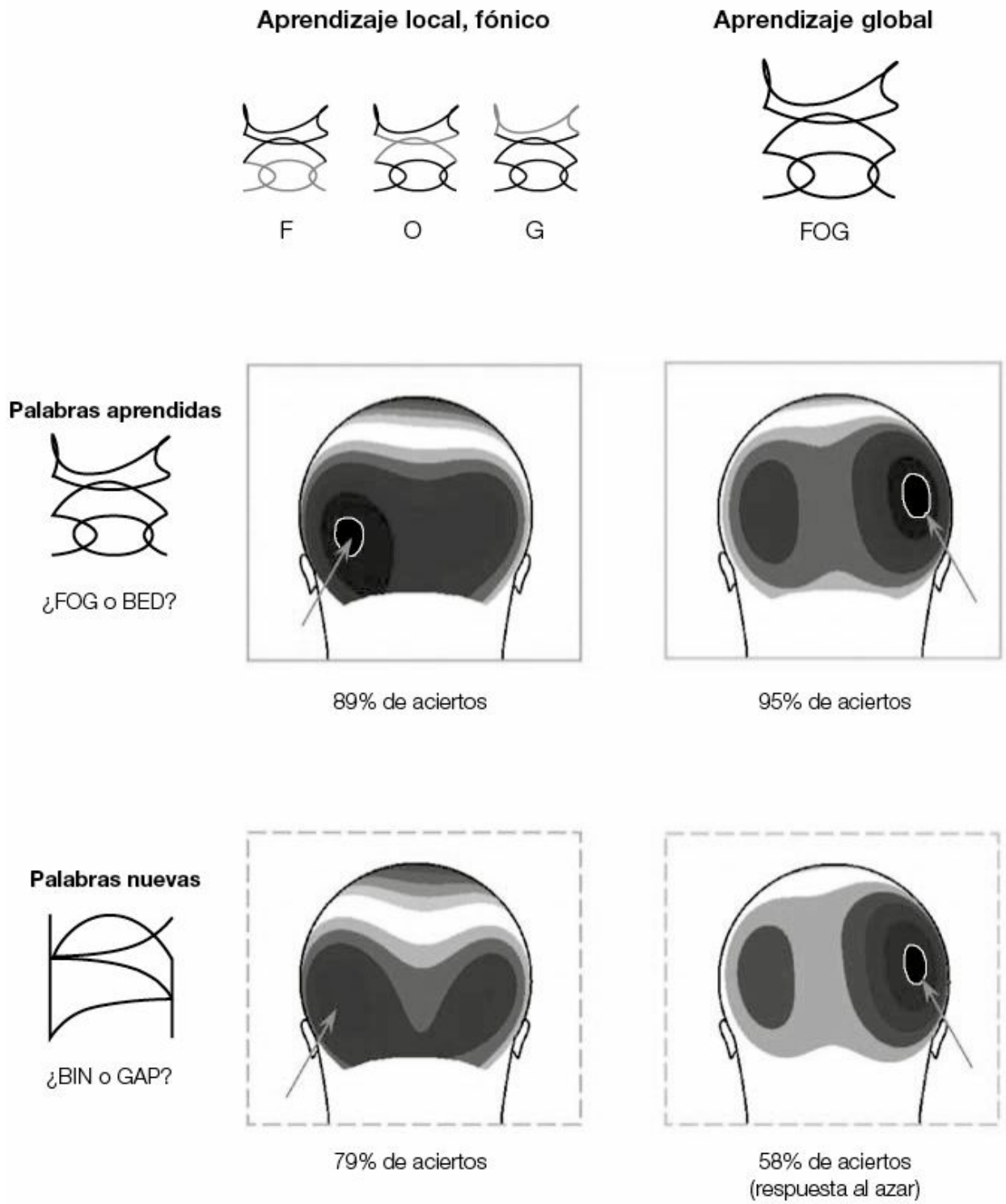


Figura 28. La orientación selectiva de la atención modula enormemente el aprendizaje. En esta experiencia, se simula el aprendizaje de la lectura de acuerdo con el método global y de

acuerdo con el método fónico. Un grupo de personas presta atención a la forma global de las palabras. Luego de trescientos ensayos, todavía no logran percibir que las palabras están formadas por letras. La atención global orienta la lectura hacia un circuito inapropiado del hemisferio derecho y les impide generalizar el aprendizaje a palabras nuevas. Solo cuando se presta atención al nivel de las letras (aprendizaje local y fónico) las personas logran descifrar el alfabeto y leer palabras nuevas, utilizando el circuito apropiado del hemisferio izquierdo.

¿Cuáles fueron los resultados? Los dos grupos lograron recordar las primeras dieciséis palabras, pero la atención generó alteraciones radicales en su capacidad para descifrar nuevas palabras. El segundo grupo, centrado en las letras, había descubierto muchas de las correspondencias entre letras y sonidos y pudo leer el 79% de las palabras nuevas. Además, un examen de sus cerebros reveló que habían activado los circuitos de lectura normales, localizados en las áreas visuales ventrales del hemisferio izquierdo. Por el contrario, en el primer grupo, prestar atención a la forma global de las palabras obstaculizó por completo la capacidad de generalizar a elementos nuevos, es decir, el aprendizaje: estos voluntarios no lograban de ninguna manera leer palabras nuevas, y activaban un circuito completamente inapropiado, situado en las áreas visuales del hemisferio derecho.

El mensaje es claro: la atención produce modificaciones drásticas en la actividad cerebral. Prestar atención a la forma global de las palabras impide descubrir el código alfabético y orienta los recursos del cerebro hacia un circuito inadecuado. Para aprender a leer, solo el entrenamiento fónico, que concentra la atención en las correspondencias entre las letras y los sonidos, activa el circuito de la lectura y permite el aprendizaje. Cada docente de primer grado debería conocer estos datos: comprendería cuán importante es dirigir la atención de los niños hacia cada letra sucesiva, y lograr que aprendan a decodificarla como sonido. Muchos datos convergentes demuestran de manera convincente la superioridad del enfoque fónico por sobre la lectura de la forma global de la palabra (Castles, Rastle y Nation, 2018; Dehaene, 2007a; Ehri y otros, 2001; National Institute of Child Health and Human Development, 2000). Cuando un niño contempla desde el nivel de letra –por ejemplo, siguiendo cada una con su dedo, de izquierda a derecha–, el aprendizaje resulta tanto más fácil. Si en cambio el niño no recibe pista alguna de atención y examina ingenuamente la palabra escrita en su conjunto, sin prestar atención a su estructura interna, no sucede nada, porque no nota qué estamos buscando transmitirle. La atención es un ingrediente clave del aprendizaje exitoso.

Por eso, y de modo primordial, una buena enseñanza requiere atención permanente a la atención de los niños. Los maestros deben elegir con sumo cuidado hacia dónde quieren dirigir la atención de los niños, porque solo los elementos que están en el foco de atención se representan en el cerebro con la fuerza suficiente para lograr un aprendizaje eficaz. Los demás estímulos (los perdedores de la competencia atencional) causan poca o ninguna agitación dentro de las sinapsis plásticas del niño.

Control ejecutivo: el cerebro sabe cómo procesar la información

El tercer y último sistema atencional determina *cómo* se procesa la información seleccionada. Lo designamos con el nombre de “ejecutivo central” o “sistema de control ejecutivo”. Está conformado por un conjunto complejo de procesos mentales que nos permite seleccionar un plan de acción y atenernos a él (D’Esposito y Grossman, 1996; Koechlin, Ody y Kouneiher, 2003; Rouault y Koechlin, 2018). Implica una estructura de regiones situadas principalmente en la corteza prefrontal, esa masa de corteza detrás de la frente, que ocupa casi un tercio de nuestros hemisferios cerebrales. Comparados con los de otros primates, los lóbulos frontales de los seres humanos son gigantescos, están mejor conectados y albergan neuronas con arborescencias dendríticas notoriamente más grandes y complejas (Elston, 2003; Sakai y otros, 2011; Schoenemann, Sheehan y Glotzer, 2005; Smaers y otros, 2017). De este modo, nuestro dispositivo de supervisión cortical está mucho más desarrollado que en los otros primates, especialmente en el nivel más alto de la jerarquía, el que nos permite supervisar las operaciones mentales y tomar conciencia de los errores: el sistema de control ejecutivo (Fleming y otros, 2010; Koechlin, Ody y Kouneiher, 2003; Rouault y Koechlin, 2018).

Imaginen que deben multiplicar mentalmente 23×8 . El sistema de control ejecutivo es el encargado de asegurar el correcto desarrollo de la serie completa de operaciones mentales pertinentes, desde el principio hasta el final: prestar atención a la cifra de la derecha (3) y multiplicarla por 8; luego almacenar el resultado en la memoria (24); prestar atención a la cifra de la izquierda (2) y multiplicarla también por 8 para obtener 16; recordar que en

realidad se trata de la columna de las decenas, por lo que el resultado es 160, y, por último, sumar para obtener el resultado final: $160 + 24 = 184$.

El control ejecutivo funciona como un panel de mando: orienta, dirige y gobierna los procesos cerebrales. Forma parte del sistema atencional porque, al igual que el alerta y la orientación, obra como un mecanismo de selección, pero en vez de cribar los estímulos que debemos procesar entre todos los que recibimos, elige entre las operaciones mentales disponibles la más adecuada para tratar la información. Así, la atención espacial y la atención ejecutiva se complementan. Cuando hacemos aritmética mental, por ejemplo, la atención espacial es el sistema que recorre la página del manual de matemáticas y pone el foco sobre el problema 23×8 , pero ese foco es guiado paso a paso por la atención ejecutiva, que en primer lugar se concentra en el 3 y el 8, luego envía estas cifras hacia los circuitos de la multiplicación, y así sucesivamente. La atención activa las operaciones pertinentes e inhibe las acciones inapropiadas; supervisa permanentemente que el programa mental se desarrolle sin inconvenientes y decide cuándo es pertinente cambiar de estrategia. También es la encargada de detectar, en un subcircuito especializado de la corteza cingulada, cuándo cometemos un error o nos desviamos de la meta, para enseguida corregir el plan de acción.

Hay un vínculo estrecho entre el control ejecutivo y lo que llamamos “memoria de trabajo”. Seguir un algoritmo mental y controlar su cumplimiento implica conservar un recuerdo de todos los elementos del programa en uso: resultados intermedios, etapas efectuadas, operaciones por realizar... En especial, la atención ejecutiva controla las entradas y las salidas de lo que llamé “espacio de trabajo neuronal global”: una memoria temporaria consciente, en la cual podemos ingresar casi cualquier información que nos parezca pertinente (Dehaene y Changeux, 2011; Dehaene y otros, 2006a; Dehaene, Kerszberg y Changeux, 1998; Dehaene y Naccache, 2001). Se trata del *router* del cerebro, el panel de mando que decide de qué manera y en qué orden enviar la información a los diferentes procesadores incluidos en el cerebro. En este nivel, las operaciones mentales son lentas y trabajan en serie: este sistema procesa una porción de información por vez, ya que es incapaz de realizar dos operaciones al mismo tiempo. En psicología, se habla de un “cuello de botella central”.

¿Realmente somos incapaces de ejecutar dos programas mentales a la vez? A veces tenemos la sensación de que podemos realizar dos tareas distintas o de que el pensamiento puede dividirse y seguir dos líneas distintas en

simultáneo, pero esto también es pura ilusión. Un experimento básico lo pone de manifiesto: podemos asignar a una persona dos tareas muy simples, por ejemplo, presionar una tecla con la mano izquierda siempre que oiga un sonido agudo y presionar otra tecla con la mano derecha si ve la letra Y. Cuando los dos objetivos se presentan en simultáneo, la persona realiza la primera tarea a la velocidad normal, pero la segunda se torna sumamente lenta, en proporción directa con el tiempo que insumió tomar la primera decisión (Chun y Marois, 2002; Marti, King y Dehaene, 2015; Marti, Sigman y Dehaene, 2012; Sigman y Dehaene, 2008). En otras palabras, la primera tarea retrasa la segunda: mientras el espacio de trabajo global está ocupado con la primera decisión, la segunda tiene que esperar. Y el retraso es enorme: alcanza fácilmente unos cientos de milisegundos. Si uno está demasiado concentrado en la primera tarea, incluso puede perder por completo el segundo objetivo. Sin embargo, para sorpresa de varios, no somos capaces de registrar este gran retraso, lo cual resulta lógico, ya que, por definición, no podemos acceder a la información que aún no ingresó en nuestro espacio de trabajo consciente. Mientras el primer estímulo es procesado por la conciencia, el segundo tiene que esperar ante las puertas hasta que el espacio de trabajo global esté libre. Lo cierto es que no tenemos introspección de ese tiempo de espera; si se nos pregunta al respecto, contestaremos con seguridad que el estímulo apareció exactamente cuando terminamos con el primero, y que lo procesamos a una velocidad normal (Corallo y otros, 2008; Marti, Sigman y Dehaene, 2012).

Una vez más, notamos que no somos conscientes de nuestros límites mentales (de hecho, sería paradójico si de alguna manera pudiéramos ser conscientes de nuestra falta de conciencia). La única explicación de que nos creamos capaces de realizar varias tareas en simultáneo es nuestra falta de conciencia acerca del gran retraso que causa. Por lo tanto, muchos de nosotros seguimos enviando mensajes de texto mientras conducimos, pese a todas las pruebas de que es una de las actividades que más distracción causa. El atractivo de la pantalla y el mito de la multitarea forman parte de la ideología más peligrosa de nuestras sociedades digitales.

¿Y si nos entrenamos? ¿Podemos convertirnos en auténticos individuos multitarea que hagan varias cosas a la vez? Quizá sí, pero solo con un entrenamiento intenso en una de las dos tareas. La automatización libera el espacio de trabajo consciente: si a fuerza de mucha repetición automatizamos una actividad, permitimos que se realice inconscientemente y, por lo tanto,

sin comprometer los recursos centrales del cerebro. Gracias a la práctica, por ejemplo, una pianista profesional puede hablar mientras toca, o un mecanógrafo logra copiar un documento mientras escucha la radio. Sin embargo, estas son raras excepciones, y los psicólogos continúan debatiéndolas, porque también es posible que la atención ejecutiva cambie rápidamente de una tarea a otra de una manera casi indetectable (Tombu y Jolicœur, 2004). La regla básica queda invariada: en cualquier situación multitarea, siempre que debamos realizar múltiples operaciones cognitivas bajo el control de la atención, al menos una de ellas desacelerará o quedará suspendida por completo.

Debido a este efecto severo de la distracción, aprender a concentrarse es un ingrediente esencial del aprendizaje. No se le puede exigir a un niño o a un adulto que aprenda dos cosas a la vez. Aprender exige dar prioridad a una tarea precisa. Y por lo tanto enseñar requiere prestar atención a los límites de la atención. Cualquier distracción lentifica o aniquila nuestros esfuerzos: si intentamos hacer varias cosas a la vez, el ejecutivo central pierde con suma rapidez el rumbo. Al respecto, los experimentos de ciencias cognitivas en el laboratorio convergen muy bien con los hallazgos educativos. Por ejemplo, los estudios de campo demuestran que un aula excesivamente decorada distrae a los niños y les impide concentrarse (Fisher, Godwin y Seltman, 2014). Otra investigación reciente demuestra que, cuando a los estudiantes se les permite usar sus teléfonos inteligentes en clase, su rendimiento se ve afectado, incluso meses después, cuando el examen se refiere específicamente al contenido trabajado ese día. Para un aprendizaje óptimo, el cerebro debe evitar cualquier distracción (Glass y Kang, 2018).

Aprender a prestar atención

La atención ejecutiva se corresponde de manera aproximada con lo que conocemos como concentración o autocontrol. Sin embargo, este sistema no se pone en funcionamiento de inmediato en el niño: harán falta quince o veinte años para que la corteza prefrontal alcance su plena madurez. El control ejecutivo surge a lo largo de la infancia y la adolescencia, a medida que el cerebro, asistido por la experiencia y la educación, aprende, con

lentitud, de modo gradual, a controlarse. Así, se requiere mucho tiempo para que el ejecutivo central del cerebro logre seleccionar sistemáticamente las estrategias apropiadas e inhibir las inadecuadas, mientras evita la distracción.

La psicología cognitiva abunda en ejemplos en que el niño supera de a poco sus errores, a medida que aumenta su capacidad de concentrarse y de inhibir las estrategias inapropiadas. El psicólogo Jean Piaget fue el primero en notarlo: los niños muy pequeños a veces cometen errores aparentemente tontos. Si, por ejemplo, escondemos muchas veces un objeto en el lugar A y luego lo escondemos en el lugar B, un bebé de menos de 1 año continuará buscándolo en el punto A (incluso si vieron perfectamente lo que sucedió). Es el famoso error “A no B”, que llevó a Piaget a la conclusión de que el bebé no domina la permanencia del objeto: el conocimiento de que un objeto no deja de existir cuando está escondido. Sin embargo, en la actualidad sabemos que esta interpretación es incorrecta. El examen de los ojos de los bebés confirma que tienen clara noción de dónde está el objeto oculto, por lo menos desde que tienen algunos meses de edad. Su dificultad, en cambio, está en resolver conflictos mentales: en la tarea A no B, la respuesta de rutina que aprendieron en ensayos anteriores les dice que se dirijan a la ubicación A, mientras que su memoria de trabajo más reciente les indica que, en ese mismo ensayo, deben inhibir esta respuesta habitual y buscar en la ubicación B. Antes de los 10 meses, es el hábito lo que prima. Alcanzada esa edad, lo que falta es el control ejecutivo, y no el conocimiento del objeto. El error “A no B” desaparece alrededor de los 12 meses, en relación directa con el desarrollo de la corteza prefrontal (Diamond y Doar, 1989; Diamond y Goldman-Rakic, 1989).

Otro error típico de los niños es la confusión entre número y tamaño. También respecto de esta cuestión, Jean Piaget hizo un descubrimiento decisivo, pero se equivocó en su interpretación. Notó que los niños pequeños, antes de los 3 años, tenían problemas para juzgar el número de un conjunto de objetos. En sus experimentos clásicos de conservación del número, Piaget primero mostró a los niños dos hileras iguales de canicas o bolitas, en correspondencia una a una, de modo que incluso los niños más pequeños estarían de acuerdo en que tenían el mismo número; más tarde, distribuyó una de las hileras colocando los objetos más separados:

000000 000000 → 0 0 0 0 0 0 000000

De manera sorprendente, los niños pasaban a decir que los dos conjuntos ya no eran iguales, y que la hilera más larga tenía mayor cantidad de objetos. A diferencia de lo que pensaba Piaget, eso no significa que los niños de esta edad no dominan la “conservación del número”: como ya vimos, los recién nacidos poseen un sentido abstracto del número, independiente del espacio o incluso de la modalidad sensorial. No, lo que sucede, una vez más, es que la dificultad proviene del control ejecutivo. El niño debe aprender a inhibir una característica destacada (el tamaño) y a amplificar un rasgo más abstracto (el número). Incluso los adultos se equivocan en este ejercicio: no es fácil decir que un conjunto de puntos tiene menos objetos, cuando esos objetos son particularmente grandes y están desplegados en el espacio, y tenemos mucha dificultad para elegir el número más alto entre 7 y 9. Lo que se desarrolla con la edad y la educación no es tanto la precisión intrínseca del sistema numérico como la capacidad de utilizarla con eficacia y concentración, más un filtrado de los indicios no pertinentes, como la densidad o el tamaño (Borst y otros, 2013; Piazza y otros, 2018; Poirel y otros, 2012). No cabe otra opción que repetirlo: el progreso en tareas como estas no interrumpe su relación estrecha con el desarrollo de las respuestas neuronales en la corteza prefrontal (Viswanathan y Nieder, 2015).

Podría multiplicar los ejemplos: en todas las edades de la vida, en todos los dominios del saber (cognitivos o emocionales), la capacidad de control ejecutivo es la que, al desarrollarse, nos permite evitar los errores (Houdé y otros, 2000; Isingrini, Perrotin y Souchay, 2008; Posner y Rothbart, 1998; Sheese y otros, 2008; Siegler, 1989). Nombren, sin equivocarse, el *color de la tinta* (negra o blanca) en que están impresas las palabras que siguen:

perro casa bien porque sofá demasiado **blanco** negro **blanco** **negro**
blanco **negro**...

La contundente desaceleración en la segunda mitad de la lista demuestra que entró en escena el control ejecutivo, que debe aprender a inhibir la lectura de las palabras para concentrarse en su color.

Ahora, intenten resolver el siguiente problema:

María usa veinticinco lápices, esto es, cinco más que Gregorio.
¿Cuántos lápices usa Gregorio?

¿Quién no sintió la necesidad de luchar contra el impulso de sumar las dos cifras? El enunciado del problema emplea la palabra “más” cuando es necesario restar; es una trampa en la cual caen muchos niños, antes de lograr autocontrol y reflexionar con mayor profundidad sobre el sentido de un problema matemático o de lógica en busca de seleccionar la operación aritmética que corresponde.

La atención y el control ejecutivo se desarrollan espontáneamente con la gradual maduración de la corteza prefrontal, que sucede durante las primeras dos décadas de nuestra vida. Sin embargo, este circuito, como los demás, es plástico, y muchas investigaciones demuestran que su desarrollo se puede acelerar con el entrenamiento y la educación (Diamond y Lee, 2011; Habibi y otros, 2018; Jaeggi y otros, 2011; Klingberg, 2010; Moreno y otros, 2011; Olesen, Westerberg y Klingberg, 2004; Rueda y otros, 2005). Dado que ese sistema interviene en tareas cognitivas muy diversas, numerosas actividades educativas, incluidas las más lúdicas, desarrollan el control ejecutivo. Michael Posner fue uno de los primeros en desarrollar *software* educativo para mejorar la capacidad de concentración de los niños muy pequeños. Uno de ellos, por ejemplo, los obliga a prestar atención a la orientación de un pez en el centro de la pantalla. En la condición crucial, el pez objetivo está rodeado por otros que miran en la dirección contraria. En el transcurso del juego, que abarca muchos niveles de dificultad creciente, el niño aprende a evitar que los peces contiguos lo distraigan, una tarea sencilla que enseña concentración e inhibición. Esta es solo una de las muchas maneras de alentar la reflexión y desalentar la respuesta inmediata, refleja.



Mucho antes de que se inventaran las computadoras, la médica y pedagoga italiana Maria Montessori ya había notado en qué medida una variedad de actividades prácticas pueden desarrollar la concentración en niños pequeños. Por ejemplo, en las escuelas Montessori de hoy en día se le pide a un niño que camine sobre una elipse trazada en el suelo, sin salirse de esa línea. Cuando el niño logra hacerlo, se aumenta la dificultad: se lo hace recomenzar con una cuchara en la boca, luego con una pelota de tenis de mesa en la cuchara, y así sucesivamente. Los estudios experimentales sugieren que el

enfoque Montessori tiene un impacto positivo en muchas facetas del desarrollo infantil (Lillard y Else-Quest, 2006; Marshall, 2017). Otros estudios muestran los beneficios atencionales de los videojuegos, de la meditación, de la práctica de un instrumento musical... Para un niño pequeño, es difícil controlar el cuerpo, la mirada, la respiración, coordinando sus gestos con los de los otros; por esto, la práctica de un instrumento musical, desde la más temprana edad, tiene efectos importantes en los circuitos atencionales del cerebro, incluido un aumento significativo y bilateral del espesor de la corteza prefrontal (figura 29; Bermúdez y otros, 2009; James y otros, 2014; Moreno y otros, 2011).

El entrenamiento del control ejecutivo puede incluso modificar el coeficiente intelectual. Esto acaso resulte sorprendente, porque en los Estados Unidos, el Reino Unido y varios otros países, suele considerarse que el célebre IQ o CI es un factor determinante del potencial mental de los niños. Sin embargo, el coeficiente intelectual es solo una capacidad de comportamiento y, como tal, está lejos de ser inmutable por la educación. Al igual que cualquiera de nuestras habilidades, su base está conformada por circuitos cerebrales específicos cuyos pesos sinápticos pueden modificarse mediante el entrenamiento. Lo que conocemos como inteligencia fluida –la capacidad de razonar y de resolver problemas nuevos– hace un uso masivo de los sistemas cerebrales del control ejecutivo: ambos movilizan una red similar de áreas cerebrales, en especial de la corteza prefrontal dorsolateral (Duncan, 2003, 2010, 2013). Sin embargo, las pruebas estandarizadas que miden la inteligencia fluida se parecen a aquellas que evalúan el control ejecutivo: unas y otras privilegian la atención, la concentración y la capacidad de pasar con rapidez de una actividad a la otra sin perder de vista el objetivo global de la tarea. De hecho, el entrenamiento de la memoria de trabajo y del control ejecutivo produce un ligero aumento de la inteligencia fluida (Au y otros, 2015). Estos resultados concuerdan con la constatación de que, aunque no carece de determinaciones genéticas, la inteligencia crece con la misma fuerza en función de los factores ambientales, especialmente de la educación. Estos efectos son enormes: el CI puede aumentar 20 puntos cuando un niño es adoptado y criado en un entorno estimulante, y el incremento ligado a la educación se sitúa entre 1 y 5 puntos por cada año de educación suplementaria (Duyme, Dumaret y Tomkiewicz, 1999; Ritchie y Tucker-Drob, 2017).

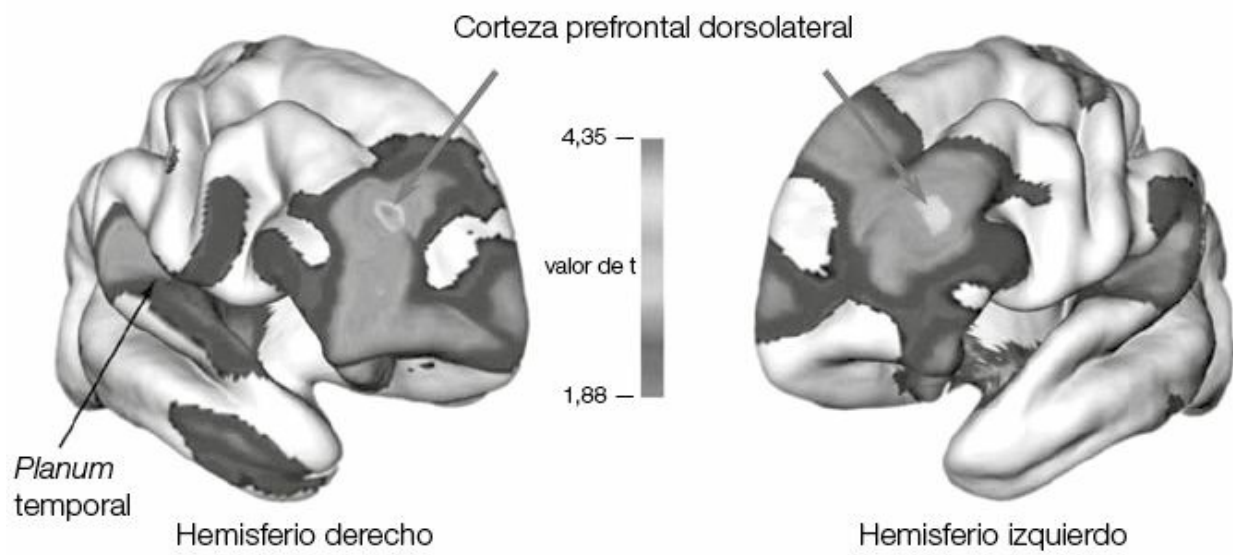


Figura 29. La atención ejecutiva (es decir, la capacidad de concentrarse y de controlarse) se

desarrolla con la edad y la educación. La práctica de un instrumento musical es una de las numerosas maneras de aprender a regular, desde la edad más temprana, la concentración y el control de uno mismo. Cuando se los compara con quienes no saben tocar un instrumento, en situaciones iguales, los músicos presentan un aumento del espesor de la corteza prefrontal dorsolateral, un área que desempeña un papel importante en el control ejecutivo.

El desafío de la investigación actual consiste en optimizar los efectos del entrenamiento cognitivo y en volver más nítidos sus límites. ¿Cómo podría aumentarse la duración de los efectos observados? Y, sobre todo, ¿cómo asegurarnos de que no se vean limitados a la tarea específica que se entrenó, sino que se expandan más allá de ella, y a lo largo de toda la vida? Ese es el desafío al que nos enfrentamos, porque es cierto que, de por sí, el cerebro tiende a desarrollar trucos específicos para cada tarea, caso por caso. La solución probablemente radique en diversificar los aprendizajes; los mejores resultados parecen ser los obtenidos por los programas educativos que estimulan el conjunto de habilidades cognitivas que constituyen la memoria de trabajo y la atención ejecutiva en una gran variedad de tareas y contextos.

Algunos hallazgos me vuelven particularmente optimista. El entrenamiento precoz de la memoria de trabajo, en especial si comienza desde el jardín de infantes, parece tener efectos positivos sobre la concentración y el éxito en numerosos dominios, incluidos aquellos de estricta pertinencia para la escuela: la lectura y las matemáticas (Bergman-Nutley y Klingberg, 2014; Blair y Raver, 2014; Klingberg, 2010; Spencer-Smith y Klingberg, 2015). Esto no es para nada sorprendente, dado que sabemos, desde hace años, que la memoria de trabajo es uno de los mejores predictores del éxito en matemáticas años más tarde (Dumontheil y Klingberg, 2011; Gathercole y otros, 2004; Geary, 2011). Los efectos de estos entrenamientos se ven multiplicados si se combina el trabajo sobre la memoria con una enseñanza más específica del concepto de la “recta numérica”, la idea esencial de que los números se organizan sobre un eje lineal, y que sumar o restar consiste en moverse sobre esta línea (Nemmi y otros, 2016). Todas estas intervenciones educativas parecen beneficiar sobre todo a los niños que provienen de contextos desfavorecidos. Si se busca mitigar los efectos de la desigualdad en las familias de nivel socioeconómico bajo, intervenir lo más temprano posible, desde el jardín de infantes, y enseñar los fundamentos del aprendizaje y de la atención puede ser una de las mejores inversiones educativas.

Presto atención si prestas atención

ὁ ἄνθρωπος φύσει πολιτικὸν ζῷον

[El hombre es por naturaleza un animal social]

Aristóteles (384-322 a.C.), *Política*

Todas las especies de mamíferos y, por supuesto, todos los primates poseen sistemas atencionales. Pero entre los seres humanos la atención presenta una característica única que acelera aún más el aprendizaje: tiene en cuenta en gran medida el contexto social. En el *Homo sapiens* más que en cualquier otro primate, la atención y el aprendizaje dependen de señales sociales: presto atención a lo que prestas atención, y aprendo de lo que me enseñas.

Desde la edad más temprana, el niño sigue al adulto con la mirada y mira su rostro. Presta extrema atención a la región de los ojos de las personas con las cuales interactúa. Tan pronto como se le dice algo, su primer reflejo no es explorar la escena, sino cruzar la mirada de la persona que le habla. Recién cuando hay contacto visual, el niño se vuelve en la dirección que examina el adulto. Esta capacidad de compartir socialmente la atención, que denominamos “atención compartida”, determina lo que el niño aprende.

Ya les hablé de los experimentos en que se hace aprender a un bebé una palabra nueva como “wug”: si el niño interactúa con la persona que habla y puede seguir su mirada en dirección al objeto al que se refiere, no tiene dificultad alguna para aprenderla en unos pocos intentos; pero si la misma palabra llega repetidas veces de un altavoz, aunque el vínculo con el objeto sea directo no se produce aprendizaje (Baldwin y otros, 1996). Lo mismo sucede con el aprendizaje de las categorías fonéticas: un niño estadounidense que, hacia la edad de 9 meses, interactúa con una niñera china durante algunas semanas, aprende los fonemas del chino; pero si recibe exactamente la misma cantidad de estimulación lingüística de un video de muy buena calidad, tampoco hay aprendizaje (Kuhl, Tsao y Liu, 2003).

Los psicólogos húngaros Gergely Csibra y György Gergely postulan que enseñar a otras personas y aprender de otras personas son adaptaciones evolutivas fundamentales de la especie humana (Csibra y Gergely, 2009; Egyed, Király y Gergely, 2013). Como animales sociales, los ejemplares de *Homo sapiens* disponemos de un “módulo pedagógico” que se pone en funcionamiento desde el momento en que prestamos atención a aquello que

los otros buscan enseñarnos. El buen éxito de la especie humana a escala planetaria se debe, al menos en parte, a un factor evolutivo específico: la capacidad de compartir la atención con los otros. Debemos a los demás, antes que a nuestra experiencia personal, la mayor parte de la información que aprendemos. Así, la cultura colectiva de los seres humanos se eleva muy por encima de lo que cada uno puede descubrir por sí solo. Esto es lo que el psicólogo Michael Tomasello llama el “efecto de trinquete” [*ratchet effect*] en la cultura: así como un trinquete evita que un montacargas se caiga, el intercambio social evita que la cultura retroceda. Cuando una persona realiza un descubrimiento útil, este se propaga rápidamente al grupo entero. Gracias al aprendizaje social, es muy inusual que el montacargas cultural vuelva a descender y que un invento importante caiga en el olvido.

Nuestro sistema atencional se adaptó a este contexto cultural. Las investigaciones de Csibra y Gergely demuestran que, desde la edad más temprana, la orientación de la atención del niño ya está muy en sintonía con las señales que le envían los adultos. La presencia de un tutor humano, que mira al niño antes de hacer una demostración específica, modula en gran medida el aprendizaje. No solo atrae la atención del niño, sino que también señala que el adulto tiene intención de enseñarle algo importante. Y a esto es inmediatamente sensible el bebé: el contacto visual induce en él una “postura pedagógica” que lo alienta a interpretar la información como importante y generalizable.

Tomemos un ejemplo: una mujer joven se vuelve hacia un objeto A haciendo una mueca, y luego hacia un objeto B con una gran sonrisa. Un bebé de 18 meses mira la escena. ¿Qué conclusión deriva? Todo depende de las señales que el adulto y el niño hayan intercambiado. Si sus miradas no se cruzaron, entonces el niño solo recuerda algo puntual: a esta persona no le gusta el objeto A, prefiere el objeto B. Pero si hubo contacto visual entre ellos, entonces el niño deduce mucho más: retiene que el objeto A es malo, y que el objeto B es agradable, y no solo para esa persona, sino para todas (figura 30). Los bebés prestan extrema atención a los indicios de comunicación voluntaria. Si alguien envía señales ostensibles de querer comunicarse con ella, una niña concluye que esa persona busca enseñarle información abstracta, no solo sus propias preferencias idiosincráticas.

Atención compartida:

Un adulto mira a un niño a los ojos...



... luego hace una mueca, mirando a un objeto...



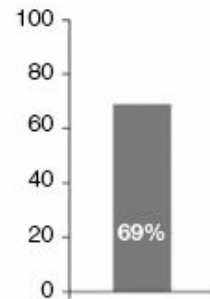
... luego se da vuelta hacia otro objeto, mientras sonríe.



Otra persona pide uno de los dos objetos.



La mayoría de los niños elige el objeto esperado.



Comprensión de las intenciones:

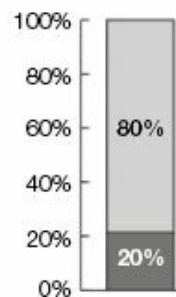
Un adulto, con las manos **ocupadas**, realiza una acción extraña con la cabeza.



Un adulto, con las manos **libres**, realiza una acción extraña con la cabeza.



El 80% de los niños imita inteligentemente con la **mano**.



El 70% de los niños imita servilmente con la **cabeza**.

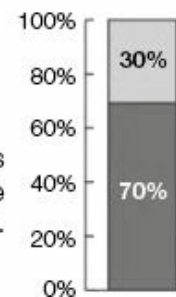


Figura 30. En la especie humana, el aprendizaje es social: tiene una fuerte dependencia de la atención y de la comprensión de las intenciones de las demás personas. Un bebé de 18 meses comprende que, si se lo mira a los ojos, se busca transmitirle información importante.

Aprende, entonces, con mayor eficacia y logra generalizar la información para utilizarla con otra persona; en cambio, no lo hace si no se lo miró (arriba). Desde los 14 meses, un bebé ya interpreta las intenciones de su entorno (abajo): luego de ver a una persona encender una luz presionándola con la cabeza, va a imitar su gesto, salvo que la persona tenga las manos ocupadas (en ese caso comprende que se puede simplemente encender la luz con la mano).

Por cierto, esto no se limita a la mirada: los niños también comprenden muy rápido qué significa señalar con el dedo (mientras que en verdad los chimpancés jamás lo logran). Incluso un bebé nota si alguien busca llamar su atención para transmitirle información importante. Por ejemplo, si un bebé de 9 meses ve a alguien que llama su atención y que luego apunta hacia un objeto, recuerda más tarde la identidad de ese objeto, porque comprende que es información que su interlocutor juzga importante; en cambio, si esa misma persona que tiende el brazo hacia el objeto no entabla contacto visual con el bebé, él recuerda tan solo la posición del objeto, no su identidad (Yoon, Johnson y Csibra, 2008).

Padres, madres, docentes, por favor, tengan siempre en mente este hecho crucial: la actitud y la mirada cambian por completo el mensaje que reciben cada niña y cada niño. Captar su atención con el contacto visual y verbal garantiza que compartirán su atención y multiplicará por igual las posibilidades de que retengan la información que ustedes buscan transmitirles.

Enseñar es prestar atención a la atención del otro

Ninguna otra especie animal sabe enseñar como la nuestra. El motivo es sencillo: acaso seamos los únicos que poseemos una “teoría” de la mente de las otras personas, la capacidad de prestar atención a la atención de los demás, de imaginar sus pensamientos, incluido lo que piensan que los otros piensan, y así sucesivamente, en una espiral infinita. Este tipo de representación recursiva, incrustada una dentro de otra, es típico del cerebro humano y tiene un papel esencial en la relación pedagógica. Los adultos que enseñan deben pensar todo el tiempo en lo que sus alumnos no saben: adaptan las palabras y eligen los ejemplos con el objetivo de transformar, cuanto antes, el conocimiento del niño. A la inversa, los alumnos saben que

su maestro sabe que ellos no saben: una vez que están comprometidos con la “postura pedagógica”, interpretan cada acto del docente como un intento de transferirles conocimientos. Y la espiral continúa *ad infinitum*: el adulto sabe que el niño sabe que el adulto sabe que él no sabe... esto permite que el adulto elija sus ejemplos sabiendo que el niño intentará generalizarlos.

Esta relación pedagógica solo existe en nuestra especie. En un artículo emblemático publicado por la revista *Science* (Thornton y McAuliffe, 2006), investigadores del área de la etología describieron con precisión una forma de enseñanza en la suricata, pequeño mamífero de la familia de las mangostas que vive en Sudáfrica (y actualmente célebre gracias a films documentales y de animación); pero anticipo que, a mi entender, en ese estudio falla la definición misma de enseñanza. ¿De qué se trataba? Del gran tema de las familias: ¡aprender a prepararse la comida! La dificultad que enfrentan las mangostas se debe a que consumen presas extremadamente peligrosas: escorpiones con aguijones mortales que es indispensable quitar antes de comer. Sin embargo, el artículo demuestra que las suricatas adultas ayudan a sus bebés proponiéndoles alimentos “preparados”, es decir, escorpiones de los que las adultas han quitado el aguijón. A medida que el pequeño crece, el adulto lo provee de una proporción cada vez mayor de escorpiones vivos y, desde luego, eso ayuda al pequeño a convertirse en un cazador autónomo. Así, de acuerdo con los autores del artículo, se cumplen tres criterios de la enseñanza: el adulto realiza un comportamiento específico en presencia de los niños; este comportamiento implica un costo para el adulto (y los niños se benefician al adquirir conocimientos más rápidamente que si el adulto no hubiera intervenido).

El caso de las suricatas es impactante: la evolución logró instalar un mecanismo singular que, a todas luces, facilita su supervivencia. Pero ¿se trata en verdad de enseñanza? En el sentido en que yo la entiendo, no. Estos datos no permiten llegar a la conclusión de que las suricatas en efecto les enseñan a sus crías, porque falta un ingrediente crucial: la atención compartida. No hay prueba alguna de que los adultos presten atención a lo que saben las crías o, a la inversa, de que los ejemplares jóvenes tengan en cuenta la postura pedagógica de los adultos. Las mangostas adultas se contentan con presentarles a las pequeñas, en función de su edad, animales más o menos peligrosos, pero nada prueba que le presten atención al conocimiento de sus cachorros. Por lo que sabemos hasta ahora, podría tratarse de un comportamiento totalmente previsto, especificado por el

aprendizaje del consumo de escorpiones; una conducta compleja, pero limitada, comparable a la famosa danza de las abejas o al desfile nupcial de los flamencos rosados.

Las mangostas y los escorpiones nos muestran, como en un espejo, lo que la enseñanza posee de único y precioso en nuestra especie. Cualquier relación pedagógica auténtica implica un vínculo mental fuerte entre el docente y el alumno o la alumna. Una buena maestra se construye un modelo mental de sus alumnos, de sus habilidades y de sus errores, y hace lo mejor que puede para lograr que sus mentes progresen. Esta definición ideal excluye al docente (humano o computadora) que se contenta con dar mecánicamente una lección estereotipada, sin haberla adaptado a las expectativas y los saberes previos de su auditorio; esta enseñanza unidireccional, sin sentido, no funciona. El alumno, por su parte, sabe que el docente sabe y que hace lo mejor posible para transmitir su saber. Cualquier relación pedagógica sana debe basarse sobre la atención, la escucha, el respeto y la confianza, en los dos sentidos.

La modesta pedagogía de las mangostas tampoco le hace justicia al rol que tiene la educación en las sociedades humanas. “Cada hombre es una humanidad, una historia universal”, decía Jules Michelet. Gracias a la educación, llevamos en nosotros lo mejor de los pensamientos de miles de generaciones que nos precedieron. Cada palabra, cada concepto que aprendemos, es una pequeña conquista que nos transmiten nuestros ancestros. Sin lenguaje, sin transmisión cultural, sin educación en colectividad, jamás habríamos podido descubrir por nuestra cuenta las herramientas mentales de las que se nos dotó en este presente.

Pero esa dependencia del *Homo sapiens* respecto de la comunicación social y la educación es una maldición tanto como es un don. La otra cara de la moneda es que por su culpa se propagan con tanta facilidad en las sociedades humanas las quimeras de los gurúes, los mitos de las religiones, las *fake news*, los rumores infundados y demás fabulaciones. Desde la edad más temprana, el cerebro absorbe ciegamente los cuentos, verdaderos o falsos, que circulan a nuestro alrededor. En un contexto social, baja la guardia y deja de comportarse como un científico en ciernes para convertirse en un loro repetidor, un imitador irreflexivo.

Un experimento célebre revela en qué medida, desde los primeros años de vida, el aprendizaje puede convertirnos en esclavos del entorno social. Desde los 14 meses, los bebés imitan de forma servil las acciones de una persona,

incluso si no les parecen provistas de sentido (Gergely, Bekkering y Kiraly, 2002). Si ven que alguien toca un botón con la cabeza mientras tiene las manos ocupadas, la razón les dicta que es suficiente con apoyar la mano, y eso hacen. Pero si ven a la misma persona apoyar la cabeza sin motivo específico alguno, ya que sus manos están libres y a la vista, entonces parecen abdicar de cualquier razonamiento y confiar ciegamente en el adulto: se ponen a imitar con fidelidad la acción que acaban de ver, aunque esté desprovista de sentido (figura 30). En la adultez, este conformismo social persiste y se amplifica: si los congéneres tienen un juicio diferente del nuestro, aun en las situaciones perceptivas más banales, revisamos a menudo nuestro juicio para sintonizarlo con el de ellos, aunque esto vaya contra cualquier plausibilidad (véase, por ejemplo, Bond y Smith, 1996).

En síntesis, el cerebro del *Homo sapiens* dispone de dos modos de aprendizaje: un modo activo, en el cual, como buenos científicos, ponemos a prueba hipótesis sobre el mundo exterior; y un modo receptivo, en el cual, sin verificación personal, absorbemos lo que otros nos transmiten. El segundo modo, mediante la acumulación cultural, permitió la expansión de las sociedades humanas desde hace cincuenta mil años. Pero sin la mente crítica que caracteriza al primer modo, el segundo es vulnerable a la propagación de todas las *fake news* habidas y por haber. La verificación activa de los conocimientos, el rechazo del simple “decir que sí” y la construcción personal del sentido son filtros esenciales para protegernos de las supersticiones y de los gurús. Es necesario, entonces, encontrar un punto medio entre los dos modos: formar estudiantes atentos, que confíen en los demás, pero que no dejen de ser autónomos, capaces de un pensamiento crítico, protagonistas de su propio aprendizaje.

Y de inmediato pasaremos a analizar este segundo pilar del aprendizaje: el compromiso activo.

[19] Véanse los programas de acceso libre *The Number Race*, www.thenumberrace.com –que desde luego cuenta con versión francesa, www.lacourseauxnombres.com/nr/home.php– y *The Number Catcher*, www.thenumbercatcher.com; en francés: www.attrapenombres.com/an/home.php. Para el aprendizaje de la lectura, véanse los programas *GraphoGame* grapholearn.fr y *ÉLAN pour la lecture*.

8. El compromiso activo

Imaginemos dos gatitos. El primero está sujeto con un collar y una correa. El segundo está sentado en un carrito que le impide moverse a su voluntad. Por lo demás, los dos están atados a un carrusel que asegura que los movimientos de uno y otro están vinculados (figura 31). La idea es que los dos animales reciban estímulos visuales idénticos, pero que uno este activo y el otro pasivo. El primero explorará el entorno por su cuenta, mientras que el segundo se moverá exactamente de la misma manera, pero sin poder controlarlo por sus medios.

Esta es la clásica experiencia del carrusel que realizaron Richard Held y Alan Hein en 1963, ¡en una época en que la ética de experimentación con animales claramente no estaba tan desarrollada como en la actualidad! Durante algunas semanas, los dos gatitos vivieron, tres horas por día, en un gran cilindro tapizado por barras verticales. Los dos recibían idénticos estímulos visuales; sin embargo, su desarrollo fue muy diferente. Esta experiencia tan sencilla desembocó en un descubrimiento fundamental: la exploración activa del mundo es esencial para el buen desarrollo de la visión (Held y Hein, 1963). Pese a la pobreza de su entorno, constituido únicamente por barras verticales, el gatito activo desarrolló enseguida una visión normal. El gatito pasivo, en cambio, perdió sus capacidades visuales, y, al final de la experiencia, fracasó en pruebas elementales de exploración visual. En la prueba del acantilado, por ejemplo, se situó al animal sobre un “puente” del cual podía salir por un costado que desembocaba en un alto acantilado, o por un costado poco profundo. El animal normal no dudó un segundo y saltó por el costado fácil. En cambio, el animal criado en la pasividad eligió al azar. Otras pruebas demostraron que no percibía bien el espacio ni lo exploraba con la pata, como hace cualquier felino.

El carrusel de Held y Hein

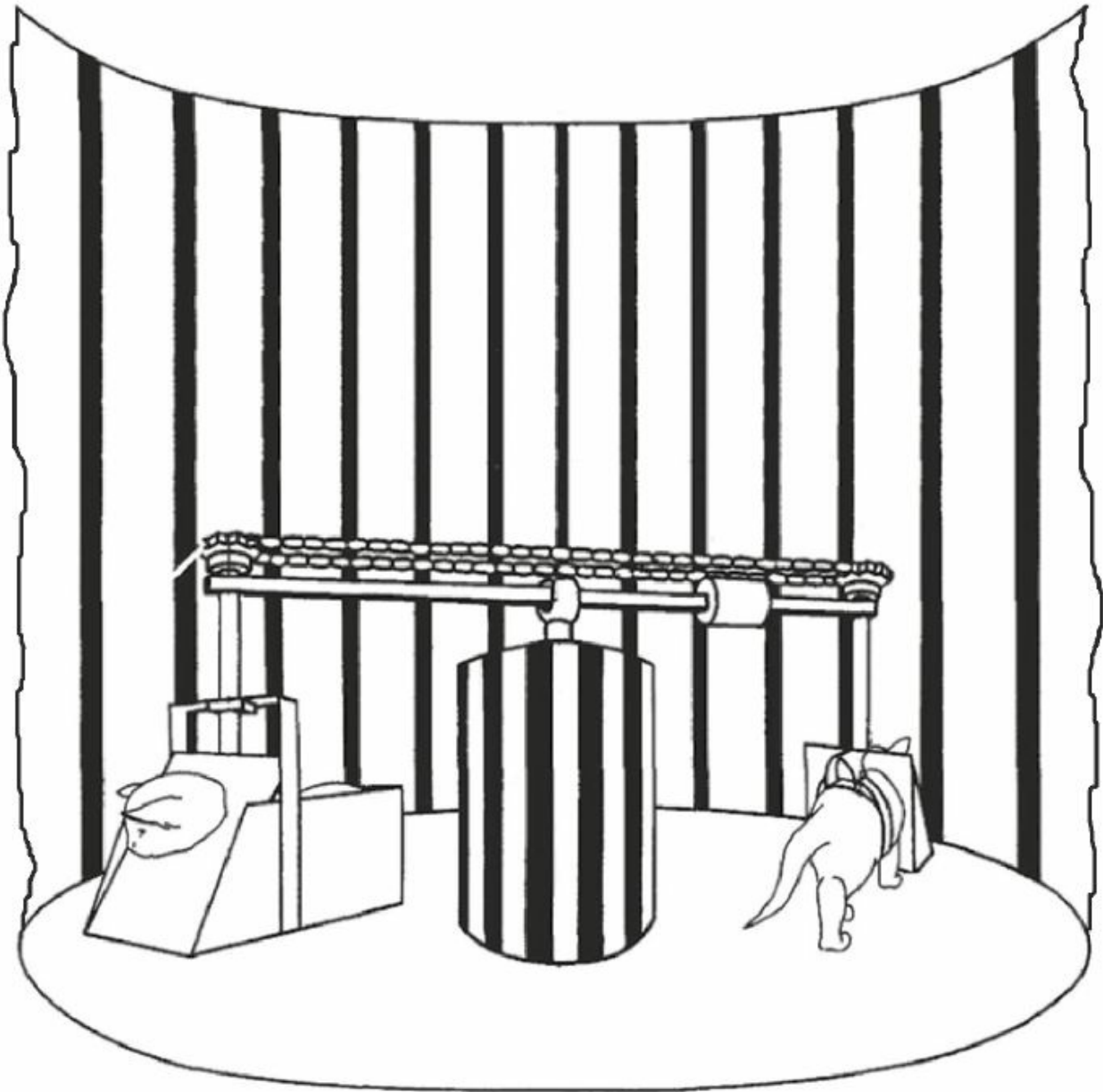


Figura 31. El compromiso activo es el segundo pilar del aprendizaje: un organismo pasivo no aprende. En la experiencia clásica de Held y Hein (1963), un gato explora activamente el mundo mientras el otro recibe de sus sentidos tan solo la estimulación pasiva. Luego de varias decenas de horas, el animal pasivo presenta importantes problemas para la exploración visual. Aprender con eficacia entraña rechazar la pasividad, comprometerse, explorar con curiosidad, generar activamente hipótesis y ponerlas a prueba.

Un organismo pasivo no aprende

La experiencia del carrusel de Held y Hein funciona como metáfora del segundo pilar del aprendizaje: el compromiso activo. Una gran cantidad de investigaciones, de los campos más diversos, sugiere que un organismo pasivo aprende poco o nada. Aprender con eficacia requiere rechazar la pasividad, comprometerse, explorar, generar hipótesis activamente y ponerlas a prueba en comparación y contraste con el mundo con el cual interactuamos.

Para aprender, el cerebro debe, en un comienzo, formarse un modelo mental hipotético del mundo exterior, y luego proyectarlo sobre su entorno y comparar sus predicciones con lo que recibe de los sentidos. Esto implica una postura activa, comprometida, atenta. La motivación es esencial: solo aprenderemos bien si tenemos una idea clara del objetivo que queremos alcanzar y nos involucramos plenamente.

Por favor, no me malinterpreten: ¡la participación activa no significa que el niño deba moverse todo el tiempo en clase! Un día visité un centro de formación en que el director, con cierto orgullo, me explicó cómo aplicaba mis ideas: había equipado todos los pupitres con pedales para que sus alumnos permanecieran activos durante la clase de matemáticas... Esto nada tiene que ver con lo que quiero decir, y aquí tocamos el límite de la metáfora de la experiencia del carrusel. Estar activo y comprometerse no quiere decir que el cuerpo deba moverse.

Es en la cabeza, no en los pies, donde hace su trabajo el compromiso activo. El cerebro aprende bien únicamente si está atento, concentrado y en plena actividad de generación de modelos mentales. Para digerir mejor lo que aprende, un estudiante activo lo reformula permanentemente en sus propias palabras o en sus pensamientos. Un estudiante pasivo (o, todavía peor, distraído) no aprovecha ninguna clase, ya que su cerebro no actualiza sus modelos mentales del mundo. En esto no incide que estén quietos o se muevan. Los dos estudiantes pueden permanecer inmóviles, pero presentan diferencias radicales en el movimiento interno del pensamiento: uno sigue activamente la clase, mientras que el otro se desconecta.

La experiencia demuestra que rara vez aprendemos si nos limitamos a acumular pasivamente estadísticas de los estímulos que recibimos. Esto puede ocurrir, pero solo en el nivel más bajo de los sistemas sensoriales o motores. Ya vimos, por ejemplo, esas experiencias en las cuales un niño escucha cientos de sílabas y termina por detectar la presencia de palabras

como *ba-be-ro*; este tipo de aprendizaje, que es implícito, parece persistir incluso mientras los bebés duermen (Hay y otros, 2011; Saffran, Aslin y Newport, 1996; también aportan material de prueba los trabajos actuales del equipo de G. Dehaene-Lambertz sobre el aprendizaje en el lactante durante el sueño). Sin embargo, esa es la excepción que confirma la regla: en la inmensa mayoría de los casos, ya que el aprendizaje atañe a propiedades cognitivas de nivel alto, como la memoria explícita del sentido de las palabras y no solo su forma, el aprendizaje parece ocurrir si y solo si la o el aprendiz presta atención, piensa, anticipa y presenta hipótesis con el riesgo de equivocarse. Sin atención, sin esfuerzo, sin reflexión profunda, la lección se desvanece sin dejar mucho rastro en el cerebro.

Procesar en profundidad para aprender mejor

Tomemos un ejemplo clásico de la psicología cognitiva: el efecto de la profundidad del procesamiento de las palabras. Imaginen que presento sesenta palabras a tres grupos de estudiantes. Al primer grupo le pido que decida si estas palabras están escritas en minúsculas o en mayúsculas; al segundo grupo, si riman con “silla”, y al tercero, si se trata de nombres de animales o no. Luego, sin que se lo esperen, les hago realizar una prueba de memoria. ¿Qué grupo recuerda mejor las palabras? La memoria resulta ser tanto mejor en el tercer grupo, que procesaba las palabras en profundidad, en el nivel del significado (un 75% de aciertos), que en los otros dos grupos, que se ocupaban de los aspectos formales, es decir, más superficiales, uno en el nivel de la letra (un 33% de respuestas correctas), y el otro en el de la rima (un 52% de aciertos; Craik y Tulving, 1975; Jacoby y Dallas, 1981). Por supuesto, en todos los grupos encontramos una débil huella implícita, inconsciente, de las sesenta palabras: el aprendizaje dejó su rastro subliminal en los sistemas ortográficos y fonológicos. Sin embargo, solo el procesamiento semántico en profundidad indujo un recuerdo explícito, detallado, de las palabras percibidas.

El mismo fenómeno ocurre en el nivel de las frases o las oraciones: los estudiantes que hacen el esfuerzo de comprenderlas por sí solos, sin que el docente les dé la solución, presentan una mejor retención de la información

en la memoria (Auble y Franks, 1978; Auble, Franks y Soraci, 1979). Es una regla general, que el psicólogo estadounidense Henry Roediger enunció de este modo: “Hacer que las condiciones de aprendizaje sean más difíciles, lo que requiere mayor esfuerzo cognitivo por parte de los estudiantes, a menudo redundante en una mayor retención” (Zaromb, Karpicke y Roediger, 2010).

Las neuroimágenes comenzaron a esclarecer los orígenes de este efecto de profundidad de procesamiento (Kapur y otros, 1994). Que el procesamiento profundo deje una huella más fuerte en la memoria se debe a que activa áreas de la corteza prefrontal, que están asociadas al procesamiento consciente de las palabras, y porque estas regiones forman potentes bucles con el hipocampo, que almacena la información en forma de recuerdos episódicos y explícitos.

En un cortometraje que se convirtió en una película de culto, *La Jetée* (1962; conocida en castellano como *El muelle*) del director francés Chris Marker, una voz en *off* enuncia este aforismo que suena como una verdad profunda: “Nada distingue los recuerdos del resto de los momentos: solo más tarde se harán reconocibles, a partir de sus cicatrices”. Bello adagio... pero falso proverbio, dado que, desde el momento de la codificación, los acontecimientos de nuestra vida que permanecerán grabados en la memoria ya son diferentes de aquellos que no dejarán huella alguna: fueron objeto de un procesamiento más activo y más profundo. Cuando se escanea a una persona mientras se la expone a una lista de palabras o de imágenes, se puede predecir cuáles de esos estímulos serán olvidados y cuáles serán retenidos media hora más tarde. La clave de la predicción consiste en verificar si esos *inputs* inducen actividad en la corteza frontal, el hipocampo y las regiones vecinas de la corteza parahipocámpica (Brewer y otros, 1998; Paller, McCarthy y Wood, 1988; Sederberg y otros, 2006; Sederberg y otros, 2003; Wagner y otros, 1998). El compromiso activo de estas regiones, reflejo directo de la profundidad con que las palabras y las imágenes viajaron dentro del cerebro, predice la fuerza del recuerdo. Una imagen inconsciente puede entrar en las áreas sensoriales, pero solo crea una onda de actividad muy pequeña en la profundidad de la corteza prefrontal. La atención, la concentración, la toma de conciencia y el procesamiento en profundidad transforman esa pequeña onda en un tsunami neuronal que invade la corteza prefrontal y maximiza la memorización (Dehaene y otros, 2001).

Tomemos otro ejemplo muy diferente: el aprendizaje de la física en la universidad. Los estudiantes deben aprender conceptos abstractos como los

de momento cinético o angular y par motor. Separémoslos en dos grupos: uno dispone de diez minutos de experimentación con una rueda de bicicleta y el otro, de diez minutos de explicaciones verbales y de observaciones del resto de los estudiantes. El resultado no da pie a equívocos: el aprendizaje es tanto mejor en el grupo que se beneficia de una interacción activa con los objetos físicos (Kontra, Goldin-Meadow y Beilock, 2012; Kontra y otros, 2015). Volver la enseñanza más profunda, más atrapante, es una garantía de éxito.

Una revisión reciente de más de doscientos estudios pedagógicos lo confirma: la clase o lección magistral, en que el alumno permanece pasivo mientras el docente despliega su peroración (o perorata) durante cincuenta minutos, es cosa del pasado (Freeman y otros, 2014). Cuando comparamos la clase magistral tradicional con las pedagogías que promueven el compromiso activo, el efecto es evidente: en todas las disciplinas –desde las matemáticas hasta la psicología, sin omitir la biología o la informática–, un estudiante activo tiene mayor éxito. Con el compromiso activo, los resultados progresan en media desviación estándar –como dirían los estadísticos–, lo que es una cantidad considerable, y la tasa de fracaso disminuye cerca del 10%. Pero ¿cuáles son las estrategias que más involucran a los alumnos? En esta cuestión no existe un método único y milagroso, sino una generosa gama de actividades que obligan (¡o invitan o incitan!) a las y los estudiantes a reflexionar por sí mismos: actividades prácticas, discusiones en que cada cual interviene, trabajo en pequeños grupos, o profesoras que interrumpen su clase para plantear una pregunta difícil sobre la que dejan a los estudiantes reflexionar un largo rato... Todas las soluciones son buenas, siempre y cuando inciten a renunciar al apoltronamiento, a la cómoda e improductiva pasividad.

El fracaso de las pedagogías del descubrimiento

Nada de esto es demasiado nuevo, me dirán ustedes, y muchos docentes ya aplican estas ideas espontáneamente. Sin embargo, en el campo pedagógico, no podemos confiar en la tradición ni en la intuición: necesitamos verificar con métodos científicos qué herramientas realmente mejoran la comprensión y la retención de los materiales del curso por parte de las y los estudiantes, y

cuáles no. Y esta es una oportunidad para aclarar una distinción muy importante. La visión de que los niños deben participar de manera activa y atenta en su propio aprendizaje, cuyos fundamentos son correctos, no debe confundirse con el constructivismo clásico o con los métodos de aprendizaje por descubrimiento, que son ideas seductoras pero, por desgracia, se demostraron ineficaces una y otra vez. La distinción es elemental, aunque pocas veces se comprende, en parte porque también estas últimas se dan en llamar “pedagogías activas”, lo cual es fuente de numerosas confusiones.

Cuando hablamos de pedagogías del descubrimiento, ¿a qué nos referimos? A una nebulosa de ideas que se remontan a Jean-Jacques Rousseau y que llegaron a nosotros por obra de pedagogos como John Dewey, Ovide Decroly, Célestin Freinet, Maria Montessori o, más cerca en el tiempo, Jean Piaget y Seymour Papert. Jean-Jacques Rousseau en *Emilio o De la educación* nos increpa con lograda retórica: “¿Me atreveré a exponer aquí la mayor, la más importante, la más útil regla de toda la educación? No se trata de ganar tiempo, sino, por el contrario, de perderlo”. Para Rousseau y sus sucesores, siempre es mejor dejar que la niña o el niño descubra por su cuenta y que construya su propio saber, con la libertad de perder horas andando a tientas, explorando... Según ellos, este tiempo no está perdido: contribuye a formar un espíritu autónomo, que piensa por sí mismo en lugar de engullir todo sin reflexionar y que sabe resolver verdaderos problemas en lugar de aprender de memoria soluciones ya elaboradas. Escuchemos otra vez a Rousseau: “Haced que vuestro alumno esté atento a los fenómenos de la naturaleza, y enseguida despertaréis su curiosidad, pero para sujetarla no os deis prisa a satisfacerla. Poned a su alcance las cuestiones y dejad que él las resuelva”.

La teoría es seductora... Pero, por desgracia, una serie de estudios, extendidos a lo largo de varias décadas, demuestra que su valor pedagógico es cercano al ominoso cero. Si se los deja solos, los niños pasan por las mayores dificultades para descubrir las reglas abstractas que rigen un área del conocimiento, aprenden tanto peor, y en ocasiones no aprenden cosa alguna. ¿Deberíamos sorprendernos por esto? ¿Cómo podríamos imaginar que, en unas pocas horas y sin orientación externa, el niño puede redescubrir lo que a la humanidad le llevó siglos? Los fracasos de este método son rotundos en todos los campos (Hattie, Brodeur y St-Cyr, 2017; Kirschner, Sweller y Clark, 2006; Kirschner y Van Merriënboer, 2013; Mayer, 2004; Mottint, 2018):

- *En lectura:* la simple exposición a las palabras escritas entra en vía muerta si no se explica a las niñas y los niños la presencia de las letras. Son pocos los pequeños que logran por sí solos correlacionar el código escrito con el oral y descubrir que todas las palabras que comienzan con el sonido /R/ incluyen, a su izquierda, el símbolo **R** o **r**. Descubrir sin ayuda el principio alfabético sería una tarea desmesurada si el docente no se tomara el trabajo de seleccionar con cuidado ejemplos, palabras sencillas o letras aisladas que la simplifiquen.
- *En matemáticas:* se dice que a los 7 años el genial matemático Carl Gauss descubrió por sí solo cómo sumar rápidamente todos los números del 1 al 100 (la solución está en la nota al pie).[\[20\]](#) Pero lo que tal vez sea válido para Gauss no se aplica a las demás criaturas. La investigación es inequívoca al respecto: en matemáticas, el aprendizaje funciona mejor cuando el docente explica un ejemplo en detalle, antes de dejar que el niño se enfrente a otros. Los niños a los que primero se les explica cómo resolver un problema logran resolver mejor casos similares que aquellos que descubrieron la solución por sí mismos.
- *En informática:* en su libro *Desafío a la mente*, el informático Seymour Papert explica por qué inventó el lenguaje Logo (en el cual una suerte de tortuga dibuja y recibe órdenes de los distintos programas). Su idea era dejar que los niños exploraran las computadoras por sí mismos, sin enseñanza, poniendo las manos en la masa. La experiencia fue un fracaso: luego de unos meses, los niños no sabían escribir más que programas breves y elementales. Todos los conceptos abstractos se les escabullían, de modo que en una prueba de resolución de problemas no obtenían mejores resultados que otros niños sin entrenamiento: lo poco de informática que habían aprendido no se generalizaba a otros ámbitos. Por el contrario, la investigación demuestra que una enseñanza explícita, que alterne explicaciones claras con su puesta a prueba, les permite desarrollar una comprensión mucho más profunda del lenguaje Logo y de la informática.

Por mi parte, vivencíé la llegada de la primera computadora personal a mi casa, cuando tenía 15 años. Por consiguiente, formo parte de la generación

que aprendió a programar en lenguaje Basic, sin profesor ni curso, aunque no sin ayuda: mi hermano y yo sacamos jugo a todas las revistas, manuales, libros y ejemplos que lográbamos conseguir. Me convertí en un programador razonablemente eficaz... Pero cuando entré a la universidad, tomé conciencia de lo grandes que eran mis lagunas en lo que hace al campo de la informática: improvisaba sin haber comprendido la estructura lógica profunda de los programas ni las prácticas correctas que los vuelven claros y legibles. Sin lugar a dudas, esto es lo peor de la pedagogía del descubrimiento: deja a los estudiantes con la ilusión de que dominan determinado tema, sin darles jamás los medios para acceder a los conceptos profundos de la disciplina.

En síntesis, el principio del compromiso activo enuncia que es crucial que el estudiante esté motivado, activo, involucrado en el aprendizaje, pero de ningún modo eso significa que deba quedar librado a su propia suerte. El fracaso del constructivismo lo deja en claro: el docente debe responder a esta demanda aportando un entorno de aprendizaje gradual, estructurado, explícito, concebido para guiarlo lo antes posible hacia la cumbre. Las estrategias pedagógicas más eficientes son aquellas que alientan a los estudiantes a asumir un compromiso activo, pero que están guiadas de cerca por el docente. En palabras del psicólogo Richard Mayer, el mejor desempeño se logra con “métodos de instrucción que involucran actividad cognitiva en lugar de actividad conductual, orientación instructiva en lugar de descubrimiento puro y enfoque curricular en lugar de exploración no estructurada”. Dentro de este marco, las maestras y los profesores proporcionan una secuencia clara y rigurosa que comienza por las bases; evalúan constantemente el dominio de sus alumnos, y desde ese punto de partida los guían para construir una pirámide de significado. Eso practican, hoy en día y en su gran mayoría, las escuelas Montessori: no dejan a los niños a la deriva sin hacer nada, sino que les proponen una serie de actividades racionales, jerarquizadas, que fueron objeto de escrutinio minucioso por parte del docente antes de ser realizadas por los niños de manera autónoma. Compromiso, placer y autonomía acompañados por una pedagogía explícita basada sobre un material estimulante: es la receta de un cóctel ganador cuya eficacia ya quedó demostrada (Lillard y Else-Quest, 2006; Marshall, 2017).

La pura pedagogía del descubrimiento, esa idea de que el niño puede autoeducarse, forma parte de los mitos educativos, muchas veces denunciados, pero que curiosamente todavía son muy populares. Pertenece a

las leyendas urbanas que marcan el campo educativo, pero no es la única; al menos otras dos ideas falsas están unidas a ella (Kirschner y Van Merriënboer, 2013):

- *El mito del nativo digital*: los niños de las generaciones recientes, a diferencia de sus padres y sus madres, serían campeones del mundo digital solo por estar rodeados de computadoras y dispositivos electrónicos desde sus primeros años. Al estar sumergidos en la informática, se habrían convertido en *nativos digitales*, destacados ejemplares de *Homo zappiens* para quienes los bits y los bytes ya no tendrían secretos. Esto es falso: el conocimiento que poseen de la tecnología suele ser superficial. Solo los más educados saben hacer un uso crítico de las fuentes de información de internet, y en ningún caso son mejores para realizar varias tareas a la vez (como vimos, el cuello de botella central que nos impide hacer dos cosas al mismo tiempo es una propiedad fundamental de la arquitectura de nuestro cerebro).
- *El mito de los estilos de aprendizaje*: cada niño poseería su propio estilo de aprendizaje, por lo que sería natural dejarlo aprender a su modo. Esto es falso: no existe prueba alguna que demuestre que algunos niños son más visuales y otros más auditivos o más táctiles (salvo los ciegos, desde luego, como nos señalaría Georges Brassens en su deliciosa canción “La mauvaise réputation” [La mala reputación]). Lo cierto es que algunas estrategias funcionan mejor que otras –por ejemplo, resulta más fácil memorizar una imagen que una palabra–, pero esto es así para todos los niños. No hay material o situación alguna que pruebe que existan subtipos de niñas o de niños con estilos de aprendizaje radicalmente diferentes, de modo que los niños de tipo A aprendan mejor con la estrategia A, y que los niños de tipo B aprendan mejor con la estrategia B (Pashler y otros, 2008). Hasta donde sabemos, todos los seres humanos compartimos el mismo algoritmo de aprendizaje.

“Pero ¿cómo...?”. Sí, ya mismo puedo escuchar las exclamaciones y las preguntas: “¿Qué pasa con todos esos libros y *software* de educación especial que dicen adaptar la educación a las necesidades de cada niño? ¿Acaso son inútiles, sin más?”. No necesariamente. Deberíamos hacer un claro deslinde

entre la estrategia de enseñanza (visual, auditiva, basada sobre recompensas contra castigos, etc.) y los contenidos de una lección. Por supuesto, hay una enorme variación entre los contenidos adquiridos a determinada edad. En primer grado, por ejemplo, el 10% de los niños, el que está en la parte superior de la curva de rendimiento, lee más de 4.000.000 de palabras por año, mientras que el 10% en la parte inferior de la misma curva, lee menos de 60.000, con obvias consecuencias en su dominio del vocabulario sofisticado y las estructuras gramaticales. En el caso extremo de una dislexia, es posible que lisa y llanamente no sean capaces de leer y que durante unos años más sigan sin lograrlo. Los déficits de desarrollo, como la dislexia y la discalculia, pueden presentarse con diferentes perfiles, e indudablemente vale la pena diagnosticar al detalle la índole exacta de la discapacidad para poder adaptar el contenido de las lecciones. Todos los niños se benefician de las intervenciones pedagógicas cuyos contenidos se adaptan a sus dificultades específicas. Por ejemplo, muchos niños, incluso en matemáticas avanzadas, no entienden cómo funcionan las fracciones, en cuyo caso es bastante útil volver a este punto (todos los niños se benefician de la metáfora de la recta numérica, que explica, por ejemplo, cómo puede ser visualizado $\frac{5}{6}$ al fragmentar un segmento en seis partes iguales y luego extraer cinco de esas porciones). Pero esta obviedad no resta valor a la afirmación, también correcta, de que todos aprendemos esos hechos, contenidos o metáforas con la misma maquinaria básica: una que prefiere la atención concentrada en vez de la doble tarea, el compromiso activo en vez de la lectura pasiva, la corrección detallada de errores en vez del elogio falso, y también la enseñanza explícita antes que el constructivismo o el aprendizaje por descubrimiento.

Sobre la curiosidad, y cómo despertarla

πάντες ἄνθρωποι τοῦ εἰδέναι ὀρέγονται φύσει

[Todos los hombres tienen naturalmente el deseo de saber]

Aristóteles, *Metafísica* (~335 a.C.)

No tengo talento especial alguno, sino que tan solo soy apasionadamente curioso.

Albert Einstein, en carta a Carl Seelig, 11 de marzo de 1952

Una de las bases del compromiso activo es la curiosidad, el deseo de aprender, la sed de conocimiento. Estimular la curiosidad de los niños significa haber ganado ya la mitad del partido. Una vez que su atención se moviliza y su mente está en busca de una explicación, no queda otra opción más que guiarlos hacia allí. En igualdad de condiciones, desde el jardín de infantes, los alumnos más curiosos son quienes tienen mejores resultados en lectura y en matemáticas (Shah y otros, 2018). Lograr mantener la curiosidad de los niños, entonces, es uno de los factores clave de una educación exitosa. Pero ¿qué es exactamente la curiosidad? ¿A qué necesidad darwiniana obedece? ¿A qué tipo de algoritmo corresponde?

Jean-Jacques Rousseau escribió en *Emilio o De la educación*: “Solo somos curiosos en la medida en que somos instruidos”. También en eso se equivocaba: la curiosidad no es un efecto de la enseñanza, una función que debamos adquirir. Está presente desde la edad más temprana y forma parte integral de nuestra biología de hombres neuronales; es uno de los ingredientes de nuestro algoritmo de aprendizaje. No esperamos, sin más, de forma pasiva, que nos llegue la información nueva, como hacen, con su característica necesidad, todas las redes neuronales actuales, simples funciones de entrada-salida sometidas a su entorno. Como señalaba Aristóteles, los seres humanos nacemos con una pasión por conocer y buscamos permanentemente la novedad, mediante la exploración activa del entorno para descubrir cosas que podamos aprender.

La curiosidad es un recurso fundamental del organismo: una fuerza propulsora que nos hace reaccionar, al igual que el hambre, la sed, la necesidad de seguridad o el deseo de reproducirnos. ¿Qué función cumple la curiosidad en la supervivencia? Las especies como la nuestra (y gran parte de los mamíferos, además de numerosos pájaros y peces) se interesan por explorar el ambiente para así manejarlo mejor. Sería peligroso instalar el nido o el lecho sin antes dominar un mapa del entorno. En un universo inestable, poblado de predadores, realizar periódicamente un reconocimiento del terreno, verificar si el ambiente se modificó, investigar el origen de un ruido inesperado puede hacer toda la diferencia entre la vida y la muerte. La curiosidad es la determinación que empuja a los animales a salir de su zona de confort para adquirir conocimiento. En este sentido, cualquier información posee valor en la moneda propia de Darwin: la supervivencia.

Prosigamos. La curiosidad se parece a la búsqueda de alimento o de parejas sexuales, pero está motivada por un valor inmaterial: la adquisición de información nueva. De hecho, las investigaciones en neurobiología demuestran que en el cerebro humano el descubrimiento de información antes desconocida conlleva su propia recompensa: activa el circuito de la dopamina. Recuerden que este circuito se enciende en respuesta al alimento, las drogas o el sexo. Sin embargo, en los primates, y acaso en todos los mamíferos, este circuito ya no responde solo a las recompensas materiales, sino también a la información nueva. Algunas neuronas dopaminérgicas señalan la cantidad de información que vendrá, como si el incremento de la información nueva aportara su propia gratificación (Bromberg-Martin y Hikosaka, 2009). De este modo, una rata puede ser condicionada no solo por la droga, sino también por la novedad: podemos inducirla a preferir un lugar por el solo motivo de que allí suele haber objetos novedosos, de modo que encontrará algo para satisfacer su curiosidad (Bevins, 2001). Los seres humanos no actuamos distinto cuando nos trasladamos a una gran ciudad con el objetivo de encontrar nuevas distracciones o cuando, ávidos de la última comidilla de la farándula o de nuestros amigos y enemigos, consumimos frenéticamente Facebook o nos volvemos visitantes seriales de Twitter.

También en nuestra especie el apetito de saber pasa por el circuito de la dopamina, incluso cuando se trata de una curiosidad puramente intelectual. Imaginen que están tumbados en un resonador magnético y se les hace una pregunta al estilo de juegos como Preguntados o Trivial Pursuit, por ejemplo: “¿En qué año se jugó el primer Mundial de Fútbol?” (Gruber, Gelman y Ranganath, 2014; véase también Kang y otros, 2009). Antes de aplacar la curiosidad, se les pregunta con qué grado de exactitud quieren conocer la respuesta. ¿Resultado? El grado de curiosidad que ponen de manifiesto predice la actividad del núcleo accumbens y del área tegmental ventral, dos regiones esenciales del circuito cerebral de la dopamina. A mayor curiosidad, más se encienden estas regiones. Eso no es todo: sus señales se desencadenan por anticipación. Incluso antes de conocer la respuesta, el simple hecho de saber que pronto van a acceder a ella excita sus circuitos dopaminérgicos y aporta su propia recompensa. Y todas estas señales de la curiosidad son útiles porque permiten predecir el aprendizaje: retendrán mejor los hechos que suscitaban su curiosidad. Incluso recordarán otros elementos anecdóticos, por ejemplo, el rostro de alguien que estaba presente o de la persona que les

enseñó esa información que tanto deseaban conocer. La avidez por saber determina la profundidad con que se procesará en la memoria.

Por obra del circuito de la dopamina, la satisfacción de nuestro apetito por aprender, o incluso de anticipar que vamos a hacerlo, constituye una recompensa en sí misma. El aprendizaje posee un valor intrínseco para el sistema nervioso. Lo que llamamos “curiosidad” no es otra cosa que la explotación de ese valor. Una vez alcanzada esta instancia, la especie humana es singular, porque, más que cualquier otra, posee la capacidad de aprender. Durante el transcurso de la hominización, nuestra capacidad de representación del mundo cambió. Somos acaso los únicos animales con capacidad de construir teorías formales del mundo en un lenguaje del pensamiento. La ciencia se convirtió en nuestro nicho ecológico: el *Homo sapiens* es la única especie que no tiene un hábitat específico, pero que, gracias al aprendizaje, logra adaptarse a casi cualquier ambiente.

Junto con esta extraordinaria expansión de la facultad del aprendizaje, la curiosidad humana parece haberse multiplicado por diez. A lo largo de la evolución, adquirimos una forma extendida de curiosidad que llamamos “curiosidad epistémica”: el puro deseo de saber, en todos los campos, incluidos los más abstractos. Como otros mamíferos, jugamos y exploramos, pero no solamente al movernos, sino también al pensar. Del mismo modo en que otros animales exploran el espacio que los rodea, nosotros exploramos los conceptos. Y nuestra especie siente emociones epistémicas específicas, vale decir, destinadas a guiar la sed de conocimiento. Nos regocijamos, por ejemplo, ante la simetría y la belleza puras de las estructuras matemáticas: un magnífico teorema puede conmovernos tanto más que una tableta de chocolate. La risa también parece ser una de esas emociones sociales que guían el aprendizaje. En efecto, muchas veces se desencadena cuando, de pronto, descubrimos que una de las hipótesis implícitas es falsa, lo que nos obliga a revisar nuestro modelo del mundo. De acuerdo con el filósofo Daniel Dennett, la hilaridad sería una señal social contagiosa que compartimos con las demás personas para así llamar la atención sobre cierta información inesperada que, de pronto, nos hace notar que debemos revisar nuestras creencias (Hurley, Dennett y Adams, 2011). Y, efectivamente, en igualdad de condiciones, el simple hecho de reír durante el aprendizaje aumenta la curiosidad y mejora la memoria (Esseily y otros, 2016).

Saber qué y cuánto sabemos multiplica la curiosidad

Varios psicólogos intentaron definir con precisión a qué algoritmo corresponde la curiosidad humana. En efecto, si lográramos comprenderlo, no solo dominaríamos mejor esta variable esencial para el aprendizaje, sino que también podríamos reproducirla en una máquina que imitara a la especie humana: un robot curioso.

Este enfoque algorítmico comenzó a dar sus frutos. Los psicólogos más importantes, desde William James hasta Jean Piaget o Donald Hebb, especularon sobre la índole de las operaciones mentales que sostienen la curiosidad. Para ellos, la curiosidad es la expresión directa de la motivación de los niños para comprender el mundo y para intentar construir un modelo de él (Loewenstein, 1994). Se desencadena cada vez que el cerebro detecta un desfase entre lo que ya conocemos y lo que nos gustaría conocer, una zona de aprendizaje potencial. A cada instante y entre las acciones que tenemos a disposición, elegiríamos aquellas con más posibilidades de reducir esa brecha de conocimiento y permitirnos adquirir información útil. La curiosidad funcionaría entonces como un pequeño sistema cibernético, parecido al famoso regulador de Watt que, en las locomotoras a vapor, abre o cierra una válvula de escape para así conservar una presión de vapor constante. En el caso de la curiosidad, se trataría de un regulador que permanentemente busca preservar determinada presión de aprendizaje. La curiosidad nos guía hacia lo que nos parece que podemos aprender, mientras que nos aleja de aquello que ya sabemos, o de las áreas que, según dicta nuestra experiencia, ya nada tienen para enseñarnos.

Así, la curiosidad no tiene una relación directa con el grado de sorpresa o de novedad, sino que traza una curva de campana (Kang y otros, 2009; Kidd, Piantadosi y Aslin, 2012, 2014; Loewenstein, 1994). No sentimos curiosidad alguna por aquello que ya vimos mil veces. Tampoco sentimos ni una brizna de atracción por las cosas demasiado nuevas y confusas, cuya estructura resulta inabordable: presentan una complejidad tal que nos disuade. Entre la monotonía de lo demasiado simple y el rechazo de lo demasiado complejo, nuestra curiosidad nos orienta naturalmente hacia los ámbitos nuevos y accesibles. Pero esta atracción cambia a cada instante. A medida que los dominamos, perdemos interés en los objetos que nos parecían atractivos, y reorientamos la curiosidad hacia otros desafíos novedosos. Así, una beba que inicialmente se apasiona por las cosas más triviales (jugar con los dedos del

pie, cerrar los ojos, esconderse detrás de la mano...), porque todo le resulta una fuente potencial de aprendizaje, una vez que absorbió todo el saber posible de estas experiencias, pierde el interés. Por ese mismo motivo ya ningún científico reproduce las experiencias de Galileo: lo conocido se vuelve aburrido.

Pero el algoritmo mismo explica también que a veces nos desviemos de un campo que nos atraía pero que luego se revela demasiado difícil: el cerebro evalúa la velocidad de aprendizaje y la curiosidad se extingue si detecta que no avanzamos lo suficiente según nuestras expectativas. Por eso, un niño puede ingresar a una orquesta en alas de su pasión por el violín para alejarse semanas más tarde, cuando nota que aprender a tocarlo no es algo que se logre en cuestión de días.

Dos especialistas en robótica franceses, Frédéric Kaplan y Pierre-Yves Oudeyer, lograron concretar estas ideas en un robot (Gottlieb y otros, 2013; Kaplan y Oudeyer, 2007), mediante un algoritmo que abarca varios módulos. El primero es un sistema clásico de inteligencia artificial, que permanentemente intenta predecir el estado del mundo exterior. El segundo, más innovador, evalúa el desempeño del primero: mide la velocidad del aprendizaje reciente y la utiliza para predecir en qué campos el robot asimila más información nueva por segundo. Por último, el tercer módulo es un circuito de recompensa, que asigna más valor a las acciones que, según se supone, deben llevar a un mejor aprendizaje. Resultado: de por sí, el sistema se orienta hacia los campos en que estima que aprenderá más, lo que constituye la definición misma de la curiosidad según Kaplan y Oudeyer.

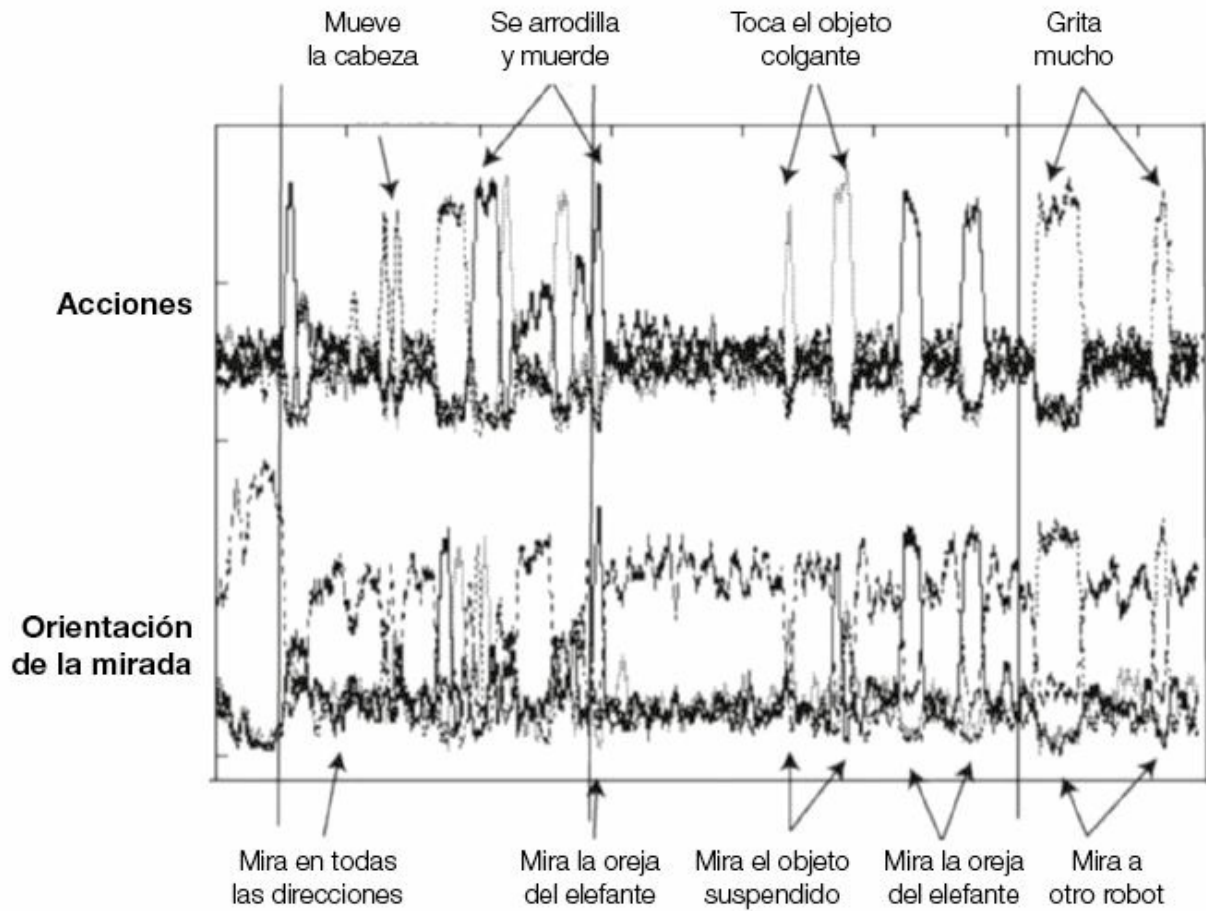


Figura 32. La curiosidad es un ingrediente esencial del algoritmo de aprendizaje que emplea nuestro cerebro. Actualmente se lo reproduce en las máquinas. Aquí, un pequeño robot explora una manta para bebés. Lo guía una función de recompensa que a cada instante le hace elegir la acción que maximiza el aprendizaje. Tan pronto como el robot domina un parámetro, una porción del mundo, esa zona pierde su interés y él reorienta su atención hacia otro lugar.

Cuando se lo posa sobre una manta, este robot curioso, provisto de este algoritmo, se comporta exactamente como un niño pequeño (figura 32). Durante algunos minutos, se apasiona por un elemento en especial y, por ejemplo, pasa todo su tiempo levantando la oreja de un elefante de tela. Pero a medida que aprende a utilizarlo, su curiosidad disminuye. En un momento, se desvía y busca activamente otras fuentes de estimulación. Luego de una hora, deja de explorar la manta: aparece una forma digital de aburrimiento, ya que el robot considera que todo lo aprendible le resulta conocido.

La analogía con un niño pequeño es impactante. Incluso los bebés de pocos meses se orientan hacia los estímulos de complejidad intermedia, ni demasiado sencillos ni demasiado complejos, pero lo suficientemente estructurados como para que puedan aprenderlos (lo que fue descrito como “efecto Ricitos de Oro”; Kidd, Piantadosi y Aslin, 2012, 2014). Para maximizar su aprendizaje, será necesario enriquecer a cada instante el ambiente con objetos novedosos, lo bastante estimulantes como para que no resulten desalentadores. Es responsabilidad de los adultos proporcionarles una jerarquía pedagógica bien concebida, que los conduzca de modo sostenido y gradual hacia la cumbre, incentivando constantemente su apetito por el saber y la novedad.

Este enfoque lleva a una predicción interesante: para ser curiosa, una criatura debe ser consciente de lo que no sabe, y, según Kaplan y Oudeyer, quizá también de su velocidad de aprendizaje. En otras palabras: desde la edad más temprana debe poseer facultades que se dan en llamar “metacognitivas”. La metacognición es la cognición sobre la cognición, vale decir, el conjunto de sistemas que supervisan los aprendizajes y evalúan todo el tiempo qué sabemos y qué no, si nos equivocamos o no, si somos rápidos o lentos, y así sucesivamente: todo lo que conocemos sobre nuestra propia mente.

La metacognición desempeña un papel determinante en la curiosidad. En efecto, ser curioso es querer saber, y eso supone también saber lo que todavía ignoramos. También en este punto, existen experiencias muy recientes que

confirman que, al cumplir 1 año de vida, y tal vez incluso antes, los niños comprenden que hay cosas que no saben (Dehaene, Lau y Koudier, 2017; Goupil, Romand-Monnier y Kouider, 2016; Lyons y Ghetti, 2011). En efecto, a esa edad, los bebés acuden rápidamente a la persona adulta que los acompaña cuando no logran resolver un problema solos. El hecho de saber que no saben los impulsa a pedir más información. Esto ya es la manifestación de una forma de curiosidad epistémica: el deseo irresistible de saber.

Tres maneras de atentar contra la curiosidad en la escuela

Todos los padres y las madres sienten nostalgia de la infancia temprana, la etapa en que la mente de su hijo brillaba de curiosidad. Entre los 2 y los 5 años, la niña siente curiosidad por todo. Su frase favorita es “¿por qué?”: no deja de experimentar con el mundo y de interrogar a los adultos con el objetivo de aplacar su sed de conocimiento. Sin embargo, este apetito que parecía insaciable termina por decaer, muchas veces luego de unos pocos años en la escuela. Algunos niños no pierden la curiosidad por todo, pero muchos otros se cierran a esa intriga. El compromiso activo se vuelve triste pasividad. ¿La ciencia de la curiosidad puede explicar por qué? Todavía no tenemos todas las respuestas, pero me gustaría proponer aquí algunas hipótesis que no son por completo arbitrarias.

- *Primera hipótesis: la falta de estimulación apropiada al nivel del niño.* De acuerdo con el algoritmo que acabamos de describir, es normal que la curiosidad disminuya con el transcurso del aprendizaje: cuanto mejor dominamos un área, más nos encontramos con los límites del aprendizaje que nos ofrece y menos nos interesa. Para sostener la curiosidad, entonces, es necesario que la escuela continúe proveyendo a la supercomputadora que es el cerebro estímulos a la altura de la inteligencia de los niños. Sin embargo, eso no siempre sucede. A los alumnos más avanzados puede faltarles estimulación: luego de algunos meses, su curiosidad se desvanece y ya no esperan gran cosa de la escuela, porque sus sistemas metacognitivos

asimilaron que, por desgracia, no aprenderán mucho más allí. Existe una solución: estimular su curiosidad dándoles, otra vez, grano para moler: idiomas para descifrar, rompecabezas para resolver, verdaderos desafíos para su inteligencia. En el otro extremo de la curva, los alumnos con dificultades pueden decaer por el motivo contrario. La causa siempre reside en la metacognición: ya no tienen motivos para ser curiosos, porque han aprendido... que no lograrán aprender. Su experiencia pasada dejó grabada, en lo más profundo de sus circuitos metacognitivos, una regla simple (aunque falsa): “Soy incapaz de aprender tal o cual campo del plan escolar” (ya sea matemáticas, lectura, historia o cualquier otro). Semejante consternación no es inusual: algunas niñas se convencen de que las matemáticas no son para ellas (Spencer, Steele y Quinn, 1999; Steele y Aronson, 1995), mientras que los niños de barriadas marginales llegan a creer que la escuela es hostil con ellos y que no les enseñará nada útil para su futuro. Este tipo de juicios metacognitivos son desastrosos, porque desmotivan a los estudiantes y aniquilan su curiosidad desde el principio. La única solución es volver a despertar la confianza de estos niños demostrándoles, por medio de problemas adaptados a su nivel, que son perfectamente capaces de aprender, a su ritmo, y que llegar a esto es en sí una recompensa. La teoría de la curiosidad postula que, cuando un niño está desmotivado, sin importar si es brillante o tiene dificultades, antes que nada hay que volver a transmitirle el placer del aprendizaje proponiéndole estímulos adaptados a sus capacidades.

- *Segundo disuasor: el castigo a la curiosidad.* El apetito de descubrimiento de los niños se puede aniquilar desde el principio por obra de una organización escolar demasiado rígida. La enseñanza tradicional, que se vale de la clase magistral, puede disuadirlos de intervenir o incluso de reflexionar. Puede causarles la impresión de que lo único que se les pide es que se sienten y se queden tranquilos hasta el final de la clase. La interpretación neurofisiológica de esta situación es sencilla: dentro del circuito de la dopamina, la señal de recompensa ligada a la curiosidad entra en competencia con las otras recompensas y penalidades exteriores. Entonces, es posible desalentar la curiosidad sancionando cada intento de exploración con una penalización. Imaginen a una niña a quien, cada vez que intenta intervenir, se la regaña, se la burla o se la castiga: “Pregunta idiota”,

“Mejor sería que te quedaras callada”, “Te quedarás media hora más a estudiar”... Su cerebro aprenderá muy rápido a inhibir la curiosidad y dejar de participar en clase: la recompensa ligada a la curiosidad que anticipaba su sistema de dopamina –el placer de aprender algo nuevo– se ve enormemente compensada por las señales negativas que recibe ese mismo circuito. El castigo repetido implica un síndrome de impotencia adquirida [*learned helplessness*], un tipo de parálisis física y mental asociada al estrés y a la ansiedad, que, según demuestra la investigación en animales, inhibe los aprendizajes (Caroni, Donato y Muller, 2012; Donato, Rompani y Caroni, 2013; Kim y Diamond, 2002; Noble, Norman y Farah, 2005). ¿La solución? Es muy conocida para la mayoría de los docentes. Es cuestión de recompensar la curiosidad y no de penalizarla: alentar las preguntas, pedirles a los niños que hagan exposiciones sobre lo que los apasiona, felicitar a cada alumno por sus iniciativas, aunque sean torpes... Las neurociencias de la motivación son claras en extremo: para tener ganas de realizar la acción X, hace falta anticipar que traerá aparejada una recompensa, que puede ser directa (alimentación, sexo, confort) o cognitiva (adquisición de información). Muchos niños y niñas perdieron cualquier tipo de curiosidad, todas y cada una de las motivaciones para aprender, porque se acostumbraron a no esperar recompensas por parte de la escuela (las calificaciones o notas numéricas, sobre las que hablaremos más adelante, suelen contribuir a este triste estado de las cosas).

- *Tercer factor que puede desalentar la curiosidad: la transmisión social del conocimiento.* Recuerden que la especie humana dispone de dos modos de aprendizaje: el modo activo, en que el niño experimenta constantemente y se interroga como un científico en ciernes; y el modo receptivo, pedagógico, en que se contenta con registrar lo que otras personas le enseñan. La escuela solo alienta el segundo modo, y acaso desaliente gradualmente el primero, a medida que el niño descubre que el docente siempre sabe más que él. Una experiencia reciente demuestra que este efecto es real: la actitud del docente puede pulverizar la curiosidad del niño (Bonawitz y otros, 2011). En su laboratorio de cognición infantil en el MIT, la psicóloga del desarrollo estadounidense Laura Schultz les presenta a alumnos del jardín de infantes un extraño artilugio: un conjunto de tubos de plástico que

incluyen juguetes inesperados de todo tipo (allí están escondidos un espejo, una bocina, un juego de luces y una caja de música). Si se le entrega este artefacto a una niña sin hacer mayor comentario, se desencadena inmediatamente un comportamiento de exploración: busca, husmea y rebusca en cada rincón hasta que encuentra todos los tesoros escondidos. Ahora, tomen un nuevo grupo de jardín de infantes e induzcan a los niños al modo pedagógico pasivo, receptivo. Todo lo que tienen que hacer es darles el objeto diciéndoles: “Miren, les voy a mostrar mi juguete, hace así”... y enciendan, por ejemplo, la caja de música. Podríamos pensar que esta frase estimula la curiosidad de los niños... pero el efecto es inverso: la exploración disminuye enormemente. El niño parece construir la hipótesis (muchas veces correcta) de que el docente intenta ayudarlo al máximo, y que por ese motivo le presentó todas las funciones interesantes del aparato. En este contexto, ya no hay necesidad de buscar: la curiosidad se ve inhibida. Otros experimentos revelan que los niños tienen en cuenta la actitud previa del docente. Si este siempre hace demostraciones exhaustivas, entonces los niños dejan de buscar. Cuando se ven confrontados a un juguete nuevo, del cual el docente muestra y explica una función, ya no exploran todas sus facetas, porque piensan que el docente agotó cuanto había que saber al respecto. Si, en cambio, el maestro los deja ver que no siempre sabe todo, entonces los niños no desisten de buscar.

¿Cuál es entonces el método correcto? Yo recomiendo tener siempre presente el concepto de compromiso activo. Involucrar al máximo la inteligencia del niño significa tenerlo alerta con preguntas y observaciones que estimulen su imaginación, y que despierten sus ganas de ir más allá. No se trata de dejar que el alumno descubra todo por sí solo, lo que equivaldría a caer en el escollo de las pedagogías del descubrimiento. Lo ideal es proponer una pedagogía estructurada, pero que aliente la creatividad del niño haciéndole saber que aún quedan mil cosas por descubrir. Recuerdo a un docente que, justo antes de las vacaciones de verano, me dijo: “Acabo de leer un pequeño problema de matemáticas que no pude resolver”. Y así fue cómo pasé todo el verano reflexionando sobre ese problema, en un intento por superar al profesor...

Suscitar el compromiso activo de los niños va de la mano de otro

imperativo: tolerar el error y corregirlo con rapidez. Este será el tercer pilar del aprendizaje que desarrollaremos.

[20] ¿Cómo sumar todos los números del 1 al 100? Se une el 1 con el 100, el 2 con el 99, el 3 con el 98, de modo tal que el resultado siempre sea 101; por último, se ve que la suma total abarca 50 pares de 101, es decir, 5050. [Al respecto, puede verse también Dehaene (2010). N. de E.]

9. El error es productivo y dar un buen *feedback* es garantía de mejores aprendizajes

Sería necesario aprender a equivocarse con buen humor. [...] Pensar es ir de error en error.

Alain (Émile Chartier), *Charlas sobre educación* (1932)

El único hombre que nunca comete un error es el hombre que nunca hace nada.

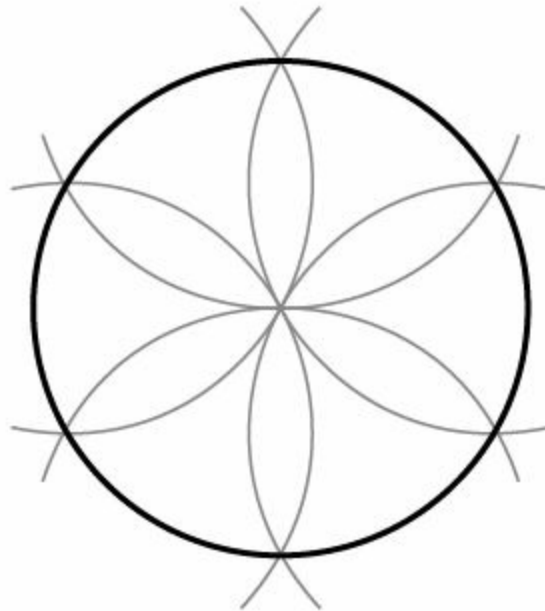
Atribuido a Theodore Roosevelt (1900)

En 1940, el joven Alexandre Grothendieck no tenía más que 11 o 12 años. No sabía que se convertiría en uno de los matemáticos más influyentes del siglo XX ni que inspiraría a una generación entera (a instancias de él se fundó, en 1958, el célebre Instituto de Altos Estudios Científicos de Bures-sur-Yvette, que respaldó a tantos futuros galardonados con la Medalla Fields). Pero el joven Alexandre ya trabajaba con las matemáticas... con mayor o menor alegría. Prestemos atención a sus memorias:

Hacia la edad de 11 o 12 años, cuando estaba recluido en el campo de concentración de Rieucros (cerca de Mende), descubrí los juegos de trazos con el compás, y quedé particularmente encantado por las rosetas de seis secciones que se obtienen al dividir la circunferencia en partes iguales con ayuda de la apertura del compás mismo, trasladada a la circunferencia seis veces, lo que lleva de regreso al punto exacto de partida. Esta constatación experimental me convenció de que la longitud de la circunferencia era exactamente igual a *seis veces* la del radio. Cuando luego [...] vi en un manual que se suponía que la relación era tanto más compleja, que teníamos

$L = 2 \pi r$, en que $\pi = 3,14\dots$, me convencí de que el libro se equivocaba y seguramente los autores del libro [...] nunca habían hecho este ejercicio de trazado tan sencillo, que mostraba claramente que $\pi = 3$.

Esta confianza que un niño puede tener en sus propias ideas, dejándose llevar por sus facultades en vez de dar por sentado lo aprendido en la escuela o lo leído en los libros, es valiosísima. Sin embargo, quienes lo rodean no dejan de desalentarla a cada instante. Muchos verán en la experiencia que reseño aquí el ejemplo de una presunción infantil, que luego debió rendirse ante el saber recibido, ya que al final su cuota de ridículo salió a relucir por el peso mismo de los hechos. Sin embargo, el modo en que yo viví este episodio en nada me daba la sensación de un desacierto, de un ridículo, sino antes bien la de un nuevo descubrimiento: [...] el de un error (Grothendieck, 1986).



Qué confesión extraordinaria y qué lección de humildad del mejor matemático del mundo: admite haber creído durante mucho tiempo que el número π era igual a 3... Y sin embargo, Grothendieck tiene razón en algo más: los errores desempeñan un papel crucial en el aprendizaje. Cometer errores es la forma más natural de aprender; así, aprendizaje y error se tornan casi sinónimos, porque cada equivocación ofrece una oportunidad.

Los Shadoks, pajarracos regordetes del espacio exterior que protagonizaban

una serie francesa de animación muy popular durante mis años de infancia, elevaron caprichosamente este concepto al rango de precepto, con el humor absurdo que los caracterizaba: “Al intentar constantemente terminamos por tener éxito. Luego, cuanto más falla algo, más posibilidades tenemos de que funcione”. Y con esta lógica perfecta e implacable, dado que el cohete que intentaban lanzar tenía solo una posibilidad en un millón de despegar, los Shadoks apuraron los 999.999 fracasos iniciales para finalmente alcanzar el éxito...

Más allá y más acá del humor, es cierto que resulta casi imposible progresar si no empezamos por fallar. Siempre que recibamos comentarios sobre cómo mejorar, siempre que tengamos una señal que nos indique el camino correcto, los errores tenderán a disminuir. Por eso, lograr un buen *feedback* –vale decir, una devolución amable que detecte, explique y corrija el error– es el tercer pilar del aprendizaje, y uno de los parámetros educativos más influyentes: la calidad y la precisión de la devolución (la señal que recibimos sobre nuestro error) son cruciales en la velocidad con la cual aprendemos.[\[21\]](#)

La sorpresa, motor del aprendizaje

Recordemos los algoritmos de aprendizaje que ya analizamos en el primer capítulo, y que permiten a un tirador ajustar su puntería o a una red neuronal artificial aprender al corregirse. La premisa es sencilla: en primer lugar intentemos, aunque eso conlleve fallar, porque el tamaño y la orientación del error nos indican cómo mejorar en el siguiente intento. Apuntamos, disparamos, evaluamos por cuánto nos alejamos del objetivo y utilizamos esta información para corregir el próximo disparo. Así, un tirador aprende a apuntar y (en una escala más amplia) una red neuronal artificial logra ajustar los millones de parámetros que definen su modelo interno del mundo exterior.

Pero ¿el cerebro funciona de la misma manera? Desde la década de 1970, se acumula la evidencia en favor de esta hipótesis. Dos investigadores estadounidenses, Robert Rescorla y Allan Wagner, proponen que el cerebro aprende únicamente si percibe un desfase entre lo que predice y lo que recibe. Aprender resulta imposible en ausencia de una señal de error: “Los

organismos solo aprenden cuando los acontecimientos contrarían sus expectativas” (Rescorla y Wagner, 1972). En otras palabras, la sorpresa es uno de los *motores* fundamentales del aprendizaje.

La teoría de Rescorla y Wagner explica muy bien los detalles de un paradigma de aprendizaje llamado “condicionamiento clásico”. Todos conocemos el caso del perro de Pávlov. En los experimentos de condicionamiento pavloviano, un perro oye una campana, que al principio es un estímulo neutro y sin efecto alguno. Con todo, después de que el tañido de la campana, por repetición, queda asociado con la comida, termina por provocar un reflejo condicionado. El perro saliva cada vez que lo oye, porque ya aprendió que este sonido precede sistemáticamente a la llegada de la comida. Al respecto, la regla de Rescorla y Wagner supone que el cerebro utiliza los estímulos sensoriales (las sensaciones generadas por la campana) para predecir la probabilidad de que los acompañe otro estímulo (el alimento). Estipula que:

- El cerebro se forja una predicción al ponderar los estímulos sensoriales.
- Luego calcula la diferencia entre la predicción y el estímulo realmente obtenido: se trata de lo que se conoce como “error de predicción”, concepto fundamental de la teoría, que determina el grado de sorpresa asociado a un estímulo.
- Después, el cerebro corrige su representación interna, en proporción directa con la fuerza del estímulo y el valor del error de predicción, con el objetivo de que su próxima predicción esté más cerca de la realidad.

Sin ambages, esta teoría incluye los tres pilares que ya postulamos: el aprendizaje solo se produce si el cerebro amplifica las entradas sensoriales apropiadas (atención), si las utiliza para generar una predicción (compromiso activo) y si logra determinar su precisión (*feedback*).

La ecuación que Rescorla y Wagner formularon en 1972 se demostró notablemente profética. Es casi idéntica a la que luego se utilizaría en las redes neuronales artificiales con el nombre de “regla delta”. Una y otra son casos peculiares de la regla de retropropagación de errores que en la práctica

emplean todos los sistemas actuales de aprendizaje supervisado (en que enseñamos explícitamente a la red una respuesta bien precisa). Además, en el caso del aprendizaje por recompensa (en el cual se indica a la red en qué grado exacto se equivoca) se utiliza otra ecuación similar: la red predice la recompensa y la diferencia entre su predicción y la recompensa efectiva se utiliza para actualizar la representación interna.

Por lo tanto, podemos afirmar que las máquinas de aprendizaje actuales dependen de ecuaciones directamente inspiradas en las ciencias del cerebro. Pero, como vimos, el cerebro humano llega todavía más lejos: para obtener un máximo de información de cada episodio de aprendizaje, es probable que se valga de un lenguaje del pensamiento y de modelos estadísticos tanto más refinados que las redes neuronales artificiales de la actualidad. Sin embargo, la idea fundamental de Rescorla y Wagner aún es exacta: el cerebro intenta anticipar sobre la base de los estímulos que recibe, y ajusta estas predicciones de acuerdo con la sorpresa, la improbabilidad o el error que comete. Aprender es reducir lo impredecible.

La teoría de Rescorla y Wagner tuvo una influencia considerable, porque constituyó un progreso muy importante respecto de las teorías previas, basadas sobre el concepto de aprendizaje asociativo. En épocas pasadas, existía la muy divulgada creencia de que el cerebro simplemente aprendía a “asociar” la campana y el alimento, en vez de predecir la llegada de la comida a partir de la campanada: se contentaba con un registro pasivo de todas las coincidencias entre los estímulos y las respuestas. Ahora bien, esta concepción se demostró por completo falsa incluso para el condicionamiento pavloviano.[\[22\]](#) Ni siquiera el cerebro de un perro es un órgano pasivo que absorbe asociaciones. El aprendizaje es activo y depende del grado de sorpresa causado por la violación de nuestras expectativas.

Una de las experiencias más notables que echan por tierra la noción de aprendizaje asociativo es la del bloqueo (Beckers y otros, 2006; Fanselow, 1998; Waelti, Dickinson y Schultz, 2001). En este caso, se expone al animal no a uno, sino a dos índices sensoriales –por ejemplo, una campana y una luz–, como anuncio de la inminente llegada del alimento. El secreto consiste en presentarlos uno después del otro. Se comienza con la luz: el animal aprende que cuando esta se enciende predice la llegada del alimento. Luego se presentan ensayos dobles, en que la campana y la luz predicen el alimento. Por último, se prueba el efecto de la campana sola. Sorpresa: ¡no tiene efecto alguno! Cuando la oye, el animal no saliva, parece que no retuvo ni una

brizna de la asociación, tantas veces repetida, entre la campana y la recompensa. ¿Qué sucedió? Este hallazgo es incompatible con el asociacionismo, pero puede explicarse perfectamente desde la teoría de Rescorla y Wagner. La idea clave es que la adquisición de la primera asociación (luz → alimento) bloquea la segunda (campana → alimento), porque la predicción que se debe solo a la luz explica todo. El animal ya sabe que la luz predice la comida, por lo que su cerebro no genera errores de predicción durante la segunda parte de la prueba, cuando la campana y la luz juntas predicen la comida. Cero error, cero aprendizaje; por lo tanto, el perro no adquiere conocimiento alguno de la asociación entre el sonido y el alimento. Cualquier regla incorporada primero bloquea el aprendizaje de la segunda.

Este experimento de bloqueo directo demuestra con claridad que el aprendizaje no funciona por asociación. A fin de cuentas, la asociación de la campanada con la comida se repitió cientos de veces, pero, como señalábamos, no logró inducir aprendizaje alguno. El experimento también revela que, en ausencia de sorpresa, no se produce aprendizaje: el error de predicción resulta esencial para el aprendizaje, al menos para los perros. De hecho, cada vez hay evidencia más convincente de que las señales de error de predicción son omnipresentes; ocurren en todo tipo de especies y en todo el cerebro.

Pero ¡atención! Hablamos de una señal de error interna, no necesariamente de un error efectivo. La teoría no implica que el animal tenga que cometer un error real para aprender. Supongamos que debo descubrir la respuesta correcta entre dos opciones posibles; por ejemplo, si la capital de Sudáfrica es Pretoria o Ciudad del Cabo. ¿La teoría da por sentado que si tengo la suerte de aventurar una suposición correcta en el primer intento (al decir “Pretoria”), no aprendo nada? Desde luego que no: incluso si respondí correctamente, mi confianza fue endeble y mi predicción, incierta. Por puro azar, sin ayuda, tenía una probabilidad del 50% de estar en lo cierto. En este caso, el *feedback* (los comentarios que recibí) me aporta una información nueva: la certeza de que mi respuesta al azar era correcta. De acuerdo con la ecuación de Rescorla y Wagner, esta información nueva genera una señal de error: mide el desfase entre la predicción (un 50% de posibilidades de tener razón) y lo que se consigue saber al final (el 100% de certidumbre de conocer la respuesta correcta). En mi cerebro, esta señal de error se propaga y actualiza mis conocimientos, lo que aumenta mis posibilidades de responder “Pretoria” la

próxima vez que me pregunten. Por lo tanto, sería erróneo creer que lo importante para el aprendizaje es cometer muchos errores, ¡como los Shadoks, que fallan al lanzar sus primeros 999.999 cohetes! Lo importante es la sorpresa, el desfase entre la predicción y la realidad; en definitiva, en eso consiste lo que llamamos “señal de error”. Así, resulta fundamental recibir comentarios explícitos que reduzcan la incertidumbre de quien aprende.

Si no hay sorpresa, el aprendizaje es poco o nulo: actualmente, esta regla parece validada en organismos de todo tipo, incluido el niño muy pequeño. Recuerden que la sorpresa es uno de los parámetros reveladores de las competencias precoces de los bebés, cuando abren de par en par los ojos y miran durante más tiempo si se les presentan eventos sorprendentes, que por arte de magia violan las leyes de la física, la aritmética, las probabilidades o la psicología (véanse las figuras 7 y 8). Pero el niño no se contenta con fruncir el ceño; cada vez que se sorprende, aprende.

La psicóloga estadounidense Lisa Feigenson llegó a esta conclusión a partir de numerosas experiencias que demuestran que cada vez que el niño percibe un evento imposible o improbable, se facilita el aprendizaje (Stahl y Feigenson, 2015). De este modo, cuando los bebés ven un objeto que atraviesa misteriosamente un muro, retienen mejor el ruido que hace o el verbo que un adulto acaba de emplear (“Como ves, acabo de blicar el juguete”). Cuando se les entrega el objeto en la mano, lo exploran durante más tiempo que a otro que no violó las leyes de la física. Este comportamiento sugiere que intentan comprender: como científicos en pañales, realizan experimentos para intentar reproducir lo que acaban de ver. Si el objeto atravesó un muro, lo tocan, como para constatar su solidez, mientras que, si lo vieron transgredir las leyes de la gravedad y permanecer misteriosamente suspendido en el aire, lo hacen caer de la mesa, para verificar si es capaz de levitar. En otras palabras, la índole del error que el niño acaba de observar define cómo actuará después para ajustar sus hipótesis. Exactamente eso predice la teoría de la propagación de errores: cada evento imprevisto trae aparejado un ajuste de los modelos internos del mundo.

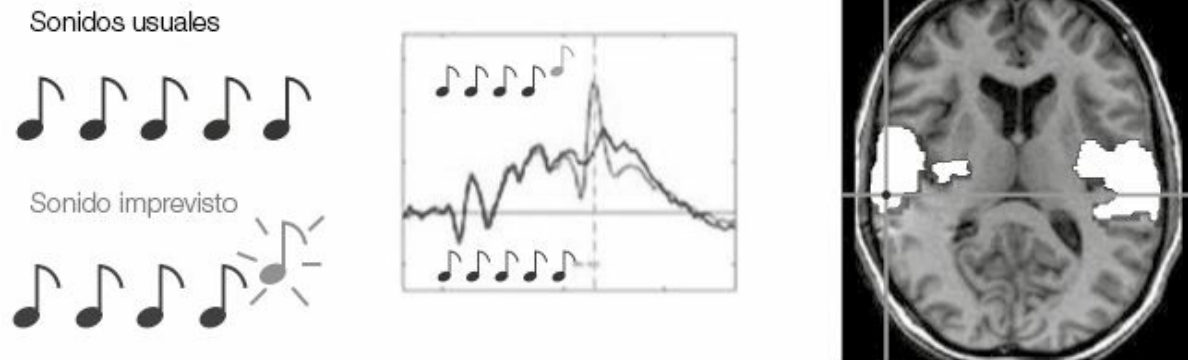
Todos estos fenómenos fueron documentados en bebés de 11 meses, pero probablemente estén presentes desde mucho más temprano. El aprendizaje por corrección de errores tiene validez universal en el mundo animal y existen muchos motivos para pensar que las señales de error rigen el aprendizaje desde el principio mismo de la vida.

El cerebro está repleto de mensajes de error

Las señales de error tienen un papel tan fundamental en el aprendizaje que casi todas las áreas cerebrales emiten e intercambian mensajes de ese tipo (figura 33; Friston, 2005; Naatanen y otros, 2007; Schultz, Dayan y Montague, 1997). Tomemos un ejemplo sencillo: imaginen que oyen una serie de notas musicales, siempre las mismas, *do do do...* Cuando se repite la nota, las respuestas cerebrales disminuyen: es la *adaptación*, fenómeno que revela que el cerebro aprende a predecir cómo seguirá la serie. De pronto, de modo imprevisible, la nota cambia: *do do do re*. El área auditiva primaria responde de inmediato con una reacción de sorpresa: la adaptación se desvanece y otras neuronas más comienzan a activarse vigorosamente ante lo inesperado. Así, la adaptación se debe tanto a la repetición como a la simple predictibilidad de las notas. Efectivamente, si se presenta una alternancia de notas como *do re do re do re...*, el cerebro se habitúa a este vaivén sistemático, su respuesta disminuye, y una secuencia inesperada (*do re do re do **do***) desencadena la sorpresa (Strauss y otros, 2015; Todorovic y De Lange, 2012).

Al parecer, el área auditiva realiza un cálculo sencillo: utiliza el pasado reciente para predecir el futuro. Dado que poco antes se repitió una nota o un grupo de notas, esta región llega a la conclusión de que otro tanto sucederá en el futuro. Esta predicción resulta útil, porque evita prestar atención a las señales previsibles y aburridas. Cualquier sonido que se repite queda relegado al nivel de entrada, ya que su actividad es cancelada por una predicción esmerada. Como la señal sensorial que ingresa es casi idéntica a la predicción que el cerebro realizó, la diferencia es nula y no se propagan señales de error a las regiones cerebrales de niveles superiores. La sustracción de la predicción cancela la información que ingresa, pero únicamente en la medida en que es previsible. Por el contrario, todo cuanto viole las expectativas será amplificado. En efecto, el área auditiva reacciona como un filtro: únicamente transmite a los niveles superiores de la corteza la información sorprendente, imprevisible, que no logra explicar por sí sola.

Corteza auditiva: detección de una nota anómala



Corteza prefrontal: detección de una violación de la melodía

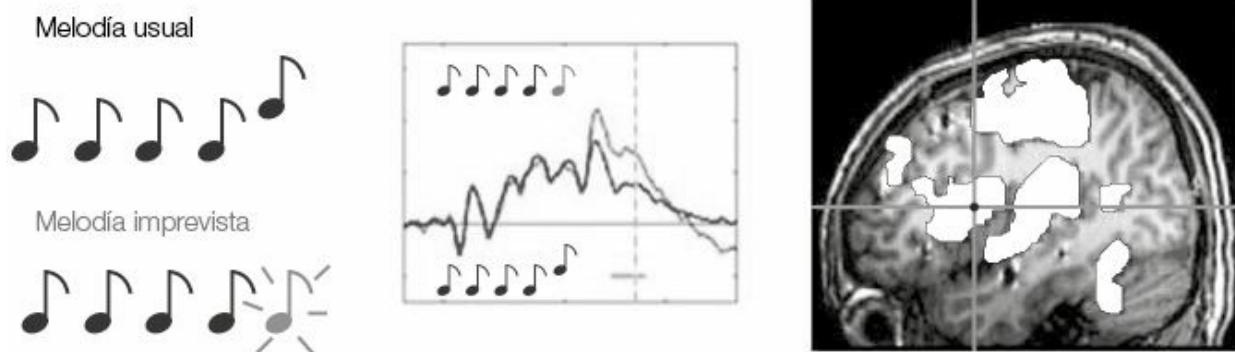


Figura 33. El error es fundamental para el aprendizaje. Al detectar los errores, el cerebro logra corregir sus modelos del mundo. Casi todas las regiones cerebrales emiten e intercambian señales de error. En esta experiencia se trata de un error auditivo. En primer lugar, se hace escuchar muchas veces una breve frase musical de cinco notas. Cuando la secuencia cambia sin aviso, una respuesta de sorpresa (en gris, tanto en el diagrama como en las corcheas) señala el error a otras regiones del cerebro, con lo cual les permite corregirse. Las áreas auditivas detectan las falsas notas locales (arriba), mientras que una red extendida, que llega hasta la corteza prefrontal, detecta las violaciones globales de la melodía (abajo).

Cada región cerebral deriva al nivel siguiente lo que no puede explicar, para que este intente, a su vez, comprenderlo. El proceso da lugar a una jerarquía de mensajes de error. Por ejemplo, como acabo de explicar, en un nivel bajo la melodía *do do do re* suscita una señal de error en la corteza auditiva cuando recibe el *re*, que difiere de los sonidos previos. Pero en un nivel más alto, es una melodía conocida, “Au clair de la lune”. La sorpresa causada por el *re*, entonces, es solo periférica; enseguida se explica por una representación de nivel superior y la señal se frena allí: el *re*, pese a su novedad, no implica sorpresa alguna en la corteza prefrontal inferior, que codifica las frases. En cambio, la repetición *do do do do* tiene el efecto inverso: como es monótona –en todos los sentidos del término–, no genera señal de error alguna en las áreas auditivas. Sin embargo, ese unísono causa sorpresa en las áreas de nivel más alto que codifican la melodía global: ustedes predijeron que subiría un tono hacia el *re* –eso que los músicos llaman “intervalo de segunda mayor”– pero a sus oídos llegó un nuevo *do*. Aquí, ¡la sorpresa es que no hay sorpresa! Hasta los monos macacos presentan, como nosotros, estos dos niveles de tratamiento auditivo: el tratamiento local de las notas en la corteza auditiva y la representación global de la melodía en la corteza prefrontal (Bekinschtein y otros, 2009; Strauss y otros, 2015; Uhrig, Dehaene y Jarraya, 2014; Wang y otros, 2015).

Este tipo de señales de error se observa en todas las regiones del cerebro: cuando los eventos se repiten y se vuelven previsibles, las descargas neuronales se habitúan; la sorpresa, es decir, la llegada repentina de un acontecimiento inesperado, desencadena un incremento drástico de la respuesta de las neuronas. Lo único que cambia, entre una región y otra, es el carácter de la predicción y de la novedad detectada. En la corteza visual, la aparición de una imagen inesperada relanza la actividad (Meyer y Olson, 2011); las áreas del lenguaje, por su parte, reaccionan ante una palabra anómala en una frase u oración.

Lean, por ejemplo, la siguiente oración:

Prefiero comer con tenedor y martillo.

Su cerebro acaba de generar una onda N400, una señal de error evocada por una palabra o una imagen incompatible con el contexto precedente (Curran y otros, 1993; Kutas y Federmeier, 2011; Kutas y Hillyard, 1980). Como su

nombre lo indica, se trata de una respuesta negativa que ocurre cerca de los 400 milisegundos (ms) después de la anomalía y surge de la actividad de poblaciones neuronales de la corteza temporal izquierda, que son sensibles al significado de las palabras. Por otro lado, el área de Broca, en la corteza prefrontal inferior, responde a los errores de sintaxis, cuando el cerebro predice determinada categoría de palabras y recibe otra, como en la oración que sigue: “No se olviden de tomar su cuando medicamento se sientan mal” (Friederici, 2002; Hahne y Friederici, 1999; pero véase un análisis crítico en Steinhauer y Drury, 2012).

Esta vez, sus áreas del lenguaje, especializadas en la sintaxis, emitieron una onda negativa e inmediatamente después una onda P600, una cota positiva que sucede cerca de 600 ms después de registrar la palabra “cuando”, que pertenece a una categoría inesperada, dado que después del “su” se espera un sustantivo. Esto constituye una evidencia de que el cerebro detectó un lapsus de gramática e intenta enmendarlo.

El circuito cerebral en que mejor se demostraron las predicciones y las señales de error es el de la recompensa (Pessiglione y otros, 2006; Schultz, Dayan y Montague, 1997; Waelti, Dickinson y Schultz, 2001). La red de la dopamina no solo responde a las recompensas reales, sino que las anticipa permanentemente. Las neuronas dopaminérgicas, situadas en un pequeño núcleo de células llamado “área tegmental ventral”, no responden de modo directo a los placeres del sexo, el alimento o la bebida; en realidad, se activan ante la diferencia entre la recompensa esperada y la obtenida, vale decir, ante el error de predicción. Por eso, cuando sin aviso previo le damos al animal, por ejemplo, una gota de agua azucarada, esta sorpresa placentera conllevará una descarga neuronal. Pero si ese mismo almíbar va precedido por una señal que lo torna previsible, entonces ya no suscitará la menor reacción. En ese caso, la señal en sí provocará un salto de actividad en las neuronas de dopamina: el aprendizaje desplaza la respuesta más próxima a la señal que precede la recompensa.

Gracias a estos mecanismos de aprendizaje predictivo, diferentes señales arbitrarias pueden convertirse en portadoras de recompensas y desencadenar una respuesta de dopamina. Este efecto secundario de recompensa se demostró en humanos ante el ofrecimiento de dinero o se confirma incluso cuando un toxicómano simplemente ve una jeringa. En esos dos casos, el cerebro anticipa constantemente recompensas futuras. Como ya señalamos en el primer capítulo, esta señal predictiva resulta de suma utilidad para el

aprendizaje, porque permite que el sistema se critique a sí mismo y prevea el error o el éxito de una acción sin esperar una confirmación externa. Por eso, la arquitectura actor-crítico, en que una red neuronal aprende a criticar y a predecir el resultado de las acciones de otra, es de uso universal en inteligencia artificial para resolver los problemas más complejos, como aprender a jugar go. Predecir, detectar el error y corregirse son los fundamentos mismos de un aprendizaje eficaz.

Feedback no es sinónimo de castigo

Frecuentemente me ha chocado el hecho de que los profesores de ciencias, aún más que los otros si cabe, no comprendan que no se comprende. Son poco numerosos los que han sondeado la psicología del error, de la ignorancia y de la irreflexión.

Gaston Bachelard, *La formación del espíritu científico* (1938)

¿Cómo podemos aprovechar al máximo las señales de error que nuestras neuronas intercambian constantemente? Para que un niño o un adulto aprenda de modo eficaz, es necesario que el entorno (según el caso, los padres, la escuela, la universidad... o un videojuego) le proporcione, con la mayor rapidez y precisión posibles, una devolución que le permita revisar el error. Una simple señal binaria (“correcto” o “incorrecto”) ya puede resultar útil. Este es el principio del aprendizaje *no supervisado*. Más allá de todo, para acelerar el aprendizaje, lo ideal es obtener comentarios detallados que consignen con precisión qué debía hacerse. Este es el principio del aprendizaje *supervisado*. Si el docente aporta la mayor exactitud posible en una devolución rápida y minuciosa sobre el error, enriquece considerablemente la información de que dispone el alumno para corregirse. En inteligencia artificial, este tipo de aprendizaje demostró ser el más eficaz, ya que ningún otro permite a la máquina detectar de inmediato el origen de la falla y corregirse, en el buen sentido de la palabra.

Sin embargo, es importante comprender que este *feedback* sobre el error nada tiene que ver con una sanción. No se aplica un “castigo” a una red neuronal, sino que simplemente se le informa en qué se equivocó, se le ofrece

una señal lo más informativa posible sobre el carácter y el signo de sus errores.

Informáticos y pedagogos se dan la mano en este aspecto. En efecto, los metaanálisis realizados por el australiano John Hattie, director del Melbourne Education Research Institute, demuestran que la calidad de los comentarios que reciben los estudiantes es uno de los determinantes de su éxito académico (Hattie, 2008). Fijar metas claras para el aprendizaje y permitir que los estudiantes lo encaren gradualmente, sin dramatizar los errores inevitables, son las claves del éxito.

Día tras días, los maestros, ya conscientes de estas ideas, son testigos del *dictum* latino *errare humanum est*: errar es humano. Con un ojo sensible, observan amablemente los errores de sus alumnos, porque se dan cuenta de que nadie aprende sin equivocarse. Saben que deben diagnosticar, de la manera más desapasionada posible, las áreas exactas de dificultad de sus estudiantes y ayudarlos a encontrar la mejor solución. Al contar con experiencia, crean un catálogo de errores, porque todos los estudiantes caen repetidamente en las mismas viejas trampas. Así, los docentes encuentran, con cuidadoso esmero, las palabras acertadas para reconfortar, tranquilizar y restaurar la confianza en sí mismos de sus alumnos, a la vez que les permiten corregir sus representaciones mentales erróneas. Están allí, a disposición, para decir la verdad, no para juzgar.

Por supuesto, los más racionales de ustedes me urgirán con comentarios como “Decirle a esa alumna lo que debería haber hecho, ¿no es lo mismo que decirle que se equivocó?”. Bueno, no exactamente. Desde un punto de vista lógico, claro que sí; cuando una pregunta solo tiene dos respuestas posibles, A o B, y un estudiante elige A, decirle “La respuesta correcta era B” es lo mismo que decirle “Te equivocaste”. Pero estamos olvidando que los estudiantes no son lógicos consumados. Para ellos, el paso adicional de deducir “si elegí A y me equivoqué, entonces la respuesta correcta debe haber sido B” no es tan inmediato. En cambio, cuando el mensaje principal es “te equivocaste”, que siempre tiene un tufillo a sanción, lo reciben perfectamente: “Metí la pata”. De hecho, cuando se realizó este experimento, los adultos, con una indiferencia cartesiana, lograban acceder a igual cantidad de información a partir de las recompensas que de las sanciones, pero los adolescentes fracasaban: aprendían más y mejor de sus éxitos que de sus errores (Palminteri y otros, 2016). *Ergo*, ahorremosles ese padecimiento,

démosles una devolución lo más neutra e informativa posible. Evitemos la confusión entre corregir y castigar.

La calificación, ese penoso sucedáneo del *feedback*

Llegada esta instancia, se vuelve indispensable que destine unas palabras a reseñar una institución escolar tan arraigada en nuestra tradición que no podemos imaginarnos la escuela sin ella, aunque sea el colmo de lo contraproducente: la nota numérica o calificación. De acuerdo con la teoría del aprendizaje, la nota es una señal de recompensa (¡o de castigo!). Sin embargo, una de sus peculiaridades consiste en que está completamente desprovista de precisión. A modo de simple cúmulo de nuestras faltas, resume, sin distinguirlas, diferentes fuentes de error. Entonces, no proporciona suficiente información, ya que por sí sola no permite saber *por qué* uno se equivocó o *cómo* puede corregirse. Incluso, hay casos en que nunca consigue ganar impulso y apartarse del cero absoluto: ¡está vacía de información! Solo es portadora del chirriante estigma social de incompetencia, deplorable etiqueta o sambenito.

Por ende, la nota, sin el acompañamiento de apreciaciones detalladas y constructivas, constituye un *feedback* muy pobre. No solo es equívoca, sino que suele llegar con varias semanas de demora, cuando el alumno ya sepultó en el olvido el detalle de qué motivó ese error. Por añadidura, la nota puede resultar rotundamente injusta, en especial para los alumnos que no logran ponerse al día porque el nivel de exigencia de los exámenes no deja de aumentar semana tras semana.

Hagamos una analogía con los videojuegos: cuando ustedes descubren un juego nuevo, al principio no saben cuál es la forma eficaz de subir de nivel, de progresar y progresar en ese mundo que acaban de conocer. Y sobre todas las cosas, ¡no tienen ganas de que se les recuerde permanentemente lo torpes que son! Por eso, los diseñadores de videojuegos presentan al comienzo niveles y pantallas más fáciles, que dan la satisfacción de ganar casi siempre. Poco a poco, se incrementa la dificultad y, con ella, el riesgo de fracaso y de frustración, que los ya fogueados programadores saben mitigar, mezclando lo fácil y lo difícil, y dando completa libertad para intentar el mismo nivel tantas

veces como sea necesario. De a poco, ustedes ven aumentar su puntaje... y al fin llega el gran día de la consagración, cuando logran superar el maldito obstáculo que durante tanto tiempo parecía insalvable. Comparen este panorama virtuoso con el boletín de notas (o registro de calificaciones, como prefieran llamarlo) de un “mal alumno” (también esa designación da mucho que pensar): comienza el año con una mala nota, y, en lugar de motivarlo y permitirle que haga la misma prueba hasta que logre resolverla, cada semana se le impone un nuevo ejercicio, que invariablemente está más allá de sus capacidades. Semana tras semana, su “puntuación” sigue tendiendo a cero. En el mercado del videojuego, un diseño así de desastroso sería un rotundo fracaso.

Con mucha frecuencia, en la escuela, la nota sirve como una sanción. No podemos ignorar sus enormes efectos sobre los sistemas emocionales del cerebro: desaliento, sentimiento de impotencia y de estigmatización... Escuchemos la voz de un experto, Daniel Pennac, un escritor consagrado, quien recibió en 2007 el Premio Renaudot por su libro *Mal de escuela*, pero que durante mucho tiempo fue un mal alumno:

Mis boletines escolares me lo confirmaban todos los meses: si yo era un tonto, era absolutamente por mi culpa. Eso me hacía odiarme a mí mismo, causándome un complejo de inferioridad y sobre todo de culpa... Me consideraba menos que nulo. Porque quien es un bueno para nada –que es lo que me repetían los profesores uno tras de otro– es nada... Yo no me veía ningún futuro, no tenía ninguna representación posible de mí mismo como adulto. No porque no deseara nada, sino porque me creía inepto en todo (Pennac, 2017).

Si bien el alumno Pennac demostró una resiliencia a prueba de balas, para la mayoría de los niños un estado de ánimo (¡o desánimo!) como este resulta desastroso, con todas las letras (y todos los números). Los efectos del estrés son estudiados sobre todo en el campo de las matemáticas, campeonas mundiales de la ansiedad escolar. Y precisamente en las horas de Matemáticas, algunos niños sufren un auténtico malestar, rayano con la depresión: saben que, sin importar lo que hagan, serán sancionados con el fracaso. La ansiedad matemática es un síndrome reconocido, medido, cuantificado: los niños que la padecen muestran activaciones de los circuitos

del dolor y del miedo, particularmente de la amígdala, un núcleo de neuronas situado por debajo de la corteza e involucrado en las emociones negativas (Ashcraft, 2002; Lyons y Beilock, 2012; Maloney y Beilock, 2012; Young, Wu y Menon, 2012). No son necesariamente menos inteligentes que los demás, pero el tsunami emocional que vivencian arrasa sus capacidades de cálculo, de memoria de corto plazo y, sobre todo, de aprendizaje.

Numerosas investigaciones, realizadas tanto en seres humanos como en otros animales, lo confirman: el estrés y la ansiedad son nocivos para la facultad de aprender. En el hipocampo del ratón, por ejemplo, el condicionamiento por el miedo bloquea literalmente la plasticidad neuronal. Una vez que el roedor fue traumatizado aleatoriamente mediante descargas eléctricas imprevisibles, el circuito se encuentra en un estado similar al del fin del período crítico, en que las sinapsis se vuelven inamovibles, fijadas en sus redes perineurales. A la inversa, si se lo inserta en un ambiente enriquecido, estimulante, la plasticidad neuronal se reabre y las neuronas recuperan su movilidad, como si se tratara de un regreso a la infancia (Caroni, Donato y Muller, 2012; Donato, Rompani y Caroni, 2013).

Por ende, asignar una mala nota mientras se la presenta como una sanción equivale a multiplicar el riesgo de inhibir los progresos del niño, ya que el estrés y el desaliento le impedirán aprender. A más largo plazo, acarrea otras consecuencias indeseables por nocivas, como cambios en la personalidad y en la imagen de sí mismo. La psicóloga estadounidense Carol Dweck estudió durante mucho tiempo los efectos negativos de la disposición mental que consiste en atribuir los fracasos (o los éxitos) a un ingrediente fijo, inmutable, de la personalidad, lo que ella llama una “mentalidad fija” [*fixed mindset*]: “Soy pésima en matemáticas”, “Las lenguas extranjeras no son mi fuerte”, y así sucesivamente. Para contrarrestar este enfoque, propone la idea, acertada en esencia, de que todos los niños pueden progresar, lo que ella llama una “mentalidad de crecimiento” [*growth mindset*].

Sus investigaciones sugieren que, a igual desempeño, la mentalidad tiene gran incidencia en el aprendizaje.^[23] De por sí, la idea de un progreso posible es una fuente de progreso. En cambio, los niños –pero también los docentes–, convencidos de que las habilidades están fijadas de modo definitivo, para siempre, y de que estamos dotados o no lo estamos, obtienen menor cantidad de buenos resultados. En efecto, el fijismo desmotiva: no alienta la atención ni el compromiso activo y lleva a interpretar el error como señal certera de una inferioridad intrínseca... ¡mientras que equivocarse es

natural y simplemente revela que esa persona lo intentó! ¿Adónde habría ido a parar Grothendieck si, a sus 11 años, si hubiera llegado a la conclusión de que era un desastre en matemáticas? La investigación muestra que la mentalidad fija incapacita por igual a los alumnos “malos” y a los “buenos”, quienes también necesitan trabajar para conservar su motivación: no les hacemos favor alguno si los dejamos creer que, como son “dotados”, no necesitan esforzarse.

Favorecer el desarrollo de una mentalidad de crecimiento no significa decirles a todos los niños que son los mejores con el simple pretexto de alimentar su autoestima, sino, antes bien, notar y señalar sus progresos cotidianos, alentar su participación, estimular su atención... y, por qué no, explicarles los fundamentos del aprendizaje: que cada cual debe esforzarse, que siempre hay que justificar una respuesta, asumiendo el riesgo de cometer un error, y que equivocarse (y corregir los errores) es la única forma de aprender.

Dejémosle la última palabra a Daniel Pennac: “Los profesores no están allí para darles miedo, sino para ayudarlos a vencer el miedo de aprender. Una vez superado este miedo, los alumnos son insaciables”.

Evaluarse para aprender mejor

Dado que la nota no tiene eficacia alguna, ¿cómo pondremos en práctica, en nuestras aulas, los conocimientos científicos acerca del procesamiento del error en el cerebro? Las reglas del juego son simples. En primer lugar, se debe lograr que el aprendiz se comprometa, aventure una respuesta, genere activamente una hipótesis, incluso si es incierta; luego, resulta indispensable ofrecerle información objetiva, no punitiva, que le permita corregirse. Hay una estrategia que responde a estos criterios, y todos los docentes la conocen: en el sector pedagógico, la llamamos... ¡la evaluación! Sin embargo, en la opinión pública no tiene mucha divulgación algo demostrado en decenas de publicaciones científicas: evaluar periódicamente los conocimientos, eso que los hablantes de inglés llaman *retrieval practise*, es una de las estrategias pedagógicas más eficaces (Carrier y Pashler, 1992; Karpicke y Roediger, 2008; Roediger y Karpicke, 2006; Szpunar, Khan y Schacter, 2013; Zaromb

y Roediger, 2010).[24] La evaluación periódica maximiza el aprendizaje a largo plazo. El simple hecho de poner a prueba la memoria la fortalece, como efecto directo del compromiso activo y del buen *feedback* acerca del error, que ya describimos aquí. Al realizar una prueba, uno se ve obligado a darse un baño de realidad y a darse cuenta de que no sabe (o todavía no sabe lo suficiente).

La enorme validez de las pruebas como instancia clave del aprendizaje está muy lejos de ser evidente para los principales actores involucrados. En su gran mayoría, los docentes y los alumnos consideran las pruebas y los exámenes como simples medios de calificación, una evaluación de conocimientos adquiridos en otra parte, durante la clase o durante los repasos. Sin embargo, la investigación deja en claro que la prueba desempeña un papel al menos tan importante como la clase misma. En una serie de experiencias que se volvieron famosas, el psicólogo estadounidense Henry Roediger y sus colaboradores propusieron a sus alumnos que memorizaran palabras durante un tiempo fijo, pero por medio de distintas estrategias. Un grupo recibió la consigna de pasar todo el tiempo estudiando, en ocho sesiones breves. A un segundo grupo se le presentaron seis sesiones de aprendizaje con dos evaluaciones intermedias. Por último, al tercer grupo se le pidió que alternara cuatro sesiones breves de estudio con cuatro evaluaciones. Como ustedes ya notaron, estas pruebas suplementarias reducían el tiempo dedicado al aprendizaje, sin que variase la duración total del programa. Y, sin embargo, el resultado fue palmario: cuarenta y ocho horas más tarde, el recuerdo de la lista de palabras mejoraba rotundamente en los alumnos que habían tenido la posibilidad de evaluarse con mayor frecuencia. Alternar periódicamente el aprendizaje y la prueba los obligaba a permanecer activos y a recibir comentarios enriquecedores, vale decir, una retroalimentación explícita: conozco tal palabra, pero jamás logro recordar tal otra. Este conocimiento de uno mismo, esta “metamemoria”, resulta útil, porque permite hacer foco sobre las palabras difíciles durante la segunda sesión de aprendizaje (Robey, Dougherty y Buttaccio, 2017). Así, quedó confirmado: cuanto más uno se evalúa, mejor retiene la clase.

Veamos otro ejemplo: imaginen que deben aprender palabras en una lengua extranjera, por ejemplo, “trineo” en lengua inuit: *qamutiik*. Pueden presentar las dos palabras juntas, una al lado de la otra, con el objetivo de asociarlas mentalmente. O bien pueden anteponer una de ellas a la otra (figura 34). La segunda opción, paradójicamente, reduce la cantidad de información

disponible: durante los primeros cinco segundos, ustedes solo ven la palabra *qamutiik*, sin que se les recuerde qué significa. Sin embargo, esa segunda opción es la que mejor funciona (Carrier y Pashler, 1992). ¿Por qué? Porque los fuerza a reflexionar, a intentar encontrar el significado de la palabra, antes de darles una devolución. El compromiso activo y el *feedback* adecuado acerca del error maximizan el aprendizaje.

Lo llamativo es que ni los estudiantes ni los docentes tienen conciencia de estos efectos. Si se les pide una opinión, todos consideran que lo importante es el estudio, no la evaluación. Así, predicen lo contrario de lo que la experimentación demuestra: en su opinión, cuanto más se estudia, mejores resultados se obtienen. Espontáneamente, los estudiantes dedican su tiempo a leer y releer las clases, resaltando cada línea con los colores del arcoíris... Se trata de estrategias mucho menos eficaces que ponerse a prueba con una breve evaluación.

¿Por qué tenemos la ilusión de que desvivirse estudiando es suficiente para aprender? Porque no vemos la diferencia entre los diversos compartimientos de nuestra memoria. Inmediatamente después de leer la lección, la información está presente en la mente, en la memoria de trabajo consciente, de modo activo. Tenemos la impresión de conocerla... pero esta memoria inmediata nada tiene que ver con la de largo plazo que necesitaremos para recuperar la información días más tarde. Luego de cinco minutos, la memoria de trabajo empieza a disiparse y, al cabo de unos días, este efecto se vuelve enorme: vale mucho más haber estudiado y haberse evaluado que haber pasado todo el tiempo estudiando (figura 34).

Nada más fácil que poner estas ideas en práctica por nuestra cuenta. Los invito a hacerlo. Todo se reduce preparar tarjetas o fichas de revisión: de un lado, escriben una pregunta, y del otro, la respuesta. Para ponerse a prueba, saquen una tras otra las tarjetas; en cada caso, intenten recordar la respuesta (predicción) y luego verifiquen al dorso (corrección del error). Si fallan, vuelvan a poner la tarjeta entre las primeras al alcance de la mano, lo que los forzarán a revisar esa misma noción en muy poco tiempo. Si esta vez la respuesta es acertada, dejen esa cartulina en una posición menos accesible: no hay necesidad de revisar inmediatamente su contenido, que sin embargo reaparecerá más tarde, cuando el olvido haya comenzado a ejercer su efecto. Este es el sistema de las *flash cards* que está en boga en las universidades estadounidenses y reside en la base de numerosos programas de aprendizaje, como la famosa plataforma de enseñanza de idiomas Duolingo.

La regla de oro: planificar intervalos entre los aprendizajes

¿Por qué la alternancia entre estudio y evaluación tiene efectos tan positivos? Porque aprovecha una de las estrategias más eficaces descubiertas por las ciencias de la educación: planificar intervalos entre los aprendizajes. Esta es la regla de oro: distribuir los períodos de entrenamiento en lugar de acumularlos. Al no agrupar todo el aprendizaje en una sola sesión, se alternan los períodos de estudio y los períodos de prueba, con intervalos cada vez más amplios, asegurando así la retención a largo plazo.

Décadas de investigaciones en psicología experimental confirman que sumar intervalos es una estrategia tanto más eficaz que la enseñanza de una sola vez (Cepeda y otros, 2009; Cepeda y otros, 2006; Rohrer y Taylor, 2006; Schmidt y Bjork, 1992). Si se dispone de un tiempo fijo para aprender algo, siempre es mejor segmentar las lecciones que aprender todo junto en un bloque. Distribuir el aprendizaje en varios días tiene grandes efectos: la experiencia demuestra que la memoria puede multiplicarse por tres si se revisa la información a intervalos regulares en lugar de intentar aprender todo de una sola vez. La regla es simple, y todos los músicos la conocen: valen más quince minutos de práctica todos los días de la semana que dos horas concentradas en una sola jornada.

¿Por qué esa estrategia es tan eficiente? Las neuroimágenes (Bradley y otros, 2015; Callan y Schweighofer, 2010) dejan de manifiesto que agrupar los problemas en una sola sesión disminuye la actividad cerebral, tal vez porque la información repetida pierde gradualmente su novedad. Como señalamos, la repetición también parece crear una ilusión de saber, un exceso de confianza debido a la presencia de información en la memoria de trabajo: parece estar disponible, la tenemos en la cabeza, por lo tanto, perdemos el interés de seguir trabajando sobre ella. En cambio, la distribución del aprendizaje aumenta la actividad cerebral: parece crear un efecto de “dificultad deseable”, que inhibe el simple almacenamiento en la memoria de trabajo y fuerza a los circuitos requeridos a trabajar más.

¿Cuál es el intervalo de tiempo más eficaz entre dos repeticiones de la misma lección? Al respecto, se observa una firme mejoría cuando el intervalo es de veinticuatro horas, probablemente porque, como veremos en un instante, el sueño desempeña un papel protagónico en la consolidación de los aprendizajes. Sin embargo, el psicólogo estadounidense Hal Pashler y sus colegas demostraron que el intervalo óptimo depende del tiempo que

deseemos retener el conocimiento en la memoria. Si ustedes necesitan recordar una información solo unos pocos días o semanas, entonces es ideal que la revisen todos los días durante cerca de una semana. Si, en cambio, desean que los conocimientos perduren varios meses o años, necesitarán incrementar el intervalo de revisión en una proporción directa. El efecto es sustancial: ¡una sola repetición de una lección, semanas después de la primera, multiplica por tres la cantidad de elementos que se recuerdan meses más tarde! Para almacenar la información en la memoria el mayor tiempo posible, lo mejor es aumentar gradualmente los intervalos de tiempo: se comienza con lecciones todos los días, luego se hace una revisión al final de una semana, de un mes, de todo un año... Esta estrategia garantiza una memoria óptima a cada instante (Kang y otros, 2014).

La figura 34 muestra por qué: cada revisión afianza el aprendizaje, aumenta la fuerza de las representaciones mentales y combate el olvido exponencial que caracteriza a nuestra memoria. Sobre todo, la distribución del aprendizaje a lo largo del tiempo parece seleccionar, entre los diferentes circuitos de la memoria disponibles en el cerebro, aquel cuya curva de olvido demuestra ser la más lenta, vale decir, aquel que proyecta la información más lejos hacia el futuro.

La importancia de evaluarse con regularidad

Puro y simple estudio:



10 segundos

Recuerdo de 5,7 palabras sobre 40

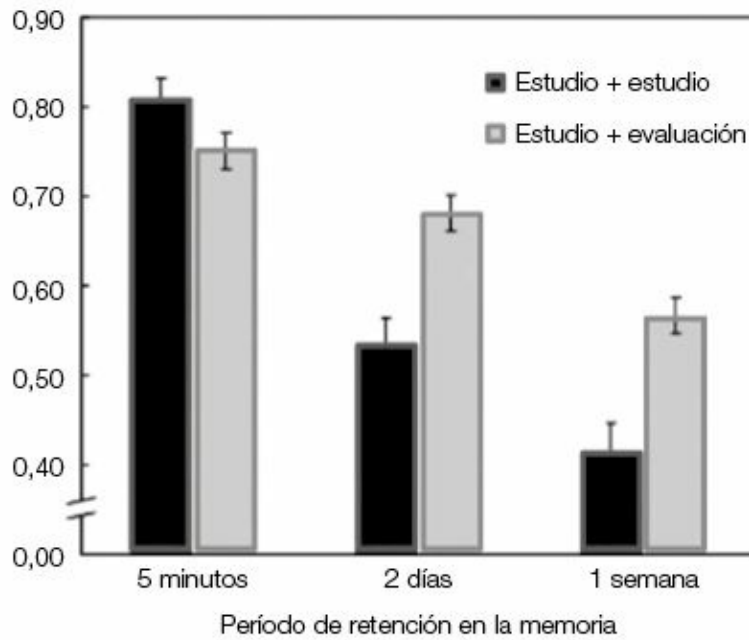
Predicción y revisión del error:



5 segundos + 5 segundos

Recuerdo de 6,4 palabras sobre 40

Porcentaje de recuerdo correcto



La importancia de planificar intervalos entre aprendizajes

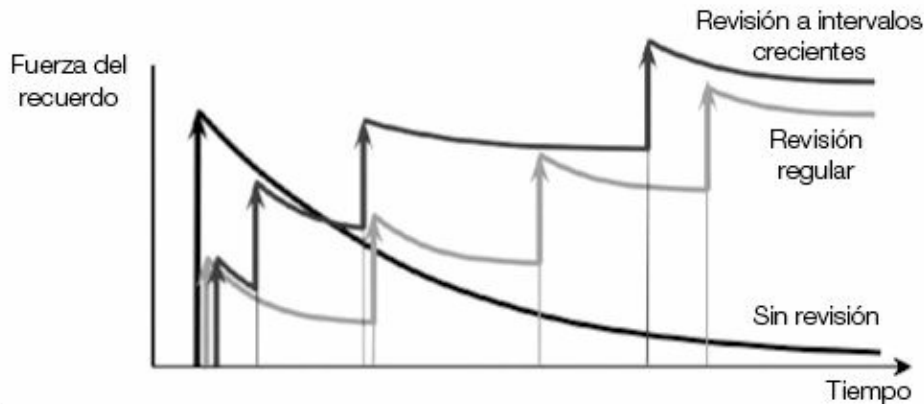


Figura 34. Evaluarnos con regularidad es una de las mejores estrategias de aprendizaje, porque nos hace tomar conciencia de nuestros errores. De este modo, si debemos aprender pares de palabras, es mejor intentar recordar en primer lugar y luego revisar el error (arriba). La experiencia demuestra que es preferible alternar los períodos de estudio y de evaluación, antes que pasar todo el tiempo estudiando (en el medio). A largo plazo, el recuerdo es mayor cuando se intercalan períodos de revisión, sobre todo si gradualmente se aumentan los intervalos de tiempo (abajo).

En efecto, todos nos equivocamos cuando pensamos en la función de la memoria: no es un sistema vuelto hacia el pasado, sino hacia el futuro. Su trabajo consiste no en mirar atrás sino, por el contrario, en enviar información hacia delante, porque estimamos que nos será útil. Al repetirla muchas veces, mediada por largos intervalos, ayudamos a nuestro cerebro a convencerse de que esta información es valiosa, que lo será durante mucho tiempo y que por eso merece que la preservemos.

De estos trabajos, Hal Pashler deriva muchas lecciones prácticas. En primer lugar, el aprendizaje siempre sale beneficiado si se lo reparte en varias sesiones. En segundo lugar, en la escuela, no basta con una revisión días o semanas después. Si se aspira a memorizar algo a largo plazo, hay que revisarlo luego de un intervalo de meses, como mínimo. Desde esta perspectiva, tenemos que repensar por completo los manuales escolares, en su mayoría organizados en capítulos que hacen foco sobre un tema específico (lo cual está bien) y una coda de ejercicios que solo se relacionan con esa lección (lo cual no está tan bien). Esta organización tiene dos consecuencias negativas: no hay suficiente tiempo entre las revisiones y los ejercicios están predigeridos, lo que impide que el alumno aprenda a detectar por sí solo qué conocimientos o qué estrategias permiten dar una respuesta plausible a determinado problema. La experiencia demuestra que es mejor mezclar los ejercicios, sin limitarse a lo que se acaba de estudiar, para que periódicamente vuelvan a ponerse en juego todos los conocimientos (Rohrer y Taylor, 2006, 2007).

¿Qué ocurre con los parciales o los exámenes de fin de año? Lo poco que sabemos sobre el aprendizaje sugiere que no son el método de evaluación ideal, porque no incitan a la revisión periódica. Sin embargo, constituyen una herramienta útil para evaluar los conocimientos adquiridos. Si bien estos exámenes estimulan el estudio de último minuto, no se trata de una estrategia necesariamente ineficaz: siempre que el alumno haya hecho un esfuerzo de aprendizaje durante los meses previos, un estudio intenso en la víspera del

examen refresca los conocimientos y ayuda a recordarlos de forma duradera. De todos modos, una revisión periódica, año tras año, reporta probablemente los mayores beneficios para el alumno. Los exámenes a corto plazo, que solo se centran en lo adquirido pocas semanas antes, no alcanzan para garantizar un recuerdo a largo plazo. Un examen acumulativo, que abarque todo el programa desde el comienzo del año, funciona mucho mejor.

Los invito a reflexionar: ¿qué relevancia tiene que una alumna revise a lo largo de todo el año los mismos temas? ¿Por qué es importante que repita muchas veces, con diferentes intervalos de tiempo, un ejercicio que ya logró hacer bien? Si aprueba perfectamente sus exámenes, sin errores, ¿aprenderá algo? Por supuesto que sí. Esto puede parecer paradójico en un capítulo dedicado a los beneficios del error; pero los aportes del *feedback* no se limitan a los tópicos en que el alumno se equivoca. Muy por el contrario, recibir una devolución sobre el error mejora la memoria, incluso cuando se eligió la respuesta correcta (Butler, Karpicke y Roediger, 2008). Como expliqué antes, mientras el conocimiento no está perfectamente consolidado, el cerebro continúa aprendiendo, aunque lo haga débilmente. Si hay incertidumbre, hay sorpresa y señales de error que se propagan por el cerebro. La incertidumbre funciona como un error virtual que habríamos podido cometer y que, por supuesto, nos da ocasión de aprender.

Lo que llamamos “sobreaprendizaje”, entonces, siempre supone ventajas. En la medida en que los conocimientos no sean absolutamente certeros, revisarlos y volver a ponerlos a prueba permite mejorar cada vez nuestros desempeños, sobre todo en el largo plazo. Además, la repetición presenta otros beneficios para el cerebro: automatiza las operaciones mentales hasta volverlas inconscientes. Esta consolidación constituye el cuarto pilar del aprendizaje; llegó el momento de analizarlo.

[21] El metaanálisis de John Hattie estima que el *feedback* tiene una magnitud de efecto de 0,73 de desviación estándar, que lo vuelve uno de los factores decisivos en el campo del aprendizaje (Hattie, 2008).

[22] Balsam y Gallistel (2009) y Gallistel (1990) critican exhaustivamente el enfoque del aprendizaje por asociación.

[23] Estas actitudes se analizan en Claro, Paunesku y Dweck (2016), Dweck (2006), Rattan y otros (2015). Sin embargo, la dimensión de estos efectos, y su consiguiente importancia práctica para la escuela fue refutada: Sisk y otros (2018).

[24] Una excelente reseña de la eficacia relativa de varias prácticas pedagógicas consta

en Dunlosky y otros (2013).

10. La consolidación

Busquen a una alumna destacada al final de primer grado. Su primer año escolar transcurrió bien. Sostenida por los tres primeros pilares del aprendizaje, aprendió rápido a leer. Con curiosidad y entusiasmo, se comprometió activamente con la lectura. Aprendió a prestarles atención a cada palabra, a cada letra, de izquierda a derecha. Con el paso de los meses, logró corregir sus errores para descifrar con fidelidad las correspondencias entre las letras y los sonidos; también consiguió reconocer las palabras irregulares. Sin embargo, aún no lee con fluidez, sino lentamente y con esfuerzo. Le falta desplegar el cuarto pilar de cualquier aprendizaje: la consolidación. Su lectura, que en esta etapa todavía acapara toda la atención, debe volverse automática e inconsciente.

El análisis de los tiempos de respuesta de la niña es revelador: cuanto más largas son las palabras, más tiempo necesita para descifrarlas (figura 35). La curva es lineal, algo característico de las operaciones seriales, que siguen un orden paso a paso: cada letra suplementaria agrega 200 ms al total. Esto es completamente normal (a esa edad, leer todavía consiste en descifrar las letras y las sílabas una por una; Zoccolotti y otros, 2005), pero no definitivo: en el transcurso de los dos años siguientes, con la práctica, su lectura ganará fluidez y velocidad. Luego de dos o tres años de práctica intensiva, el efecto de la longitud de las palabras habrá desaparecido. La niña se habrá convertido en una lectora eficaz, a quien le insumirá la misma cantidad de tiempo leer una palabra de tres letras o de ocho. El reconocimiento visual dejará de utilizar procesamiento secuencial y comenzará a operar en paralelo, dado que decodificará simultáneamente de todas las letras de la palabra.

Este es un excelente ejemplo de la consolidación, que ocurre en todos los ámbitos: pasar de un tratamiento lento, consciente, laborioso, a un funcionamiento rápido, inconsciente, automático. Nuestro cerebro nunca deja de aprender; incluso cuando domina una habilidad, continúa sobreaprendiéndola. Dispone de mecanismos de automatización que

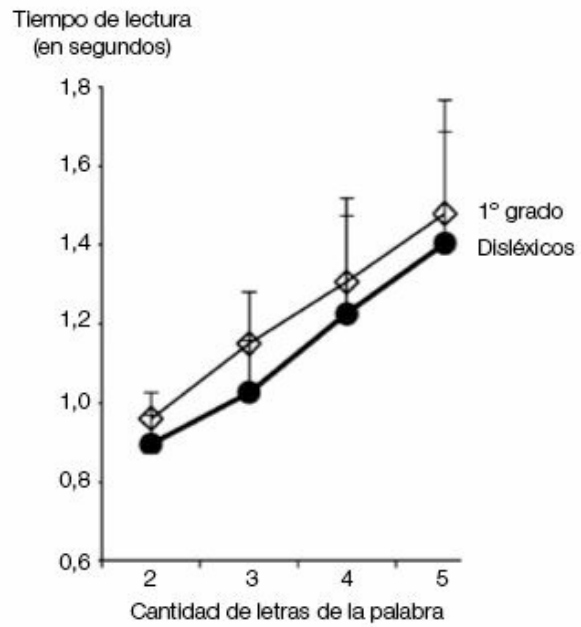
“compilan” las operaciones que utilizamos periódicamente a modo de rutinas más eficaces. Las transfiere a otras regiones del cerebro, fuera del área consciente, donde podrán desarrollarse con total autonomía, sin interrumpir las otras operaciones en pleno avance.

Liberar los recursos cerebrales

Cuando se escanea el cerebro de lectores principiantes, ¿qué se ve? El circuito normal de la lectura (que abarca las áreas visuales del reconocimiento de las letras y las regiones temporales del tratamiento de los fonemas, sílabas y palabra) ya está activado, pero también está acompañado por una activación masiva de las regiones parietales y prefrontales que reflejan el esfuerzo, la atención espacial y las estrategias conscientes (Dehaene-Lambertz, Monzalvo y Dehaene, 2018). Esta actividad intensa consume mucha energía y desaparecerá conforme el aprendizaje se consolide (figura 35). En una persona menuda que ya adquirió esa pericia, dichas áreas dejan de involucrarse en la lectura; solo las veremos activarse si se perturba la operación, por ejemplo, cuando se suman espacios entre las l e t r a s, o se las hace *l e t r a s*, lo que obliga al cerebro experto a volver al modo principiante (Cohen y otros, 2008; Vinckier y otros, 2006).

Convertir la lectura en una rutina implica poner en funcionamiento un circuito acotado y especializado para el procesamiento eficaz de las cadenas de letras que vemos con mayor frecuencia. A medida que aprendemos a leer, desarrollamos un circuito de extraordinaria eficacia para reconocer los caracteres y sus combinaciones más comunes (Binder y otros, 2006; Dehaene y otros, 2005; Grainger y Whitney, 2004; Vinckier y otros, 2007). Nuestro cerebro compila estadísticas y detecta qué letras son las más habituales, dónde aparecen más a menudo y a cuáles están asociadas. Aun la corteza visual primaria se adapta a la forma y posición de las letras más usuales (Chang y otros, 2015; Dehaene y otros, 2010; Sigman y otros, 2005; Szwed y otros, 2011, 2014). Luego de años de sobreaprendizaje, este circuito pasa al modo de rutina y logra funcionar sin la menor intervención consciente (Dehaene y otros, 2001, 2004). En esta etapa, la activación de la corteza parietal y prefrontal desaparece: podemos leer sin esfuerzo.

Lectura con esfuerzo



Lectura automatizada

El mismo niño al final de 2º grado

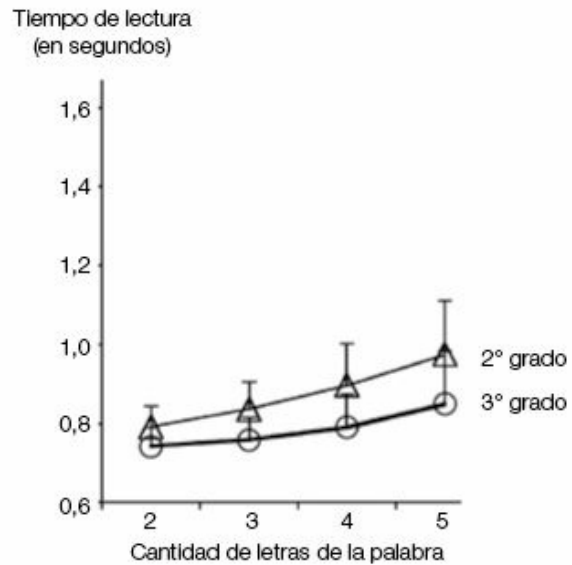


Figura 35. La consolidación es el cuarto pilar del aprendizaje. Al principio, cualquier aprendizaje exige esfuerzos intensivos, y activa las regiones parietales y frontales asociadas a

la atención espacial y ejecutiva. En un lector principiante, la decodificación de una palabra requiere grandes esfuerzos. La lectura es secuencial: cuantas más letras tiene la palabra, más tarda en leerla (arriba). Con la práctica, el aprendizaje se automatiza: se vuelve rápido, paralelo e inconsciente (abajo). Se pone en funcionamiento un circuito especializado, que libera los recursos de la corteza para otras tareas.

Lo que es cierto para la lectura vale para los demás campos del aprendizaje. Cuando aprendemos a tocar un instrumento musical, a conducir un vehículo o a utilizar la pantalla de nuestra tablet, al principio nuestros gestos están bajo el control de la corteza prefrontal: los producimos de manera lenta y consciente, uno por uno. Luego de algunas sesiones, el esfuerzo desaparece; para entonces, podemos realizar esas acciones mientras hablamos o pensamos en otra cosa: la tarea se trasladó a la corteza motora y en especial a los núcleos grises centrales, un grupo de circuitos subcorticales que registran los comportamientos automáticos y rutinarios (¡incluidas las oraciones y las imprecaciones!). Lo mismo ocurre en el ámbito de la aritmética: desde la perspectiva de un principiante, cada cálculo parece un monte Everest, y para escalarlo concentra grandes esfuerzos y moviliza los circuitos de la corteza prefrontal. En esta etapa, el cálculo es secuencial: para resolver $6 + 3$, es habitual que el niño cuente las etapas una por una hasta llegar a la cumbre: “¡Siete, ocho... nueve!”. A medida que la consolidación se abre camino, la actividad prefrontal se desvanece en provecho de los circuitos especializados de la corteza parietal y temporal ventral (Ansari y Dhital, 2006; Rivera y otros, 2005).[\[25\]](#)

¿Por qué es tan importante la automatización? Porque libera los recursos de la corteza. Recuerden que las redes del control ejecutivo de la corteza parietal y prefrontal imponen un cuello de botella cognitivo: no pueden hacer dos cosas a la vez. Cuando se concentran en la realización de determinada tarea, las demás decisiones conscientes se desaceleran o quedan a un lado. Así, en la medida en que un aprendizaje no está automatizado, acapara los preciosos recursos de la atención ejecutiva e impide que el niño se concentre en cualquier otra cosa. Consolidar un aprendizaje es muy importante, porque permite que los recursos del cerebro queden disponibles para otros objetivos.

Tomemos un ejemplo concreto. Imaginen que deben concentrarse en la lectura de un texto y, al mismo tiempo, resolver un problema de matemáticas, como un lector principiante:

Un ci-clis-ta sa-le de Nan-tes ha-cia Pa-rís al me-dio-dí-a. La dis-

tan-cia es de tres-cien-tos ki-ló-me-tros. Lle-ga a des-ti-no a las vein-te ho-ras de e-se mis-mo dí-a. ¿Cuál es su ve-lo-ci-dad pro-me-dio?

La complejidad es evidente: resulta casi imposible hacer las dos cosas a la vez. Y esa dificultad de lectura limita cualquier capacidad de reflexión aritmética. Para progresar, es indispensable que las herramientas mentales que nos son más útiles, como la lectura o la aritmética, se conviertan en una segunda naturaleza, que operen de forma inconsciente y sin esfuerzos. No podemos construir los niveles más altos de la pirámide educativa sin primero consolidar sus bases.

El sueño, un ingrediente clave

Como ya comentamos, el aprendizaje se beneficia mucho cuando sucede a intervalos regulares: en lugar de engullir una lección completa en una sola jornada, es preferible distribuir el aprendizaje en pequeños bocados y degustarlos en diferentes días. Esto ocurre por un motivo muy sencillo: todas las noches, nuestro cerebro consolida lo que aprendió durante la jornada. Sin lugar a dudas, las neurociencias hicieron uno de los descubrimientos más importantes de los últimos treinta años al verificar que el sueño no es un simple período de inactividad ni de limpieza de los desechos que las neuronas acumularon durante la vigilia. Todo lo contrario; mientras dormimos, el cerebro repasa los acontecimientos importantes que registró durante el día y, poco a poco, los muda a un compartimiento más eficaz de nuestra memoria.

A decir verdad, el descubrimiento se remonta a las primeras décadas del siglo XX, cuando los psicólogos estadounidenses John Jenkins y Karl Dallenbach (1924) retomaron los estudios clásicos sobre la memoria y volvieron a analizar los trabajos del investigador pionero en el estudio de la memoria, el alemán Hermann Ebbinghaus: ya a finales del siglo XIX había demostrado que el olvido seguía una curva exponencial (cuanto más tiempo dejamos pasar, menos recordamos lo aprendido). Jenkins y Dallenbach notaron que la curva del olvido planteada por Ebbinghaus –hermosa, monótonamente decreciente– presentaba una anomalía: no reflejaba una

pérdida de memoria entre las ocho y las catorce horas posteriores a un nuevo aprendizaje. Fue toda una revelación. En la experiencia de Ebbinghaus, el período de ocho horas correspondía a pruebas realizadas en una misma jornada, y el de catorce horas, a pruebas entre las cuales mediaba una noche de descanso. Para comprobarlo, concibieron una nueva experiencia que discriminaba las dos variables: por un lado, el período de tiempo transcurrido antes de evaluar la memoria y, por otro, el hecho de haber dormido o no. Así, les enseñaron a sus estudiantes un conjunto de sílabas aleatorias, tanto cerca de las 0 hs, justo antes de dormir, como por la mañana. El resultado fue contundente: el aprendizaje obtenido durante la mañana se desplomó con el paso del tiempo, en concordancia con la ley exponencial de Ebbinghaus; en cambio, el aprendizaje de la medianoche permaneció estable a lo largo del tiempo (siempre y cuando los participantes hubieran disfrutado de dos horas de sueño como mínimo). En otras palabras: según parece, dormir nos impide olvidar.

Se nos ocurren muchas interpretaciones alternativas para estos resultados. Quizá el problema se deba a la vigilia: mientras está despierto, el cerebro tal vez acumule toxinas que se eliminan al dormir; o bien la memoria podría sufrir una interferencia por obra de los acontecimientos que se suceden en el lapso entre el aprendizaje y la evaluación, no durante el sueño. Pero todas estas ideas fueron rechazadas definitivamente en 1994, cuando un grupo de investigadores demostró que el sueño *mejora* el aprendizaje (Karni y otros, 1994): sin entrenamiento adicional, el desempeño cognitivo y motor se incrementa luego de un período de sueño. La experiencia dista de ser compleja: durante un día, un voluntario aprende a detectar una barra proyectada en un lugar preciso de la retina. Su desempeño mejora lentamente, pero luego de unas horas de entrenamiento ya no exhibe mejoría; parece que se alcanzó el límite. Entonces, se lo deja dormir, para evaluarlo nuevamente una vez despierto. ¡Oh, sorpresa! Su desempeño resulta tanto mejor después del descanso, y así sucede en esa jornada y las siguientes. En efecto, el sueño es el causante del incremento del aprendizaje, porque si se despierta a la persona durante la noche, cada vez que ingresa en el período de sueño REM – en cuyo transcurso la actividad cerebral es rápida y más parecida a la de la vigilia–, no presenta ninguna mejoría cuando se despierta.

Más tarde, numerosos trabajos confirmaron estos descubrimientos (Huber y otros, 2004; Stickgold, 2005; Walker y otros, 2003; Walker y Stickgold, 2004). La intensidad del aprendizaje varía directamente en función de la

duración del sueño, y sobre todo de su profundidad. Cuando se ponen electrodos sobre el cráneo, se puede evaluar la calidad del sueño y así predecir cuánto mejorará el desempeño al despertar. La relación también parece verificarse en sentido inverso: la necesidad de sueño estaría definida por los acontecimientos que se produjeron durante la jornada. En los animales, un gen relacionado con la plasticidad cerebral, el *zif268*, aumenta su expresión en el hipocampo y la corteza durante el sueño REM, sobre todo en ratas previamente expuestas a un ambiente enriquecido: el crecimiento del estímulo implica un aumento de la plasticidad nocturna (Ribeiro y otros, 1999).

Todavía no se confirmaron los roles de las diferentes etapas del sueño, pero, según parece, el sueño profundo permite la consolidación y la generalización de los conocimientos (lo que los psicólogos llaman “memoria semántica” o “declarativa”), mientras que el sueño REM consolida el aprendizaje perceptivo y motor (memoria procedural).

El cerebro dormido revive los episodios de la víspera

Si bien las demostraciones psicológicas de los efectos del sueño fueron bastante convincentes, aún no se había detectado el mecanismo neuronal por el cual un cerebro dormido podía aprender incluso mejor que mientras estaba despierto. En 1994, los neuropsicólogos Matt Wilson y Bruce McNaughton realizaron un descubrimiento notable en experimentos con ratas: sin ningún tipo de estímulo exterior, las neuronas del hipocampo se activan espontáneamente durante el sueño (Ji y Wilson, 2007; Louie y Wilson, 2001; Skaggs y McNaughton, 1996; Wilson y McNaughton, 1994). Además, esta actividad no es aleatoria: ¡vuelve a trazar los pasos que el animal dio durante el día!

En efecto, como vimos en el capítulo 4, en el hipocampo hay células de lugar, vale decir, neuronas que se disparan cuando el animal está (o cree estar) en determinado punto del espacio (y hay una gran diversidad, ya que cada una prefiere un lugar diferente). Si investigamos una cantidad suficiente de ellas, veremos que abarcan el espacio completo por donde se pasea el animal. Por ejemplo, cuando se desplaza por un pasillo, algunas neuronas se activan a la

entrada, otras en el medio, y otras más cerca del final. Así, el paseo de la rata se refleja en la sucesiva activación de una serie de células de lugar: el movimiento en el espacio real se vuelve una secuencia temporal en el espacio neuronal.

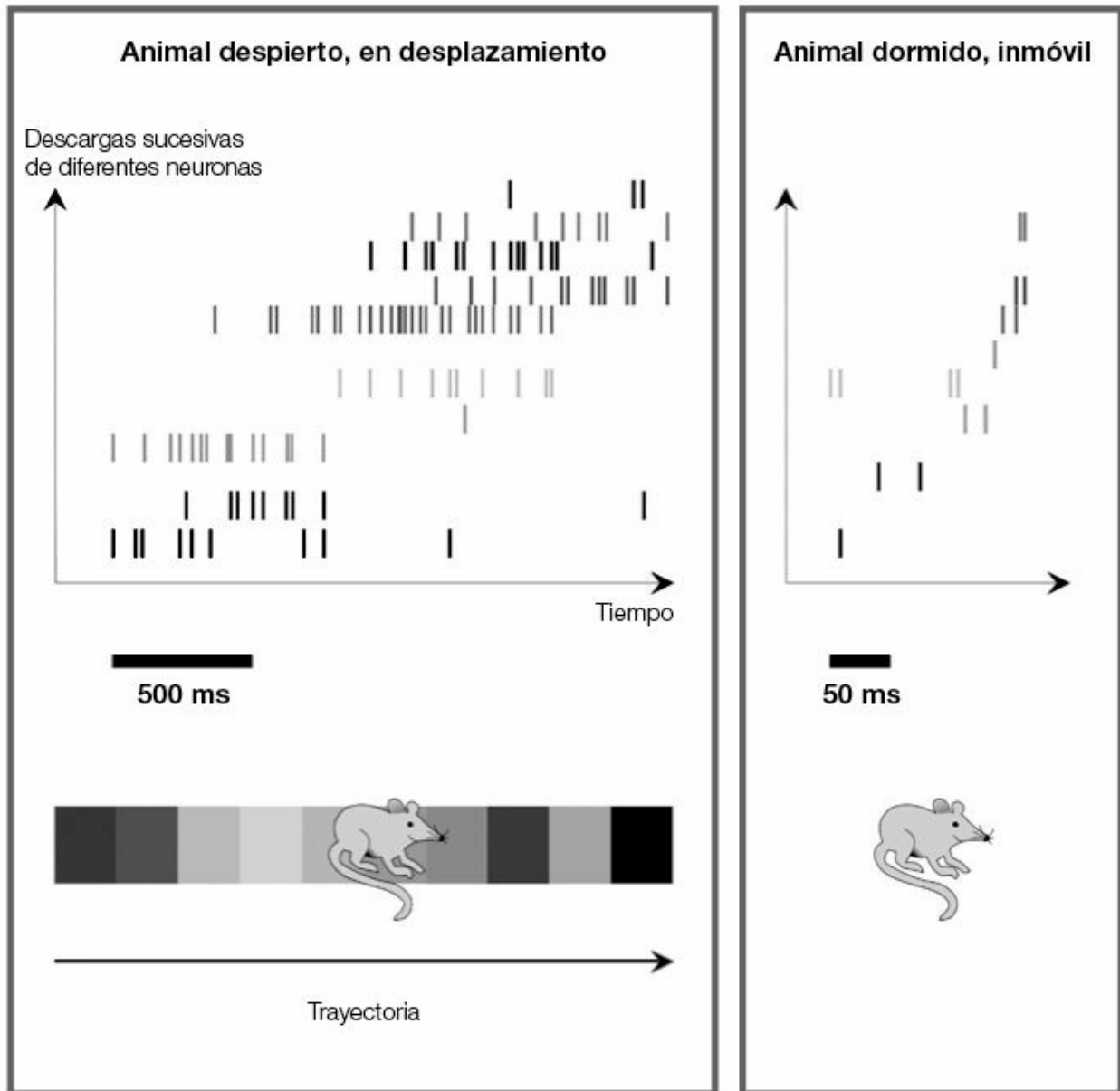


Figura 36. El sueño desempeña un papel central en la consolidación del aprendizaje. Cuando una rata se duerme, las neuronas del hipocampo vuelven a repasar muchas veces, a una velocidad acelerada, las secuencias de movimientos que realizó cuando estaba despierta. Esta actividad, que se extiende hasta la corteza, se repite cientos de veces durante la noche. Al despertar, el aprendizaje de la jornada precedente está mejor consolidado y automatizado. Nuestro cerebro (animal o humano) incluso puede descubrir, durante el sueño, regularidades

que se nos habían escapado durante la vigilia.

El descubrimiento del equipo de Wilson y McNaughton es elocuente: cuando la rata se queda quieta y se duerme, vuelven a activarse las mismas neuronas del hipocampo, y lo hacen en el mismo orden. Reproducen tal cual las trayectorias del período de vigilia anterior. La única diferencia es la velocidad: durante el sueño, las descargas neuronales pueden sucederse aceleradas, multiplicadas por veinte. Cuando duerme, ¡la rata sueña con una carrera de alta velocidad!

La relación entre la actividad de las neuronas del hipocampo y la posición del animal es tan fiable que los científicos lograron invertir el proceso y, sobre la base de las descargas neuronales, decodificar el contenido del sueño (Chen y Wilson, 2017; Horikawa y otros, 2013). Para esto, en primer lugar, se utilizan los datos registrados en el transcurso de la jornada, cuando la rata se paseó por el mundo real. Esta información permite entrenar un decodificador: un programa informático que descubre la relación sistemática entre las descargas neuronales y la posición del animal. Luego este decodificador se aplica a los datos registrados durante el sueño y se constata que, mientras duerme, el cerebro traza trayectorias virtuales por el espacio.

Así, el cerebro de la rata repasa a alta velocidad los patrones de actividad que vivenció durante la jornada. Mientras duerme, renacen en él los recuerdos de las horas de vigilia. En efecto, este fenómeno no se limita al hipocampo, sino que se extiende a la corteza, donde desempeña un papel determinante en la plasticidad sináptica y en la consolidación del aprendizaje. Gracias a la reactivación durante el descanso, incluso un suceso único de nuestra vida podrá ser repasado una centena de veces en la medida en que se lo registre en la memoria episódica. Esta transferencia podría incluso ser la función principal del sueño (Diekelmann y Born, 2010). El hipocampo acumularía los recuerdos de la jornada en una memoria rápida, mientras que, en el transcurso del descanso, la reactivación de estas señales permitiría ejercitar otras redes neuronales adicionales, principalmente aquellas situadas en la corteza: una memoria más lenta, pero capaz de obtener el máximo de información de cada episodio. Efectivamente, en la corteza de una rata que aprende a realizar una tarea nueva, cuanto más se reactiva una neurona mientras el animal duerme, más se incrementa su participación en la tarea una vez que despierta (Ramanathan, Gulati y Ganguly, 2015).[\[26\]](#) La reactivación del hipocampo desemboca en la automatización cortical.

¿Ocurre el mismo fenómeno en los seres humanos? Sí. Las neuroimágenes del cerebro humano muestran que durante el sueño se reactivan los circuitos utilizados en el transcurso de la jornada (Horikawa y otros, 2013; Jiang y otros, 2017; Peigneux y otros, 2004). Un estudio escaneó durante la noche a un grupo de jugadores luego de varias horas de enfrentar al célebre Tetris, videojuego en que las piezas no dejan de caer desde lo alto de la pantalla: en sueños, los participantes alucinaban con una cascada de formas geométricas, y sus ojos hacían los movimientos correspondientes, de arriba abajo. Es más, en un experimento reciente, se hizo dormir a voluntarios dentro de un resonador magnético, y se los despertó en el momento en que sus electroencefalogramas sugerían que estaban soñando. Las resonancias (MRI) mostraron que, justo antes de que se los despertara, numerosas regiones de sus cerebros se habían activado espontáneamente, y que esa actividad predecía el contenido de los sueños. De este modo, cuando uno de los participantes, por ejemplo, decía haberse encontrado con personas en el sueño, se podía ver que el área cortical del reconocimiento de rostros se había activado mientras dormía. Otras experiencias revelan que la extensión de esta activación predice no solo el contenido del sueño, sino también el grado de consolidación en la memoria que tendrá esta información luego de despertar. Algunos neurocirujanos, incluso, comenzaron a tomar registros de neuronas individuales del cerebro humano y observaron que, tal como en las ratas, los patrones de activación replican la secuencia de acontecimientos vividos durante la jornada.

El vínculo entre el sueño y el aprendizaje es sólido. Numerosos experimentos confirman que las variaciones espontáneas de la profundidad del sueño correlacionan con el desempeño el día posterior. Por ejemplo, la frecuencia y la intensidad de las ondas lentas del sueño en alguien que durante la vigilia haya aprendido a utilizar un *joystick* aumentarán en función de las regiones parietales del cerebro involucradas en este aprendizaje: a mayor incremento, más progresa el desempeño de la persona (Huber y otros, 2004). Del mismo modo, luego de un aprendizaje motor, las neuroimágenes muestran un incremento en la actividad de la corteza motora, el hipocampo y el cerebelo, en simultáneo con una disminución en determinadas áreas frontales, parietales y temporales (Walker, Stickgold y otros, 2005). Todos estos resultados sugieren que el sueño facilita la automatización. Es evidente que luego de dormir la actividad cerebral se desplaza: se afianza una parte de

los conocimientos adquiridos durante la jornada y se traslada hacia circuitos más automáticos y más especializados.

Si bien la automatización y el sueño están entrelazados, cualquier científico sabe que correlación no equivale necesariamente a causalidad. Para saber más al respecto, hoy en día se puede aprovechar la posibilidad de aumentar artificialmente la profundidad del sueño mediante un efecto de resonancia en el cerebro. Durante el sueño, la actividad cerebral oscila espontáneamente con una frecuencia lenta, del rango de los 40 a 50 ciclos por minuto. Al inyectar un pequeño impulso, justo en la frecuencia correspondiente, se logra que estos ritmos entren en resonancia y aumenten su intensidad, de modo similar a lo que ocurre cuando se empuja una hamaca en el momento exacto para hacerla oscilar con una amplitud enorme. Precisamente esto hizo el equipo del científico alemán Jan Born, especialista en sueño, de dos maneras diferentes: por un lado, inyectó minúsculas corrientes a través del cráneo y, por otro, con un recurso tanto más sencillo, reprodujo por un altavoz un ruido difuso sincronizado con las ondas cerebrales del sueño. Ya sea exaltado o aplacado por el sonido de las olas, el cerebro de quien duerme se deja llevar por este ritmo irresistible y produce una cantidad notoriamente incrementada de ondas lentas, características del sueño profundo. En los dos casos, al día siguiente, este entrenamiento redundaba en una mejor consolidación del aprendizaje (Marshall y otros, 2006; Ngo y otros, 2013).

Una empresa francesa comenzó a sacar réditos económicos de este efecto: vende una vincha que supuestamente ayuda a conciliar el sueño y aumenta su profundidad mediante pequeños sonidos que estimulan los ritmos lentos del cerebro durante el descanso. Otros investigadores contemplaron la posibilidad de incrementar el aprendizaje forzando al cerebro a reactivar, durante la noche, algunos recuerdos más que otros. Imaginen que, mientras aprenden determinados hechos, el aire del aula está saturado de aroma a rosas. Durante su sueño profundo, se rocía esta fragancia en la habitación. El experimento revela que, a la mañana siguiente, lo que aprendieron estará tanto más consolidado que si hubieran dormido sin exposición al perfume (Rasch y otros, 2007). En este caso, el aroma sirve como una clave inconsciente que orienta al cerebro a reactivar un episodio específico de la jornada, lo que aumenta su consolidación en la memoria.

Se puede obtener el mismo efecto con claves auditivas. Supongan que se les pide memorizar la posición de cincuenta imágenes, cada una asociada a un sonido o, antes bien, una onomatopeya (el gato hace “miau”, la vaca hace

“mu”, etc.). Es una tarea difícil... pero siempre es bueno consultar las cosas con la almohada. En un experimento, un grupo de investigadores estimuló el cerebro de sus voluntarios, durante el sueño, repitiendo algunos de esos sonidos. Al oírlos, los participantes (aun de manera inconsciente, en la profundidad del sueño) guiaron la reactivación neuronal nocturna, y a la mañana siguiente recordaban tanto mejor la ubicación de las respectivas imágenes (Antony y otros, 2012; Bendor y Wilson, 2012; Rudoy y otros, 2009).

En el futuro, ¿todos manipularemos nuestro sueño, en busca de aprender mejor? Muchos estudiantes lo hacen ya de forma espontánea: revisan una lección importante justo antes de dormirse; sin saberlo, intentan guiar su repaso nocturno. Pero no debemos confundir estos descubrimientos, cuya eficacia está demostrada, con la falsa idea de que se pueden adquirir nuevas habilidades mientras se duerme. Algunos charlatanes venden grabaciones que supuestamente pueden hacernos aprender una lengua extranjera durante el transcurso de la noche, pero las investigaciones demuestran que no tienen efecto alguno (Arzi y otros, 2012). El cerebro dormido no aprende nada nuevo: lo único que puede hacer es reactivar lo que ya experimentó. Para adquirir una habilidad tan compleja como un buen desempeño en un idioma que no se dominaba previamente, lo único que funciona es estudiar durante el día y luego dormir para dejar que el sueño reactive y consolide lo aprendido.

Sueño de un descubrimiento de verano

¿Dormir simplemente fortalece el recuerdo de lo que aprendimos? Numerosos científicos piensan que no: dan cuenta de descubrimientos que hicieron mientras dormían. El caso más famoso es el del químico alemán Friedrich August Kekulé von Stradonitz, quien descubrió en un sueño la estructura del benceno, una molécula inusual, ya que sus seis átomos de carbono forman un bucle cerrado, como un anillo... o una serpiente que se muerde la cola. Sí, en esos términos describió Kekulé el sueño de aquella presagiosa noche:

Una vez más, los átomos formaban un torbellino ante mí. [...] Mi

ojo mental, entrenado por las repetidas visiones de este tipo, ya podía distinguir grandes formas extrañas y largas cadenas. Largas filas se retorcían como serpientes. De pronto, ocurrió algo. Una de las serpientes había mordido su propia cola; la forma giró burlonamente ante mis ojos. Como iluminado por un relámpago, me desperté.

Y Kekulé concluyó: “Si aprendiéramos a soñar, señoras y señores, ¿entonces quizá encontraríamos la verdad!”.

¿Es cierto que el sueño tiene el poder de llevarnos por el camino de la verdad y aumentar la creatividad? Los pareceres de los historiadores de la ciencia discrepan en cuanto a la autenticidad del uróboro de Kelulé, pero la idea de la incubación nocturna tiene gran difusión entre los científicos y los artistas. El diseñador Philippe Starck, creador de centenas de objetos de formas inéditas, lo expresó con humor en una entrevista reciente: “Cuando me voy a acostar, le digo a mi mujer: ‘¡Bueno, me voy al trabajo!’”.[27] Yo mismo tuve muchas veces la vivencia de descubrir durante el sueño la solución para un problema difícil que se me había planteado en la jornada. Sin embargo, un conjunto de anécdotas, por amplio que sea, carece de valor probatorio. Hace falta experimentar, y exactamente eso hizo el ya mencionado equipo de Jan Born (Wagner y otros, 2004). Durante el día, estos investigadores enseñaron a sus voluntarios un algoritmo complejo, del cual debían valerse para aplicar una serie de cálculos sucesivos a determinada cifra. Pero el problema tenía un atajo escondido que los participantes desconocían: había un truco que permitía reducir considerablemente el tiempo de cálculo. Antes de dormir, muy pocos participantes lo habían descubierto. Sin embargo, una buena noche de sueño duplicó la cantidad de voluntarios que descubrían el truco, mientras que entre quienes no habían dormido las exclamaciones de *eureka* brillaban por su ausencia. Los resultados no dependían de la hora del día a la cual fueran evaluados, de modo que el tiempo transcurrido no era el factor determinante: solo el sueño genuino traía aparejada la comprensión.

Así, la consolidación nocturna no se reduce al simple refuerzo de los conocimientos existentes. Los descubrimientos de la jornada no solo quedan almacenadas, sino que también se los recodifica en una forma más abstracta y general. Indudablemente, la reactivación neuronal nocturna desempeña un papel crucial en este proceso. Cada noche, nuestras ideas de la jornada se

reproducen centenas de veces, a un ritmo acelerado, lo que multiplica las posibilidades de que la corteza descubra allí reglas provistas de sentido. Además, la aceleración de las descargas neuronales comprime la información. La reproducción a alta velocidad implica que las neuronas que durante el día se activan separadas por largos intervalos muestran activaciones adyacentes en la secuencia comprimida de la noche. Este mecanismo parece ideal para agrupar, sintetizar, comprimir y “convertir la información bruta en conocimientos útiles y explotables”, la definición misma de la inteligencia según el especialista en inteligencia artificial Demis Hassabis.

En el futuro, ¿las máquinas inteligentes tendrán que dormir como nosotros? La pregunta bien puede parecer descabellada, pero creo que en cierto sentido ese momento llegará: los algoritmos informáticos de aprendizaje probablemente incorporarán una fase de consolidación similar a lo que llamamos “dormir”. De hecho, los informáticos ya diseñaron varios algoritmos de aprendizaje computarizado que imitan la alternancia de sueño y vigilia (Hinton y otros, 1995; Hinton, Osindero y Teh, 2006) y aportan ejemplos inspiradores para el nuevo enfoque que defiendo en este libro: aprender consiste en construir un modelo generativo interno del mundo exterior. Recuerden que nuestro cerebro alberga numerosos modelos internos, capaces de resintetizar imágenes mentales más verdaderas que la naturaleza, diálogos realistas y deducciones provistas de sentido. En el estado de vigilia, ajustamos estos modelos al entorno: los datos sensoriales nos sirven para seleccionar los modelos que mejor se adaptan al mundo que nos rodea. Durante esta etapa, el aprendizaje es ante todo una operación de tipo ascendente: las señales sensoriales inesperadas, cuando son confrontadas con las predicciones de nuestro modelo interno, generan señales de error de predicción, que ascienden en la jerarquía cortical y ajustan los pesos estadísticos en cada paso, con lo cual paulatinamente los modelos descendentes ganan en precisión.

La nueva idea es que, en el transcurso del sueño, el cerebro funciona en el sentido inverso, de arriba abajo. Durante la noche, utilizamos los modelos generativos para sintetizar nuevas imágenes, nunca antes vistas, y una parte del cerebro se entrena para reconocer estas imágenes creadas a partir de fragmentos. Mientras soñamos, las conexiones que se ajustan son aquellas que van de los niveles más bajos a los más altos: detectan qué regularidades del mundo exterior apuntan hacia tal o cual modelo interno. Así, nos

volvemos cada vez más eficaces: después de una buena noche de sueño, incluso una pista ínfima nos basta para encontrar el mejor modelo mental de la realidad, sin importar que a veces sea excesivamente abstracto.

Este enfoque considera que los sueños no son otra cosa que imágenes de entrenamiento, recreaciones de la realidad que utilizamos para multiplicar las experiencias (necesariamente limitadas) que tuvimos en las horas de vigilia que los precedieron. El sueño resolvería un problema con el cual deben lidiar todos los algoritmos de aprendizaje: la poca cantidad de datos disponibles. En efecto, las redes neuronales artificiales necesitan enormes conjuntos de datos para aprender, pero la vida es demasiado corta y el cerebro debe contentarse con la cantidad limitada de información que logra recopilar durante el día. Acaso el sueño haya sido la solución que nuestro cerebro encontró para simular de forma acelerada, en cuestión de horas, innumerables eventos que no habría podido vivenciar en la realidad empírica en una vida.

Durante estos experimentos mentales, ocasionalmente hacemos descubrimientos. Aquí no hay siquiera un atisbo de magias o alquimias: a medida que nuestro motor de simulación mental funciona, a veces llega a resultados inesperados, de modo similar a un jugador de ajedrez que, una vez dominadas las reglas, puede pasar años explorando sus consecuencias. En efecto, la humanidad debe a las imágenes mentales algunos de sus mayores descubrimientos científicos: por ejemplo, cuando Einstein soñaba con capturar un rayo de luz (preanuncio del fotón) o cuando Newton imaginaba que la Luna caía hacia la Tierra como una manzana. Incluso es probable que el experimento más célebre y celebrado de Galileo, en que arrojó objetos desde la Torre de Pisa para demostrar que la velocidad de la caída libre no depende de la masa, jamás haya sucedido en la realidad concreta. Una experiencia del pensamiento pudo haber sido suficiente: Galileo imaginó que dos esferas, una liviana y otra pesada, caían desde lo alto de la torre, supuso que la más pesada caería más rápido, y usó sus modelos mentales para demostrar que esto llevaría a una contradicción. Y este notable científico, bien pudo señalar algo por el estilo:

Supongamos que uno las dos esferas con un hilo de masa despreciable. El conjunto de dos esferas, que ahora forma un objeto más pesado, debería caer aún más rápido. Pero esto es absurdo, porque la esfera liviana, que cae a menor velocidad, debería desacelerar a la más pesada. Estas contradicciones infinitas solo

tienen una salida: todos los objetos caen a la misma velocidad, independientemente de su masa.

Este es el tipo de argumentación de que es capaz nuestro simulador mental, tanto en la vigilia como en el sueño. El hecho mismo de que sepamos cómo desplegar semejantes conjuntos imaginarios pone de relieve la extraordinaria variedad de algoritmos de que dispone el cerebro. Por supuesto, aprendemos durante el día, pero la actividad neuronal nocturna multiplica nuestro potencial. Tal vez allí resida uno de los secretos de la especie humana: nuestro sueño podría ser uno de los más profundos y eficaces en comparación con el de los demás primates (Samson y Nunn, 2015).

El sueño, la infancia y la escuela

Ahora bien, ¿qué ocurre con los niños? Todos saben que los bebés pasan la mayor parte del tiempo durmiendo y que el tiempo de sueño se acorta con la edad (queda muy lejos esa noche perpetua, de solo dormir, que el poeta Catulo prometía a Lesbia). Esto es lógico: la primera infancia es un período de aprendizaje privilegiado durante el cual es probable que los algoritmos del cerebro deban funcionar tanto más. De hecho, las investigaciones demuestran que, incluso con la misma duración, el sueño del niño es de dos a tres veces más eficaz que el del adulto. Después de un aprendizaje intensivo, un niño de 10 años se sumerge mucho más rápido en el sueño profundo que un adulto. Sus ondas lentas del sueño son más intensas, y el resultado es claro: cuando estudia una lección cae en el sueño y al día siguiente se despierta renovado y bien dispuesto; luego constatamos que descubrió muchas más regularidades que un adulto en la misma situación (Wilhelm y otros, 2013).

La consolidación nocturna funciona desde los primeros meses de vida. En el bebé de menos de 1 año, por ejemplo, el aprendizaje de las palabras depende lisa y llanamente del sueño. Un bebé que duerme la siesta, aunque solo dure una hora y media, retiene mucho mejor las palabras que se le enseñaron antes de descansar (Friedrich y otros, 2015; Seehagen y otros, 2015). Pero el detalle sobresaliente es que no solo las aprende, sino que las generaliza mejor: la primera vez que escucha la palabra “caballo”, la asocia únicamente

con una o dos imágenes específicas; pero luego de haber dormido, su cerebro logra asociar esta palabra a nuevos ejemplares nunca antes vistos. Como un Kekulé en pañales, realiza descubrimientos mientras duerme y se despierta con una teoría mucho más elaborada de la palabra “caballo”.

¿Y qué ocurre durante la edad escolar? También en este caso las experiencias son muy claras: en el jardín de infantes, hasta una breve siesta por la tarde mejora el aprendizaje de la mañana (Kurdziel, Duclos y Spencer, 2013). Para obtener el máximo beneficio, el sueño debe ocurrir en las horas inmediatamente posteriores al aprendizaje. Este beneficio, sin embargo, solo se da en niños que duermen la siesta a menudo. Por ende, no parece conveniente obligar a los niños a dormir, sino tan solo dejar dormir a los que desean hacerlo: el cerebro regula naturalmente su necesidad de dormir de acuerdo con la cantidad de estímulos nuevos que conoció durante la jornada.

Por desgracia, con la actual sobreabundancia de exposición a televisión, *smartphones* e internet, el sueño de los niños, al igual que el de los adultos, se ve asediado, amenazado, dificultado por todos los frentes. ¿Cuáles son las consecuencias de esta realidad? ¿Es cierto que la privación crónica de sueño acarrea trastornos específicos del aprendizaje que, según se dice, están aumentando? De momento, esta posibilidad no va más allá de la pura hipótesis, pero hay algunos indicios en ese sentido: es posible que una fracción no desdeñable de los niños hiperactivos con trastornos de atención solo sufran de una privación crónica de sueño. En algunos casos, por ejemplo, se detectan problemas respiratorios que les impiden entrar en el sueño profundo, y una simple limpieza de las vías respiratorias a veces basta y sobra para eliminar no solo los problemas de sueño, sino también los de atención. Una experiencia reciente incluso sugiere que la estimulación eléctrica del cerebro, al aumentar la profundidad de las ondas lentas del sueño, también podría corregir el déficit de aprendizaje en los niños hiperactivos (Avior y otros, 2004; Cortese y otros, 2013; Hiscock y otros, 2015; Prehn-Kristensen y otros, 2014).

Dejémoslo bien claro: estos datos aún deben someterse a repetidas pruebas, y de ninguna manera intento negar la existencia de auténticos trastornos de atención (se trata de niños que pueden beneficiarse mucho con un entrenamiento de su atención, o en quienes a veces un fármaco como la ritalina puede tener efectos muy positivos). Sea como fuere, en el plano educativo ya no cabe duda de que mejorar la duración y la calidad del sueño constituye una intervención efectiva para *todos* los niños y muy

especialmente para aquellos que presentan dificultades de rendimiento escolar o problemas de aprendizaje.

La idea pasó por pruebas piloto en adolescentes. En su caso, la cronobiología demostró que el ciclo del sueño cambia durante la pubertad: los adolescentes no sienten la necesidad de acostarse temprano, pero, como todos pueden haber vivenciado, tienen grandes dificultades para levantarse. Esto no supone mala voluntad, sino que se trata de la simple consecuencia de una reorganización de las redes neuronales y hormonales que controlan el ciclo de sueño-vigilia. Por desgracia, nadie parece haber informado al respecto a las autoridades escolares, quienes todavía exigen a los estudiantes que hagan acto de presencia a primera hora de la mañana. ¿Qué ocurriría si cambiáramos esta convención tan arbitraria? El experimento se realizó con éxito: si la entrada a la escuela se posterga entre media hora y una hora, todo cambia: los adolescentes duermen mejor, su ausentismo se reduce, la atención en clase aumenta y el desempeño escolar mejora (American Academy of Pediatrics, 2014). Y la lista de efectos positivos podría continuar: la Academia Estadounidense de Pediatría recomienda enfáticamente retrasar los horarios de inicio escolar como una medida eficiente para combatir la obesidad, la depresión y los accidentes (entre otros, por conducir con sueño). El hecho de que el bienestar físico y mental de los niños pueda mejorarse con tanta facilidad y sin costo alguno sirve de magnífico ejemplo de los beneficios de adaptar el sistema educativo a las características de la biología cerebral.

[25] El hipocampo también parece hacer un robusto aporte a la memorización de los hechos aritméticos (Qin y otros, 2014).

[26] Véanse también los resultados que revelan un efecto directo del sueño en la plasticidad sináptica en Norimoto y otros (2018).

[27] Philippe Starck, *France Inter*, viernes 8 de junio de 2018.

Conclusión

Entiendo que la mayor y principal dificultad de la ciencia humana es la acertada dirección y educación de los niños.

Montaigne, *Ensayos* (1580)

La pedagogía es como la medicina: un arte, pero que se apoya –o debería apoyarse– sobre conocimientos científicos precisos.

Jean Piaget, “La pedagogía moderna” (1949)

Ahora que nuestro viaje llega a su fin, tengo la esperanza de haberlos convencido de que, gracias a los avances de la psicología cognitiva, las neurociencias, la inteligencia artificial y las ciencias de la educación, disponemos de conocimientos detallados sobre el modo en que nuestro cerebro aprende. Estos conocimientos no nos resultan obvios, lo que da mayor motivo para que revisemos la mayoría de nuestras ideas sobre el aprendizaje:

- *No, el bebé no es una tabula rasa carente de saberes:* desde el primer año de vida, dispone de un enorme conjunto de conocimientos sobre los objetos, los números, las probabilidades, el espacio y las personas.
- *No, el niño no es una esponja que absorbe dócilmente la estructura de su entorno.* Recuerden a Felipe, el niño ciego y tetrapléjico que descollaba como cuentacuentos en Brasil, o a Nicholas Saunderson, el matemático ciego que sucedió a Newton en cátedra: esas dos experiencias nos muestran que la recepción de información sensorial puede estar alterada o ausente, sin que eso elimine la capacidad de abstracción.
- *No, el cerebro no es únicamente una red neuronal maleable que se deja modelar por la información que recibe:* todos los grandes haces

de conexiones están presentes desde el nacimiento, y la plasticidad cerebral, aunque indispensable, solo ajusta los últimos milímetros de nuestras conexiones.

- *No, el aprendizaje no se produce de forma pasiva, por simple exposición a datos, clases magistrales o conferencias.* Por el contrario, la psicología cognitiva y las neuroimágenes demuestran que el niño es un científico en ciernes que constantemente genera nuevas hipótesis, y que el cerebro es un órgano siempre alerta, que aprende poniendo a prueba los modelos que proyecta sobre el mundo exterior.
- *No, el error no es la marca de los “malos” alumnos:* equivocarse es parte integrante del aprendizaje, porque solo gracias al error el cerebro puede ajustar sus modelos cuando descubre una discrepancia entre sus expectativas y la realidad.
- *No, el sueño no es apenas un período de descanso:* forma parte de nuestro algoritmo de aprendizaje como momento privilegiado en que el cerebro reproduce sus modelos una y otra vez, en bucle, y amplifica, por un factor de diez o de cien, lo adquirido durante el día.
- *Y, por último, no, las terminales informáticas actuales no están cerca de superar al cerebro humano.* Al menos por ahora, nuestro cerebro sigue siendo el dispositivo de procesamiento de la información más rápido, más eficaz y el que menos energía consume. Como una verdadera máquina probabilística, logra extraer de cada episodio de la jornada aun los fragmentos ínfimos de información y transformarlos durante la noche en conocimientos abstractos y generales, de un modo que todavía no sabemos reproducir en las computadoras.

En la batalla prometeica que libran el microchip y la neurona, la máquina y el cerebro, este último todavía conserva la ventaja. En principio, en la mecánica cerebral no hay elemento que una máquina no pueda imitar. Por lo pronto, todas las ideas que enuncié en este libro ya funcionan como herramientas en las manos de informáticos cuyas investigaciones se inspiran abiertamente en las neurociencias (Hassabis y otros, 2017; Lake y otros, 2016). Sin embargo, en la práctica, las máquinas todavía deben transitar un largo camino. Para avanzar, necesitarán varios de los ingredientes que analizamos aquí: un lenguaje del pensamiento interno que permita recombinar los conceptos con

flexibilidad; algoritmos capaces de razonar a partir de distribuciones de probabilidad; una función de curiosidad; sistemas eficaces de gestión de la atención y de la memoria; y, tal vez, como ya vimos, un algoritmo de sueño-vigilia que expanda el entrenamiento y multiplique las oportunidades de descubrimiento. Hoy en día, comienzan a aparecer algoritmos de este tipo, pero todavía están a años luz de alcanzar el desempeño de un bebé recién nacido. Nuestro cerebro aún lleva las de ganar frente a las máquinas, y presumo que seguirá siendo así durante mucho tiempo.

Trece recomendaciones para optimizar el potencial de los niños

Cuanto más estudio el cerebro humano, más me impresiona. Pero también sé que su desempeño es frágil, porque depende mucho del entorno donde se desarrolla. Muchos niños (incluso demasiados) no alcanzan a realizar su potencial de aprendizaje con plenitud porque la familia, la escuela o la sociedad no les proveen las condiciones ideales.

Las comparaciones internacionales son alarmantes: desde hace quince o veinte años, los rendimientos escolares de muchos países occidentales – incluido mi propio país, Francia – se desplomaron, mientras que ganaron terreno muchos países y ciudades asiáticos, como Singapur, Shanghái o Hong Kong.^[28] En matemáticas, que solía ser el punto fuerte de Francia, los resultados descendieron drásticamente entre 2003 y 2015. La tendencia se afianzó en 2018: se confirmó que mi país ocupaba el último lugar de Europa en la prueba Timms, que evalúa el aprendizaje de los alumnos de 15 años en matemáticas y en ciencias; de todos los países participantes registró la mayor baja en veinte años. En 2016 ocurrió algo similar ocurrió con la prueba Pirls, que evalúa la lectura en cuarto grado: en menos de dos décadas, Francia, patria de muchos matemáticos notables, sufrió un naufragio y actualmente ocupa la parte más baja del *ranking* europeo, en especial cuando es cuestión de interpretar textos más complejos. Para peor, el país cuyo lema patrio clama Igualdad y Fraternidad presenta los peores resultados en el área de integración escolar: Francia es el país con el mayor impacto del nivel social de las familias en sus resultados escolares.

Frente a estos pésimos resultados con demasiada frecuencia tiende a

levantarse el dedo para acusar a los docentes. En realidad, nadie sabe cuáles son las causas de esta caída reciente: ¿la escuela, los padres, o la sociedad como un todo? ¿Debemos culpar a la falta de sueño, la falta de atención, los videojuegos, las familias reducidas? En cualquiera de los casos, estoy seguro de que los avances recientes de las ciencias del aprendizaje pueden ayudar a revertir esta oscura tendencia. Hoy en día sabemos mucho más sobre las condiciones que maximizan el aprendizaje y la memoria. Como padres, madres y docentes, debemos aprender a implementarlas en la vida cotidiana, tanto en casa como en la escuela.

A lo largo de este libro, presenté una serie de resultados científicos que convergen en ideas muy simples y fáciles de aplicar. Revisémoslas juntos:

- *No subestimemos a los niños y a las niñas.* Desde el nacimiento, cada niño posee un rico conjunto de habilidades y conocimientos. El concepto de objeto, el sentido de los números, el don de las lenguas, el conocimiento de las personas y de sus intenciones... Tantos módulos cerebrales que ya están presentes en los más pequeños y que la educación reciclará más tarde durante los cursos de física, matemáticas, lengua o filosofía. Apoyémonos sobre las intuiciones precoces de los niños: cada palabra o cada símbolo que aprenden, aunque sea abstracto, debe conectarse con conocimientos previos. Esta conexión les dará sentido.
- *Aprovechemos los períodos sensibles.* En los primeros años de vida, millares de sinapsis se crean y se destruyen todos los días. Esta efervescente actividad vuelve al cerebro particularmente receptivo a cierto tipo de aprendizaje; en primer lugar, el de las lenguas. Sepamos aprovechar este período sensible para exponer a los niños a una segunda lengua tan pronto como sea posible, sin perder de vista que la plasticidad se prolonga por lo menos hasta la adolescencia. Durante esta etapa, la inmersión en una lengua extranjera puede transformar el cerebro en el lapso de pocos meses.
- *Enriquezcamos el entorno.* Como consumado genio del aprendizaje, el cerebro del niño sigue siendo la más poderosa de las supercomputadoras. Respetémoslo dándole, desde la edad más temprana, estímulos que estén a su altura: juegos de palabras o de

construcción, historias, desafíos, acertijos, rompecabezas... No dudemos en tener conversaciones serias con los niños, respondamos a sus preguntas, incluso a las más difíciles, con un vocabulario elaborado para explicarle lo que nosotros entendemos sobre el mundo que nos rodea. Darles a los pequeños un entorno enriquecido, especialmente en lo que respecta al lenguaje, maximizará el crecimiento de su cerebro y preservará, durante el mayor tiempo posible, su plasticidad juvenil.

- *Abandonemos la idea de que todos los niños son diferentes.* La idea de que cada uno de nosotros posee su propio estilo de aprendizaje es un mito. Las neuroimágenes dejan de manifiesto que todos poseemos circuitos y reglas de aprendizaje muy similares. Las áreas cerebrales responsables de la lectura, el cálculo mental o las matemáticas son casi las mismas en todos, milímetros más o menos, incluso sin distinción en el caso de niñas y niños ciegos. Todos los niños se encuentran con dificultades similares, y los mismos métodos de enseñanza permiten resolverlas. Las diferencias individuales, cuando existen, versan sobre el conocimiento previo, la motivación y la velocidad del aprendizaje. Sepamos diagnosticar cuidadosamente el nivel de los niños para estimularlos con problemas adaptados a su nivel y a sus ganas, pero, sobre todas las cosas, asegurémonos de que todas las criaturas adquieran los fundamentos del lenguaje, la alfabetización y las matemáticas que, sin excepción, necesitan.
- *Prestemos atención a la atención.* La atención es la puerta de entrada al aprendizaje: casi ninguna información será memorizada si antes no fue amplificadas por la atención y la toma de conciencia. En la escuela, los docentes deben lograr cautivar la atención de sus alumnos y orientarla hacia lo que realmente importa. Esto implica eliminar con cuidado cualquier fuente de distracción: los manuales demasiado ilustrados, las aulas con decoración excesiva, las letras o los dígitos deformados u ornamentados con personajes no hacen otra cosa que distraer al niño de su tarea principal y le impiden concentrarse.
- *Hagamos que el niño se vuelva activo, curioso, comprometido y autónomo.* Un alumno pasivo no aprende nada. Volvámoslo más activo. Provoquemos la demanda permanente de su inteligencia para que su mente resplandezca de curiosidad y genere permanentemente

nuevas hipótesis. Pero no esperemos que descubra todo por sí solo: guiémoslo mediante contenidos y secuencias pedagógicas estructurados.

- *Hagamos de cada día de escuela algo placentero.* Los circuitos de la recompensa son moduladores esenciales de la plasticidad cerebral. Utilicémoslos recompensando cada esfuerzo y haciendo que cada hora de clase sea un momento divertido. Ningún niño es insensible a las recompensas materiales, pero su cerebro social responde de la misma manera a las sonrisas y los mensajes de aliento. El sentimiento de ser apreciado, pero también la conciencia de que se está progresando, aportan su propia recompensa. Como contrapartida, erradiquemos la ansiedad y el estrés que bloquean el aprendizaje, sobre todo en matemáticas.
- *Alentemos los esfuerzos.* Una escuela placentera no es sinónimo de una escuela “sin esfuerzo”. Al contrario, las habilidades más interesantes, como la lectura, las matemáticas o tocar un instrumento exigen meses o años de aprendizaje. Evitemos hacerles creer a los niños que todo será fácil, porque además esto traerá aparejado que, cuando no logren los objetivos, piensen que son burros o toscos. En cambio, expliquémosles que todos los alumnos deben hacer esfuerzos y que, cuando los hacen, consiguen progresar. Adoptemos una mentalidad progresista, no una mentalidad rígida.
- *Ayudemos a los alumnos a profundizar su pensamiento.* Cuanto más profundamente procese la información nuestro cerebro, mejor podremos recordarla. Nunca nos contentemos con el aprendizaje superficial; busquemos siempre una comprensión más profunda. Y recordemos las palabras de Henry Roediger: “Hacer que las condiciones de aprendizaje sean más difíciles, lo que requiere mayor esfuerzo cognitivo por parte de los estudiantes, a menudo redundan en una mayor retención”.
- *Fijemos objetivos claros.* Los alumnos aprenden mejor cuando les enunciamos con claridad el objetivo de la enseñanza y pueden constatar que todo lo que se les propone converge allí. Expliquémosles sin rodeos qué esperamos de ellos y permanezcamos enfocados en esa meta.

- *Aceptemos y corriamos los errores.* Para actualizar los modelos mentales, nuestras áreas cerebrales deben intercambiar mensajes de error. El error es, entonces, la condición misma del aprendizaje. No lo sancionemos; pero corriámoslo rápidamente, dándole al niño una devolución detallada y sin presión sobre el error que cometió. De acuerdo con el informe de la Education Endowment Foundation, la calidad del *feedback* del docente a sus alumnos es el trampolín más eficaz del progreso escolar.
- *Repasemos, una y otra vez.* Aprender de una sola vez no es suficiente, también es necesario consolidar los contenidos del aprendizaje con el objetivo de que se vuelvan automáticos, inconscientes y reflexivos. Solo la automatización libera la corteza prefrontal, que, de este modo, queda disponible para otras actividades. La estrategia más eficaz consiste en segmentar el aprendizaje: un poco todos los días. Por eso, distribuir las sesiones de estudio y práctica permitirá que la información se imprima en la memoria de modo definitivo.
- *Dejemos dormir a los niños.* El sueño es un ingrediente esencial de nuestro algoritmo de aprendizaje. Cada período de sueño, incluida la siesta, aporta un beneficio suplementario. Asegurémonos, entonces, de que los niños duerman mucho y con profundidad. Para obtener el máximo beneficio del período de trabajo inconsciente nocturno, repasar una lección o volver a leer un problema justo antes de dormir podría resultar una estrategia útil. No olvidemos que la cronobiología indica que durante la adolescencia los ciclos del sueño se desplazan ¡y no programemos el despertador para un horario tan temprano!

Solo si aprendemos a conocernos mejor obtendremos las mayores ventajas de los poderosos algoritmos con que está equipado el cerebro. Sin lugar a duda, todos los niños se verían beneficiados si se conocieran los cuatro pilares del aprendizaje: la atención, el compromiso activo, la importancia de un buen *feedback* sobre el error y la consolidación. Cuatro eslóganes los resumen con eficacia: “Concéntrense por completo”; “Participen en clase”; “Hagan ejercicios todos los días”; “Aprovechen cada noche”. Son mensajes muy sencillos que cada cual debería conocer y poner en práctica, desde la edad más temprana.

Una alianza para la escuela del mañana

¿Sabremos lograr una armonía entre nuestras escuelas y los descubrimientos de las neurociencias y de las ciencias cognitivas? Para conseguirlo, se necesita una nueva alianza. Así como la medicina descansa actualmente en una pirámide de investigación biológica y farmacológica, estoy seguro de que, en el futuro, la educación se valdrá de la investigación basada sobre evidencia, incluidos experimentos de laboratorio pero también pruebas a escala real y estudios de implementación en el aula. Solo si unimos nuestras fuerzas podremos lograrlo. La escuela del mañana deberá convocar a los docentes, las familias y los científicos alrededor de una causa común: revivir en las niñas y los niños la curiosidad y la alegría de aprender, para así ayudarlos a optimizar su potencial cognitivo.

Expertos del aula, los maestros tienen la invaluable tarea de criar a nuestros hijos e hijas, quienes pronto tendrán el futuro del mundo en sus manos. Sin embargo, a menudo ejercen esa tarea en condiciones difíciles y con herramientas escasas. Se merecen mucho más respeto y una enorme inversión. Los docentes hoy enfrentan desafíos cada vez más severos, que incluyen la disminución de los recursos, el aumento del tamaño de las clases, la creciente violencia y la implacable tiranía del plan de estudios. Sorprendentemente, la mayoría de ellos recibe poca capacitación profesional y casi ninguna introducción a la ciencia del aprendizaje. Creo que debemos cambiar urgentemente este estado de cosas. Espero que este libro contribuya a que las universidades, los institutos de formación y las escuelas superiores del profesorado ofrezcan un plan de estudios a la altura del compromiso que los docentes tienen con nuestros niños.

También espero que los docentes estén de acuerdo en que de ningún modo el avance de las neurociencias del aprendizaje puede coartar su libertad pedagógica. Por el contrario, el objetivo de este libro es permitirles ejercer mejor esa libertad. “Pienso en un héroe”, señaló el actual premio Nobel Bob Dylan cuando en 1985 presentó su álbum de *rarities* e inéditos *Biograph*, “como alguien que entiende el grado de responsabilidad que su libertad conlleva”. Por eso, la creatividad pedagógica genuina solo puede provenir de una plena conciencia de la gama de estrategias disponibles, de la posibilidad de elegir cuidadosamente entre ellas, con conocimiento de su impacto en los estudiantes. Los principios que enuncié a lo largo de este libro son compatibles con múltiples enfoques pedagógicos, y queda todo por inventar

para traducirlos en actos dentro del aula. Espero mucho de la imaginación de los maestros, porque la considero esencial para desencadenar el entusiasmo de los niños.

En mi opinión, las escuelas del futuro también deberían involucrar más a las familias y los adultos responsables, ya que son actores primordiales en el desarrollo de un niño, cuyas acciones preceden y prolongan las de la escuela. El entorno inmediato es donde los niños tienen la oportunidad de expandir, a través del trabajo y los juegos, el conocimiento que adquirieron en clase. La familia está abierta los siete días de la semana y, por lo tanto, puede, mejor que la escuela, aprovechar al máximo la alternancia entre vigilia y sueño, entre aprendizaje y consolidación. Las escuelas deberían dedicar más tiempo a la capacitación de los padres, porque es una de las intervenciones más eficaces: pueden ser relevos invaluable para los maestros, además de observadores astutos y detallistas de las dificultades de sus niños y niñas.

Por último, los científicos deben comprometerse junto con los docentes y la escuela para consolidar el creciente campo de las ciencias de la educación. En comparación con el gran progreso que tuvieron las ciencias cognitivas y las neurociencias en los últimos treinta años, la investigación en educación todavía es un área relativamente descuidada. Los entes estatales y privados deberían incitar a los científicos a realizar grandes programas de investigación básica en todos los campos de las ciencias del aprendizaje, desde las neurociencias y las neuroimágenes hasta la neuropsicología de los trastornos del desarrollo, la psicología cognitiva y la sociología de la educación. El cambio de escala, del laboratorio al aula, no es tan fácil como parece. Tenemos una gran necesidad de experimentación en las escuelas, en espacios a escala real. Las ciencias cognitivas pueden ayudar a concebir y a evaluar herramientas educativas innovadoras. Del mismo modo en que la medicina se apoya sobre la biología, el campo de la educación debe dar vida a un nuevo ecosistema de investigación, más sistemático, más riguroso, que propicie la colaboración entre docentes e investigadores en la búsqueda incesante de una educación más eficaz, basada en datos fácticos y correctamente probados.

Docentes, familias, científicos: todos juntos para hacer progresar las ciencias de la educación y su aplicación práctica en la escuela. La fórmula me parece clara, eficaz, y la urgencia, real.

[28] Véanse las pruebas PISA (Program for International Student Assessment; www.oecd.org/pisa-fr), Timms (Trends in International Mathematics and Science Study) y Pirls (Progress in International Reading Literacy Study; timssandpirls.bc.edu).

Agradecimientos

El desarrollo de este libro se vio estimulado por numerosos encuentros. Los primeros, con Michael Posner y Bruce McCandliss, de la Universidad de Oregón, me convencieron de que las ciencias cognitivas son pertinentes para la educación. Debo mucho a los numerosos encuentros científicos organizados con ayuda de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) y de Bruno Della Chiesa. Más tarde, Marcela Peña, Sidarta Ribeiro, Mariano Sigman, Alejandro Maiche y Juan Valle Lisboa formaron a una generación de jóvenes científicos en los inolvidables encuentros anuales de la Latin American School for Education, Cognitive and Neural Sciences. Tuve la suerte de participar en todas sus ediciones y les debo mucho.

Otra persona que tuvo esta suerte es mi esposa y colega Ghislaine Dehaene-Lambertz. Hace treinta y dos años que debatimos sobre el desarrollo cerebral y, por añadidura, sobre la educación de nuestros hijos. No hace falta decir que le debo todo a ella, incluida una relectura atenta de las páginas precedentes.

Otro aniversario: hace treinta y tres años que me uní a los laboratorios de Jacques Mehler y de Jean-Pierre Changeux. Su influencia sobre mi pensamiento es inmensa, y ellos reconocerán aquí un buen número de sus temas predilectos, como también lo harán otros colegas y amigos muy cercanos como Lúcia Braga, Laurent Cohen, Naama Friedmann, Véronique Izard, Régine Kolinsky, José Morais, Lionel Naccache, Christophe Pallier, Mariano Sigman, Elizabeth Spelke y Josh Tenenbaum.

Mi amigo Antonio Battro me alentó constantemente en mis investigaciones sobre el cerebro, la educación y el aprendizaje. Le estoy igualmente agradecido por haberme hecho conocer a Nico, un artista con una personalidad notable, que me autorizó con gentileza a reproducir aquí algunos de sus cuadros. Gracias también a Yoshua Bengio, Alain Chédotal, Guillaume y David Dehaene, Molly Dillon, Jessica Dubois, György Gergely,

Eric Knudsen, Leah Krubitzer, Bruce McCandliss, Josh Tenenbaum, Fei Xu y Robert Zatorre por haberme autorizado a reproducir las numerosas figuras que dan color a este texto.

También me gustaría agradecerles a todas las instituciones que sustentan mis investigaciones desde hace años con una fidelidad a prueba de todo, y particularmente al Instituto Nacional de Salud e Investigación Médica (Inserm), la Comisión de Energía Atómica y Energías Renovables, el Collège de France, la Universidad París-Sur, el European Research Council (ERC) y la Fundación Bettencourt-Schueller. Gracias a ellos me he podido rodear de estudiantes y colaboradores brillantes y enérgicos. Son demasiado numerosos para ser citados aquí, pero se reconocerán en la larga lista de publicaciones que sigue. Vaya una mención particular para Anna Wilson, Dror Dotan y Cassandra Potier-Watkins, con quienes desarrollé programas educativos e intervenciones en las aulas.

Jean-Michel Blanquer, ministro de Educación de Francia, me honró con su confianza al proponerme presidir su primer Consejo Científico, algo que quiero agradecerle con vehemencia aquí. Les debo mi reconocimiento a todos los miembros de este Consejo, y particularmente a Esther Duflo, Michel Fayol, Marc Gurgand, Caroline Huron, Elena -Pasquinelli, Franck Ramus, Elizabeth Spelke y Jo Ziegler, así como a su secretario general Nelson Vallejo-Gómez, por su compromiso y por todo lo que me enseñaron.

La preparación de esta edición se vio beneficiada por una relectura atenta de Marie-Lorraine Colas.

Por último, menciono que este es el sexto libro que publico con mi editora Odile Jacob y su marido Bernard Gottlieb. Su amistad y su confianza me conmueven profundamente. Poco después, el ojo crítico de Wendy Wolf y Terezia Cichel, mis editoras en Penguin Viking, siguió haciendo aportes al texto. Y no habría llegado a sus manos sin la constante ayuda de mi agente, Brockman Inc. Gracias, John y Max, por su inagotable apoyo y su invaluable *feedback*. La versión actual renueva y confirma el vínculo de años con Siglo XXI Editores, que publica mi obra en castellano bajo el cuidado de Yamila Sevilla. Agradezco a Carlos Díaz y a Diego Golombek la inclusión de este nuevo libro en la Serie Mayor de Ciencia que ladra, y a Josefina D'Alessio y Luciano Padilla López la traducción y el exquisito cuidado del texto.

S. D., París, septiembre de 2019

Bibliografía

- Abboud, S., Maidenbaum, S., Dehaene, S., Amedi, A. (2015), “A number-form area in the blind”, *Nature Communications*, 6, 6026.
- Adibpour, P., Dubois, J., Dehaene-Lambertz, G. (2018), “Right but not left hemispheric discrimination of faces in infancy”, *Nature Human Behaviour*, 2 (1), 67-79.
- Ahissar, M., Hochstein, S. (1993), “Attentional control of early perceptual learning”, *Proc. Natl Acad. Sci. USA*, 90 (12), 5718-5722.
- Almas, A. N., Degnan, K. A., Radulescu, A., Nelson, C. A., Zeanah, C. H., Fox, N. A. (2012), “Effects of early intervention and the moderating effects of brain activity on institutionalized children’s social skills at age 8”, *Proc. Natl Acad. Sci. USA*, 109 (suppl. 2), 17.228-17.231.
- Amalric, M., Dehaene, S. (2016), “Origins of the brain networks for advanced mathematics in expert mathematicians”, *Proceedings of the National Academy of Sciences*, disponible en www.pnas.org.
- (2017), “Cortical circuits for mathematical knowledge: Evidence for a major subdivision within the brain’s semantic networks”, *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 373 (1740).
- Amalric, M., Denghien, I., Dehaene, S. (2017), “On the role of visual experience in mathematical development: Evidence from blind mathematicians”, *Developmental Cognitive Neuroscience*.
- Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., Dehaene, S. (2017), “The language of geometry: Fast comprehension of geometrical primitives and rules in human adults and preschoolers”, *PLoS Computational Biology*, 13 (1), e1.005.273.
- Amedi, A., Raz, N., Pianka, P., Malach, R., Zohary, E. (2003), “Early

- “visual” cortex activation correlates with superior verbal memory performance in the blind”, *Nature Neuroscience*, 6 (7), 758-766.
- American Academy of Pediatrics (2014), “School start times for adolescents”, *Pediatrics*, 134 (3), 642-649.
- Amunts, K., Lenzen, M., Friederici, A. D., Schleicher, A., Morosan, P., Palomero-Gallagher, N., Zilles, K. (2010), “Broca’s region: Novel organizational principles and multiple receptor mapping”, *PLoS Biology*, 8 (9).
- Amunts, K., Zilles, K. (2015), “Architectonic mapping of the human brain beyond Brodmann”, *Neuron*, 88 (6), 1086-1107.
- Ansari, D., Dhital, B. (2006), “Age-related changes in the activation of the intraparietal sulcus during nonsymbolic magnitude processing: An event-related functional magnetic resonance imaging study”, *J. Cogn. Neurosci.*, 18 (11), 1820-1828.
- Antony, J. W., Gobel, E. W., O’Hare, J. K., Reber, P. J., Paller, K. A. (2012), “Cued memory reactivation during sleep influences skill learning”, *Nature Neuroscience*, 15 (8), 1114-1116.
- Arnould, L. (1900), *Une âme en prison. Histoire de l’éducation d’une aveuglesourde-muette de naissance*, Paris, Oudin.
- Arzi, A., Shedlesky, L., Ben-Shaul, M., Nasser, K., Oksenberg, A., Hairston, I. S., Sobel, N. (2012), “Humans can learn new information during sleep”, *Nature Neuroscience*, 15 (10), 1460-1465.
- Ashcraft, M. H. (2002), “Math anxiety: Personal, educational, and cognitive consequences”, *Current Directions in Psychological Science*, 11, 181-185.
- Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., Jaeggi, S. M. (2015), “Improving fluid intelligence with training on working memory: A metaanalysis”, *Psychonomic Bulletin and Review*, 22 (2), 366-377.
- Auble, P. M., Franks, J. J. (1978), “The effects of effort toward comprehension on recall”, *Memory and Cognition*, 6 (1), 20-25.

- Auble, P. M., Franks, J. J., Soraci, S. A., (1979), "Effort toward comprehension: Elaboration or 'aha'?", *Memory and Cognition*, 7 (6), 426-434.
- Avior, G., Fishman, G., Leor, A., Sivan, Y., Kaysar, N., Derowe, A. (2004), "The effect of tonsillectomy and adenoidectomy on inattention and impulsivity as measured by the Test of Variables of Attention (TOVA) in children with obstructive sleep apnea syndrome", *Otolaryngology*, 131 (4), 367-371.
- Bahdanau, D., Cho, K., Bengio, Y. (2014), "Neural machine translation by jointly learning to align and translate", arxiv.org/abs/1409.0473.
- Baillargeon, R., DeVos, J. (1991), "Object permanence in young infants: Further evidence", *Child Development*, 62 (6), 1227-1246.
- Baillargeon, R., Needham, A., Devos, J. (1992), "The development of young infants' intuitions about support", *Early Development and Parenting*, 1 (2), 69-78.
- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., Tidball, G. (1996), "Infants' reliance on a social criterion for establishing word-object relations", *Child Development*, 67 (6), 3135-3153.
- Balsam, P. D., Gallistel, C. R. (2009), "Temporal maps and informativeness in associative learning", *Trends in Neurosciences*, 32 (2), 73-78.
- Banino, A., Barry, C., Uria, B., Blundell, C., Lillicrap, T., Mirowski, P., ..., Kumaran, D. (2018), "Vector-based navigation using grid-like representations in artificial agents", *Nature*, 557 (7705), 429-433.
- Bao, S., Chan, V. T., Merzenich, M. M. (2001), "Cortical remodelling induced by activity of ventral tegmental dopamine neurons", *Nature*, 412 (6842), 79-83.
- Bavelier, D., Green, C. S., Han, D. H., Renshaw, P. F., Merzenich, M. M., Gentile, D. A. (2011), "Brains on video games", *Nature Reviews Neuroscience*, 12 (12), 763-768.

- Beckers, T., Miller, R. R., De Houwer, J., Urushihara, K. (2006), "Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference", *Journal of Experimental Psychology: General*, 135 (1), 92.
- Bedny, M. (2017), "Evidence from blindness for a cognitively pluripotent cortex", *Trends Cogn. Sci.*, 21 (9), 637-648.
- Bedny, M., Pascual-Leone, A., Dodell-Feder, D., Fedorenko, E., Saxe, R. (2011), "Language processing in the occipital cortex of congenitally blind adults", *Proc. Natl Acad. Sci. USA*, 108 (11), 4429-4434.
- Behne, T., Carpenter, M., Call, J., Tomasello, M. (2005), "Unwilling versus unable: infants' understanding of intentional action", *Developmental Psychology*, 41 (2), 328.
- Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., Naccache, L. (2009), "Neural signature of the conscious processing of auditory regularities", *Proc. Natl Acad. Sci. USA*, 106 (5), 1672-1677.
- Belle, M., Godefroy, D., Couly, G., Malone, S. A., Collier, F., Giacobini, P., Chédotal, A. (2017), "Tridimensional visualization and analysis of early human development", *Cell*, 169 (1), 161-173.e12.
- Bendor, D., Wilson, M. A. (2012), "Biasing the content of hippocampal replay during sleep", *Nature Neuroscience*, 15 (10), 1439-1444.
- Berens, A. E., Nelson, C. A. (2015), "The science of early adversity: Is there a role for large institutions in the care of vulnerable children?", *The Lancet*, 386 (9991), 388-398.
- Bergman-Nutley, S., Klingberg, T. (2014), "Effect of working memory training on working memory, arithmetic and following instructions", *Psychological Research*, 78 (6), 869-877.
- Berkes, P., Orban, G., Lengyel, M., Fiser, J. (2011), "Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment", *Science*, 331 (6013), 83-87.
- Bermúdez, P., Lerch, J. P., Evans, A. C., Zatorre, R. J. (2009), "Neuroanatomical correlates of musicianship as revealed by cortical

thickness and voxel-based morphometry”, *Cereb. Cortex*, 19 (7), 1583-1596.

Bernal, S., Dehaene-Lambertz, G., Millotte, S., Christophe, A. (2010), “Two-year-olds compute syntactic structure on-line”, *Dev. Sci.*, 13 (1), 69-76.

Bessa, C., Maciel, P., Rodrigues, A. J. (2013), “Using *C. elegans* to decipher the cellular and molecular mechanisms underlying neurodevelopmental disorders”, *Molecular Neurobiology*, 48 (3), 465-489.

Bevins, R. A. (2001), “Novelty seeking and reward: Implications for the study of high-risk behaviors”, *Current Directions in Psychological Science*, 10 (6), 189-193.

Bialystok, E., Craik, F. I. M., Green, D. W., Gollan, T. H. (2009), “Bilingual minds”, *Psychological Science in the Public Interest*, 10 (3), 89-129.

Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., Buchanan, L. (2006), “Tuning of the human left fusiform gyrus to sublexical orthographic structure”, *NeuroImage*, 33 (2), 739-748.

Blair, C., Raver, C. C. (2014), “Closing the achievement gap through modification of neurocognitive and neuroendocrine function: Results from a Cluster randomized controlled trial of an innovative approach to the education of children in kindergarten”, *PLoS ONE*, 9 (11).

Blair, K. P., Rosenberg-Lee, M., Tsang, J. M., Schwartz, D. L., Menon, V. (2012), “Beyond natural numbers: Negative number representation in parietal cortex”, *Frontiers in Human Neuroscience*, 6, 7.

Bliss, T. V., Lømo, T. (1973), “Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path”, *The Journal of Physiology*, 232 (2), 331-356.

Bock, A. S., Binda, P., Benson, N. C., Bridge, H., Watkins, K. E., Fine, I. (2015), “Resting-state retinotopic organization in the absence of retinal input and visual experience”, *J. Neurosci.*, 35 (36), 12366-

12.382.

- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., Schulz, L. (2011), "The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery", *Cognition*, 120 (3), 322-330.
- Bond, R., Smith, P. B. (1996), "Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task", *Psychological Bulletin*, 119 (1), 111.
- Borst, G., Poirel, N., Pineau, A., Cassotti, M., Houdé, O. (2013), "Inhibitory control efficiency in a Piaget-like class-inclusion task in school-age children and adults: A developmental negative priming study", *Developmental Psychology*, 49 (7), 1366-1374.
- Bouhali, F., Thiebaut de Schotten, M., Pinel, P., Poupon, C., Mangin, J.-F., Dehaene, S., Cohen, L. (2014), "Anatomical connections of the visual word form area", *J. Neurosci.*, 34 (46), 15.402-15.414.
- Bradley, M. M., Costa, V. D., Ferrari, V., Codispoti, M., Fitzsimmons, J. R., Lang, P. J. (2015), "Imaging distributed and massed repetitions of natural scenes: Spontaneous retrieval and maintenance", *Human Brain Mapping*, 36 (4), 1381-1392.
- Braga, L. W., Amemiya, E., Tauil, A., Sugueida, D., Lacerda, C., Klein, E., ..., Dehaene, S. (2017), "Tracking adult literacy acquisition with functional MRI: A single-case study", *Mind, Brain and Education*.
- Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., Gabrieli, J. D. (1998), "Making memories: Brain activity that predicts how well visual experience will be remembered", *Science*, 281 (5380), 1185-1187.
- Brodmann, K. (1909), *Localisation in the Cerebral Cortex*, Londres, Springer.
- Bromberg-Martin, E. S., Hikosaka, O. (2009), "Midbrain dopamine neurons signal preference for advance information about upcoming rewards", *Neuron*, 63 (1), 119-126.
- Brun, V. H., Leutgeb, S., Wu, H. Q., Schwarcz, R., Witter, M. P., Moser, E. I., Moser, M. B. (2008), "Impaired spatial representation in CA1

after lesion of direct input from entorhinal cortex”, *Neuron*, 57 (2), 290-302.

Buon, M., Jacob, P., Margules, S., Brunet, I., Dutat, M., Cabrol, D., Dupoux, E. (2014), “Friend or foe? Early social evaluation of human interactions”, *PLoS ONE*, 9 (2).

Butler, A. C., Karpicke, J. D., Roediger, H. L. (2008), “Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses”, *J. Exp. Psychol. Learn. Mem. Cogn.*, 34 (4), 918-928.

Byers-Heinlein, K., Werker, J. F. (2009), “Monolingual, bilingual, trilingual: Infants’ language experience influences the development of a word-learning heuristic”, *Dev. Sci.*, 12 (5), 815-823.

Callan, D. E., Schweighofer, N. (2010), “Neural correlates of the spacing effect in explicit verbal semantic encoding support the deficient-processing theory”, *Human Brain Mapping*, 31 (4), 645-659.

Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., ..., Ramey, C. T. (2012), “Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up”, *Developmental Psychology*, 48 (4), 1033-1043.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., Pan, Y. (2014), “Early childhood investments substantially boost adult health”, *Science*, 343 (6178), 1478-1485.

Cantlon, J. F., Brannon, E. M., Carter, E. J., Pelphrey, K. A. (2006), “Functional imaging of numerical processing in adults and 4-y-old children”, *PLoS Biology*, 4 (5).

Cantlon, J. F., Li, R. (2013), “Neural activity during natural viewing of *Sesame Street* statistically predicts test scores in early childhood”, *PLoS Biology*, 11 (1).

Cardoso-Leite, P., Bavelier, D. (2014), “Video game play, attention, and learning: How to shape the development of attention and influence learning?”, *Current Opinion in Neurology*, 27 (2), 185-191.

Carey, S., Bartlett, E. (1978), “Acquiring a single new word”,

Proceedings of the Stanford Child Language Conference, 15, 17-29.

Caroni, P., Donato, F., Muller, D. (2012), "Structural plasticity upon learning: Regulation and functions", *Nature Reviews. Neuroscience*, 13 (7), 478-490.

Carreiras, M., Seghier, M. L., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., Price, C. J. (2009), "An anatomical signature for literacy", *Nature*, 461 (7266), 983-986.

Carrier, M., Pashler, H. (1992), "The influence of retrieval on retention", *Memory and Cognition*, 20 (6), 633-642.

Castles, A., Rastle, K., Nation, K. (2018), "Ending the Reading Wars: Reading Acquisition From Novice to Expert", *Psychological Science in the Public Interest*, 19 (1), 5-51.

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., Ingvar, M. (1998), "The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain", *Brain*, 121 (Pt 6), 1053-1063.

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., Pashler, H. (2009), "Optimizing distributed practice: Theoretical analysis and practical implications", *Experimental Psychology*, 56 (4), 236-246.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., Rohrer, D. (2006), "Distributed practice in verbal recall tasks: A review and quantitative synthesis", *Psychological Bulletin*, 132 (3), 354-380.

Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., Bonatti, L. L. (2018), "Precursors of logical reasoning in preverbal human infants", *Science*, 359 (6381), 1263-1266.

Chafee, M. V. (2013), "A scalar neural code for categories in parietal cortex: Representing cognitive variables as 'more' or 'less'", *Neuron*, 77 (1), 7-9.

Chakraborty, M., Jarvis, E. D. (2015), "Brain evolution by brain pathway duplication", *Philos. Trans. R. Soc. B*, 370 (1684).

- Chang, C. H. C., Pallier, C., Wu, D. H., Nakamura, K., Jobert, A., Kuo, W. J., Dehaene, S. (2015), "Adaptation of the human visual system to the statistics of letters and line configurations", *NeuroImage*, 120, 428-440.
- Changeux, J.-P. (1983), *L'homme neuronal*, París, Fayard [ed. cast.: *El hombre neuronal*, Madrid, Espasa-Calpe, 1986].
- Chao, Z. C., Takaura, K., Wang, L., Fujii, N., Dehaene, S. (2018), "Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain", *Neuron*, 100 (5).
- Chen, Z., Wilson, M. A. (2017), "Deciphering neural codes of memory during sleep", *Trends in Neurosciences*, 40 (5), 260-275.
- Chiao, J. Y. (2010), "Neural basis of social status hierarchy across species", *Current Opinion in Neurobiology*, 20 (6), 803-809.
- Cho, K., Courville, A., Bengio, Y. (2015), "Describing multimedia content using attention-based encoder-decoder networks", *IEEE Transactions on Multimedia*, 17 (11), 1875-1886.
- Chun, M. M., Marois, R. (2002), "The dark side of visual attention", *Current Opinion in Neurobiology*, 12 (2), 184-189.
- Clark, E. V. (1988), "On the logic of contrast", *Journal of Child Language*, 15 (2), 317-335.
- Claro, S., Paunesku, D., Dweck, C. S. (2016), "Growth mindset tempers the effects of poverty on academic achievement", *Proc. Natl Acad. Sci USA*, 113 (31), 8664-8668.
- Cohen, L., Dehaene, S., McCormick, S., Durant, S., Zanker, J. M. (2016), "Brain mechanisms of recovery from pure alexia: A single case study with multiple longitudinal scans", *Neuropsychologia*, 91, 36-49.
- Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., Montavont, A. (2008), "Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways" *NeuroImage*, 40 (1), 353-366.
- Conel, J. L. (1939), *The Postnatal Development of the Human Cerebral Cortex*, Cambridge (Mass.), Harvard University Press, vol. 1-8.

- Constantinescu, A. O., O'Reilly, J. X., Behrens, T. E. J. (2016), "Organizing conceptual knowledge in humans with a gridlike code", *Science*, 352 (6292), 1464-1468.
- Corallo, G., Sackur, J., Dehaene, S., Sigman, M. (2008), "Limits on introspection: Distorted subjective time during the dual-task bottleneck", *Psychological Science*, 19 (11), 1110-1117.
- Cortese, S., Brown, T. E., Corkum, P., Gruber, R., O'Brien, L. M., Stein, M., ..., Owens, J. (2013), "Assessment and management of sleep problems in youths with attention-deficit/hyperactivity disorder", *Journal of the American Academy of Child and Adolescent Psychiatry*, 52 (8), 784-796.
- Costa, A., Sebastián-Gallés, N. (2014), "How does the bilingual experience sculpt the brain?", *Nature Reviews. Neuroscience*, 15 (5), 336-345.
- Courchesne, E., Pierce, K., Schumann, C. M., Redcay, E., Buckwalter, J. A., Kennedy, D. P., Morgan, J. (2007), "Mapping early brain development in autism", *Neuron*, 56 (2), 399-413.
- Courtney, S. M., Ungerleider, L. G., Keil, K., Haxby, J. V. (1997), "Transient and sustained activity in a distributed neural system for human working memory", *Nature*, 386 (6625), 608-611.
- Craik, F. I., Tulving, E. (1975), "Depth of processing and the retention of words in episodic memory", *Journal of Experimental Psychology: General*, 104 (3), 268.
- Csibra, G., Gergely, G. (2009), "Natural pedagogy", *Trends Cogn. Sci.*, 13 (4), 148-153.
- Çukur, T., Nishimoto, S., Huth, A. G., Gallant, J. L. (2013), "Attention during natural vision warps semantic representation across the human brain", *Nature Neuroscience*, 16 (6), 763-770.
- Curran, T., Tucker, D. M., Kutas, M., Posner, M. I. (1993), "Topography of the N400: Brain electrical activity reflecting semantic expectancy", *Electroencephalography and Clinical Neurophysiology*, 88, 188-209.

- Cyr, M., Shi, R. (2013), "Development of abstract grammatical categorization in infants", *Child Development*, 84 (2), 617-629.
- Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., ..., Saxe, R. (2017), "Organization of high-level visual cortex in human infants", *Nature Communications*, 8, 13.995.
- Dehaene, S. (2003), "The neural basis of the Weber-Fechner law: A logarithmic mental number line", *Trends Cogn. Sci.*, 7 (4), 145-147.
- (2005), "Evolution of human cortical circuits for reading and arithmetic: The "neuronal recycling" hypothesis", en S. Dehaene, J. R. Duhamel, M. Hauser, G. Rizzolatti (eds.), *From Monkey Brain to Human Brain*, Cambridge (Mass.), MIT Press, 133-157.
- (2006), *Vers une science de la vie mentale*, París, Collège de France - Fayard [ed. cast.: *En busca de la mente*, Buenos Aires, Siglo XXI, 2018].
- (2007a), *Les neurones de la lecture*, París, Odile Jacob (trad. ing.: *Reading in the brain*, Nueva York, Penguin Viking, 2009 [ed. cast.: *El cerebro lector*, Buenos Aires, Siglo XXI, 2014]).
- (2007b), "Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation", en P. Haggard, Y. Rossetti (eds.), *Attention and Performance XXII. Sensorimotor foundations of higher cognition*, Cambridge (Mass.), Harvard University Press, 527-574.
- (2010), *La bosse des maths. Quinze ans après*, París, Odile Jacob [ed. cast.: *El cerebro matemático*, Buenos Aires, Siglo XXI, 2016].
- (2014), *Le code de la conscience*, París, Odile Jacob [ed. cast.: *La conciencia en el cerebro*, Buenos Aires, Siglo XXI, 2015].
- Dehaene, S., Akhavein, R. (1995), "Attention, automaticity and levels of representation in number processing", *J. Exp. Psychol. Learn. Mem. Cogn.*, 21, 314-326.
- Dehaene, S., Bossini, S., Giraux, P. (1993), "The mental representation of parity and numerical magnitude", *Journal of Experimental Psychology: General*, 122, 371-396.
- Dehaene, S., Changeux, J.-P. (2011), "Experimental and theoretical approaches to conscious processing", *Neuron*, 70 (2), 200-227.

- Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., Sergent, C. (2006), "Conscious, preconscious, and subliminal processing: A testable taxonomy", *Trends Cogn. Sci.*, 10 (5), 204-211.
- Dehaene, S., Cohen, L. (2007), "Cultural recycling of cortical maps", *Neuron*, 56 (2), 384-398.
- Dehaene, S., Cohen, L., Morais, J., Kolinsky, R. (2015), "Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition", *Nature Reviews. Neuroscience*, 16 (4), 234-244.
- Dehaene, S., Cohen, L., Sigman, M., Vinckier, F. (2005), "The neural code for written words: A proposal", *Trends Cogn. Sci.*, 9 (7), 335-341.
- Dehaene, S., Dupoux, E., Mehler, J. (1990), "Is numerical comparison digital: Analogical and symbolic effects in two-digit number comparison", *J. Exp. Psychol. Hum. Percep. Perf.*, 16, 626-641.
- Dehaene, S., Izard, V., Pica, P., Spelke, E. (2006), "Core knowledge of geometry in an Amazonian indigene group", *Science*, 311, 381-384.
- Dehaene, S., Izard, V., Spelke, E., Pica, P. (2008), "Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures", *Science*, 320 (5880), 1217-1220.
- Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., Cohen, L. (2004), "Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence", *Psychological Science*, 15 (5), 307-313.
- Dehaene, S., Kerszberg, M., Changeux, J.-P. (1998), "A neuronal model of a global workspace in effortful cognitive tasks", *Proc. Natl Acad. Sci. USA*, 95 (24), 14.529-14.534.
- Dehaene, S., Lau, H., Kouider, S. (2017), "What is consciousness, and could machines have it?", *Science*, 358 (6362), 486-492.
- Dehaene, S., Marques, J. F. (2002), "Cognitive Euroscience: Scalar variability in price estimation and the cognitive consequences of switching to the Euro", *Quarterly Journal of Experimental Psychology*, 55 (3), 705-731.

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., Pallier, C. (2015), “The neural representation of sequences: From transition probabilities to algebraic patterns and linguistic trees”, *Neuron*, 88 (1), 2-19.

Dehaene, S., Naccache, L. (2001), “Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework”, *Cognition*, 79, 1-37.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., Rivière, D. (2001), “Cerebral mechanisms of word masking and unconscious repetition priming”, *Nature Neuroscience*, 4 (7), 752-758.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., ..., Cohen, L. (2010), “How learning to read changes the cortical networks for vision and language”, *Science*, 330 (6009), 1359-1364.

Dehaene-Lambertz, G., Dehaene, S., Hertz-Pannier, L. (2002), “Functional neuroimaging of speech perception in infants”, *Science*, 298 (5600), 2013-2015.

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., Roche, A., Sigman, M., Dehaene, S. (2006), “Functional organization of perisylvian activation during presentation of sentences in preverbal infants”, *Proc. Natl Acad. Sci. USA*, 103 (38), 14240-14245.

Dehaene-Lambertz, G., Monzalvo, K., Dehaene, S. (2018), “The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition”, *PLoS Biology*, 16 (3).

Dehaene-Lambertz, G., Spelke, E. S. (2015), “The infancy of the human brain”, *Neuron*, 88 (1), 93-109.

Denison, S., Xu, F. (2010), “Integrating physical constraints in statistical inference by 11-month-old infants”, *Cogn. Sci.*, 34 (5), 885-908.

Dennett, D. C. (1996), *Darwin's Dangerous Idea. Evolution and the Meaning of Life*, Nueva York, Simon & Schuster.

- Desimone, R., Duncan, J. (1995), “Neural mechanisms of selective visual attention”, *Annu. Rev. Neurosci.*, 18, 193-222.
- D’Esposito, M., Grossman, M. (1996), “The physiological basis of executive function and working memory”, *The Neuroscientist*, 2, 345-352.
- Diamond, A., Doar, B. (1989), “The performance of human infants on a measure of frontal cortex function the delayed response task”, *Developmental Psychobiology*, 22.
- Diamond, A., Goldman-Rakic, P. S. (1989), “Comparison of human infants and rhesus monkeys on Piaget’s A-not-B task: Evidence for dependence on dorsolateral prefrontal cortex”, *Experimental Brain Research*, 74, 24-40.
- Diamond, A., Lee, K. (2011), “Interventions shown to aid executive function development in children 4 to 12 years old”, *Science*, 333 (6045), 959-964.
- Diekelmann, S., Born, J. (2010), “The memory function of sleep”, *Nature Reviews. Neuroscience*, 11 (2), 114-126.
- Diester, I., Nieder, A. (2007), “Semantic associations between signs and numerical categories in the prefrontal cortex”, *PLoS Biology*, 5 (11).
— (2010), “Numerical values leave a semantic imprint on associated signs in monkeys”, *J. Cogn. Neurosci.*, 22 (1), 174-183.
- Ditz, H. M., Nieder, A. (2015), “Neurons selective to the number of visual items in the corvid songbird endbrain”, *Proc. Natl Acad. Sci. USA*, 112 (25), 7827-7832.
- Doeller, C. F., Barry, C., Burgess, N. (2010), “Evidence for grid cells in a human memory network”, *Nature*, 463 (7281), 657-661.
- Donato, F., Rompani, S. B., Caroni, P. (2013), “Parvalbumin-expressing basket-cell network plasticity induced by experience regulates adult learning”, *Nature*, 504 (7479), 272-276.
- Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., May, A. (2004), “Neuroplasticity: Changes in grey matter induced by

training”, *Nature*, 427 (6972), 311-312.

Dubois, J., Dehaene-Lambertz, G., Perrin, M., Mangin, J.-F., Cointepas, Y., Duchesnay, E., ..., Hertz-Pannier, L. (2007), “Asynchrony of the early maturation of white matter bundles in healthy infants: Quantitative landmarks revealed noninvasively by diffusion tensor imaging”, *Human Brain Mapping*.

Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J.-F., Le Bihan, D., Dehaene-Lambertz, G. (2009), “Structural asymmetries in the infant language and sensori-motor networks”, *Cereb. Cortex*, 19 (2), 414-423.

Dubois, J., Poupon, C., Thirion, B., Simonnet, H., Kulikova, S., Leroy, F., Hertz-Pannier, L., Dehaene-Lambertz, G. (2015), “Exploring the early organization and maturation of linguistic pathways in the human infant brain”, *Cereb. Cortex*, 26 (5), 2283-2298.

Dumontheil, I., Klingberg, T. (2011), “Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later”, *Cereb. Cortex*, 22 (5), 1078-1085.

Duncan, J. (2003), “Intelligence tests predict brain response to demanding task events”, *Nature Neuroscience*, 6 (3), 207-208.

— (2010), “The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour”, *Trends Cogn. Sci.*, 14 (4), 172-179.

— (2013), “The structure of cognition: Attentional episodes in mind and brain”, *Neuron*, 80 (1), 35-50.

Dundas, E. M., Plaut, D. C., Behrmann, M. (2013), “The joint development of hemispheric lateralization for words and faces”, *Journal of Experimental Psychology: General*, 142 (2), 348-358.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., Willingham, D. T. (2013), “Improving students’ learning with effective learning techniques: Promising directions from cognitive and educational psychology”, *Psychological Science in the Public Interest. A Journal of the American Psychological Society*, 14 (1), 4-58.

Duyme, M., Dumaret, A. C., Tomkiewicz, S. (1999), “How can we boost IQs of ‘dull children’?: A late adoption study”, *Proc. Natl Acad. Sci.*

USA, 96 (15), 8790-8794.

Dweck, C. S. (2006), *Mindset: The New Psychology of Success*, Nueva York, Random House.

Egyed, K., Király, I., Gergely, G. (2013), “Communicating shared knowledge in infancy”, *Psychological Science*, 24 (7), 1348-1353.

Ehri, L. C., Nunes, S. R., Stahl, S. A., Willows, D. M. M. (2001), “Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel’s meta-analysis”, *Review of Educational Research*, 71, 393-447.

Ellis, A. W., Ralph, L., Matthew, A. (2000), “Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks”, *J. Exp. Psychol. Learn. Mem. Cogn.*, 26 (5), 1103.

Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., Plunkett, K. (1996), *Rethinking Innateness: A Connectionist Perspective on Development*, Cambridge (Mass.), MIT Press.

Elsayed, G. F., Shankar, S., Cheung, B., Papernot, N., Kurakin, A., Goodfellow, I., Sohl-Dickstein, J. (2018), “Adversarial examples that fool both human and computer vision”, arxiv.org/abs/180208195.

Elston, G. N. (2003), “Cortex, cognition and the cell: New insights into the pyramidal neuron and prefrontal function”, *Cereb. Cortex*, 13 (11), 1124-1138.

Epelbaum, M., Milleret, C., Buisseret, P., Duffer, J. L. (1993), “The sensitive period for strabismic amblyopia in humans”, *Ophthalmology*, 100 (3), 323-327.

Esseily, R., Rat-Fischer, L., Somogyi, E., O’Regan, K. J., Fagard, J. (2016), “Humour production may enhance observational learning of a new tool-use action in 18-month-old infants”, *Cognition and Emotion*, 30 (4), 817-825.

Ester, E. F., Sprague, T. C., Serences, J. T. (2015), “Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory”, *Neuron*, 87 (4), 893-905.

- Everaert, M. B. H., Huybregts, M. A. C., Chomsky, N., Berwick, R. C., Bolhuis, J. J. (2015), "Structures, not strings: Linguistics as part of the cognitive sciences", *Trends Cogn. Sci.*, 19 (12), 729-743.
- Fanselow, M. S. (1998), "Pavlovian conditioning, negative feedback, and blocking: Mechanisms that regulate association formation", *Neuron*, 20 (4), 625-627.
- Fattal, I., Friedmann, N., Fattal-Valevski, A. (2011), "The crucial role of thiamine in the development of syntax and lexical retrieval: A study of infantile thiamine deficiency", *Brain*, 134 (Pt 6), 1720-1739.
- Fawcett, S. L., Wang, Y.-Z., Birch, E. E. (2005), "The critical period for susceptibility of human stereopsis", *Investigative Ophthalmology and Visual Science*, 46 (2), 521-525.
- Fischer, M. H. (2003), "Cognitive representation of negative numbers", *Psychological Science*, 14 (3), 278-282.
- Fisher, A. V., Godwin, K. E., Seltman, H. (2014), "Visual environment, attention allocation, and learning in young children when too much of a good thing may be bad", *Psychological Science*, 25 (7), 1362-1370.
- Fitzgerald, J. K., Freedman, D. J., Fanini, A., Bennur, S., Gold, J. I., Assad, J. A. (2013), "Biased associative representations in parietal cortex", *Neuron*, 77 (1), 180-191.
- Fitzsimonds, R. M., Song, H. J., Poo, M. M. (1997), "Propagation of activity-dependent synaptic depression in simple neural networks", *Nature*, 388 (6641), 439-448.
- Flechsig, P. (1876), *Die Leitungsbahnen im Gehirn und Rückenmark des Menschen auf Grund entwicklungsgeschichtlicher Untersuchungen dargestellt*, Léipzig, Engelmann.
- Flege, J. E., Munro, M. J., MacKay, I. R. (1995), "Factors affecting strength of perceived foreign accent in a second language", *The Journal of the Acoustical Society of America*, 97 (5), 3125-3134.
- Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., Rees, G. (2010), "Relating introspective accuracy to individual differences in brain

structure”, *Science*, 329 (5998), 1541-1543.

Fodor, J. A. (1975), *The Language of Thought*, Nueva York, Thomas Y. Crowell.

Fodor, J. A., Pylyshyn, Z. W. (1988), “Connectionism and cognitive architecture: A critical analysis”, *Cognition*, 28, 3-71.

Fodor, J., McLaughlin, B. P. (1990), “Connectionism and the problem of systematicity: Why Smolensky’s solution doesn’t work”, *Cognition*, 35 (2), 183-204.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., Wenderoth, M. P. (2014), “Active learning increases student performance in science, engineering, and mathematics”, *Proc. Natl Acad. Sci. USA*, 111 (23), 8410-8415.

Friederici, A. D. (2002), “Towards a neural basis of auditory sentence processing”, *Trends Cogn. Sci.*, 6 (2), 78-84.

Friedmann, N., Kerbel, N., Shvimer, L. (2010), “Developmental attentional dyslexia”, *Cortex: a Journal Devoted to the Study of the Nervous System and Behavior*, 46 (10), 1216-1237.

Friedmann, N., Rusou, D. (2015), “Critical period for first language: The crucial role of language input during the first year of life”, *Current Opinion in Neurobiology*, 35, 27-34.

Friedrich, M., Wilhelm, I., Born, J., Friederici, A. D. (2015), “Generalization of word meanings during infant sleep”, *Nature Communications*, 6, 6004.

Friston, K. (2005), “A theory of cortical responses”, *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 360 (1456), 815-836.

Froemke, R. C., Merzenich, M. M., Schreiner, C. E. (2007), “A synaptic memory trace for cortical receptive field plasticity”, *Nature*, 450 (7168), 425-429.

Fukuchi-Shimogori, T., Grove, E. A. (2001), “Neocortex patterning by the secreted signaling molecule FGF8”, *Science*, 294 (5544), 1071-1074.

- Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., Moser, M.-B. (2004), "Spatial representation in the entorhinal cortex", *Science*, 305 (5688), 1258-1264.
- Galaburda, A. M., LoTurco, J., Ramus, F., Fitch, R. H., Rosen, G. D. (2006), "From genes to behavior in developmental dyslexia", *Nature Neuroscience*, 9 (10), 1213-1217.
- Galgali, A. R., Mante, V. (2018), "Set in one's thoughts", *Nature Neuroscience*, 21 (4), 459-460.
- Gallistel, C. R. (1990), *The Organization of Learning*, Cambridge (Mass.), MIT Press.
- Gaser, C., Schlaug, G. (2003), "Brain structures differ between musicians and non-musicians", *J. Neurosci.*, 23 (27), 9240-9245.
- Gathercole, S. E., Pickering, S. J., Knight, C., Stegmann, Z. (2004), "Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age", *Applied Cognitive Psychology*, 18 (1), 1-16.
- Geary, D. C. (2011), "Cognitive predictors of achievement growth in mathematics: A five year longitudinal study", *Developmental Psychology*, 47 (6), 1539-1552.
- Genzel, L., Rossato, J. I., Jacobse, J., Grieves, R. M., Spooner, P. A., Battaglia, F. P., Fernández, G., Morris, R. G. M. (2017), "The yin and yang of memory consolidation: Hippocampal and neocortical", *PLoS Biology*, 15 (1).
- George, D., Lehrach, W., Kansky, K., Lázaro-Gredilla, M., Laan, C., Marthi, B., ..., Phoenix, D. S. (2017), "A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs", *Science*, 358 (6368).
- Gerber, P., Schlaffke, L., Heba, S., Greenlee, M. W., Schultz, T., Schmidt-Wilcke, T. (2014), "Juggling revisited A voxel-based morphometry study with expert jugglers", *NeuroImage*, 95, 320-325.
- Gergely, G., Bekkering, H., Kiraly, I. (2002), "Rational imitation in

preverbal infants”, *Nature*, 415 (6873), 755.

Gergely, G., Csibra, G. (2003), “Teleological reasoning in infancy: The naïve theory of rational action”, *Trends Cogn. Sci.*, 7 (7), 287-292.

Gerhand, S., Barry, C. (1999), “Age of acquisition, word frequency, and the role of phonology in the lexical decision task”, *Memory and Cognition*, 27 (4), 592-602.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., ..., Grantham-McGregor, S. (2014), “Labor market returns to an early childhood stimulation intervention in Jamaica”, *Science*, 344 (6187), 998-1001.

Glass, A. L., Kang, M. (2018), “Dividing attention in the classroom reduces exam performance”, *Educational Psychology*, publicado online.

Goldman-Rakic, P. S. (1995), “Cellular basis of working memory”, *Neuron*, 14 (3), 477-485.

Golestani, N., Molko, N., Dehaene, S., Le Bihan, D., Pallier, C. (2007), “Brain structure predicts the learning of foreign speech sounds”, *Cereb. Cortex*, 17 (3), 575-582.

Golub, M. D., Sadtler, P. T., Oby, E. R., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., ..., Yu, B. M. (2018), “Learning by neural reassociation”, *Nature Neuroscience*, 21 (4), 607-616.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ..., Bengio, Y. (2014), “Generative adversarial networks”, [arxiv.org/abs/14062661](https://arxiv.org/abs/1406.2661).

Goodman, C. S., Shatz, C. J. (1993), “Developmental mechanisms that generate precise patterns of neuronal connectivity”, *Neuron*, 10, 77-98.

Goodman, N. D., Ullman, T. D., Tenenbaum, J. B. (2011), “Learning a theory of causality”, *Psychological Review*, 118 (1), 110-119.

Gopnik, A. (2000), *The Scientist in the Crib: What Early Learning Tells Us About the Mind*, Nueva York, William Morrow.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., Danks, D. (2004), “A theory of causal learning in children: Causal maps

and Bayes nets”, *Psychological Review*, 111 (1), 3-32.

Gottlieb, J., Oudeyer, P.-Y., Lopes, M., Baranes, A. (2013), “Information-seeking, curiosity, and attention: Computational and neural mechanisms”, *Trends Cogn. Sci.*, 17 (11), 585-593.

Goupil, L., Romand-Monnier, M., Kouider, S. (2016), “Infants ask for help when they know they don’t know”, *Proc. Natl Acad. Sci. USA*, 113 (13), 3492-3496.

Grainger, J., Whitney, C. (2004), “Does the huamn mnid raed wrods as a wlohe?” *Trends Cogn. Sci.*, 8 (2), 58-59.

Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., Himes, J. H. (1991), “Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: The Jamaican Study”, *The Lancet*, 338 (8758), 1-5.

Green, C. S., Bavelier, D. (2003), “Action video game modifies visual selective attention”, *Nature*, 423 (6939), 534-537.

Groen, G. J., Parkman, J. M. (1972), “A chronometric analysis of simple addition”, *Psychological Review*, 79, 329-343.

Grothendieck, A. (1986), *Récoltes et semailles*, www.quarante-deux.org/archives/klein/prefaces/Romans_1965-1969/Recoltes_et_semailles.pdf.

Gruber, M. J., Gelman, B. D., Ranganath, C. (2014), “States of curiosity modulate hippocampus: Dependent learning via the dopaminergic circuit”, *Neuron*, 84 (2), 486-496.

Guerguiev, J., Lillicrap, T. P., Richards, B. A. (2017), “Towards deep learning with segregated dendrites”, *ELife*, 6, e22901.

Gullick, M. M., Wolford, G. (2013), “Understanding less than nothing: Children’s neural response to negative numbers shifts across age and accuracy”, *Frontiers in Psychology*, 4, 584.

Gweon, H., Tenenbaum, J. B., Schulz, L. E. (2010), “Infants consider both the sample and the sampling process in inductive generalization”, *Proc. Natl Acad. Sci. USA*, 107 (20), 9066-9071.

- Habibi, A., Damasio, A., Ilari, B., Elliott Sachs, M., Damasio, H. (2018), "Music training and child development: A review of recent findings from a longitudinal study", *Annals of the New York Academy of Sciences*.
- Hafting, T., Fyhn, M., Molden, S., Moser, M. B., Moser, E. I. (2005), "Microstructure of a spatial map in the entorhinal cortex", *Nature*, 436 (7052), 801-806.
- Hahne, A., Friederici, A. D. (1999), "Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes", *J. Cogn. Neurosci.*, 11 (2), 194-205.
- Halberda, J., Feigenson, L. (2008), "Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults", *Dev. Psychol.*, 44 (5), 1457-1465.
- Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G., Dehaene, S. (2015), "Origins of the specialization for letters and numbers in ventral occipitotemporal cortex", *Trends Cogn. Sci.*, 19 (7), 374-382.
- Hannagan, T., Nieder, A., Viswanathan, P., Dehaene, S. (2017), "A random-matrix theory of the number sense", *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 373 (1740).
- Hartshorne, J. K., Tenenbaum, J. B., Pinker, S. (2018), "A critical period for second language acquisition: Evidence from 2/3 million English speakers", *Cognition*, 177, 263-277.
- Hassabis, D., Kumaran, D., Summerfield, C., Botvinick, M. (2017), "Neuroscience-inspired artificial intelligence", *Neuron*, 95 (2), 245-258.
- Hattie, J. (2008), *Visible Learning*, Londres-Nueva York, Routledge.
- Hattie, J., Brodeur, M., St-Cyr, C. (2017), *L'apprentissage visible pour les enseignants. Connaître son impact pour maximiser le rendement des élèves*, Quebec, Presses de l'Université du Québec.
- Hauser, M. D., Chomsky, N., Fitch, W. T. (2002), "The faculty of language: What is it, who has it, and how did it evolve?", *Science*, 298 (5598), 1569-1579.

- Hauser, M. D., Watumull, J. (2017), "The Universal Generative Faculty: The source of our expressive power in language, mathematics, morality, and music", *Journal of Neurolinguistics*, 43 (B), 78-94.
- Hay, J. F., Pelucchi, B., Graf Estes, K., Saffran, J. R. (2011), "Linking sounds to meanings: Infant statistical learning in a natural language", *Cognitive Psychology*, 63 (2), 93-106.
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., Yavitz, A. (2010), "The rate of return to the HighScope Perry Preschool Program", *Journal of Public Economics*, 94 (1), 114-128.
- Held, R., Hein, A. (1963), "Movement-produced stimulation in the development of visually guided behavior", *Journal of Comparative and Physiological Psychology*, 56, 872-876.
- Hensch, T. K. (2005), "Critical period plasticity in local cortical circuits", *Nature Reviews. Neuroscience*, 6 (11), 877.
- Hespos, S. J., Baillargeon, R. (2008), "Young infants' actions reveal their developing knowledge of support variables: Converging evidence for violation-of-expectation findings", *Cognition*, 107 (1), 304-316.
- Hinton, G. E., Dayan, P., Frey, B. J., Neal, R. M. (1995), "The 'wake-sleep' algorithm for unsupervised neural networks", *Science*, 268 (5214), 1158-1161.
- Hinton, G. E., Osindero, S., Teh, Y. W. (2006), "A fast learning algorithm for deep belief nets", *Neural Computation*, 18 (7), 1527-1554.
- Hiscock, H., Sciberras, E., Mensah, F., Gerner, B., Efron, D., Khano, S., Oberklaid, F. (2015), "Impact of a behavioural sleep intervention on symptoms and sleep in children with attention deficit hyperactivity disorder, and parental mental health: Randomised controlled trial", *BMJ (Clinical Research Ed.)*, 350, h68.
- Holtmaat, A., Caroni, P. (2016), "Functional and structural underpinnings of neuronal assembly formation in learning", *Nature Neuroscience*, 19 (12), 1553-1562.
- Horikawa, T., Tamaki, M., Miyawaki, Y., Kamitani, Y. (2013), "Neural

decoding of visual imagery during sleep”, *Science*, 340 (6132), 639-642.

Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B., Tzourio-Mazoyer, N. (2000), “Shifting from the perceptual brain to the logical brain: The neural impact of cognitive inhibition training”, *J. Cogn. Neurosci.*, 12 (5), 721-728.

Huber, R., Ghilardi, M. F., Massimini, M., Tononi, G. (2004), “Local sleep and learning”, *Nature*, 430 (6995), 78-81.

Hurley, M. M., Dennett, D. C., Adams, R. B. (2011), *Inside Jokes. Using Humor to Reverse-Engineer the Mind*, Cambridge (Mass.), MIT Press.

Huttenlocher, P. R., Dabholkar, A. S. (1997), “Regional differences in synaptogenesis in human cerebral cortex”, *J. Comp. Neurol.*, 387 (2), 167-178.

Hutton, J. S., Horowitz-Kraus, T., Mendelsohn, A. L., DeWitt, T., Holland, S. K., C-MIND Authorship Consortium (2015), “Home reading environment and brain activation in preschool children listening to stories”, *Pediatrics*, 136 (3), 466-478.

Hutton, J. S., Phelan, K., Horowitz-Kraus, T., Dudley, J., Altaye, M., DeWitt, T., Holland, S. K. (2017), “Shared reading quality and brain activation during story listening in preschool-age children”, *Journal of Pediatrics*, 191, 204-211.e1.

Iriki, A. (2005), “A prototype of Homo-Faber: A silent precursor of human intelligence in the tool-using monkey brain”, en S. Dehaene, J. R. Duhamel, M. Hauser, G. Rizzolatti (eds.), *From Monkey Brain to Human Brain*, Cambridge (Mass.), MIT Press, 133-157.

Isingrini, M., Perrotin, A., Souchay, C. (2008), “Aging, metamemory regulation and executive functioning”, *Prog. Brain Res.*, 169, 377-392.

Izard, V., Dehaene-Lambertz, G., Dehaene, S. (2008), “Distinct cerebral pathways for object identity and number in human infants”, *PLoS Biology*, 6 (2), 275-285.

Izard, V., Sann, C., Spelke, E. S., Streri, A. (2009), “Newborn infants perceive abstract numbers”, *Proc. Natl Acad. Sci. USA*, 106 (25),

10.382-10.385.

Jacob, S. N., Nieder, A. (2009), "Notation-independent representation of fractions in the human parietal cortex", *J. Neurosci.*, 29 (14), 4652-4657.

Jacoby, L. L., Dallas, M. (1981), "On the relationship between autobiographical memory and perceptual learning", *Journal of Experimental Psychology: General*, 110 (3), 306-340.

Jaeggi, S. M., Buschkuhl, M., Jonides, J., Shah, P. (2011), "Shortand longterm benefits of cognitive training", *Proc. Natl Acad. Sci. USA*, 108 (25), 10.081-10.086.

James, C. E., Oechslin, M. S., Van De Ville, D., Hauert, C.-A., Descloux, C., Lazeyras, F. (2014), "Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks", *Brain Structure and Function*, 219 (1), 353-366.

Jaynes, E. T. (2003), *Probability Theory: The Logic of Science*, Cambridge, Cambridge University Press.

Jenkins, J. G., Dallenbach, K. M. (1924), "Obliviscence during sleep and waking", *The American Journal of Psychology*, 35 (4), 605-612.

Ji, D., Wilson, M. A. (2007), "Coordinated memory replay in the visual cortex and hippocampus during sleep", *Nature Neuroscience*, 10 (1), 100-107.

Jiang, X., Long, T., Cao, W., Li, J., Dehaene, S., Wang, L. (2018), "Production of supra-regular spatial sequences by macaque monkeys", *Current Biology*, 28 (12), 1851-1859.

Jiang, X., Shamie, I., K. Doyle, W., Friedman, D., Dugan, P., Devinsky, O., ..., Halgren, E. (2017), "Replay of large-scale spatio-temporal patterns from waking during subsequent NREM sleep in human cortex", *Scientific Reports*, 7 (1), 17.380.

Jo, J., Bengio, Y. (2017), "Measuring the tendency of CNNs to Learn surface statistical regularities", arxiv.org/abs/171111561.

Johansson, F., Jirenhed, D. A., Rasmussen, A., Zucca, R., Hesselow, G.

(2014), “Memory trace and timing mechanism localized to cerebellar Purkinje cells”, *Proc. Natl Acad. Sci. USA*, 111 (41), 14.930-14.934.

Johnson, J. S., Newport, E. L. (1989), “Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language”, *Cognitive Psychology*, 21 (1), 60-99.

Josselyn, S. A., Köhler, S., Frankland, P. W. (2015), “Finding the engram”, *Nature Reviews. Neuroscience*, 16 (9), 521-534.

Kaminski, J., Call, J., Fischer, J. (2004), “Word learning in a domestic dog: Evidence for ‘fast mapping’”, *Science*, 304 (5677), 1682-1683.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., Camerer, C. F. (2009), “The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory”, *Psychological Science*, 20 (8), 963-973.

Kang, S. H. K., Lindsey, R. V., Mozer, M. C., Pashler, H. (2014), “Retrieval practice over the long term: should spacing be expanding or equalinterval?”, *Psychonomic Bulletin and Review*, 21 (6), 1544-1550.

Kanjlia, S., Lane, C., Feigenson, L., Bedny, M. (2016), “Absence of visual experience modifies the neural basis of numerical thinking”, *Proc. Natl Acad. Sci. USA*, 113 (40), 11.172-11.177.

Kano, T., Brockie, P. J., Sassa, T., Fujimoto, H., Kawahara, Y., Iino, Y., ..., Maricq, A. V. (2008), “Memory in *Caenorhabditis elegans* is mediated by NMDA-type ionotropic glutamate receptors”, *Current Biology*, 18 (13), 1010-1015.

Kaplan, F., Oudeyer, P.-Y. (2007), “In search of the neural circuits of intrinsic motivation”, *Frontiers in Neuroscience*, 1 (1), 225.

Kapur, S., Craik, F. I., Tulving, E., Wilson, A. A., Houle, S., Brown, G. M. (1994), “Neuroanatomical correlates of encoding in episodic memory: Levels of processing effect”, *Proc. Natl Acad. Sci. USA*, 91 (6), 2008-2011.

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J., Sagi, D. (1994), “Dependence on REM sleep of overnight improvement of a perceptual skill”, *Science*, 265 (5172), 679-682.

- Karpicke, J. D., Roediger, H. L. (2008), "The critical importance of retrieval for learning", *Science*, 319 (5865), 966-968.
- Kastner, S., Ungerleider, L. G. (2000), "Mechanisms of visual attention in the human cortex", *Annu. Rev. Neurosci.*, 23, 315-341.
- Keller, H. (1903), *The Story of My Life*, Nueva York, Simon & Schuster [ed. cast.: *La historia de mi vida*, México, Editores Asociados, 1984].
- Kellman, P. J., Spelke, E. S. (1983), "Perception of partly occluded objects in infancy", *Cognitive Psychology*, 15, 483-524.
- Kemp, C., Tenenbaum, J. B. (2008), "The discovery of structural form", *Proc. Natl Acad. Sci. USA*, 105 (31), 10.687-10.692.
- Kerkoerle, T. van, Self, M. W., Roelfsema, P. R. (2017), "Layer-specificity in the effects of attention and working memory on activity in primary visual cortex", *Nature Communications*, 8, 13.804.
- Kidd, C., Piantadosi, S. T., Aslin, R. N. (2012), "The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex", *PLoS ONE*, 7 (5).
- (2014), "The Goldilocks effect in infant auditory attention", *Child Development*, 85 (5), 1795-1804.
- Kilgard, M. P., Merzenich, M. M. (1998), "Cortical map reorganization enabled by nucleus basalis activity", *Science*, 279 (5357), 1714-1718.
- Kim, J. J., Diamond, D. M. (2002), "The stressed hippocampus, synaptic plasticity and lost memories", *Nature Reviews. Neuroscience*, 3 (6), 453-462.
- Kim, W. B., Cho, J. H. (2017), "Encoding of discriminative fear memory by input-specific LTP in the amygdala", *Neuron*, 95 (5), 1129-1146.e5.
- Kirschner, P. A., Sweller, J., Clark, R. E. (2006), "Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching", *Educational Psychologist*, 41 (2), 75-86.
- Kirschner, P. A., Van Merriënboer, J. J. G. (2013), "Do learners really know best? Urban legends in education", *Educational Psychologist*, 48

(3), 169-183.

Kitamura, T., Ogawa, S. K., Roy, D. S., Okuyama, T., Morrissey, M. D., Smith, L. M., ..., Tonegawa, S. (2017), "Engrams and circuits crucial for systems consolidation of a memory", *Science*, 356 (6333), 73-78.

Klingberg, T. (2010), "Training and plasticity of working memory", *Trends Cogn. Sci.*, 14 (7), 317-324.

Knops, A., Thirion, B., Hubbard, E. M., Michel, V., Dehaene, S. (2009), "Recruitment of an area involved in eye movements during mental arithmetic", *Science*, 324 (5934), 1583-1585.

Knops, A., Viarouge, A., Dehaene, S. (2009), "Dynamic representations underlying symbolic and nonsymbolic calculation: Evidence from the operational momentum effect", *Atten. Percept. Psychophys.*, 71 (4), 803-821.

Knudsen, E. I., Knudsen, P. F. (1990), "Sensitive and critical periods for visual calibration of sound localization by barn owls", *J. Neurosci.*, 10 (1), 222-232.

Knudsen, E. I., Zheng, W., DeBello, W. M. (2000), "Traces of learning in the auditory localization pathway", *Proc. Natl Acad. Sci. USA*, 97 (22), 11.815-11.820.

Koechlin, E., Dehaene, S., Mehler, J. (1997), "Numerical transformations in five month old human infants", *Mathematical Cognition*, 3, 89-104.

Koechlin, E., Ody, C., Kouneiher, F. (2003), "The architecture of cognitive control in the human prefrontal cortex", *Science*, 302 (5648), 1181-1185.

Koepp, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., ..., Grasby, P. M. (1998), "Evidence for striatal dopamine release during a video game", *Nature*, 393, 266-268.

Kolinsky, R., Morais, J., Content, A., Cary, L. (1987), "Finding parts within figures: A developmental study", *Perception*, 16 (3), 399-407.

Kolinsky, R., Verhaeghe, A., Fernandes, T., Mengarda, E. J., Grimm-

- Cabral, L., Morais, J. (2011), "Enantiomorphy through the Looking-Glass: Literacy effects on mirror-image discrimination", *Journal of Experimental Psychology: General*, 140 (2), 210-238.
- Kontra, C., Goldin-Meadow, S., Beilock, S. L. (2012), "Embodied learning across the life span", *Topics in Cognitive Science*, 4 (4), 731-739.
- Kontra, C., Lyons, D. J., Fischer, S. M., Beilock, S. L. (2015), "Physical experience enhances science learning", *Psychological Science*, 26 (6), 737-749.
- Kouider, S., Stahlhut, C., Gelskov, S. V., Barbosa, L. S., Dutat, M., Gardelle, V. de, Christophe, A., Dehaene, S., Dehaene-Lambertz, G. (2013), "A neural marker of perceptual consciousness in infants", *Science*, 340 (6130), 376-380.
- Krause, M. R., Zanos, T. P., Csorba, B. A., Pilly, P. K., Choe, J., Phillips, M. E., ..., Pack, C. C. (2017), "Transcranial direct current stimulation facilitates associative learning and alters functional connectivity in the primate brain", *Current Biology*, 27 (20), 3086-3096.
- Kropff, E., Treves, A. (2008), "The emergence of grid cells: Intelligent design or just adaptation?", *Hippocampus*, 18 (12), 1256-1269.
- Krubitzer, L. (2007), "The magnificent compromise: Cortical field evolution in mammals", *Neuron*, 56 (2), 201-208.
- Kuhl, P. K., Tsao, F. M., Liu, H. M. (2003), "Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning", *Proc. Natl Acad. Sci. USA*, 100 (15), 9096-9101.
- Kurdziel, L., Duclos, K., Spencer, R. M. C. (2013), "Sleep spindles in midday naps enhance learning in preschool children", *Proc. Natl Acad. Sci. USA*, 110 (43), 17.267-17.272.
- Kushnir, T., Xu, F., Wellman, H. M. (2010), "Young children use statistical sampling to infer the preferences of other people", *Psychological Science*, 21 (8), 1134-1140.
- Kutas, M., Federmeier, K. D. (2011), "Thirty years and counting:

- Finding meaning in the N400 component of the event-related brain potential (ERP)", *Annu. Rev. Psychol.*, 62, 621-647.
- Kutas, M., Hillyard, S. A. (1980), "Reading senseless sentences: Brain potentials reflect semantic incongruity", *Science*, 207 (4427), 203-205.
- Kutter, E. F., Bostroem, J., Elger, C. E., Mormann, F., Nieder, A. (2018), "Single Neurons in the Human Brain Encode Numbers", *Neuron.*, 100 (3), 753-761.
- Kwan, K. Y., Lam, M. M. S., Johnson, M. B., Dube, U., Shim, S., Rašin, M. R., ..., Sestan, N. (2012), "Species-dependent posttranscriptional regulation of NOS1 by FMRP in the developing cerebral cortex", *Cell*, 149 (4), 899-911.
- Lake, B. M., Salakhutdinov, R., Tenenbaum, J. B. (2015), "Human-level concept learning through probabilistic program induction", *Science*, 350 (6266), 1332-1338.
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., Gershman, S. J. (2016), "Building machines that learn and think like people", *Behav. Brain Sci.*, 1-101.
- (2017), "Building machines that learn and think like people", *Behav. Brain Sci.*, 40, e253.
- Landau, B., Gleitman, H., Spelke, E. (1981), "Spatial knowledge and geometric representation in a child blind from birth", *Science*, 213 (4513), 1275-1278.
- Lane, C., Kanjlia, S., Omaki, A., Bedny, M. (2015), "'Visual' cortex of congenitally blind adults responds to syntactic movement", *J. Neurosci.*, 35 (37), 12.859-12.868.
- Langston, R. F., Ainge, J. A., Couey, J. J., Canto, C. B., Bjerknes, T. L., Witter, M. P., ..., Moser, M. B. (2010), "Development of the spatial representation system in the rat", *Science*, 328 (5985), 1576-1580.
- LeCun, Y., Bengio, Y., Hinton, G. (2015), "Deep learning", *Nature*, 521 (7553), 436-444.
- LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998), "Gradient-based learning applied to document recognition", *Proceedings of the IEEE*, 86

(11), 2278-2324.

Lefèvre, J., Mangin, J.-F. (2010), “A reaction-diffusion model of human brain development”, *PLoS Comput. Biology*, 6 (4).

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., Niv, Y. (2017), “Dynamic interaction between reinforcement learning and attention in multidimensional environments”, *Neuron*, 93 (2), 451-463.

Leppanen, P. H., Richardson, U., Pihko, E., Eklund, K. M., Guttorm, T. K., Aro, M., Lyytinen, H. (2002), “Brain responses to changes in speech sound durations differ between infants with and without familial risk for dyslexia”, *Dev. Neuropsychol.*, 22 (1), 407-422.

Lerner, Y., Honey, C. J., Silbert, L. J., Hasson, U. (2011), “Topographic mapping of a hierarchy of temporal receptive windows using a narrated story”, *J. Neurosci.*, 31 (8), 2906-2915.

Leroy, F., Cai, Q., Bogart, S. L., Dubois, J., Coulon, O., Monzalvo, K., ..., Dehaene-Lambertz, G. (2015), “New human-specific brain landmark: The depth asymmetry of superior temporal sulcus”, *Proc. Natl Acad. Sci. USA*, 112 (4), 1208-1213.

Li, P., Legault, J., Litcofsky, K. A. (2014), “Neuroplasticity as a function of second language learning: Anatomical changes in the human brain”, *Cortex*, 58, 301-324.

Li, S., Lee, K., Zhao, J., Yang, Z., He, S., Weng, X. (2013), “Neural competition as a developmental process: Early hemispheric specialization for word processing delays specialization for face processing”, *Neuropsychologia*, 51 (5), 950-959.

Lillard, A., Else-Quest, N. (2006), “The early years. Evaluating Montessori education”, *Science*, 313 (5795), 1893-1894.

Lisman, J., Buzsáki, G., Eichenbaum, H., Nadel, L., Ranganath, C., Redish, A. D. (2017), “Viewpoints: How the hippocampus contributes to memory, navigation and cognition”, *Nature Neuroscience*, 20, 1434-1447.

Liu, S., Ullman, T. D., Tenenbaum, J. B., Spelke, E. S. (2017), “Ten-month-old infants infer the value of goals from the costs of actions”,

Science, 358 (6366), 1038-1041.

Livingstone, M. S., Vincent, J. L., Arcaro, M. J., Srihasam, K., Schade, P. F., Savage, T. (2017), "Development of the macaque face-patch system", *Nature Communications*, 8, 14.897.

Loewenstein, G. (1994), "The psychology of curiosity: A review and reinterpretation", *Psychological Bulletin*, 116 (1), 75-98.

Lømo, T. (2018), "Discovering long-term potentiation (LTP) Recollections and reflections on what came after", *Acta Physiologica* (Óxford), 222 (2).

Louie, K., Wilson, M. A. (2001), "Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep", *Neuron*, 29 (1), 145-156.

Lyons, I. M., Beilock, S. L. (2012), "When math hurts: Math anxiety predicts pain network activation in anticipation of doing math", *PLoS ONE*, 7 (10).

Lyons, K. E., Ghetti, S. (2011), "The development of uncertainty monitoring in early childhood", *Child Development*, 82 (6), 1778-1787.

Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T., Kulju, P., Laakso, M. L., ..., Viholainen, H. (2004), "Early development of children at familial risk for dyslexia-follow-up from birth to school age", *Dyslexia*, 10 (3), 146-178.

Ma, L., Xu, F. (2013), "Preverbal infants infer intentional agents from the perception of regularity", *Developmental Psychology*, 49 (7), 1330-1337.

Mack, A., Rock, I. (1998), *Inattentional Blindness*, Cambridge (Mass.), MIT Press.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., Frith, C. D. (2000), "Navigation-related structural change in the hippocampi of taxi drivers", *Proc. Natl Acad. Sci. USA*, 97 (8), 4398-4403.

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R.

- S., Burgess, N. (2003), "Navigation expertise and the human hippocampus: A structural brain imaging analysis", *Hippocampus*, 13 (2), 250-259.
- Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Godjil, S., Dubois, J., ..., Wallois, F. (2013), "Syllabic discrimination in premature human infants prior to complete formation of cortical layers", *Proc. Natl Acad. Sci. USA*, 110 (12), 4846-4851.
- Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., Caramazza, A. (2009), "Category-specific organization in the human brain does not require visual experience", *Neuron*, 63 (3), 397-405.
- Maloney, E. A., Beilock, S. L. (2012), "Math anxiety: Who has it, why it develops, and how to guard against it", *Trends Cogn. Sci.*, 16 (8), 404-406.
- Markman, E. M., Wachtel, G. F. (1988), "Children's use of mutual exclusivity to constrain the meanings of words", *Cognitive Psychology*, 20 (2), 121-157.
- Markman, E. M., Wasow, J. L., Hansen, M. B. (2003), "Use of the mutual exclusivity assumption by young word learners", *Cognitive Psychology*, 47 (3), 241-275.
- Marques, J. F., Dehaene, S. (2004), "Developing intuition for prices in euros: Rescaling or relearning prices?", *J. Exp. Psychol. Appl.*, 10 (3), 148-155.
- Marshall, C. (2017), "Montessori education: A review of the evidence base", *NPJ Science of Learning*, 2 (1), 11.
- Marshall, L., Helgadottir, H., Molle, M., Born, J. (2006), "Boosting slow oscillations during sleep potentiates memory", *Nature*, 444 (7119), 610-613.
- Marti, S., King, J.-R., Dehaene, S. (2015), "Time-resolved decoding of two processing chains during dual-task interference", *Neuron*, 88 (6), 1297-1307.
- Marti, S., Sigman, M., Dehaene, S. (2012), "A shared cortical bottleneck underlying Attentional Blink and Psychological Refractory Period",

NeuroImage, 59 (3), 2883-2898.

Martin, S. L., Ramey, C. T., Ramey, S. (1990), "The prevention of intellectual impairment in children of impoverished families: Findings of a randomized trial of educational day care", *American Journal of Public Health*, 80 (7), 844-847.

Maye, J., Werker, J. F., Gerken, L. (2002), "Infant sensitivity to distributional information can affect phonetic discrimination", *Cognition*, 82 (3).

Mayer, R. E. (2004), "Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction", *The American Psychologist*, 59 (1), 14-19.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., McClelland, J. L. (2002), "Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception", *Cognitive, Affective and Behavioral Neuroscience*, 2 (2), 89-108.

McCloskey, M., Rapp, B. (2000), "A visually based developmental reading deficit", *Journal of Memory and Language*, 43, 157-181.

McCrink, K., Wynn, K. (2004), "Large-number addition and subtraction by 9-month-old infants", *Psychological Science*, 15 (11), 776-781.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., Amiel-Tison, C. (1988), "A precursor of language acquisition in young infants", *Cognition*, 29 (2), 143-178.

Meyer, T., Olson, C. R. (2011), "Statistical learning of visual transitions in monkey inferotemporal cortex", *Proc. Natl Acad. Sci. USA*, 108 (48), 19.401-19.406.

Millum, J., Emanuel, E. J. (2007), "Ethics. The ethics of international research with abandoned children", *Science*, 318 (5858), 1874-1875.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ..., Hassabis, D. (2015), "Human-level control through deep reinforcement learning", *Nature*, 518 (7540), 529-533.

- Mongelli, V., Dehaene, S., Vinckier, F., Peretz, I., Bartolomeo, P., Cohen, L. (2017), "Music and words in the visual cortex: The impact of musical expertise", *Cortex*, 86, 260-274.
- Mongillo, G., Barak, O., Tsodyks, M. (2008), "Synaptic theory of working memory", *Science*, 319 (5869), 1543-1546.
- Monzalvo, K., Fluss, J., Billard, C., Dehaene, S., Dehaene-Lambertz, G. (2012), "Cortical networks for vision and language in dyslexic and normal children of variable socio-economic status", *NeuroImage*, 61 (1), 258-274.
- Morais, J. (2017), "Literacy and democracy", *Language, Cognition and Neuroscience*, 33 (3), 351-371.
- Morais, J., Bertelson, P., Cary, L., Alegria, J. (1986), "Literacy training and speech segmentation", *Cognition*, 24, 45-64.
- Morais, J., Kolinsky, R. (2005), "Literacy and cognitive change", en M. Snowling, C. Hulme (eds.), *The Science of Reading. A Handbook*, Óxford, Blackwell, 188-203.
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., Chau, T. (2011), "Short-term music training enhances verbal intelligence and executive function", *Psychological Science*, 22 (11), 1425-1433.
- Morrison, C. M., Ellis, A. W. (1995), "Roles of word frequency and age of acquisition in word naming and lexical decision", *J. Exp. Psychol. Learn. Mem. Cogn.*, 21 (1), 116.
- Morton, J., Johnson, M. H. (1991), "CONSPEC and CONLERN: A two-process theory of infant face recognition", *Psychol. Rev.*, 98 (2), 164-181.
- Mottint, O. (2018), "Faut-il renoncer aux pédagogies actives?", démocratisation-scolaire.fr, 29 de marzo, www.democratisation-scolaire.fr/spip.php?article287.
- Moyer, R. S., Landauer, T. K. (1967), "Time required for judgements of numerical inequality", *Nature*, 215, 1519-1520.

- Muckli, L., Naumer, M. J., Singer, W. (2009), “Bilateral visual field maps in a patient with only one hemisphere”, *Proc. Natl Acad. Sci. USA*, 106 (31), 13.034-13.039.
- Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buchel, C., Weiller, C. (2003), “Broca’s area and the language instinct”, *Nature Neuroscience*, 6 (7), 774-781.
- Naatanen, R., Paavilainen, P., Rinne, T., Alho, K. (2007), “The mismatch negativity (MMN) in basic research of central auditory processing: A review”, *Clin. Neurophysiol.*, 118 (12), 2544-2590.
- National Institute of Child Health and Human Development (2000), *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*, NIH Publications n° 00-4769, Washington, U.S. Government Printing Office.
- Nau, M., Navarro Schröder, T., Bellmund, J. L. S., Doeller, C. F. (2018), “Hexadirectional coding of visual space in human entorhinal cortex”, *Nature Neuroscience*, 21 (2), 188-190.
- Nelson, C. A., Zeanah, C. H., Fox, N. A., Marshall, P. J., Smyke, A. T., Guthrie, D. (2007), “Cognitive recovery in socially deprived young children: The Bucharest Early Intervention Project”, *Science*, 318 (5858), 1937-1940.
- Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., ..., Dehaene, S. (2017), “Neurophysiological dynamics of phrase-structure building during sentence processing”, *Proc. Natl Acad. Sci. USA*, 114 (18), E3669-E3678.
- Nemmi, F., Helander, E., Helenius, O., Almeida, R., Hassler, M., Räsänen, P., Klingberg, T. (2016), “Behavior and neuroimaging at baseline predict individual response to combined mathematical and working memory training in children”, *Developmental Cognitive Neuroscience*, 20, 43-51.
- Ngo, H.-V. V., Martinetz, T., Born, J., Mölle, M. (2013), “Auditory closedloop stimulation of the sleep slow oscillation enhances memory”,

Neuron, 78 (3), 545-553.

Nieder, A., Dehaene, S. (2009), "Representation of number in the brain", *Annu. Rev. Neurosci.*, 32, 185-208.

Noble, K. G., Norman, M. F., Farah, M. J. (2005), "Neurocognitive correlates of socioeconomic status in kindergarten children", *Dev. Sci.*, 8 (1), 74-87.

Norimoto, H., Makino, K., Gao, M., Shikano, Y., Okamoto, K., Ishikawa, T., ..., Ikegaya, Y. (2018), "Hippocampal ripples down-regulate synapses", *Science*, 359 (6383), 1524-1527.

Obayashi, S., Suhara, T., Kawabe, K., Okauchi, T., Maeda, J., Akine, Y., ..., Iriki, A. (2001), "Functional brain mapping of monkey tool use", *NeuroImage*, 14 (4), 853-861.

Oechslin, M. S., Gschwind, M., James, C. E. (2018), "Tracking training-related plasticity by combining fMRI and DTI: The right hemisphere ventral stream mediates musical syntax processing", *Cereb. Cortex*, 28 (4), 1209-1218.

Olah, C., Mordvintsev, A., Schubert, L. (2017), "Feature visualization: How neural networks build up their understanding of images", *Distill*, 7 de noviembre, distill.pub/2017/feature-visualization.

Olesen, P. J., Westerberg, H., Klingberg, T. (2004), "Increased prefrontal and parietal activity after training of working memory", *Nature Neuroscience*, 7 (1), 75-79.

Orbán, G., Berkes, P., Fiser, J., Lengyel, M. (2016), "Neural variability and sampling-based probabilistic representations in the visual cortex", *Neuron*, 92 (2), 530-543.

Paller, K. A., McCarthy, G., Wood, C. C. (1988), "ERPs predictive of subsequent recall and recognition performance", *Biological Psychology*, 26 (1-3), 269-276.

Pallier, C., Dehaene, S., Poline, J. B., LeBihan, D., Argenti, A. M., Dupoux, E., Mehler, J. (2003), "Brain imaging of language plasticity in adopted adults: Can a second language replace the first?", *Cereb. Cortex*, 13 (2), 155-161.

- Pallier, C., Devauchelle, A. D., Dehaene, S. (2011), "Cortical representation of the constituent structure of sentences", *Proc. Natl Acad. Sci. USA*, 108 (6), 2522-2527.
- Palminteri, S., Kilford, E. J., Coricelli, G., Blakemore, S.-J. (2016), "The computational development of reinforcement learning during adolescence", *PLoS Computational Biology*, 12 (6).
- Pashler, H., McDaniel, M., Rohrer, D., Bjork, R. (2008), "Learning styles concepts and evidence", *Psychological Science in the Public Interest*, 9 (3), 105-119.
- Pegado, F., Comerlato, E., Ventura, F., Jobert, A., Nakamura, K., Buiatti, M., ..., Dehaene, S. (2014), "Timing the impact of literacy on visual processing", *Proc. Natl Acad. Sci. USA*, 111 (49), E5233-5242.
- Pegado, F., Nakamura, K., Braga, L. W., Ventura, P., Nunes Filho, G., Pallier, C., ..., Dehaene, S. (2014), "Literacy breaks mirror invariance for visual stimuli: A behavioral study with adult illiterates", *Journal of Experimental Psychology: General*, 143 (2), 887-894.
- Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., ..., Maquet, P. (2004), "Are spatial memories strengthened in the human hippocampus during slow wave sleep?", *Neuron*, 44 (3), 535-545.
- Penn, D. C., Holyoak, K. J., Povinelli, D. J. (2008), "Darwin's mistake: Explaining the discontinuity between human and nonhuman minds", *Behav. Brain Sci.*, 31 (2), 109-130; discussion, 130-178.
- Pennac, D. (2017), "Daniel Pennac: 'J'ai été d'abord et avant tout professeur'" (entrevista con Catherine Vincent), *Le Monde*, 12 de febrero, sección "Culture", disponible en lemonde.fr.
- Peña, M., Werker, J. F., Dehaene-Lambertz, G. (2012), "Earlier speech exposure does not accelerate speech acquisition", *J. Neurosci.*, 32 (33), 11.159-11.163.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., Frith, C. D. (2006), "Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans", *Nature*, 442 (7106), 1042-1045.

- Piantadosi, S. T., Tenenbaum, J. B., Goodman, N. D. (2012), “Bootstrapping in a language of thought: A formal model of numerical concept learning”, *Cognition*, 123 (2), 199-217.
- (2016), “The logical primitives of thought: Empirical foundations for compositional cognitive models”, *Psychological Review*, 123 (4), 392-424.
- Piazza, M., DeFeo, V., Panzeri, S., Dehaene, S. (2018), “Learning to focus on number”, *Cognition*, 181, 35-45.
- Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., ..., Zorzi, M. (2010), “Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia”, *Cognition*, 116 (1), 33-41.
- Piazza, M., Pica, P., Izard, V., Spelke, E. S., Dehaene, S. (2013), “Education enhances the acuity of the nonverbal approximate number system”, *Psychological Science*, 24 (6), 1037-1043.
- Pica, P., Lemer, C., Izard, V., Dehaene, S. (2004), “Exact and approximate arithmetic in an Amazonian indigene group”, *Science*, 306 (5695), 499-503.
- Pierce, L. J., Klein, D., Chen, J.-K., Delcenserie, A., Genesee, F. (2014), “Mapping the unconscious maintenance of a lost first language”, *Proc. Natl Acad. Sci. USA*, 111 (48), 17.314-17.319.
- Pinheiro-Chagas, P., Dotan, D., Piazza, M., Dehaene, S. (2017), “Finger tracking reveals the covert stages of mental arithmetic”, *Open Mind*, 1 (1), 30-41.
- Pittenger, C., Kandel, E. R. (2003), “In search of general mechanisms for longlasting plasticity: Aplysia and the hippocampus”, *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 358 (1432), 757-763.
- Poirel, N., Borst, G., Simon, G., Rossi, S., Cassotti, M., Pineau, A., Houdé, O. (2012), “Number conservation is related to children’s prefrontal inhibitory control: An fMRI study of a piagetian task”, *PLoS ONE*, 7 (7).
- Poo, M.-M., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T.,

- Martin, K. C., ..., Stevens, C. (2016), "What is memory? The present state of the engram", *BMC Biology*, 14, 40.
- Posner, M. I. (1994), "Attention: The mechanisms of consciousness", *Proc. Natl Acad. Sci. USA*, 91, 7398-7403.
- Posner, M. I., Rothbart, M. K. (1998), "Attention, self-regulation and consciousness", *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 353 (1377), 1915-1927.
- Prado, E. L., Dewey, K. G. (2014), "Nutrition and brain development in early life", *Nutrition Reviews*, 72 (4), 267-284.
- Prehn-Kristensen, A., Munz, M., Göder, R., Wilhelm, I., Korr, K., Vahl, W., ..., Baving, L. (2014), "Transcranial oscillatory direct current stimulation during sleep improves declarative memory consolidation in children with attention-deficit/hyperactivity disorder to a level comparable to healthy controls", *Brain Stimulation*, 7 (6), 793-799.
- Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D. C., Menon, V. (2014), "Hippocampal-neocortical functional reorganization underlies children's cognitive development", *Nature Neuroscience*, 17 (9), 1263-1269.
- Quartz, S. R., Sejnowski, T. J. (1997), "The neural basis of cognitive development: A constructivist manifesto", *Behav. Brain Sci.*, 20 (4), 537-556; discussion, 556-596.
- Rakic, P., Bourgeois, J. P., Eckenhoff, M. F., Zecevic, N., Goldman-Rakic, P. S. (1986), "Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex", *Science*, 232 (4747), 232-235.
- Ramanathan, D. S., Gulati, T., Ganguly, K. (2015), "Sleep-dependent reactivation of ensembles in motor cortex promotes skill consolidation", *PLoS Biology*, 13 (9).
- Ramirez, S., Liu, X., Lin, P.-A., Suh, J., Pignatelli, M., Redondo, R. L., ..., Tonegawa, S. (2013), "Creating a false memory in the hippocampus", *Science*, 341 (6144), 387-391.
- Ramirez, S., Liu, X., MacDonald, C. J., Moffa, A., Zhou, J., Redondo, R. L., Tonegawa, S. (2015), "Activating positive memory engrams

- suppresses depression-like behaviour”, *Nature*, 522 (7556), 335-339.
- Rankin, C. H. (2004), “Invertebrate learning: What can’t a worm learn?”, *Current Biology*, 14 (15), R617-618.
- Rasch, B., Büchel, C., Gais, S., Born, J. (2007), “Odor cues during slow-wave sleep prompt declarative memory consolidation”, *Science*, 315 (5817), 1426-1429.
- Rasmussen, A., Jirenhed, D. A., Hesslow, G. (2008), “Simple and complex spike firing patterns in Purkinje cells during classical conditioning”, *Cerebellum*, 7 (4), 563-566.
- Rattan, A., Savani, K., Chugh, D., Dweck, C. S. (2015), “Leveraging mindsets to promote academic achievement: Policy recommendations”, *Perspectives on Psychological Science*, 10 (6), 721-726.
- Reich, L., Szwed, M., Cohen, L., Amedi, A. (2011), “A ventral visual stream reading center independent of visual experience”, *Current Biology*, 21 (5), 363-368.
- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., Reissland, N. (2017), “The human fetus preferentially engages with face-like visual stimuli”, *Current Biology*, 27 (12), 1825-1828.e3.
- Rescorla, R. A., Wagner, A. R. (1972), “A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement”, *Classical Conditioning II: Current Research and Theory*, 2, 64-99.
- Ribeiro, S., Goyal, V., Mello, C. V., Pavlides, C. (1999), “Brain gene expression during REM sleep depends on prior waking experience”, *Learning and Memory*, 6 (5), 500-508.
- Ritchie, S. J., Tucker-Drob, E. (2017), “How much does education improve intelligence? A meta-analysis”, *PsyArXiv Preprints*, <https://psyarxiv.com/kymhp>.
- Rivera, S. M., Reiss, A. L., Eckert, M. A., Menon, V. (2005), “Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex”, *Cereb. Cortex*, 15 (11), 1779-1790.

- Robey, A. M., Dougherty, M. R., Buttaccio, D. R. (2017), “Making retrospective confidence judgments improves learners’ ability to decide what not to study”, *Psychological Science*, 28 (11).
- Roediger, H. L., Karpicke, J. D. (2006), “Test-enhanced learning: Taking memory tests improves long-term retention”, *Psychological Science*, 17 (3), 249-255.
- Rohrer, D., Taylor, K. (2006), “The effects of overlearning and distributed practise on the retention of mathematics knowledge”, *Applied Cognitive Psychology*, 20 (9), 1209-1224.
- (2007), “The shuffling of mathematics problems improves learning”, *Instructional Science*, 35 (6), 481-498.
- Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., Gabrieli, J. D. E. (2018), “Beyond the 30-million-word gap: Children’s conversational exposure is associated with language-related brain function”, *Psychological Science*, 29 (5), 700-710.
- Rouault, M., Koechlin, E. (2018), “Prefrontal function and cognitive control: From action to language”, *Current Opinion in Behavioral Sciences*, 21, 106-111.
- Rudoy, J. D., Voss, J. L., Westerberg, C. E., Paller, K. A. (2009), “Strengthening individual memories by reactivating them during sleep”, *Science*, 326 (5956), 1079.
- Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W. J., Bick, A., Frost, S. J., ..., Frost, R. (2015), “Universal brain signature of proficient reading: Evidence from four contrasting languages”, *Proc. Natl Acad. Sci. USA*, 112 (50), 15.510-15.515.
- Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., Posner, M. I. (2005), “Training, maturation, and genetic influences on the development of executive attention”, *Proc. Natl Acad. Sci. USA*, 102 (41), 14.931-14.936.
- Rugani, R., Fontanari, L., Simoni, E., Regolin, L., Vallortigara, G. (2009), “Arithmetic in newborn chicks”, *Proc. Biol. Sci.*, 276 (1666), 2451-2460.

- Rugani, R., Vallortigara, G., Priftis, K., Regolin, L. (2015), "Animal cognition. Number-space mapping in the newborn chick resembles humans' mental number line", *Science*, 347 (6221), 534-536.
- Sabbah, N., Authié, C. N., Sanda, N., Mohand-Saïd, S., Sahel, J.-A., Safran, A. B., ..., Amedi, A. (2016), "Increased functional connectivity between language and visually deprived areas in late and partial blindness", *NeuroImage*, 136, 162-173.
- Sackur, J., Dehaene, S. (2009), "The cognitive architecture for chaining of two mental operations", *Cognition*, 111 (2), 187-211.
- Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., ..., Batista, A. P. (2014), "Neural constraints on learning", *Nature*, 512 (7515), 423-426.
- Saffran, J. R., Aslin, R. N., Newport, E. L. (1996), "Statistical learning by 8-month-old infants", *Science*, 274 (5294), 1926-1928.
- Sakai, T., Mikami, A., Tomonaga, M., Matsui, M., Suzuki, J., Hamada, Y., ..., Matsuzawa, T. (2011), "Differential prefrontal white matter development in chimpanzees and humans", *Current Biology*, 21 (16), 1397-1402.
- Salimpoor, V. N., Van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., Zatorre, R. J. (2013), "Interactions between the nucleus accumbens and auditory cortices predict music reward value", *Science*, 340 (6129), 216-219.
- Samson, D. R., Nunn, C. L. (2015), "Sleep intensity and the evolution of human cognition", *Evolutionary Anthropology*, 24 (6), 225-237.
- Sangrigoli, S., Pallier, C., Argenti, A.-M., Ventureyra, V. A., Schonen, S. de (2005), "Reversibility of the other-race effect in face recognition during childhood", *Psychological Science*, 16 (6), 440-444.
- Saygin, Z. M., Norton, E. S., Osher, D. E., Beach, S. D., Cyr, A. B., Ozernov-Palchik, O., ..., Gabrieli, J. D. E. (2013), "Tracking the roots of reading ability: White matter volume and integrity correlate with phonological awareness in prereading and early-reading kindergarten children", *J. Neurosci.*, 33 (33), 13.251-13.258.

- Saygin, Z. M., Osher, D. E., Koldewyn, K., Reynolds, G., Gabrieli, J. D., Saxe, R. R. (2012), "Anatomical connectivity patterns predict face selectivity in the fusiform gyrus", *Nature Neuroscience*, 15 (2), 321-327.
- Saygin, Z. M., Osher, D. E., Norton, E. S., Youssofian, D. A., Beach, S. D., Feather, J., ..., Kanwisher, N. (2016), "Connectivity precedes function in the development of the visual word form area", *Nature Neuroscience*, 19 (9), 1250-1255.
- Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., Botvinick, M. M. (2016), "Statistical learning of temporal community structure in the hippocampus", *Hippocampus*, 26 (1), 3-8.
- Schlaug, G., Jancke, L., Huang, Y., Staiger, J. F., Steinmetz, H. (1995), "Increased corpus callosum size in musicians", *Neuropsychologia*, 33 (8), 1047-1055.
- Schmidt, R. A., Bjork, R. A. (1992), "New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training", *Psychological Science*, 3 (4), 207-217.
- Schoenemann, P. T., Sheehan, M. J., Glotzer, L. D. (2005), "Prefrontal white matter volume is disproportionately larger in humans than in other primates", *Nature Neuroscience*, 8 (2), 242-252.
- Schultz, W., Dayan, P., Montague, P. R. (1997), "A neural substrate of prediction and reward", *Science*, 275, 1593-1599.
- Schweinhart, L. J. (1993), *Significant Benefits. The High/Scope Perry Preschool Study through Age 27*, Ypsilanti, High/Scope Educational Research Foundation, ERIC.
- Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J., Madsen, J. R. (2003), "Theta and gamma oscillations during encoding predict subsequent recall", *J. Neurosci.*, 23 (34), 10.809-10.814.
- Sederberg, P. B., Schulze-Bonhage, A., Madsen, J. R., Bromfield, E. B., McCarthy, D. C., Brandt, A., ..., Kahana, M. J. (2006), "Hippocampal and neocortical gamma oscillations predict memory formation in humans", *Cereb. Cortex*, 17 (5), 1190-1196.

- Seehagen, S., Konrad, C., Herbert, J. S., Schneider, S. (2015), "Timely sleep facilitates declarative memory consolidation in infants", *Proc. Natl Acad. Sci.*, 112 (5), 1625-1629.
- Seitz, A., Lefebvre, C., Watanabe, T., Jolicœur, P. (2005), "Requirement for high-level processing in subliminal learning", *Current Biology*, 15 (18), R753-755.
- Senghas, A., Kita, S., Özyürek, A. (2004), "Children creating core properties of language: Evidence from an emerging sign language in Nicaragua", *Science*, 305 (5691), 1779-1782.
- Shah, P. E., Weeks, H. M., Richards, B., Kaciroti, N. (2018), "Early childhood curiosity and kindergarten reading and math academic achievement", *Pediatric Research*, 1.
- Shatz, C. J. (1996), "Emergence of order in visual system development", *Proc. Natl Acad. Sci. USA*, 93 (2), 602-608.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., Makuch, R. (1992), "Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability", *N. Engl. J. Med.*, 326 (3), 145-150.
- Sheese, B. E., Rothbart, M. K., Posner, M. I., White, L. K., Fraundorf, S. H. (2008), "Executive attention and self-regulation in infancy", *Infant. Behav. Dev.*, 31 (3), 501-510.
- Sheridan, M. A., Fox, N. A., Zeanah, C. H., McLaughlin, K. A., Nelson, C. A. (2012), "Variation in neural development as a result of exposure to institutionalization early in childhood", *Proc. Natl Acad. Sci. USA*, 109 (32), 12.927-12.932.
- Shi, R., Lepage, M. (2008), "The effect of functional morphemes on word segmentation in preverbal infants", *Dev. Sci.*, 11 (3), 407-413.
- Shipston-Sharman, O., Solanka, L., Nolan, M. F. (2016), "Continuous attractor network models of grid cell firing based on excitatory-inhibitory interactions", *Journal of Physiology*, 594 (22), 6547-6557.
- Shneidman, L. A., Arroyo, M. E., Levine, S. C., Goldin-Meadow, S.

- (2013), "What counts as effective input for word learning?", *Journal of Child Language*, 40 (3), 672-686.
- Shneidman, L. A., Goldin-Meadow, S. (2012), "Language input and acquisition in a Mayan village: How important is directed speech?", *Dev. Sci.*, 15 (5), 659-673.
- Shohamy, D., Turk-Browne, N. B. (2013), "Mechanisms for widespread hippocampal involvement in cognition", *Journal of Experimental Psychology: General*, 142 (4), 1159-1170.
- Siegler, R. S. (1989), "Mechanisms of cognitive development", *Annu. Rev. Psychol.*, 40, 353-379.
- Siegler, R. S., Opfer, J. E. (2003), "The development of numerical estimation: Evidence for multiple representations of numerical quantity", *Psychological Science*, 14 (3), 237-243.
- Siegler, R. S., Thompson, C. A., Schneider, M. (2011), "An integrated theory of whole number and fractions development", *Cognitive Psychology*, 62 (4), 273-296.
- Sigman, M., Dehaene, S. (2008), "Brain mechanisms of serial and parallel processing during dual-task performance", *J. Neurosci.*, 28 (30), 7585-7598.
- Sigman, M., Pan, H., Yang, Y., Stern, E., Silbersweig, D., Gilbert, C. D. (2005), "Top-down reorganization of activity in the visual pathway after learning a shape identification task", *Neuron*, 46 (5), 823-835.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van den Driessche, G., ..., Hassabis, D. (2016), "Mastering the game of Go with deep neural networks and tree search", *Nature*, 529 (7587), 484-489.
- Simons, D. J., Chabris, C. F. (1999), "Gorillas in our midst: Sustained inattention blindness for dynamic events", *Perception*, 28 (9), 1059-1074.
- Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., Macnamara, B. N. (2018), "To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses", *Psychological Science*, 29 (4), 549-571.

- Skaggs, W. E., McNaughton, B. L. (1996), "Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience", *Science*, 271 (5257), 1870-1873.
- Smaers, J. B., Gómez-Robles, A., Parks, A. N., Sherwood, C. C. (2017), "Exceptional evolutionary expansion of prefrontal cortex in great apes and humans", *Current Biology*, 27 (5), 714-720.
- Spelke, E. (2003), "What makes us smart? Core knowledge and natural language", en D. Gentner, S. Goldin-Meadow (eds.), *Language in Mind*, Cambridge (Mass.), MIT Press.
- Spencer, S. J., Steele, C. M., Quinn, D. M. (1999), "Stereotype threat and women's math performance", *Journal of Experimental Social Psychology*, 35 (1), 4-28.
- Spencer-Smith, M., Klingberg, T. (2015), "Benefits of a working memory training program for inattention in daily life: A systematic review and meta-analysis", *PLoS ONE*, 10 (3).
- Srihasam, K., Mandeville, J. B., Morocz, I. A., Sullivan, K. J., Livingstone, M. S. (2012), "Behavioral and Anatomical consequences of early versus late symbol training in macaques", *Neuron*, 73 (3), 608-619.
- Stahl, A. E., Feigenson, L. (2015), "Observing the unexpected enhances infants' learning and exploration", *Science*, 348 (6230), 91-94.
- Starkey, P., Cooper, R. G. Jr (1980), "Perception of numbers by human infants", *Science*, 210, 1033-1035.
- Starkey, P., Spelke, E. S., Gelman, R. (1990), "Numerical abstraction by human infants", *Cognition*, 36, 97-127.
- Steele, C. M., Aronson, J. (1995), "Stereotype threat and the intellectual test performance of African Americans", *Journal of Personality and Social Psychology*, 69 (5), 797.
- Steinhauer, K., Drury, J. E. (2012), "On the early left-anterior negativity (ELAN) in syntax studies", *Brain and Language*, 120 (2), 135-162.
- Stickgold, R. (2005), "Sleep-dependent memory consolidation", *Nature*,

437 (7063), 1272-1278.

Strauss, M., Sitt, J. D., King, J.-R., Elbaz, M., Azizi, L., Buiatti, M., ..., Dehaene, S. (2015), "Disruption of hierarchical predictive coding during sleep", *Proc. Natl Acad. Sci. USA*, 112 (11), E1353-1362.

Striem-Amit, E., Amedi, A. (2014), "Visual cortex extrastriate body-selective area activation in congenitally blind people "seeing" by using sounds", *Current Biology*, 24 (6), 687-692.

Strnad, L., Peelen, M. V., Bedny, M., Caramazza, A. (2013), "Multivoxel pattern analysis reveals auditory motion information in MT+ of both congenitally blind and sighted individuals", *PLoS ONE*, 8 (4).

Sun, T., Patoine, C., Abu-Khalil, A., Visvader, J., Sum, E., Cherry, T. J., ..., Walsh, C. A. (2005), "Early asymmetry of gene transcription in embryonic human left and right cerebral cortex", *Science*, 308 (5729), 1794-1798.

Sun, Z. Y., Klöppel, S., Rivière, D., Perrot, M., Frackowiak, R., Siebner, H., Mangin, J.-F. (2012), "The effect of handedness on the shape of the central sulcus", *NeuroImage*, 60 (1), 332-339.

Sur, M., Garraghty, P. E., Roe, A. W. (1988), "Experimentally induced visual projections into auditory thalamus and cortex", *Science*, 242 (4884), 1437-1441.

Sur, M., Rubenstein, J. L. R. (2005), "Patterning and plasticity of the cerebral cortex", *Science*, 310 (5749), 805-810.

Sutton, R. S., Barto, A. G. (1998), *Reinforcement Learning: An Introduction*, Cambridge (Mass.), MIT Press.

Szpunar, K. K., Khan, N. Y., Schacter, D. L. (2013), "Interpolated memory tests reduce mind wandering and improve learning of online lectures", *Proc. Natl Acad. Sci. USA*, 110 (16), 6313-6317.

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabregue, R., Amadon, A., Cohen, L. (2011), "Specialization for written words over objects in the visual cortex", *NeuroImage*, 56 (1), 330-344.

- Szwed, M., Qiao, E., Jobert, A., Dehaene, S., Cohen, L. (2014), "Effects of literacy in early visual and occipitotemporal areas of Chinese and French readers", *J. Cogn. Neurosci.*, 26 (3), 459-475.
- Szwed, M., Ventura, P., Querido, L., Cohen, L., Dehaene, S. (2012), "Reading acquisition enhances an early visual process of contour integration", *Dev. Sci.*, 15 (1), 139-149.
- Takeuchi, T., Duzskiewicz, A. J., Morris, R. G. M. (2014), "The synaptic plasticity and memory hypothesis: Encoding, storage and persistence", *Philos. Trans. R. Soc. B. Biol. Sci.*, 369 (1633).
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L., Goodman, N. D. (2011), "How to grow a mind: Statistics, structure, and abstraction", *Science*, 331 (6022), 1279-1285.
- Terrace, H. S., Petitto, L. A., Sanders, R. J., Bever, T. G. (1979), "Can an ape create a sentence?", *Science*, 206 (4421), 891-902.
- Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., Dehaene, S. (2014), "Learning to read improves the structure of the arcuate fasciculus", *Cereb. Cortex*, 24 (4), 989-995.
- Thornton, A., McAuliffe, K. (2006), "Teaching in wild meerkats", *Science*, 313 (5784), 227-229.
- Todorovic, A., De Lange, F. P. (2012), "Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields", *J. Neurosci.*, 32 (39), 13389-13395.
- Tombu, M., Jolicoeur, P. (2004), "Virtually no evidence for virtually perfect time-sharing", *J. Exp. Psychol. Hum. Percept. Perform.*, 30 (5), 795-810.
- Uhrig, L., Dehaene, S., Jarraya, B. (2014), "A hierarchy of responses to auditory regularities in the macaque brain", *J. Neurosci.*, 34 (4), 1127-1132.
- Van Praag, H., Kempermann, G., Gage, F. H. (2000), "Neural consequences of environmental enrichment", *Nature Reviews. Neuroscience*, 1 (3), 191-198.

- Van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., Roelfsema, P. R. (2018), “The threshold for conscious report: Signal loss and response bias in visual and frontal cortex”, *Science*, 360 (6388), 537-542.
- Ventura, P., Fernandes, T., Cohen, L., Morais, J., Kolinsky, R., Dehaene, S. (2013), “Literacy acquisition reduces the influence of automatic holistic processing of faces and houses”, *Neuroscience Letters*, 554, 105-109.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., Cohen, L. (2007), “Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system”, *Neuron*, 55 (1), 143-156.
- Vinckier, F., Naccache, L., Papeix, C., Forget, J., Hahn-Barma, V., Dehaene, S., Cohen, L. (2006), “‘What’ and ‘where’ in word reading: Ventral coding of written words revealed by parietal atrophy”, *J. Cogn. Neurosci.*, 18 (12), 1998-2012.
- Viswanathan, P., Nieder, A. (2013), “Neuronal correlates of a visual ‘sense of number’ in primate parietal and prefrontal cortices”, *Proc. Natl Acad. Sci. USA*, 110 (27), 11187-11192.
- (2015), “Differential impact of behavioral relevance on quantity coding in primate frontal and parietal neurons”, *Current Biology*, 25 (10), 1259-1269.
- Vogel, E. K., Machizawa, M. G. (2004), “Neural activity predicts individual differences in visual working memory capacity”, *Nature*, 428 (6984), 748-751.
- Voss, M. W., Vivar, C., Kramer, A. F., Van Praag, H. (2013), “Bridging animal and human models of exercise-induced brain plasticity”, *Trends Cogn. Sci.*, 17 (10), 525-544.
- Wacongne, C., Labyt, E., Van Wassenhove, V., Bekinschtein, T., Naccache, L., Dehaene, S. (2011), “Evidence for a hierarchy of predictions and prediction errors in human cortex”, *Proc. Natl. Acad. Sci. USA*, 108 (51), 20754-20759.

- Waelti, P., Dickinson, A., Schultz, W. (2001), "Dopamine responses comply with basic assumptions of formal learning theory", *Nature*, 412 (6842), 43-48.
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., ..., Buckner, R. L. (1998), "Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity", *Science*, 281 (5380), 1188-1191.
- Wagner, U., Gais, S., Haider, H., Verleger, R., Born, J. (2004), "Sleep inspires insight", *Nature*, 427 (6972), 352-355.
- Walker, M. P., Brakefield, T., Hobson, J. A., Stickgold, R. (2003), "Dissociable stages of human memory consolidation and reconsolidation", *Nature*, 425 (6958), 616-620.
- Walker, M. P., Stickgold, R. (2004), "Sleep-dependent learning and memory consolidation", *Neuron*, 44 (1), 121-133.
- Walker, M. P., Stickgold, R., Alsup, D., Gaab, N., Schlaug, G. (2005), "Sleep-dependent motor memory plasticity in the human brain", *Neuroscience*, 133 (4), 911-917.
- Walker, S. P., Chang, S. M., Powell, C. A., Grantham-McGregor, S. M. (2005), "Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: Prospective cohort study", *The Lancet*, 366 (9499), 1804-1807.
- Wang, L., Krauzlis, R. J. (2018), "Visual selective attention in mice", *Current Biology*, 28 (5), 676-685.e4.
- Wang, L., Uhrig, L., Jarraya, B., Dehaene, S. (2015), "Representation of numerical and sequential patterns in macaque and human brains", *Current Biology*, 25 (15), 1966-1974.
- Warneken, F., Tomasello, M. (2006), "Altruistic helping in human infants and young chimpanzees", *Science*, 311 (5765), 1301-1303.
- Watanabe, T., Nanez, J. E., Sasaki, Y. (2001), "Perceptual learning without perception", *Nature*, 413 (6858), 844-848.

- Weber-Fox, C. M., Neville, H. J. (1996), "Maturation constraints on functional specialization for language processing: ERP and behavioral evidence in bilingual speakers", *J. Cogn. Neurosci.*, 8, 231-256.
- Werker, J. F., Hensch, T. K. (2014), "Critical periods in speech perception: New directions", *Annu. Rev. Psychol.*, 66, 173-196.
- Werker, J. F., Tees, R. C. (1984), "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life", *Infant Behavior and Development*, 7, 49-63.
- Whitlock, J. R., Heynen, A. J., Shuler, M. G., Bear, M. F. (2006), "Learning induces long-term potentiation in the hippocampus", *Science*, 313 (5790), 1093-1097.
- Widloski, J., Fiete, I. R. (2014), "A model of grid cell development through spatial exploration and spike time-dependent plasticity", *Neuron*, 83 (2), 481-495.
- Wilhelm, I., Rose, M., Imhof, K. I., Rasch, B., Büchel, C., Born, J. (2013), "The sleeping child outplays the adult's capacity to convert implicit into explicit knowledge", *Nature Neuroscience*, 16 (4), 391-393.
- Wills, T. J., Cacucci, F., Burgess, N., O'Keefe, J. (2010), "Development of the hippocampal cognitive map in preweanling rats", *Science*, 328 (5985), 1573-1576.
- Wilson, M. A., McNaughton, B. L. (1994), "Reactivation of hippocampal ensemble memories during sleep", *Science*, 265 (5172), 676-679.
- Windsor, J., Moraru, A., Nelson, C. A., Fox, N. A., Zeanah, C. H. (2013), "Effect of foster care on language learning at eight years: Findings from the Bucharest Early Intervention Project", *Journal of Child Language*, 40 (3), 605-627.
- Wynn, K. (1992), "Addition and subtraction by human infants", *Nature*, 358, 749-750.
- Xu, F., García, V. (2008), "Intuitive statistics by 8-month-old infants",

Proc. Natl Acad. Sci. USA, 105 (13), 5012-5015.

Xu, F., Tenenbaum, J. B. (2007), "Word learning as Bayesian inference", *Psychol. Rev.*, 114 (2), 245-272.

Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhutdinov, R., ..., Bengio, Y. (2015), "Show, attend and tell: Neural image caption generation with visual attention", arxiv.org/abs/150203044.

Yang, C. (2013), "Ontogeny and phylogeny of language", *Proc. Natl Acad. Sci. USA*, 110 (16), 6324-6327.

Yoncheva, Y. N., Blau, V. C., Maurer, U., McCandliss, B. D. (2010), "Attentional focus during learning impacts N170 ERP responses to an artificial script", *Dev. Neuropsychol.*, 35 (4), 423-445.

Yoon, J. M. D., Johnson, M. H., Csibra, G. (2008), "Communication-induced memory biases in preverbal infants", *Proc. Natl Acad. Sci. USA*, 105 (36), 13.690-13.695.

Yoon, K., Buice, M. A., Barry, C., Hayman, R., Burgess, N., Fiete, I. R. (2013), "Specific evidence of low-dimensional continuous attractor dynamics in grid cells", *Nature Neuroscience*, 16 (8), 1077-1084.

Young, C. B., Wu, S. S., Menon, V. (2012), "The neurodevelopmental basis of math anxiety", *Psychological Science*, 23 (5), 492-501.

Zaromb, F. M., Karpicke, J. D., Roediger, H. L. (2010), "Comprehension as a basis for metacognitive judgments: Effects of effort after meaning on recall and metacognition", *J. Exp. Psychol. Learn. Mem. Cogn.*, 36 (2), 552-557.

Zaromb, F. M., Roediger, H. L. I. (2010), "The testing effect in free recall is associated with enhanced organizational processes", *Memory and Cognition*, 38 (8), 995-1008.

Zhu, X., Wang, F., Hu, H., Sun, X., Kilgard, M. P., Merzenich, M. M., Zhou, X. (2014), "Environmental acoustic enrichment promotes recovery from developmentally degraded auditory cortical processing", *The J. Neurosci.*, 34 (16), 5406-5415.

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A.,

Spinelli, D. (2005), “Word length effect in early reading and in developmental dyslexia”, *Brain and Language*, 93 (3), 369-373.

Zylberberg, A., Dehaene, S., Roelfsema, P. R., Sigman, M. (2011), “The human Turing machine: A neural framework for mental programs”, *Trends Cogn. Sci.*, 15 (7), 293-300.

Créditos de material gráfico

Figura 1 (arriba y abajo), gentilmente autorizado por Antonio Battro.

Figuras 2 (abajo), **8**, **13** (arriba derecha y centro), **14** (abajo derecha), **29** (arriba) y **34**, Stanislas Dehaene.

Figura 3 (arriba), tomado de ai.googleblog.com/2017/05/using-machine-learning-to-explore.html; (abajo), Olah, Mordvintsev y Schubert (2017).

Figura 4, tomado de Guerguiev, Lillicrap y Richards (2017: figura 6).

Figura 5, tomado de Tenenbaum y otros (2011), gentilmente autorizado por Joshua Tenenbaum.

Figura 6, tomado de Kemp y Tenenbaum (2008: figuras 2 y 3).

Figura 7 (arriba), gentilmente autorizado por Fei Xu; (abajo), cedido por Molly Dillon y Elizabeth Spelke.

Figura 9 (arriba), gentilmente autorizado por G. Dehaene-Lambertz y J. Dubois; (abajo), tomado de Dehaene-Lambertz y otros (2006).

Figura 10, tomado de Krubitzer (2007), con su gentil autorización.

Figura 11 (arriba), gentilmente cedido por Alain Chédotal; tomado de Belle y otros (2017); abajo, cedido por Dehaene-Lambertz y Dubois.

Figura 12, a partir de Amunts y otros (2010: figuras 1 y 7).

Figura 13 (arriba derecha), tomado de pexels.com/photo/animals-apiary-beehive-beekeeping-928978/; (abajo), a partir de Hafting y otros (2005).

Figura 14 (izquierda): © Aisa/Leemage; (arriba derecha), gentilmente

autorizado por Philip Buttery.

Figura 15, tomado de Muckli, Naumer y Singer (2009: figura 2).

Figura 16, © 1939... 1967, Harvard College; © renovado 1967... 1991.

Figura 17 (arriba), a partir de Flege, Munro y MacKay (1995); Johnson y Newport (1989); Hartshorne y otros (2018); (abajo), tomado de Pierce y otros (2014: figura 3).

Figura 18 (arriba), gentilmente cedido por Eric Knudsen; (abajo), a partir de Knudsen, Zheng y DeBello (2000: figuras 2 y 3).

Figura 19 (arriba), © Michael Carroll, a partir de Berens y Nelson (2015); (abajo), tomado de Almas y otros (2012: figura 1).

Figura 20, tomado de Amalric y Dehaene (2016).

Figura 21 (abajo), a partir de Amalric, Denghien y Dehaene (2017).

Figuras 22 y 23, tomado de Dehaene y otros (2010).

Figura 24 (arriba), © Dehaene-Lambertz; (abajo), a partir de Monzalvo y otros (2012).

Figuras 25 y 35 (izquierda), tomado de Dehaene-Lambertz, Monzalvo y Dehaene (2018).

Figura 26, a partir de Xu y otros (2015).

Figura 27 (abajo), a partir de Kilgard y Merzenich (1998).

Figura 28, gentilmente cedido por Bruce McCandliss (Yoncheva y otros, 2010).

Figura 29 (abajo), gentilmente autorizado por Robert Zatorre (Bermúdez y otros, 2009).

Figura 30, reproducido con gentil autorización de Gyorgy Gergely; (arriba), tomado de Egyed, Király y Gergely (2013); (abajo), a partir de Gergely, Bekkering y Király (2002).

Figura 31, a partir de Held y Hein (1963: figura 1).

Figura 32, a partir de Kaplan y Oudeyer (2007: figura 3).

Figura 33, adaptado de Bekinschtein y otros (2009) y Strauss y otros (2015).

Figura 35 (derecha), tomado de Zoccolotti y otros (2005).

Figura 36, a partir de Chen y Wilson (2017).

El restante material iconográfico pertenece al dominio público y está libre de derechos. En todos los casos, se invoca el *fair use*.