

Agencia y control: el rol subcortical en las buenas decisiones¹

Agency and Control: The Subcortical Role in Good Decisions

Patricia S. Churchland*; Christopher L. Suhler**

*University of California San Diego, Estados Unidos
pschurchland@ucsd.edu

**University of California San Diego, Estados Unidos
clsuhler@ucsd.edu

[Traducción al español del texto: Churchland, P. S., & Suhler, C. (2014). Agency and control: The subcortical role in good decisions. En W. Sinnott-Armstrong, *Moral Psychology. Free Will and Moral Responsibility*, 4, 309-326.]

1. Introducción

Si los procesos cerebrales no conscientes contribuyen a la toma de decisiones, ¿qué modificación supondría, si es que la hay, en nuestra concepción tradicional de la diferencia entre tener el control y ser un agente responsable? Esta es la pregunta que motiva nuestra exploración en el presente capítulo. Antes de comenzar, una aclaración preliminar.

‘Libre albedrío’ es una expresión adornada con trampas de osos semánticas. Preferimos evitarlas. Una suposición que con frecuencia enreda a los incautos, es que las acciones pueden ser clasificadas en uno u otro de estos dos contenedores: elegida libremente o no elegida libremente. La toma de decisiones real, sin embargo, es mucho más compleja. A veces una persona tiene mucho sueño, mucha hambre, estrés crónico o está terriblemente asustada (Arnsten 2009). A veces una persona puede sufrir daño cerebral. El pulcro modelo ordenado en dos contenedores no se adecua para nada a la realidad de la toma de decisiones. Otra suposición limitante es que el libre albedrío implica una libertad con respecto a todos los

¹ Traducción: E. Joaquín Suárez-Ruiz. Revisión: Martín Daguerre. Esta traducción cuenta con la autorización de la editorial original (MIT Press).



Received: 17/09/2022. Final version: 16/11/2022

eISSN 0719-4242 – © 2022 Instituto de Filosofía, Universidad de Valparaíso

This article is distributed under the terms of the

Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License



CC BY-NC-ND

antecedentes causales de la decisión. Nadie podría sostener seriamente una noción de libre albedrío en este sentido. Pretender negar el libre albedrío por esas causas que “no deben ser causadas” es como negar lo que todos suponemos de hecho: muchas personas actúan de manera controlada gran parte de sus vidas². Las trampas de osos semánticas de este tipo abundan.

La capacidad de autocontrol, como se desprende de investigaciones en psicología y neurociencia, está ligada tanto a ciertos antecedentes causales, como al funcionamiento de estructuras y vías neuronales específicas. A su vez, el autocontrol es un fenómeno graduado. Se desarrolla a medida que el bebé madura y puede disminuir a medida que la demencia lo destruye. También puede verse afectado por factores como el estrés, el hambre, el frío y el agotamiento. Sin detenernos en este punto, comprendemos que estas consideraciones motivan un cambio que va de un lenguaje centrado en el ‘libre albedrío’ a uno centrado en el ‘control’³. Discutir sobre la metafísica esotérica del libre albedrío suele ser improductivo y, en cualquier caso, los asuntos metafísicos ya han sido bien discutidos en otros trabajos (Churchland 2002; Flanagan 2003; Dennett 2004).

Provisoriamente, definimos ‘control’ como la capacidad de un individuo para actuar de manera inteligente y adaptativa dentro de un entorno particular, con el fin de mantener una meta, diferir la gratificación y suprimir impulsos desventajosos. El control resulta crucial tanto para los asuntos en el dominio legal, como también para los debates filosóficos sobre la responsabilidad. A su vez, es importante para un conjunto más general de suposiciones con respecto a cómo pensar sobre uno mismo.

Debido a que, en este punto del desarrollo, no podemos definir ‘control’ con mayor precisión, resulta de utilidad ampliar esta definición provisional señalando algunos casos prototípicos en los que existe acuerdo sobre la aplicación del término (Johnson 1993). En el caso prototípico de la ‘acción controlada’ (agencia), se afirma que un humano adulto sano que está despierto y no tiene problemas cognitivos es un agente cuando entra en una panadería y compra una barra de pan. Su acción está bajo su control. Por el contrario, a un hombre que es sonámbulo y patea a su esposa se lo considera como alguien que no tiene un control total de sus acciones. A su vez, por un lado, aunque se considera que una persona que es adicta crónica a la nicotina no tiene control total sobre su elección de comprar un paquete de cigarrillos, sí se comprende que posee cierta responsabilidad respecto de esa acción. Por otro lado, una persona que emite una respuesta de sobresalto por un ruido fuerte e inesperado, patea una linterna y, en consecuencia, incendia un granero, tiene comprometida su capacidad de control y no se la considera responsable. Los prototipos existen donde la concordancia es máxima. Los casos difíciles surgen cuando la persona aún no es madura, no está completamente despierta, ignora algún punto crucial o sufre de daño cerebral, adicción o una condición psiquiátrica. Más allá de que en contextos como el legal se requiera de una respuesta definitiva, a veces puede no

² David Hume reconoció este punto en sus argumentos contra una concepción cartesiana de la elección.

³ Patricia Churchland señaló este punto en *Brainwise* (Churchland 2002).

haber una respuesta correcta sobre si cierta persona tenía o no el control de sus acciones. El control, como ocurre con muchos otros conceptos, tiene una estructura radial con grados decrecientes de similitud desde los casos centrales hasta los que se encuentran en sus límites borrosos⁴.

Recientemente, sin embargo, un creciente cuerpo de investigaciones sobre la prevalencia de la cognición no consciente y la influencia de los factores no conscientes en el comportamiento, sugiere que incluso los casos prototípicos de agencia no son lo que parecen (Wilson 2002). ¿Cómo puede una persona tener algún tipo de control sobre factores no conscientes? ¿Cómo pueden tales factores entrar en las deliberaciones y reflexiones conscientes de una persona? ¿Cómo pueden ser parte del proceso de sopesar razones y evidencia realizado por una persona? Y si acaso no puede tener control sobre factores no conscientes, es decir, si no puede hacerlos parte de sus deliberaciones conscientes, ¿cómo podría ser responsable de alguna acción? Como algunos críticos nos han dicho alguna vez, la cognición y las influencias no conscientes son “ajenas al agente”.

Con el fin de abordar estos hechos sobre el procesamiento no consciente, se presentan tres opciones posibles:

1. Nadie es responsable de nada; el control, y por lo tanto el libre albedrío, es una ilusión
2. El control se relaciona con “responder a razones” de manera consciente. Por lo tanto, el control (y el libre albedrío) es, en realidad, independiente del procesamiento no consciente.
3. En el sistema nervioso de los mamíferos, el control involucra procesos conscientes y procesos no conscientes. De hecho, la capacidad de control se observa no solamente en todos los mamíferos, sino probablemente en muchas aves (ver también Panksepp 1998).

La primera opción es absolutamente insostenible. En primer lugar, las comunidades humanas no dejarán de responder con desaprobación y castigo a quienes cometan agresiones, robos y tráfico de información privilegiada. Sencillamente se trata de una necesidad pragmática propia de la vida en grupos sociales. Responsabilizar a las personas es un elemento que permite sesgar sus decisiones futuras, una característica no trivial para los mamíferos sociales. Si acaso se pretendiese seriamente dismantelar el sistema de justicia penal sobre la base de que el libre albedrío es una ilusión, la justicia por mano propia ocuparía su rol. No

⁴ Es esencial evitar la trampa común, aunque seductora, de tratar de definir con precisión un concepto antes de que la ciencia haya progresado de manera lo suficientemente prudente (no fantasiosa) como para permitir tal precisión. Lamentablemente, muchos buenos proyectos se han detenido en el primer paso por el esfuerzo prematuro de forjar una definición precisa, condición la cual es como el de las polillas con el fuego para aquellos que se ganan la vida filosóficamente inventando extraños y divertidos contraejemplos a definiciones prematuras. La precisión en la definición de α y el progreso en la comprensión científica de α coevolucionan. Véase Churchland (2002).

es un resultado deseable. Lidar con malhechores se tramita mejor al interior de los límites del sistema de justicia penal. La base institucional para la asignación de responsabilidades es profunda y altamente sofisticada, y debemos seguir confiando en ella. En segundo lugar, los padres seguirán dando forma a los sistemas de recompensas de sus hijos con aprobaciones y desaprobaciones. No hacerlo sería absurdo. Sesgar las inclinaciones del niño es de lo que en gran medida se trata el aprendizaje social.

Tampoco es sostenible la segunda opción. Su supuesto subyacente equipara agencia con agencia consciente, y acción controlada con acción conscientemente controlada. El problema con esta suposición es que no cuadra con la realidad del comportamiento y cognición humanos. Una vez más, hechos molestos se enfrentan a ideas filosóficas favoritas. Es posible demostrar la influencia de actividades no conscientes en las decisiones conscientes. Las habilidosas acciones de un jugador de baloncesto, por ejemplo, suceden demasiado rápido como para que involucren una gran deliberación consciente sobre si lanzar o pasar el balón. Sin embargo, las decisiones del jugador son altamente controladas y voluntarias. Un chef habilidoso puede identificar cuál es el problema con un pastel fallido en lo que dura un parpadeo; un habilidoso capitán puede determinar cómo hacer frente a una ola que se aproxima, sin pasar por un proceso de razonamiento consciente; un obstetra experto puede responder de inmediato al darse cuenta de que el bebé tiene el cordón enrollado alrededor del cuello. Sin embargo, todas estas acciones muestran un impresionante nivel de control. Tampoco pueden explicarse como meras acciones de estímulo-respuesta. Hacen uso de un vasto trasfondo de conocimiento y competencia, es decir, hacen uso de habilidades. Adherir a la opinión de que el control se restringe a un dominio 'autónomo' de 'respuestas a razones' consciente y a una deliberación consciente, es tan poco realista como lo es una pava de cera.

Debido a sus absurdas consecuencias prácticas y a su incapacidad para cuadrar con los hechos del comportamiento y la cognición humanos, estas dos primeras opciones resultan poco atractivas. Afortunadamente, sin embargo, existe un sentido de 'control' perfectamente respetable, gracias al cual es posible distinguir entre los casos en los que una persona tiene un nivel de control normal, de aquellos en los que el control está comprometido o ausente. Por ejemplo, con un nivel de control magistral, Bernard Madoff ejecutó su esquema Ponzi durante años. Por el contrario, el paciente de Damasio, E.V.R., que poseía un daño importante en su corteza prefrontal ventromedial, sufrió pérdida de control después de su cirugía, pero no antes (Damasio 1999). El control cae a lo largo de un espectro de casos que va desde comprometido hasta normal y, tomando de referencia dicho espectro, los dos casos antes mencionados se encuentran separados por mucha distancia. Por tanto, es preciso distinguir entre, por un lado, grados de control y, por otro lado, diferentes aspectos del control. Las ciencias del control están comenzando a ayudarnos a comprender dichos mecanismos. No pueden, sin embargo, dar respuestas completas, tanto por el hecho de que las ciencias en sí mismas son claramente incompletas, como porque algunas cuestiones relativas a cuándo asignar responsabilidad

también conllevan juicios sobre el valor social. Por lo tanto, como podemos ver, la tercera de las opciones anteriormente enumeradas resulta ser la más convincente y, en consecuencia, es la que exploramos a continuación.

2. Control y procesamiento no consciente

Las cuestiones relativas a la neurobiología de los procesos conscientes pueden separarse en dos temas vinculados. El primero se relaciona con los mecanismos que resultan necesarios para ser consciente de cualquier cosa. Las regiones del tronco encefálico, el tálamo central y la corteza parecen jugar un rol dominante en esta función (Llinás 2001; Schiff 2008; Edelman, Gally & Baars 2011). Suponiendo que ya se es consciente en el sentido anterior, existen otros mecanismos que permiten ser consciente de algo en particular, como de ver una cara, escuchar el canto de una calandria o recordar haber metido un gol. Si estás en un sueño profundo o en coma, no eres consciente de nada en absoluto. Suponiendo que estás despierto, los contenidos particulares de la consciencia parecen seleccionarse según ‘lo que se necesita saber’ y, habitualmente, esa selección es en sí misma una operación no consciente (Sheth et al., 2012). No todo lo que forma parte de tu interocepción o de tu entorno externo está en tu consciencia: tu cerebro selecciona lo que es relevante o muy destacado. Aunque todavía no se comprende bien cómo funciona esto, actualmente se están logrando avances.

La cognición consciente tiene una capacidad bastante limitada y solo puedes resolver de manera efectiva un problema a la vez (‘qué voy a cocinar para la cena esta noche’ frente a ‘qué voy a escribir para el resumen’). Los procesos no conscientes, por el contrario, probablemente no estén limitados por este cuello de botella atencional y operen en paralelo. No obstante, incluso en este nivel, algunas operaciones estarán limitadas a ser procesadas en serie por el requisito de una acción a la vez. Los procesos conscientes probablemente operan más lentamente que ciertos procesos no conscientes, como sucede con el reconocimiento de patrones o la evocación de memoria asociativa. Cuando una decisión resulta ser de suma importancia, es posible reevaluar y reconsiderar conscientemente la que sería una respuesta automática al problema (véase, nuevamente, Sheth et al. 2012). Aun así, muchas decisiones del día a día no son cruciales y gran parte de la toma de decisiones se lleva a cabo sin mucha reflexión. A menos que algo salga mal, un comportamiento hábil, como hacer pan o conducir, generalmente no requiere de gran concentración. Habitualmente, el reconocimiento de que una decisión es de suma importancia o de que algo salió mal es en sí mismo el producto de procesos no conscientes de evaluación y verificación de errores, de la misma manera que otras formas de reconocimiento de patrones también son producto del procesamiento no consciente (Sheth et al. 2012). Se trata de un tipo de habilidad que puede desarrollarse a medida que se cultiva el escepticismo y la vigilancia en relación con ciertos tipos de condiciones y problemas. Tal como señala Kahneman (2012), “intuición es otro nombre que se le da a la experiencia”.

Además, tan solo algunos procesamientos no conscientes son relevantes para la cuestión de si acaso alguien tiene o no la capacidad de autocontrol. Nadie, ni siquiera el más ferviente defensor de un punto de vista del control basado en la ‘respuesta a razones’, se preocupa por el preprocesamiento en la retina, por ejemplo, o por el mantenimiento de la temperatura corporal normal, los cuales son realizados sin la necesidad de consciencia. A nadie le importa que tu cerebro, de manera inconsciente, averigüe y decida hacer un movimiento sacádico cada 300 milisegundos, un comportamiento que ha sido demostrado es inteligente y dirigido a un objetivo (Yarbus 1967; Zelinsky Rao, Hayhoe y Ballard 1997). Esto sucede con una miríada de otros procesos no conscientes, muchos de los cuales involucran decisiones. Pero entonces, ¿qué subdominios de la actividad no consciente parecen afectar al autocontrol?

En primer lugar, contingencias aparentemente menores en una situación particular pueden inclinarnos hacia una acción en lugar de otra, como ayudar a una persona a la que se le cayeron sus compras o sonreír a un transeúnte. Este dominio ha resultado atractivo para aquellos filósofos que tienden a ver la influencia no consciente de las contingencias situacionales como algo que socava cualquier presunción fuerte de autocontrol, incluso en condiciones aparentemente normales (ver más abajo).

En segundo lugar, el sistema de recompensas y el aprendizaje por refuerzo en general juegan un papel muy importante en el aprendizaje sobre el mundo físico y social. El sistema de recompensa/refuerzo, incluidos los ganglios basales y otras estructuras subcorticales y corticales, es crucial en el desarrollo de hábitos y habilidades mediante los cuales las personas pueden reprimir impulsos perjudiciales, generar y evaluar opciones, manejar el estrés, ordenar preferencias y tomar decisiones bajo restricciones temporales. Todas estas operaciones son importantes en las diversas manifestaciones del control.

Un tercer subdominio, dependiente del aprendizaje por refuerzo, atañe a la capacidad de tomar decisiones sobre el nivel de relevancia de algo, como la determinación de si una percepción, un recuerdo o una idea resultan o no relevantes para resolver un problema específico. A veces este juicio puede ser consciente pero, aun ocurriendo de esa manera, muchas cuestiones irrelevantes ya son previamente excluidas de forma no consciente. Las determinaciones de relevancia son decisiones de selección. Se relacionan con a qué se le presta atención, lo cual a menudo se resuelve de manera no consciente. A su vez, los contextos pueden proporcionar un marco para nuestras interpretaciones de lo que se dice o de un evento en particular, sesgándonos de una forma u otra. Si vas caminando apurado y alguien te dice “estás pasando cerca del banco”, entiendes que ‘banco’ se refiere a un banco para sentarse con el que podrías tropezar. Si estás en el Cabildo de Buenos Aires, cerca del Banco de la Nación Argentina, vas a entender que ‘banco’ significa ‘institución financiera’⁵ (Kahneman 2012). Hay otros dominios que también pueden ser importantes, pero a los fines de nuestro capítulo nos centraremos en estos tres.

⁵ N. del T.: se optó por introducir un ejemplo más cercano que el mencionado por los autores, situado en un Bank of America de la ciudad de San Diego, Estados Unidos.

3. Situando mis elecciones

En un artículo anterior (Suhler & Churchland 2009), desafiamos la idea de que habitualmente y en general las personas tienen, en el mejor de los casos, un frágil control sobre su comportamiento, cuestión que ha sido investigada por ciertos psicólogos sociales. Según algunos filósofos (Doris 2002; Harman 1999; Appiah 2008), el comportamiento humano está en gran medida a merced de factores situacionales externos que motivan nuestras respuestas (una posición que se conoce como ‘situacionismo’). A la luz de la hipótesis situacionista, nuestras acciones son con frecuencia meras consecuencias automáticas de factores ambientales y no el resultado de un agente con una capacidad de control robusta. Nuestros hábitos, temperamentos, carácter y objetivos pueden verse sacudidos por contingencias menores que ocurren en el momento.

A medida que estudiamos los hallazgos en cuestión, llegamos a la conclusión de que, una vez que se considera la totalidad de los estudios científicos relevantes para las cuestiones vinculadas con el control del comportamiento, el enfoque situacionista resulta ser mucho menos convincente de lo que inicialmente parece. Al centrarse casi exclusivamente en los hallazgos provenientes de la psicología social, los situacionistas han omitido muchas líneas de investigación de otros subcampos que indican que el control no es tan frágil como proponen los filósofos anteriormente aludidos. Las cosas se ven de un modo muy diferente cuando se equilibra la imagen con datos científicos que revelan la robustez del control, como la capacidad de mantener una meta a pesar de las distracciones, de diferir la gratificación, de detener una acción a mitad de camino, de desarrollar hábitos ventajosos y de contener impulsos. Esto se ve en humanos, pero también en monos, ratas y, resulta plausible predecir, en muchas otras especies.

La hipótesis del control frágil se manifiesta débil al compararla con la gran cantidad de hallazgos que demuestran la omnipresencia y sofisticación de la cognición no consciente y la persecución de objetivos (Bargh et al. 1996), el trabajo sobre el mantenimiento de objetivos y su base neurobiológica (Miller & Cohen 2001), los descubrimientos sobre las vías neuronales que sustentan la supresión de impulsos de acción y pensamiento (Aron et al., 2007), y la investigación sobre las limitaciones del control ejecutivo consciente (Baumeister 2005). A su vez, la investigación sobre la “laboriosidad aprendida” resulta importante para demostrar el rol que posee el sistema de recompensas en el refuerzo de un patrón de comportamiento, lo cual da como resultado la persistencia en la persecución de una meta (Eisenberger, Kuhlman y Cotterell 1992). La hipótesis principal aquí es que a través del sistema de recompensas, la sensación de gran esfuerzo se vuelve en sí misma gratificante. Esto implica no solo que una capacidad de control robusto es real, sino que también puede fortalecerse a través del aprendizaje por refuerzo. Esto no debería ser una gran sorpresa para los situacionistas.

¿Acaso el conocimiento en una circunstancia particular de posibles sesgos, así como también la habilidad para reconocer cuándo la situación requiere ser escéptico, hacen alguna

diferencia en cuanto a si las contingencias menores pueden ejercer algún efecto? El sentido común ciertamente así lo sugiere⁶ y, según argumentaremos a continuación, la evidencia neurobiológica lo respalda.

También registramos la preocupación por el hecho de que en muchos artículos citados en apoyo al situacionismo, no todos los sujetos evidenciaban poseer los efectos sugeridos y, con frecuencia, el tamaño de la muestra era pequeño. Notamos que, en general, las acciones en cuestión no eran muy importantes. Les preocupaba la decisión de si tomar o no un lápiz, por ejemplo, pero no de si acaso se debía abandonar la facultad de derecho o cambiar la aseguradora del vehículo. En resumen, estas preocupaciones sugieren precaución respecto de sacar conclusiones contundentes sobre la falta de control en sujetos sanos en general.

Desde la publicación de nuestro artículo, los desarrollos en psicología empírica exigen aún más cautela al interpretar y utilizar los resultados de la psicología social. Actualmente hay informes de fallas en la replicación de resultados en algunos de los estudios mencionados⁷. Además, el estudio de Bakker y Wicherts (2011) de 281 artículos en psicología, encontró que en el 18% de los artículos los resultados estadísticos se informaron incorrectamente y, en el 15%, al recalcular los resultados estadísticos se revirtió el hallazgo declarado.

También suscita preocupación la confesión de un psicólogo social holandés muy respetado y ampliamente publicado, Deiderik Stapel, respecto de que inventó datos para aproximadamente dos tercios de sus publicaciones⁸. Su fraude ha socavado los resultados que aparentemente muestran que usamos mejores modales en la mesa si hay una copa de vino sobre la mesa, o que las personas discriminan más si su entorno está desordenado, o que tendemos a ser más agresivos si comemos carne. Aunque estos supuestos resultados fueron muy publicitados en los medios, actualmente ya no hay razón para creer que sean ciertos.

Un segundo psicólogo social holandés de alto nivel, Dirk Smeesters, luego de que dos de sus artículos fueran retirados debido al análisis realizado por Uri Simonsohn⁹, ha sido declarado culpable por la Universidad Erasmus de Róterdam de 'selección de datos' y de no mantener registros adecuados. Después de este hecho, Smeesters ha dimitido de la Universidad Erasmus. En su artículo más reciente, Simonsohn (2012) describe en detalle las técnicas estadísticas que se utilizaron para analizar los artículos de Smeester y también las de otro reconocido psicólogo social, Larry Sanna, antiguamente de la Universidad de Carolina del Norte. Con mucho cuidado, Simonsohn muestra que también en varios artículos de Sanna

⁶ Véase esta discusión: http://www.sciencenews.org/view/feature/id/340408/title/The_Hot_and_Cold_of_Priming

⁷ Véase Pashler, Coburn y Harris (2012); Doyen, Klein, Pichon y Cleeremans (2012). Para obtener una respuesta, consulte <http://www.psychologytoday.com/blog/the-natural-unconscious/201203/nothing-in-their-heads>. Véase más en, <http://www.nature.com/news/replication-studies-bad-copy-1.10634>

⁸ Véase <http://www.nature.com/news/2011/111101/full/479015a.html>

⁹ Consulte el informe <http://bps-research-digest.blogspot.ca/2012/07/has-psychologist-been-condemned-for.html>. Para una discusión con Simonsohn, consulte <http://www.nature.com/news/the-data-detective-1.10937>

los datos parecen estar fabricados. Al momento de escribir este capítulo, Sanna aparentemente ha renunciado. En uno de los artículos de Sanna se abordó la observación de que existe una relación metafórica entre la moral y el terreno elevado. Sanna y sus colegas sospecharon que la relación quizás podría ser literal. Por lo tanto, testearon la siguiente predicción: los sujetos en una elevación más alta (p. ej., en un escenario elevado) se comportarán de manera más prosocial que quienes estén en una elevación más baja (p. ej., en la fosa de una orquesta), y en un nivel de prosocialidad media estarán aquellos espacialmente ubicados en una altura a medio camino de los dos anteriores. Sanna y sus colegas informaron que sus datos confirmaron su predicción.

Aún más preocupantes que los casos de absoluta incompetencia estadística y fraude son las formas más sutiles, y probablemente mucho más comunes, en las que los falsos positivos (hallazgos espurios) pueden resultar de las elecciones de los psicólogos sobre cómo recopilar, analizar y reportar los datos. Este problema es abordado en un importante artículo de Simmons y colegas (Simmons, Nelson y Simonsohn 2011), el cual detalla cómo los ‘grados de libertad del investigador’ pueden aumentar drásticamente la prevalencia de falsos positivos: el hallazgo de un efecto cuando en realidad no existe. Utilizados de manera crítica, los grados de libertad que describen, como el control por género, la elección del tamaño de la muestra (y de cuándo dejar de recopilar datos) y la exclusión de ciertos sujetos (por ejemplo, con valores atípicos) o de ciertas condiciones experimentales, son a veces herramientas útiles y apropiadas para favorecer el hallazgo de efectos genuinos en los datos. Sin embargo, cuando no se emplean con la mayor de las prudencias, también pueden utilizarse para obtener un efecto en los datos que, al menos nominalmente, están por debajo del umbral de lo significativo a nivel estadístico ($p \leq 0,05$). Además, como demuestran claramente los modelos de Simmons y sus colegas, la probabilidad de un resultado falso positivo aumenta drásticamente cuando se emplea más de uno de los grados de libertad.

Las dificultades empíricas recién mencionadas no implican que ninguno de los resultados en psicología social se mantenga. El problema, en cambio, es el de la incertidumbre. En ausencia de replicación de terceros y de nuevos análisis de datos, no podemos estar seguros de cuáles funcionan y cuáles no. No sabemos si acaso los filósofos situacionistas que confían en los datos de los psicólogos sociales han vuelto a analizar los datos sin procesar o replicado los hallazgos en cuestión, o si simplemente asumen que las conclusiones son genuinas. Desde nuestro punto de vista, la reciente avalancha de cuestionamientos a los datos de ciertos psicólogos sociales brinda aún más razones para ser cautos con respecto a las conclusiones situacionistas sobre el control frágil.

Owen Flanagan, en su cuidadosa disección del situacionismo y desde una perspectiva ampliamente filosófica y empírica, ha realizado algunas de las críticas más devastadoras (Flanagan 2009). Flanagan señala correctamente que las decisiones son, por supuesto, sensibles a los acontecimientos del entorno inmediato. ¿No es eso parte de lo que involucra el ‘responder a razones’, después de todo? Sin embargo, igualmente obvio resulta que las decisiones no se ven abrumadas por eventos irrelevantes. Las decisiones están sesgadas por

hábitos, rasgos de carácter, conocimientos previos y temperamento. A su vez, las decisiones pueden verse afectadas por otros eventos del historial reciente del agente que alteraron su valencia emocional y de estrés. También pueden verse afectadas por el hambre, el agotamiento, el miedo o la hipotermia que pudiese tener dicho agente. Como acertadamente señala Flanagan, los hechos acerca de los hábitos no implican que los factores contingentes deban considerarse irrelevantes. Este tipo de factores relativos a una situación en particular puede afectar la toma de decisiones. Pero tampoco, con seguridad, los factores contingentes son primordiales en toda toma de decisiones. La toma de decisiones es un proceso de satisfacción de restricciones y los hábitos, las habilidades y la experiencia son restricciones importantes (Churchland 2011; Litt, Eliasmith y Thagard 2008).

4. El Sistema de Recompensa: “Yo, Agente”

En nuestra discusión anterior sobre el control no consciente, discutimos una variedad de enfoques empíricos que tienden a respaldar la hipótesis del control no consciente. Sin embargo, una línea de trabajo crucial que no exploramos hasta ahora fue el papel del sistema de recompensas (positivas y negativas) en el establecimiento de lo que Aristóteles denominó “buenos hábitos de pensamiento y acción”. A veces, entre los filósofos contemporáneos, el hecho de que Aristóteles haya puesto tanto énfasis en los buenos hábitos ha suscitado cierto desconcierto. Prefieren, en cambio, centrarse en la deliberación consciente sobre ‘razones’ y en los principios presentes en la consciencia al momento de la decisión. Sin embargo, como estimó Aristóteles (cosa que los filósofos contemporáneos tienden a no hacer), las habilidades (sociales, para resolver problemas, para tomar decisiones) son crucialmente importantes en todo comportamiento, incluido el comportamiento social, y para alcanzar una vida buena. Aristóteles habría disfrutado de los recientes avances científicos en la comprensión del sistema de recompensa/refuerzo y de su rol indispensable en el desarrollo de capacidades de control socialmente ventajosas y adaptativas. En lo que sigue nos centraremos en este tema.

Entre muchos otros factores, dos fuerzas han moldeado poderosamente la evolución del cerebro: las ventajas del aprendizaje (por permitir reducir la incertidumbre y promover la adaptabilidad) y las ventajas de minimizar los costos de energía y tiempo. Se trata de dos factores que interactúan entre sí.

Reducir la incertidumbre (mejorar la capacidad de hacer predicciones precisas) es una restricción fundamental en la evolución del sistema nervioso (de todos los sistemas nerviosos). Para hacer predicciones sobre lo que sucederá luego, basado en lo que sucedió antes, los cerebros logran reducir la incertidumbre. Actualizan dichos circuitos de predicción dependiendo de lo que suceda en el ahora. Siendo que permiten guiar el comportamiento subsiguiente favoreciendo así la supervivencia y la reproducción, la capacidad de hacer y usar predicciones resulta valiosa para todos los animales.

Con el fin de hacer predicciones útiles en todas las escalas de tiempo (excepto en las más cortas), los sistemas nerviosos realizan cambios estructurales en sus circuitos. El crecimiento de espinas dendríticas neuronales (especialmente las de las neuronas piramidales), junto con la poda de aquellas que no contribuyen, es un componente importante del aprendizaje. El sistema de recompensa/refuerzo del cerebro utiliza ambas estrategias. Cuando el cerebro aprende el paradero espacial de fuentes de alimentos, a andar en bicicleta o a comportarse en un tribunal, se pueden observar cambios estables relacionados con el crecimiento o la poda de dendritas. Es importante destacar que cuando hay crecimiento, también debe haber expresión génica que favorezca la producción de las proteínas que constituyen los medios para dicho crecimiento. Por lo tanto, el aprendizaje y la expresión génica están estrechamente vinculados.

Aprender a perseguir un objetivo en lugar de otro implica evaluar las consecuencias esperadas de un plan y ajustarlo en consecuencia. Esto también debe hacerse en el contexto de la coordinación homeostática de necesidades e impulsos. Todos los cerebros de mamíferos (probablemente todos los cerebros de vertebrados) hacen esto. En lo que a la capacidad de aprendizaje por refuerzo en mamíferos respecta, las principales diferencias entre especies dependen de la complejidad relativa de la corteza prefrontal y de las relaciones entre los componentes del antiguo sistema de recompensa (ver figura 8.1). He aquí un ejemplo en el que la diferencia cuantitativa puede generar diferencias cualitativas: los cerebros humanos tienen una corteza prefrontal más grande y con más neuronas, en términos absolutos, que otros mamíferos terrestres. Resulta pertinente enfatizar esto porque nuestro objetivo aquí es caracterizar el rol que cumple el sistema de recompensa cuando sesga una elección inteligente. Al adornar la antigua organización del sistema de recompensa subcortical con refinada información cortical, podremos evaluar un plan por sus posibles consecuencias. Un *input* cortical más profundo habilita predicciones y evaluaciones más profundas. Unas metas pueden anidarse dentro de otras. Los planes pueden tornarse muy elaborados y las metas muy abstractas. Al basarse en patrones de causalidad previamente aprendidos, el cerebro puede ensamblar evaluaciones de las posibles consecuencias de un plan. Se trata de modelos “si... entonces” que pueden volverse muy sofisticados.

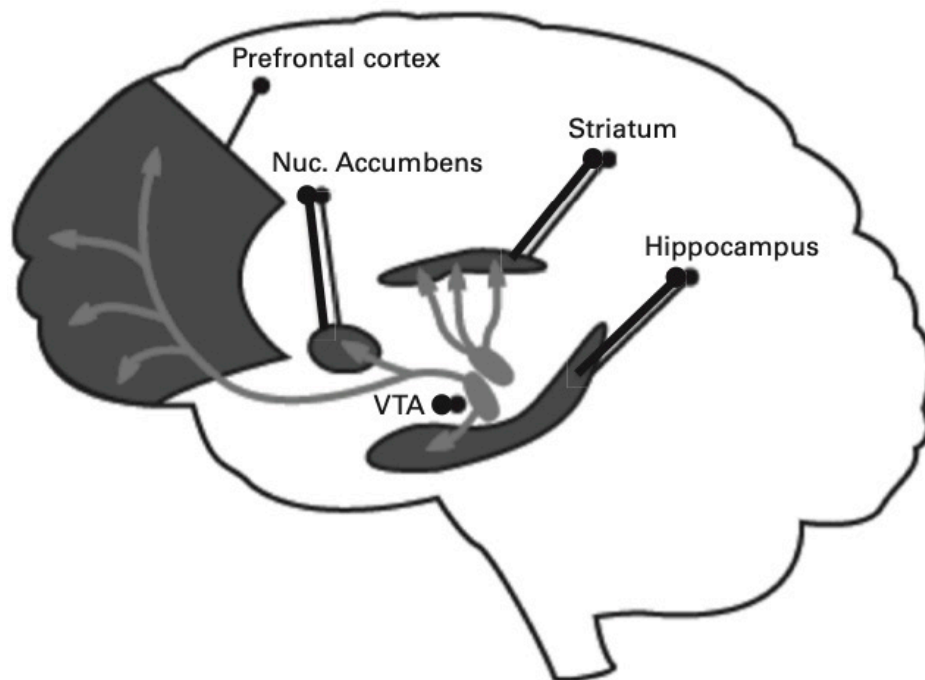


Figura 8.1. Esquema que muestra los circuitos principales del sistema de recompensas. ATV, área tegmental ventral.

Los ajustes en los circuitos reflejan si una acción produjo resultados positivos o negativos. En el sistema de recompensas también implican una predicción sobre lo que sucederá en una próxima ocasión similar. Un resultado positivo inclina al sistema de recompensas del cerebro a favor de esa acción; un resultado negativo lo inclina en contra (ver Thorndike 1911; Hikosaka 2010). No importa si posee un sistema nervioso simple o complejo, el aprendizaje de un organismo depende de que el sistema nervioso reaccione positivamente a algunos resultados de acción, negativamente a otros, y no reaccione (o reaccione poco) a otros. Hay grados de recompensa; una quemadura grave es algo realmente horrible, un pinchazo no es tan malo. Si el sistema nervioso de un organismo es insensible al dolor o al placer, o si sus circuitos neuronales no pueden modificarse (ver arriba), incluso cuando busca un resultado bueno o malo, tenderá a desaparecer. En nuestro mundo neurobiológico, la red de recompensa y aprendizaje es muy importante (Montague, Hyman y Cohen 2004).

Una predicción precisa aumenta las probabilidades de sobrevivir; desperdiciar energía no. Los animales deben gastar energía para encontrar comida y agua, para mantenerse calientes y seguros. Gramo por gramo, los sistemas nerviosos requieren muchas más energía que otros órganos. Es importante, por lo tanto, el hecho de que la modificación de los circuitos generada con el fin de reflejar las invariantes en el mundo físico y social, también ayuda a mantener bajos los costos de energía. Si tu sistema nervioso aprende que las frambuesas son sabrosas

y las cerezas silvestres son desagradables, no necesitas perder tiempo y energía juntando cerezas silvestres. El ahorro de energía significa menos tiempo cazando al azar, menos tiempo buscando comida en total y más tiempo disponible para mantenerte saludable y buscar pareja.

Así sucede en el ámbito social. Si tu sistema nervioso se ajustó para que evites tomar la comida de tu hermano mayor, dado que la última vez te golpeó, no tienes que sufrir dolor repetidamente por robarle su comida. Puedes encontrar otras estrategias alimentarias más eficientes a nivel energético. En humanos, la socialización de los jóvenes respecto de lo que es y no es aceptable implica aprobación y desaprobación, lo cual, a su vez, significa placer y dolor. Patear a tu compañero genera desaprobación; incluirlo en tu partido de fútbol genera aprobación. La ‘madre naturaleza’ podría haber puesto todo el conocimiento necesario en los genes, pero resulta mucho más fácil y rápido diseñar sistemas nerviosos que puedan sintonizarse con el medio ambiente, con el fin de permitirles aprender.

Nuestro sistema de recompensas responde a la aprobación y a la desaprobación, pero no sencillamente como una suerte de reflejo. Su abanico de respuestas se vuelve cada vez más complejo y sutil a medida que la corteza prefrontal se expande durante la evolución y se desarrolla durante la maduración de un individuo. La complejidad es aún más abrumadora cuando se tiene en cuenta su coordinación para con las respuestas homeostáticas, la capacidad de respuesta al estrés, el apego, la asociación y la toma de riesgos competitiva. A veces, si existen otras motivaciones que resultan ser lo suficientemente fuertes, puede minimizarse la desaprobación y, a veces, puede ignorarse la aprobación porque, por ejemplo, se percibe como poco confiable. Algunas personas son más tercas que otras; algunas están más dispuestas a correr riesgos que otras. De modo que los factores temperamentales también juegan un papel. A lo largo de la vida, pero especialmente en los jóvenes, la modificación lenta y adaptativa del circuito de las vías de recompensa prefrontal da forma a lo que sentimos y a cómo respondemos emocionalmente en contextos sociales (ver especialmente Del Giudice, Ellis y Shirliff 2011). En esto consiste el desarrollo del control. El control es antiguo y, especialmente en animales maduros, puede ser robusto frente a distracciones.

El condicionamiento clásico se comprende bastante bien, especialmente en especies neurológicamente simples (Carew, Walters y Kandel 1981; Carew, Hawkins y Kandel 1983; Hawkins, Abrams, Carew y Kandel 1983), pero hasta los últimos 15 años la neurobiología del aprendizaje por ensayo y error (reforzamiento) seguía siendo imprecisa. Ahora que se sabe que los ganglios basales desempeñan un rol crucial en esta forma de aprendizaje, la naturaleza exacta de dicho rol está siendo ampliamente explorada, tanto mediante registros a nivel unicelular en animales como utilizando tecnología de imágenes en humanos (Rangel, Camerer, & Montague, 2008). Aunque el campo ha florecido de manera espectacular, hasta ahora su impacto ha sido insignificante en la filosofía.

Una intuición conceptual de Read Montague contribuye a comprender, desde una perspectiva evolutiva, por qué el circuito de refuerzo/recompensa resulta tan importante en un animal que se abre camino en el mundo: En el circuito de refuerzo, algunos de los recursos

energéticos del animal se utilizan para responder a cierta recompensa (comer la sabrosa fruta que encontraste o lamerte las heridas resultantes de una pelea con un congénere), pero parte de ellos también son dedicados a modificar los circuitos con el fin de incorporar una predicción: si la próxima vez hago A, entonces la recompensa B probablemente ocurrirá nuevamente. Lo intenté, me equivoqué y mi cerebro codifica esa información a nivel estructural. Mi inversión energética genera ganancias (Montague 2006). Este resumen es una simplificación de la complejidad del proceso de aprendizaje, por supuesto, pero ilustra el punto no controvertido de que el aprendizaje por refuerzo ha sido fuertemente seleccionado en la evolución de los sistemas nerviosos.

Montague y sus colegas han explorado un aspecto adicional del aprendizaje por ensayo y error que va más allá de la recompensa y el castigo directos. Se trata del aprendizaje contrafáctico: un aprendizaje basado en la evaluación del resultado de la opción no tomada (generalmente llamado ‘error ficticio’). Consiste en aprender sobre “Lo que debería haber hecho en lugar de lo que hice” y requiere una comparación entre “Lo que obtuve” en la opción tomada con “Lo que podría haber obtenido” en la opción alternativa. Aunque esta comparación y evaluación parece ser en su mayor parte un asunto inconsciente, el sentimiento de decepción generado por la opción a la que se renuncia puede, por supuesto, sentirse conscientemente. Sus efectos pueden verse con bastante claridad en estudios de imágenes. Nótese especialmente que el registro del error ficticio es conceptualmente sutil y, según cualquier estándar razonable, inteligente. Consiste en el reconocimiento del valor que poseía el camino que no recorrí. De modo que esos recursos conceptuales en el trasfondo permiten poner en relación la opción alternativa con uno mismo y así comparar lo que no sucedió con lo que sucedió, como resultado de la elección tomada (para luego usar esa comparación, no necesariamente sino quizás, con el fin de ajustar las predicciones propias y tomar una decisión diferente la próxima vez). Los ganglios basales y la corteza prefrontal son nuevamente cruciales para estos cálculos, aunque la naturaleza precisa del rol que desempeñan en este punto las neuronas de la corteza prefrontal sigue sin comprenderse bien. A nuestro entender, el descubrimiento del error ficticio y de los roles posibles que podría tener en la forma en que los humanos aprenden a desenvolverse en sus ámbitos social y físico, es monumental.

Aquello que sea la ‘agencia’, precisa tener raíces profundas en el aprendizaje por refuerzo. Los ganglios basales, sorprendentemente, también son esenciales para la difícil tarea computacional de averiguar si fuiste tú quien causó un evento doloroso o placentero específico, y exactamente qué fue lo que hiciste para que ese evento se produjera (Redgrave, Vautrelle y Reynolds 2011). Por ejemplo, el perro de Patricia Churchland aprendió a abrir el pestillo de una puerta; tiró del cable interior que, a su vez, liberó el pestillo no visible del exterior y luego empujó para abrir la puerta accionada por resorte. ¿Cómo supo exactamente qué movimiento, entre los muchos que hizo, fue el relevante para abrir el pestillo no visible? Es probable que pienses que esta información está dada, que simplemente resulta obvia. Pues no lo es. Desde el punto de vista del cerebro, nada relacionado con la causalidad se da en ese antiguo sentido filosófico. Aunque no se sabe exactamente cómo se resuelve este tipo de problema computacional, se

sabe que la solución depende del uso de señales en *loop* cronometradas con precisión que van desde la corteza motora hasta los ganglios basales, los cuales representan la decisión motora que se acaba de tomar. Cualquier representación de uno mismo como un agente que abrió el pestillo tirando del cable, depende de cálculos que ocurren en los ganglios basales.

Estas funciones están ancladas por *inputs* fásicos, disparados por los sentidos, del neurotransmisor dopamina en el estriado dorsal, el estriado ventral (núcleo accumbens), la amígdala y la corteza prefrontal. Son las mismas estructuras que resultan importantes para el desarrollo de hábitos, de habilidades sociales y de la vida en general. Son importantes para tomar decisiones bajo incertidumbre, lo que incluye prácticamente todo lo que decides hacer. Son importantes para aprender cuándo reducir la velocidad y cuándo reflexionar más cuidadosamente sobre qué es lo que hay que hacer a continuación. Son importantes para evaluar por qué tomaste la decisión menos valiosa en lugar de la decisión alternativa. Son esenciales para reconocer qué es y qué no es relevante para resolver un problema (consulte la figura 8.2). Nuestra hipótesis es que cualquier animal con un sistema de aprendizaje de refuerzo saludable y en funcionamiento, y con un sistema saludable para mantener la homeostasis, un régimen normal de sentir placer y dolor, responder a estímulos ambientales y modificar planes, es capaz de “responder a razones” (ver de nuevo Del Giudice, Ellis, & Shirtcliff 2011). Aunque no utilice lenguaje, es capaz de generar modelos “si... entonces” de varias opciones. Una razón (una parte del modelo “si... entonces”), puede ser tanto una percepción o un recuerdo, como también el reconocimiento de un error ficticio, un sentimiento o una conexión asociativa. Aunque en los animales que usan lenguaje puede ocurrir, las razones no tienen necesariamente que expresarse en un idioma o ser parte de un silogismo lógico. A medida que la capacidad de la corteza prefrontal para generar modelos que guían el comportamiento se vuelve más compleja, las razones pueden volverse magníficamente complejas. En el fondo, sin embargo, la función esencial del sistema nervioso es coordinar todos estos factores de tal manera que permitan mantener la vida del individuo.

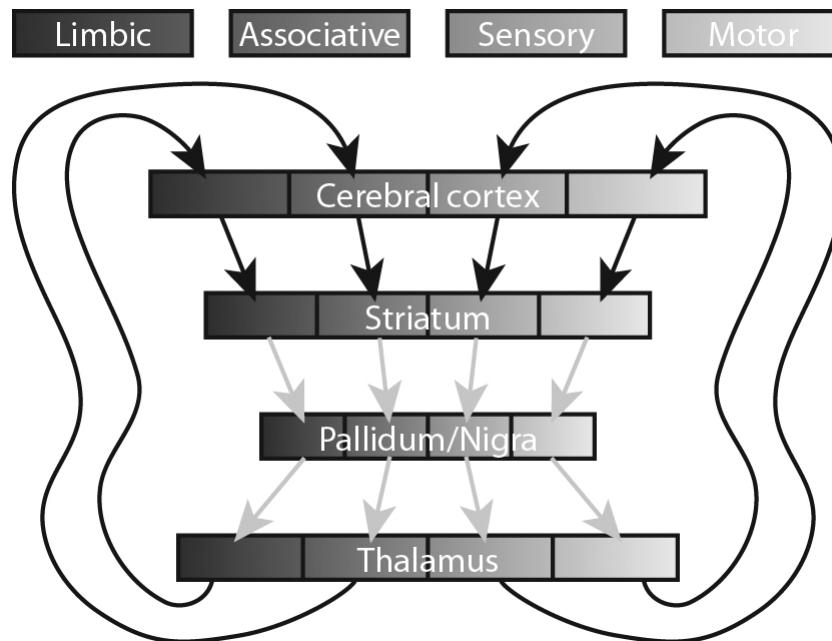


Figura 8.2. Esto ilustra, por un lado, la naturaleza circular de las vías entre la corteza y las estructuras subcorticales y, por otro lado, la naturaleza más o menos segregada de esas vías. Los tonos de gris que delimitan las estructuras subcorticales corresponden a los tonos de gris en los subcampos de la corteza. La relación entre la corteza y el cuerpo estriado es muy compleja y aún no se comprende bien (cortesía de Redgrave, Vautrelle y Reynolds 2011; con autorización).

En un artículo reciente, Kirk, Harvey y Montague (2011) reportan un resultado que vuelve sobre el postulado situacionista mientras muestra algo notable sobre las vías estriado-prefrontales. A grandes rasgos, el primer hallazgo consiste en que los no expertos pueden estar sesgados a favor de una obra de arte cuando el logotipo de apoyo de una empresa patrocinadora está cerca. El segundo hallazgo, sin embargo, es que los expertos están protegidos contra dicho sesgo. No porque se aíslen conscientemente, sino en virtud de su habilidad y experiencia pasada. A su vez, los expertos, pero no los novatos, reclutan un área de la corteza prefrontal, a saber, la corteza prefrontal dorsolateral. Kirk et al. proponen que en los expertos sucede una suerte de censura por la corteza prefrontal ventromedial de la tendencia al sesgo mencionado, la cual sí se observa en los no expertos. Esto sugiere que el cerebro experto ha evaluado la situación como propicia para el sesgo, estimulando un tipo de vigilancia y escepticismo para el que probablemente nos prepare algún componente del sistema de recompensa. Si bien el trabajo preliminar puede haber sido consciente y realizado durante las primeras etapas de aprendizaje sobre el mundo, este nos parece un maravilloso ejemplo de control no consciente. También es un fantástico ejemplo de cómo el aprendizaje pasado puede aparecer sensata y prudentemente en la forma en que actuamos en el presente. Es un ejemplo de cómo el

conocimiento y la habilidad pueden convertirnos en agentes más efectivos. Esto no necesitaría ser enfatizado si no fuese por la convicción de los situacionistas de que las contingencias de una situación determinan regular y habitualmente nuestras elecciones.

El conocimiento que despliegan los expertos en sus decisiones no les es “ajeno”; es parte de lo que los convierte en los agentes que son. Su experiencia es parte de lo que ellos son. Es lo que les permite actuar inteligentemente en una situación compleja. Cuando se toma una decisión, el conocimiento no necesita ser consciente para que sea efectivo. Este punto lo entendieron bien Aristóteles y Hume. Los datos de los niveles neurobiológico y psicológico indican que las funciones evaluativas en la generación de metas, la elección de metas, la selección de acciones, la evaluación de consecuencias, la actualización de predicciones y la actualización de metas, dependen de procesos inconscientes de evaluación/predicción, algunos de cuyos productos también se sustentan en procesos conscientes. Estos procesos conscientes están en deuda con el sistema dopaminérgico del mesencéfalo. Una cantidad asombrosa de datos desplegados por los diferentes métodos a los que tenemos acceso en la actualidad (modelos animales, modelos informáticos, estudios neurológicos en humanos, estudios de imágenes en humanos, experimentos de comportamiento), están enriqueciendo lentamente una comprensión de las estructuras dopaminérgicas del mesencéfalo como motores del aprendizaje por refuerzo. No, no resumen toda la historia. Sí, junto con la influencia de la gran corteza prefrontal humana, los resultados se vuelven enormemente sutiles y complejos. Sí, hay otros tipos de aprendizaje. Y sí, hay mucho que no entendemos sobre los mecanismos involucrados. Pero aun así.

5. Conclusiones

La evolución del cerebro de los mamíferos comprende vínculos estrechos entre el sistema de recompensa (positivo y negativo) y la corteza prefrontal. Dichos vínculos respaldan el comportamiento controlado, esto es, el comportamiento que es apropiado para el mantenimiento de la vida del animal en sus mundos físico y social.

El desarrollo de patrones de autocontrol durante la maduración es fundamental para la adquisición de habilidades que permiten favorecer elecciones prudentes en los mundos social y físico. Contrariamente a los filósofos que buscan una forma de libre albedrío milagrosamente independiente de la causalidad, este marco es muy aristotélico. Reconoce que el fin concreto reside en comprender correctamente la causalidad: en lugar de buscar de manera poco realista evitar la causalidad por completo, se trata de sesgar la selección de la meta y de suprimir los impulsos de manera que cumplan bien y a largo plazo su cometido.

Ahora bien, quizás te preguntes, ¿toda esa actividad del sistema de recompensas es realmente parte de mí? A esto respondemos, ¿cómo no va a ser parte de ti?, ¿qué serías sin él? Tus deliberaciones conscientes son lo que son debido a cómo están integradas con el resto de tu cerebro. Eres un individuo completo e integrado. Tu vida consciente es lo que es por la forma en que se relaciona con los productos de tu cerebro no consciente. Tus hábitos de

acción y de pensamiento son importantes precisamente por la razón que Aristóteles entendió tan bien. Cultívalas con cuidado, haz que funcionen a tu favor, porque son una gran parte de lo que te hace ser tú.

Agradecimientos

Agradecemos especialmente a Peter Redgrave por sus extensos y útiles comentarios a un borrador anterior. Muchas gracias también a Paul Churchland y a Owen Flanagan.

Referencias bibliográficas

- Appiah, K. A. (2008). Experimental philosophy. *Proceedings and Addresses of the American Philosophical Association*, 82, 7-22.
- Arnsten, A. F. T. (2009). Stress signaling pathways that impair prefrontal cortex structure and function. *Nature Reviews. Neuroscience*, 10, 410-422.
- Aron, A. R., Durston, S., Eagle, D. M., Logan, G. D., Stinear, C. M., & Stuphorn, V. (2007). Converging evidence for a fronto – basal – ganglia network for inhibitory control of action and cognition. *Journal of Neuroscience*, 27, 11860-11864.
- Bakker, M., & Wicherts, J. M. (2011). The (mis)reporting of statistical results in psychology. *Behavior Research Methods*, 43, 666-678.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230-244.
- Baumeister, R. F. (2005). *The cultural animal: Human nature, meaning, and social life*. Oxford: Oxford University Press.
- Carew, T., Hawkins, R., & Kandel, E. (1983). Differential classical conditioning of a defensive withdrawal reflex in *Aplysia californica*. *Science*, 219, 397-400.
- Carew, T., Walters, E., & Kandel, E. (1981). Classical conditioning in a simple withdrawal reflex in *Aplysia californica*. *Journal of Neuroscience*, 1, 1426-1437.
- Churchland, P. S. (2002). *Brainwise*. Cambridge, MA: MIT Press.
- Churchland, P. S. (2011). *Braintrust: What neuroscience tells us about morality*. Princeton: Princeton University Press.
- Damasio, A. R. (1999). *The feeling of what happens*. New York: Harcourt.
- Del Giudice, M., Ellis, B. J., & Shirliff, E. A. (2011). The Adaptive Calibration Model of stress responsivity. *Neuroscience and Biobehavioral Reviews*, 35, 1562-1592.
- Dennett, D. C. (2003). *Freedom evolves*. New York: Viking/Penguin.



- Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. New York: Cambridge University Press.
- Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS ONE*, 7(1), e29081. <https://doi.org/10.1371/journal.pone.0029081>
- Edelman, G. M., Gally, J. A., & Baars, B. J. (2011). Biology of consciousness. *Frontiers in Psychology*, 2(4), 1-7.
- Eisenberger, R., Kuhlman, D. M., & Cotterell, N. (1992). Effects of social values, effort training, and goal structure on task persistence. *Journal of Research in Psychology*, 26, 258-272.
- Flanagan, O. (2003). *The problem of the soul: Two visions of mind and how to reconcile them*. New York: Basic Books.
- Flanagan, O. (2009). Moral science? Still metaphysical after all these years. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 52-78). New York: Cambridge University Press.
- Harman, G. (1999). Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society*, 99, 315-332.
- Hawkins, R., Abrams, T., Carew, T., & Kandel, E. (1983). A cellular mechanism of classical conditioning in *Aplysia*: Activity-dependent amplification of presynaptic facilitation. *Science*, 219, 400-405.
- Hikosaka, O. (2010). The habenula: From stress evasion to value-based decisionmaking. *Nature Reviews Neuroscience*, 11, 503-513.
- Johnson, M. (1993). *Moral imagination: Implications of cognitive science for ethics*. Chicago: University of Chicago Press.
- Kahneman, D. (2012). *Intuition: Marvels and flaws: Fourteenth Annual Lynford Lecture*. <http://www.poly.edu/multimedia/videos/2012/03/14th-annual-lynford-lecture-daniel-kahneman>
- Kirk, U., Harvey, A., & Montague, P. R. (2011). Domain expertise insulates against judgment bias by monetary favors through a modulation of ventromedial prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 10332-10336.
- Litt, A., Eliasmith, C., & Thagard, P. (2008). Neural affective decision theory: Choices, brains, emotions. *Cognitive Systems Research*, 9, 252-273.
- Llinás, R. R. (2001). *I of the vortex: From neurons to self*. Cambridge, MA: MIT Press.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167-202.
- Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, 431, 760-767.
- Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience*, 29, 417-448. <https://doi.org/10.1146/annurev.neuro.29.051605.112903>



- Panksepp, J. (1998). *Affective neuroscience*. New York: Oxford University Press.
- Pashler, H., Coburn, N., & Harris, C. R. (2012). Priming of social distance? Failure to replicate effects on social and food judgments. *PLoS One*, *7*, e42510.
- Rangel, A., Camerer, C., & Montague, R. (2008). A framework for studying the neurobiology of value-based decision-making. *Nature Reviews Neuroscience*, *9*, 545-556.
- Redgrave, P., Vautrelle, N., & Reynolds, J. N. J. (2011). Functional properties of the basal ganglia's re-entrant architecture: Selection and reinforcement. *Neuroscience*, *198*, 138-151.
- Schiff, N. D. (2008). Central thalamic contributions to arousal regulation and neurological disorders of consciousness. *Annals of the New York Academy of Sciences*, *1129*, 105-118.
- Sheth, S. A., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D., et al. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature*, *488*, 218-221.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology. *Psychological Science*, *22*, 1359-1366.
- Simonsohn, U. (2012). *Just post it: The lesson from two cases of fabricated data detected by statistics alone*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2114571
- Suhler, C. L., & Churchland, P. S. (2009). Control: Conscious and otherwise. *Trends in Cognitive Sciences*, *13*, 341-347.
- Thorndike, E. L. (1911). *Animal intelligence*. New York: Macmillan.
- Wilson, T. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Harvard University Press.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.
- Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science*, *8*, 448-453.