

Derecho de la neurología: Libro blanco para el desarrollo de la
Carta de Derechos Digitales en el ámbito de la neurotecnología
Beatriz Gregoraci, Jorge Castillo Abella y Francisco Velasco Caballero (dirs.)
ISBN: 979-13-87913-93-3
Madrid, 2026
pp. 119-140
DOI: 10.37417/derecho-de-la-neurotecnologia/02
Marcial Pons Ediciones Jurídicas y Sociales
Editado bajo licencia Creative Commons Attribution 4.0 International License

CAPÍTULO II

LEER CEREBROS, LEER MENTES

Jesús VEGA ENCABO
Universidad Autónoma de Madrid
jesus.vega@uam.es

1.

Hace no muchos años lo que sigue se habría tomado como un mero experimento de pensamiento, sin atisbo alguno de que pudiera tener una posible realización; más un escenario de ciencia ficción que una trayectoria posible en nuestra cultura. Pero quizá estemos a punto de ver cómo se materializa este experimento con nuestras vidas y descubrir que los asuntos conceptuales sobre los que se apoya mi reflexión se muestran como lo que son, puros asuntos empíricos y técnicos. Lo que para muchos es el reducto último, privado, de los pensamientos, en sí mismo inaccesible, aparece como la construcción de una limitación técnica que estamos a punto de superar. El uso de técnicas de registro e inscripción de estados del cerebro¹ *podría* anular la protección última de los propios pensamientos ante la mirada y la influencia de otros. Es importante insistir sobre un punto: no nos empeñemos en una imagen de nuestras mentes como espacios esencialmente *pro-*

¹ Aunque no solo, pues hay desarrollos de técnicas fundadas en la aplicación de IA que, con base a datos conductuales, están también «accediendo» de alguna manera a los contenidos de los estados mentales. Véanse los resultados del siguiente estudio: CAZETTES, F., REATO, D., AUGUSTO, E. *et al.* (2025).

píos, pues el que lo sean en buena medida es consecuencia de una mera limitación técnica que será en algún momento superada.

He aquí el experimento. Como otros muchos experimentos de pensamiento ofrece muchas variaciones de interés. Imaginemos una sociedad en la que el avance de la neurociencia y de la neurotecnología sea tal que se convierta en rutina «leer» activaciones neuronales con todo detalle; será además posible «escribir» sobre los cerebros con base en el desarrollo de esas técnicas de «lectura». De hecho, se ha desentrañado el «código» de lectura y escritura de los cerebros, y eso ha permitido el desarrollo de ciertos dispositivos tecnológicos que registren las activaciones y las «traduzcan» en una especie de pantalla en lenguaje interpretable. Eso, lo que aparece en la pantalla en cada momento, es lo que uno piensa (se supone que uno podría igualmente «producirse» determinados estados mentales inscribiéndolos en el cerebro a través del dispositivo). Por supuesto, otros podrían —si se les permitiera hacerlo— leer mi mente/cerebro simplemente accediendo a lo que se muestra en la pantalla: es información públicamente accesible (también se podría intervenir en mi cerebro desde el dispositivo). Ciertamente, la sociedad ha promulgado leyes —o extendido y reinterpretado las ya vigentes— para proteger a los individuos de un acceso indebido a los dispositivos; solo uno mismo puede decidir sobre la accesibilidad al dispositivo y a sus contenidos. No obstante, en tus relaciones con los otros y en los intercambios más cotidianos, la *norma* es que cada uno deje «ver» los propios pensamientos: nos involucramos en ese escenario con la mente de los otros fundamentalmente mediante la lectura de estos dispositivos. Podríamos incluso conjeturar que así somos educados desde la infancia. No sería tampoco descabellado imaginar que el dispositivo también nos sirviera a nosotros mismos para leer los propios estados mentales. ¿Qué mejor forma de introspección que la inspección de las pantallas? Se podría especular incluso con el hecho de que la confusión que a veces experimentamos con relación a si estamos pensando tal o cual cosa quede disuelta tras la consulta del dispositivo, que registra la actividad cerebral correspondiente a tal o cual pensamiento (¡incluso si este es inconsciente!). En general, en esta sociedad se despliega todo un conjunto de nuevas prácticas de leer mentes *mediante el registro de información (objetiva)* sobre el funcionamiento de los cerebros. El logro es de tal magnitud que cabe esperar incluso una mayor *fiabilidad* en el proceso de la que podríamos reclamar siguiendo nuestras prácticas habituales de leer mentes.

Es este un escenario ciertamente distópico. Pero no me resulta fácil dar cuenta de qué hace de esta sociedad, que se ha dotado de las que podríamos ver como otras prácticas fiables de lectura de mentes (y de intervención y de manipulación de las mismas), un espacio distópico. Que se trata de otras prácticas es una constatación derivada del hecho indiscutible de que ya tenemos procedimientos (¿fiables?)

para hacerlo, a saber, nuestras prácticas ordinarias de leer y de manipular las mentes de los otros, y la propia mente. Regularmente sabemos lo que otros sienten o piensan, ellos mismos están dispuestos a hacérselo saber (y a veces a ocultarlo), y podemos manipular sus creencias, sus deseos, sus emociones y sentimientos *a partir de lo que creemos que son sus condiciones mentales específicas*. Compartimos un conjunto de prácticas que conforman nuestra «psicología popular», prácticas que, según algunos autores, se fundan sobre un conocimiento de principios teóricos que guían nuestra interpretación mental y, según otros, sobre una serie de habilidades (sociales) adquiridas. En cualquier caso, es una dotación de especie que nos hace ser los peculiares «animales maquiavélicos» que somos².

Toda interacción social, forma de comunicación o manifestación cultural lleva consigo el despliegue por parte de cada uno de nosotros de un repertorio de conceptos mentales y psicológicos adquiridos en el desarrollo. Los hemos aprendido fundamentalmente en contextos de interacción cara a cara³ y caracterizan de modo crucial una actitud participante y no meramente «táctica» con los otros. Una actitud participante, como bien nos recordaba Peter F. Strawson, es inseparable de estar dispuestos a determinadas actitudes reactivas ante la conducta de los demás y la propia⁴. Son prácticas que revelan el entramado normativo sobre el que se asienta la aplicación de los conceptos mentales y cómo estos permean nuestras vidas.

En lo anterior no veo nada que pudiera resultar controvertido. Más discutible resultará para muchos la siguiente afirmación, que me atrevo a calificar de ontológica: la forma de la mente humana (y seguramente de otros animales) está conformada por el particular modo en que compartimos trozos de mente con otros congéneres. Las prácticas de atribución (y autoatribución) de mente son el medio a través del cual se despliegan conceptos psicológicos cuyas condiciones de aplicación manifiestan rasgos lógicos peculiares (como la asimetría entre los usos de primera y de tercera persona)⁵. Estos no son rasgos

² Sobre la así llamada psicología popular y su relación con la «teoría de la mente», capacidad que nos permite explicar, interpretar y manipular la mente de los otros hay una extensa literatura. Un volumen ya clásico es DAVIES, M. y STONE, T. (eds.) (1995); véase igualmente CARRUTHERS, P. y SMITH, P.K. (eds.) (1996). Imprescindible es el artículo de GOLDMAN, A.I. (1993). Sobre la hipótesis de la inteligencia maquiavélica, se puede consultar también BYRNE, R.W. y WHITEN, A. (eds.) (1988).

³ Para la idea de que nuestra capacidad de mentalización se origina y se sostiene en interacciones cara a cara, véase la obra pionera de GOMILA, A. (2002). La misma idea se desarrolla también en un volumen reciente, PÉREZ, D. I. y GOMILA, A. (2021).

⁴ Véase el clásico STRAWSON, P.F. (1963). (Hay traducción castellana: *Libertad y resentimiento y otros ensayos*, Paidós, Barcelona, 1995, 37-8).

⁵ Este es un asunto tan central en la filosofía de la mente contemporánea que se encuentra discutido en numerosas obras. Entre ellas, destacaría el siguiente libro: WRIGHT, C.; SMITH, B.C. y MACDONALD, C. (eds.) (2000).

puramente contingentes del modo en que se nos dan los fenómenos mentales —capturados en la asimetría— sino que dan forma a esos fenómenos, al menos en tanto que implican vivencias que estamos dispuestos a reconocer en uno mismo y en otros.

Me gustaría reseñar, desde un inicio, que estas consideraciones, por sí mismas, no invalidan la coherencia del escenario que se dibuja en nuestro experimento de pensamiento. De hecho, el experimento permanece indefinido en algunos aspectos. Pues, en primer lugar, no está claro si la adopción de prácticas de lectura objetiva (no participante) de mentes, mediante inferencias soportadas en datos de activación cerebral, sustituiría a, o conviviría con, las prácticas ordinarias. Si, como he sugerido, estas últimas no son eliminables (pues dotación de especie), parecería obvio que deberían convivir⁶. Pero, en segundo lugar, esto no es suficiente para determinar la prioridad de una u otra al guiar nuestra conducta, pues si asumimos que existe una mayor fiabilidad en los procedimientos objetivos de lectura cerebral que en las tentativas y, a veces, frágiles figuras de la interpretación, nada impide que, con el tiempo, las prácticas ordinarias pierdan peso y se conviertan en residuales. En todo caso, y en tercer lugar, no ha de olvidarse que en las mismas prácticas ordinarias se enredan actitudes objetivas y participantes (tanto hacia uno mismo como hacia los otros) y que muchos de los aspectos de la vida mental que estaríamos dispuestos a declarar (sin mediar inferencia) se ven puestos en entredicho por datos (objetivos) de la conducta. En otras palabras, que ya vivimos en cierto sentido en un marco en el que la vida mental está abierta a la exploración objetiva de otros y de uno mismo, lo que hace que nada excluya que otorgar mayor peso a la lectura de los dispositivos que revelan, de manera abierta y transparente, datos sobre las mentes/cerebro sirva, ante todo, como un refuerzo a toda esa dimensión de acceso objetivo a la mente. Tendremos, a nuestra disposición, un mayor número de datos objetivos a partir de los cuales *realizar inferencias* que nos acerquen a una visión más ajustada y precisa de lo que es la vida mental de los otros e, incluso, propia.

Por tanto, hay varias vías abiertas en la interpretación que hagamos de este experimento de pensamiento. En la primera, la transformación es radical en la medida en que se socavan definitivamente las prácticas ordinarias de atribución. En la segunda conviven, quizá de manera tensa, pues es necesario integrar más «fuentes» objetivas de datos a la hora de interactuar con la mente de otros o tratar las propias condiciones mentales.

⁶ Aquí la cuestión se entrecruza con los debates sobre el alcance del materialismo eliminativo en filosofía de la mente, introducido en el artículo pionero de CHURCHLAND, P.M. (1981), pues uno podría defender la necesidad práctica de mantener el vocabulario mentalista (el que nos sirve para las atribuciones), pero que esto no dice nada sobre la ontología de lo mental (los estados, los acontecimientos y los procesos *qua* mentales han de desaparecer de nuestra ontología, y esto es lo que asume el programa de leer/escribir en el cerebro).

Me gustaría destacar una primera cuestión que se suscita al interpretar lo que se despliega en la pantalla del dispositivo (o en los datos básicos de registro de actividad cerebral): hemos de decidir si lo que ahí se refleja es, por así decir, la vida mental misma —mi vida mental en tanto que se manifiesta en los registros del dispositivo— o si no son más que otros datos *sobre* mí mismo, como los que indirectamente capturan mis gustos y preferencias, opiniones y decisiones: serían datos *sobre* mi vida mental en tanto que accesibles en tercera persona y que se pueden equiparar a otros muchos datos que están ya disponibles sobre mí, fisiológicos, conductuales, etc. Serían más datos personales, como aquellos que ya —de un modo u otro— invaden la infósfera y por cuya privacidad estamos seriamente preocupados⁷. Garantizar que estos datos y su difusión estén bajo mi control, tal que uno mismo sea el responsable y el gestor de qué datos sobre mi vida y bajo qué condiciones se vuelven disponibles, es una preocupación legítima a la que las legislaciones han de responder⁸.

Pero no olvidemos que nuestro experimento va un poco más allá, al menos en su versión más radical. Suponíamos que *progresivamente* las prácticas ordinarias de atribución de estados mentales iban dejando su lugar a estas prácticas de lectura, tal que, al menos por lo que concierne a la perspectiva que otros —y quizá también uno mismo— adoptarían ante lo que se despliega en la pantalla, esta no sería la de meros datos *sobre* mí, sino datos que, en la medida en que los hago accesibles, *constituyen* esa parte de mi vida mental que importa para nuestros intercambios (conversacionales o de otro tipo). Con ello quiero decir que hay una segunda cuestión a la que atender, relativa a cómo *tomamos* tales datos una vez que las prácticas de *apoyarse* en ellos se convierten en estables y dominantes. Es decir, que hemos de tener en cuenta el hecho de que la actitud ante esos datos puede ser muy diversa, pues no está claro en qué medida uno estaría o no dispuesto a *rechazar* como propio de la condición mental lo que aparece en la pantalla del dispositivo. ¡Y no digamos lo que legítimamente estarían los otros en posición de extraer de los datos que uno hace accesibles! Por ejemplo, ¿tendríamos razones para pensar que estos datos nos ofrecen una cierta *seguridad* sobre la vida mental de los otros o de uno mismo?⁹ ¿Mayor seguridad que otros datos objetivos a los

⁷ Sería ridículo por mi parte reseñar ni siquiera someramente trabajos dedicados a la protección de datos y la privacidad, con todas las derivaciones sociales, éticas, políticas y jurídicas. Me voy a limitar a recordar el capítulo que Luciano Floridi dedica al tema en su obra sobre la infósfera (FLORIDI, 2016).

⁸ Para una interesante discusión, consúltase WAJNERMAN PAZ, A. (2022).

⁹ El contraste seguridad/incertidumbre ha sido consustancial a una cierta imagen de lo mental, aquella que se articula a su vez en el contraste entre lo externo y lo interno. Wittgenstein ha dedicado reflexiones muy pertinentes a esta imagen, especialmente en sus últimos escritos sobre temas de psicología. Véase *Últimos escritos de filosofía de la psicología*, vol. 1 y 2 (WITTGENSTEIN, 2008).

que tenemos acceso? Podríamos aquí comparar el dispositivo con otros ya disponibles y que han extendido las prácticas del «autorrastreo» (*self-tracking*)¹⁰. Estas afectan directamente a cómo nos conducimos en relación con nuestros propios cuerpos, a la gestión de los tiempos de vida e, incluso, a las reacciones propias y de los otros en la vida en común. Y esto es así en la medida en que *autorizamos* los datos de alguna manera para que puedan formar parte de los procesos de decisión que nos guían en el control de nuestra propia conducta, es decir, que los tomamos como fiables y dignos de informar nuestras decisiones. Muchos de estos dispositivos sacan a la luz aspectos de nuestra vida a los que somos ciegos. Si no aceptamos que la mente es autotransparente (una asunción heredada del cartesianismo y que tiene ya escasos seguidores), ¿no sería un modo legítimo y apropiado de adentrarse en condiciones oscuras de nuestras propias mentes? ¿Y también de la mente de los otros? ¿Qué y cuánto se registrará y se revelará a través de estos dispositivos de lectura/escritura?¹¹ En el fondo de estas cuestiones, y de otras que veremos a continuación, se sitúa una pregunta sobre si podrá haber condiciones legítimas para tomar lo reflejado en los dispositivos como posibles informes *en primera persona* de lo que sucede en nuestras mentes. Solo esto acercaría, en alguna medida, las prácticas de leer mentes que estos dispositivos proporcionan a nuestras prácticas ordinarias.

No se puede olvidar que, en las prácticas ordinarias de atribución mental, un informe en primera persona se apoya sobre fundamentos muy distintos a lo que autoriza un informe en tercera persona. Es más, la presunción de verdad de los primeros (si uno es sincero) no se extiende a los segundos, ya que el propio juego del lenguaje psicológico (un aspecto de la gramática de los términos mentales) asume que desde datos «externos» no es posible reconocer con seguridad las condiciones mentales de alguien. Sin duda, uno podría pensar —legítimamente— que el hecho de que nuestros conceptos mentales exhiban esta «extraña» lógica interna, que requiere de condiciones de aplicación diferentes

¹⁰ Consúltese sobre su impacto el libro de NEFF, G. y NAFFUS, D., *Self-tracking* (2016).

¹¹ Los fenómenos mentales, como bien se sabe, son sumamente complejos y no es fácil ver cómo se podría lograr una «traducción» de las activaciones cerebrales en formas que sean cercanas a cómo se vive esta abigarrada heterogeneidad de lo mental. Pues, por ejemplo, ¿cuántos estados/eventos mentales se producen en un determinado lapso de tiempo? ¿Simultáneamente? ¿Se traducirían secuencialmente? ¿Se superponen? ¿Hay estados inconscientes? ¿Cómo «distinguirlos» a la hora de leerlos? ¿Cuánto duran? ¿Hay algunos que se puede suponer que son casi instantáneos? No quiero argumentar desde una perspectiva en la que la propia viabilidad del proyecto de leer/escribir mentes sea puesta en entredicho. Solo quiero reseñar que no me parece *obvio* que sea como tal inteligible y que, en todo caso, exigirá decisiones sobre qué y cómo será accesible o disponible de este modo. En ocasiones se muestra una excesiva preocupación sobre la posibilidad de leer pensamientos determinados aun cuando son mucho más accesibles estructuras emocionales y afectivas, que se verían modificadas fácilmente con intervenciones neurotecnológicas.

según las personas gramaticales, los hace susceptibles de eliminación en una concepción última de la realidad. Sin embargo, esto no los hace menos definitorios de una realidad humana compartida y que, en cierta medida, resulta insoslayable. ¿Es así? Una observación wittgensteiniana en sus escritos de filosofía de la psicología anticipa experimentos de pensamiento con tal alcance antropológico:

Pero piensa ahora en que llegases a una sociedad en la que, como podríamos decir, los sentimientos se reconocen con seguridad a partir de las apariencias (no usamos la figura de lo externo y de lo interno). Pero ¿no sería esto similar a llegar, desde un país en el que se llevan muchas máscaras, a otro en el que no se lleva ninguna o se llevan muy pocas? (Algo así como ir de Inglaterra a Irlanda). La vida es aquí simplemente distinta¹².

Me interesa principalmente esta última observación: ¿en qué sentido la vida sería distinta si nuestras prácticas de atribución de estados mentales no estuvieran sostenidas sobre conceptos psicológicos, contruidos de tal manera que parece asumida una imagen de separación de lo interno y lo externo, es decir, si estas prácticas suponen, a su vez, que todos los datos objetivos —desplegados en la pantalla del dispositivo— constituyen de modo suficiente un acceso a la vida mental de uno mismo y de los otros? ¿Es esta una vida distinta en el sentido de pertenecer a una diferente cultura, como sugiere la comparación wittgensteiniana? Sería así compatible con la *forma* de vida humana; bastaría con acomodarse a esas prácticas en las que hemos ido eliminando una máscara tras otra, hasta no dejar ninguna o, al menos, no dejar aquellas que se despliegan por el hecho de tener conceptos con tal forma gramatical. ¿O el experimento mental wittgensteiniano va más lejos y la transformación alcanza la forma *humana* de vida?

No puedo entrar aquí en la significación de la crítica wittgensteiniana de la figura de lo interno y lo externo. A mí no me cabe duda de que el objetivo de Wittgenstein no era dejar de aplicar esta figura sino dismantelar algunas de sus aplicaciones por ser fuente de confusiones (y no solo filosóficas)¹³. Mi apuesta es que el texto wittgensteiniano antes citado apunta a algo más que un «cambio de cultura»; desfigurar en nuestras prácticas los conceptos psicológicos afectaría al marco general de la vida humana. Los conceptos no son simple reflejo de una realidad sino que le dan forma¹⁴. Y Wittgenstein es muy consciente de lo que significaría sustituir *conceptos flexibles*, como los que pueblan nuestro vocabulario sobre lo mental, por *conceptos rígidos*, como

¹² WITTGENSTEIN (2008, p. 286-7).

¹³ Véase el § 424 de sus *Investigaciones Filosóficas* (Crítica, Barcelona, 1988, p. 305): «La figura está *ahí*; y no niego su *corrección*. Pero ¿cuál es su aplicación?».

¹⁴ Baste con esta anotación wittgensteiniana de nuevo: «El operar con conceptos impregna nuestra vida» (*Últimos escritos* II, 51, ed. Castellana, p. 318).

los que se sostienen sobre códigos descifrables. Estos conceptos articulan una forma de vida en la que las asimetrías entre la primera y la tercera persona, que solemos glosar —equivocamente— como un contraste entre seguridad interna e inseguridad externa, dibujan una imagen en la que lo mental exhibe una cierta indeterminación. Leer una mente bajo la premisa de que las señales son prueba objetiva de tal o cual estado de pensamiento o de sentimiento implica cortar lazos que tensan internamente los conceptos bajo los cuales habitamos la realidad mental. Es a esto a lo que se refiere Wittgenstein con su observación, al hecho de que la realidad no se hará habitable bajo las mismas premisas. De hecho resultará de alguna manera *inconcebible*¹⁵. Lo que se pone en juego es en qué medida el tipo de objetividad que asumiríamos con esas prácticas es algo que, dentro del marco general de la vida humana, resulta concebible. Muchos querrían afirmar que no solo es concebible sino que estamos en vías de realizarlo y que la pregunta es más bien si es *aceptable*, normativamente, dar pasos que afecten tan sustancialmente a nuestra forma de relacionarnos bajo la forma impuesta por tales o cuales conceptos¹⁶.

2.

Pero ¿qué importancia hemos de otorgar a reflexiones de este tenor, ocupadas más bien en la centralidad de los conceptos mentales en nuestro modo de existir, para encarar los desafíos —para muchos, a tomar muy en serio— planteados por los desarrollos más recientes de las neurotecnologías? La aparición y la extensión continuada de muy distintas técnicas y herramientas neurotecnológicas en tratamientos terapéuticos y en potenciales aplicaciones comerciales plantean la necesidad y/o la conveniencia de proteger la mente/cerebro de intervenciones indeseadas¹⁷. Lo que hasta hace pocas fechas parecía, como

¹⁵ Con razón se insistirá en que no es de ninguna manera inconcebible. De hecho, no parece imposible empíricamente hablando, como demuestran las investigaciones recientes. Impaciente, nuestro interlocutor podría preguntarse: ¿qué sentido tiene poner el énfasis en que esta transformación *conceptual* pudiera tener consecuencias decisivas en la forma *humana* de vida? La resistencia es legítima, más aún razonable. No obstante, animo a pensar en lo que cada uno de nosotros hacemos —en el despliegue de tal repertorio conceptual— cada vez que *compartimos trozos de mente* con otros. Es ese un modo de *habitar* una realidad que nos es común y que, en cierta manera, dejaría de estar disponible si se deforman (*¡hasta la rigidez de las herramientas objetivantes!*) nuestros frágiles y maleables conceptos mentales.

¹⁶ Desgraciadamente tampoco es este el enfoque más generalizado, pues el foco de atención se ha dirigido hacia la protección de los cerebros, asumiendo que esto es suficiente para garantizar la libertad cognitiva, en todos sus aspectos sustanciales. Algunos autores ya han señalado que encarar la protección de las mentes solo desde una perspectiva neurocéntrica está plagado de problemas o, al menos, está sujeto a muchas limitaciones. Véase por ejemplo BUBLITZ, J. C. (2024).

¹⁷ Quizá la presentación de FARAHANY, N. (2023) de las aplicaciones y sus desafíos sea la más sistemática.

hemos visto, inconcebible, es decir, un acceso e intervención directos a los pensamientos y a los sentimientos de cada uno de nosotros, resulta ser ahora una ventana abierta de oportunidad de control por medio de instrumentos que registran y que modifican la actividad cerebral¹⁸. Si la mente, como ese «lugar» al que retirarse, ya no puede ser considerada como un santuario interno al que otros no *pueden* tener acceso, en cierto sentido «inalienable», parecería que nuevos y graves riesgos se ciernen sobre nosotros. Si no de modo general, sí al menos en una cierta tradición del individualismo liberal, se ha identificado en el «yo interior» no solo lo más valioso del ser humano sino también la fuente misma del valor¹⁹, pues ahí reside nuestra capacidad de conformar nuestras ideas, de moldear nuestras emociones y sentimientos, de afirmar nuestros deseos y nuestra voluntad en toda libertad, sin intromisiones, sin interferencias indebidas. Sin garantías de que otros no puedan acceder a esa vida interna difícilmente se puede asegurar lo que se ha dado en llamar nuestra libertad cognitiva, constitutiva de todo lo que es más precioso y de valor para el ser humano²⁰.

Distintas formas de dualismo mente/cuerpo (cerebro) han intentado a lo largo de la historia dar cuenta, en sentido ontológico, de la inaccesibilidad. Recordemos cómo esa retirada al espacio interior de la subjetividad funda para Descartes el modo de ser de la mente, lo que a su vez garantiza que ni siquiera un demonio maligno podrá engañarle sobre su propia existencia. Es como si el pensamiento fundamentara sobre sí mismo su existir, sin dependencia de nada externo (salvo Dios), y no estuviera sometido a las determinaciones de otras formas de lo real. Pues ni siquiera el hecho, que el propio Descartes reconoce, de que hay una interacción (causal) entre la mente y el cuerpo/cerebro socava esa independencia ontológica. Que se produzca un estado/evento mental a través de las afecciones del cuerpo —o una intervención sobre el cerebro— no afecta a lo que es el percatarse mismo constitutivo de tal estado, que es lo único que lo caracteriza en tanto mental. Es obvio, sin embargo, que si todos los estados mentales tuvieran correlatos neuronales específicos el control externo de la vida mental de uno no nos pertenecería en la medida en que fueran resultado de la intervención del demonio maligno o del neurocientífico, por más que cada uno de nosotros los reconociera como los que uno está teniendo al estar en ellos²¹. El dualismo clásico parece así delimitar y contras-

¹⁸ De ahí que hace ya casi una década se plantearan disposiciones de naturaleza ética e incluso jurídica para reforzar la protección de la mente/cerebro. Por ejemplo, IENCA, M., ANDORNO, R. (2017) y YUSTE, R., GOERING, S., ARCAS, B. *et al.* (2017).

¹⁹ Se puede repasar la historia que ha trazado C. Taylor en su clásico ensayo *The Sources of the Self. The Making of the Modern Identity* (1989).

²⁰ Véase LAVAZZA, A. (2018).

²¹ Es este un punto importante al que no siempre se presta la debida atención. En el fondo de estas cuestiones no importa nada el que uno adopte puntos de vista dualistas o cualquier otra forma de materialismo fisicalista. Los temas de interés se plantean ya a partir del hecho funda-

tar la inaccesibilidad del yo y de sus estados (mentales) para los demás, por un lado, y la accesibilidad (y manipulabilidad) del cerebro y sus estados, así como la de otros cuerpos, por otro. Pero es ya claro que la mera constatación de correlaciones neuronales sistemáticas con estados mentales específicos quiebra cualquier atractivo de establecer el contraste en estos términos. Los estados mentales serían *accesibles* a través de la detección de sus correlatos, ¡incluso si fueran estados ontológicamente diferentes, tal y como reclama el dualista, de las activations cerebrales!

La inaccesibilidad no está garantizada ontológicamente. Y en estos momentos se vislumbra la posibilidad de que sea asaltada por medios técnicos. Esto significa que nada puede ser visto como *esencialmente* interno y que, si la imagen de lo interno/externo había contribuido a mantener la idea de que en lo interno se sostenía el valor mismo de la libertad cognitiva de los sujetos, ahora hay que clarificar cuál sea su valor y la protección de la misma se hace más urgente²². Desde mi punto de vista, un aspecto al que se presta escasa atención en las discusiones tiene que ver con una lectura adecuada de lo que implica la imagen interno/externo, y qué consecuencias tiene para una ontología de lo mental que reconozca aún lo subjetivo y las formas de valor que lo acompañan, inseparables de nuestras formas de compromiso (práctico) y de los aspectos expresivos de lo mental²³. Es esta una ontología de lo mental que se asienta sobre las exigencias de nuestras prácticas, de aquellas en las que ponemos en juego nuestros conceptos psicológicos. Mi acercamiento a estas cuestiones es ciertamente oblicuo;

mental de las correlaciones estados cerebrales/estados mentales. Quizá algunos se vean tentados a restaurar asunciones cartesianas a partir de la idea de que el núcleo del yo tiene que ver con sentirse uno mismo y sentir los estados mentales como propios. Algo así está en juego en los debates sobre la identidad y la integridad mentales, nociones poco definidas y cuya aplicación en términos jurídicos no está exenta de controversia. Algunas referencias son las siguientes: BUBLITZ J-C. (2020); LAVAZZA, A. y GIORGI, R. (2023); CASSINADRI, G. (2025).

²² Creo que los debates sobre esta noción de libertad cognitiva han sido muy fructíferos, aunque no está claro que haya una noción unívoca sobre qué se entiende por tal cosa. El término lo introdujo BOIRE, R. G., «On cognitive liberty», *Journal of Cognitive Liberties*, 4 ensayos entre 1993 y 2003 (disponibles en línea, accedido: 22 de noviembre de 2025: <https://www.cognitive-liberty.org/ccle1/1jcl/1jcl7.htm>). Interesante es también BUBLITZ, C. (2015) y SOMMAGGIO P.; MAZZOCCA, M.; GEROLA, A. y FERRO, F. (2017). FARAHANY (2023) en el libro ya citado resume las protecciones requeridas para la mente/cerebro en forma de libertad cognitiva, aunque el término podría remitirse más concretamente a las capacidades de autodeterminación que cualquier sociedad liberal considera que han de ser protegidas para ejercer incluso los propios derechos. Así lo ha argumentado ampliamente J. C. Bublitz en una serie de artículos que comienza con el pionero BUBLITZ (2013).

²³ Solo algunos de los autores que se han preocupado por cuestiones éticas y jurídicas de protección de la mente/cerebro han apuntado a hechos como este. Véase por ejemplo BUBLITZ (2023), quien reconoce que cualquier recomendación, normativa o establecimiento de derechos en relación a las neurotecnologías debe adoptar un punto de vista que atienda no solo a los aspectos que tienen que ver con los cerebros sino también con, por ejemplo, la dimensión subjetiva de lo mental.

no pretende fundamentar normativamente ni las necesarias regulaciones éticas ni los legítimos derechos que hay que salvaguardar. Es un acercamiento que incide en que hay un rasgo ontológicamente significativo de lo mental del que no puede prescindirse en cualquier argumento sobre cómo ha de reconocerse lo mental como dotado de valor: es lo que denominaré la dimensión expresiva que está en el trasfondo de aspectos valiosos a preservar, como la tan traída y llevada privacidad o la autodeterminación y libertad cognitivas, que se asientan sobre el ejercicio de la autoridad en primera persona. Son todos estos aspectos un efecto de estar insertos en estas prácticas conformadas por el tipo de conceptos que ponemos en juego y, para mí, es esto lo que está amenazado de profunda distorsión en un despliegue descontrolado de las neurotecnologías.

Hay un acuerdo generalizado en que tanto el registro de la actividad cerebral como su modificación suponen una amenaza a la privacidad mental. Tampoco debería sorprender que sea la cuestión de la privacidad lo que más discusiones haya suscitado en la filosofía a lo largo de la historia. ¿Qué forma de la privacidad se ha de proteger de las indebidas intervenciones neurotecnológicas? La discusión en las últimas décadas ha sido enconada, más aún tras las reflexiones wittgensteinianas sobre los lenguajes privados²⁴. Y no es descabellado afirmar que una cierta noción de la privacidad de la vida (mental) de cada uno ha configurado la filosofía moderna en todos sus aspectos fundamentales. La cuestión, como es bien sabido, es inseparable del problema escéptico sobre el conocimiento de la mente de otros, sobre las condiciones de su accesibilidad o inaccesibilidad. Ya he insistido en la importancia de esta cuestión en el experimento mental que he planteado, pero no estará de más volver sobre ello con un poco más de atención.

La raíz del problema no es otra que la fuerza de algunas de las intuiciones sobre las que se apoya el escéptico. Este se aprovecha para su argumento de hechos muy básicos, como el reconocimiento de que no tenemos nunca literalmente el *mismo* sentimiento o pensamiento que otro, incluso si cabe decir que podemos sentir o pensar lo mismo (estados distintos con el mismo contenido). Este hecho tan básico, al que apela el escéptico, nos interpela a causa de su significación en nuestra vida común, pues entre las cosas que nos importan especialmente está el que cuando alguien tiene dolor sea alguien en particular quien lo tenga. El escéptico juega su juego a partir de la idea de que no podemos saber lo que alguien está sintiendo porque no podemos tener el *mismo* sentir (dolor). Es el que alguien sienta dolor y que nadie lo sepa *ni pueda* saberlo lo que hace que el escéptico pueda insistir en sus afirmaciones. Como muy bien ha visto Stanley Cavell²⁵, toda

²⁴ Principalmente en su clásico *Investigaciones Filosóficas* (WITTGENSTEIN, 1988).

²⁵ Es un tema que aborda en su artículo «Knowing and Acknowledging» (CAVELL, 2002).

la discusión se enreda una vez que se sustancia en la idea de si tener el mismo dolor es o no relevante para establecer si podemos saber lo que otro siente.

Desde mi punto de vista, en el debate se entremezclan tres sentidos de privacidad, que podríamos denominar ontológico, epistémico y expresivo. Veamos cada uno de ellos por separado:

Privacidad ontológica: en esta forma de privacidad, está en juego la identidad de los estados tal que podemos decir o no que tenemos las *mismas* experiencias o los *mismos* estados mentales. Cómo se explique esta forma de privacidad es menos decisivo, bien mediante la postulación de objetos privados (punto de las críticas wittgensteinianas), bien mediante una dependencia fuerte de la identidad de un yo. Lo que sí parece ya claro es que el mero hecho de que estos sean *mis* estados mentales tal que otro no puede tenerlos no tiene por sí mismo consecuencias epistémicas de ningún tipo²⁶. Es obvio que nadie puede tener *mi* brazo, pero de ahí nada se sigue respecto de si solo yo puedo conocerlo. Esto quiere decir que del mero hecho de que yo solo pueda tener tales estados no se sigue si otros pueden o no conocerlos.

Privacidad epistémica: nos referimos aquí a si solo uno mismo o también otros pueden saber que uno tiene tales o cuales experiencias o estados mentales. La condición de privacidad epistémica se puede expresar al menos de dos maneras: o bien no puedo saber de modo alguno (es inconcebible que yo pueda saber lo que otro piensa, etc.) o bien no puedo saberlo del modo en que él lo sabe (directamente, de modo privilegiado, etc.)²⁷. Sin prejuzgar una solución al problema escéptico, no es menos cierto que nuestra experiencia común en la mutua atribución de estados y actitudes mentales se apoya sobre una convicción de que no estamos de modo generalizado en el error y de que normalmente sabemos si otro siente dolor o piensa tal o cual cosa. Además, también sin renunciar a la idea de que haya un modo diferente en que «sepamos» cuáles son los propios estados mentales, no es menos cierto que no hay una explicación obvia de su especificidad; decir que los conocemos de modo directo, por ejemplo, mediante un proceso de autoobservación está lleno de trampas filosóficas. En cualquier caso, está por ver que haya algo que los haga intrínsecamente privados en términos epistémicos.

Privacidad expresiva. Pero hay un sentido de privacidad que se revela en nuestras densas prácticas de uso de los conceptos mentales

²⁶ El escéptico, por lo que he dicho en los párrafos anteriores, no aceptaría esto sin más. Mi objetivo aquí no es responder al escéptico cuanto distinguir los diferentes sentidos de privacidad de los estados mentales.

²⁷ Se puede consultar para una discusión ya clásica el artículo de AYER, A. J., «Privacy» (1963).

y que, en cierto modo, sí les es esencial. Son los contextos de interacción en los que participamos los que muestran que está en nuestro poder mantener ocultos, sin exhibir, aspectos de nuestra vida mental. Y que son lo que son por este hecho mismo. No dominamos algunos de estos conceptos —pongamos el concepto de creencia— más que si está en su sitio esta capacidad de «reserva» mental. No es una cuestión, por tanto, de que nunca podamos conocer el sentir de otro porque nos está oculto, sino de que muchas veces no lo sabemos porque nos lo ha ocultado²⁸. Este es el sentido de privacidad que, por un lado, se funda sobre un rasgo de lo mental derivado del hecho de la constitución expresiva (al menos) de algunos de nuestros conceptos psicológicos y, por otro lado, apunta a establecer por qué resulta importante y valioso para nosotros persistir en el ocultamiento (a otros e incluso a uno mismo)²⁹. Por tanto, hay una noción de privacidad expresiva que se traduce en una capacidad de manifestar la vida mental según formas distintas y complejas de expresión, según sea requerido en las interacciones con otros. La privacidad *completa* no es sino la posibilidad del grado cero de expresión, una posibilidad que no se da nunca como tal, pero que enmarca la forma misma de la interacción con otros bajo conceptos mentales específicos, los nuestros.

Es la privacidad expresiva la que se pone en juego en el modo de tratarnos como sujetos y no meramente como fuente de datos objetivos³⁰. Incluso el hecho de revelar algo inintencionadamente, como cuando hay una expresión que se escapa a nuestro control pero revela algo de nosotros, refleja una comprensión del otro como yendo más allá de la facticidad de los hechos. Lo revela *a su pesar*, pero esa revelación manifiesta su subjetividad como un aspecto que quizá el propio sujeto estaría dispuesto a reconocer en tanto propio. Preservar la identidad de la mente (¿cerebro?) tiene mucho que ver con no forzar ni violentar la manifestación de lo que es considerado como propio, como siendo importante para uno, tal que es relativamente inaccesible de acuerdo a ciertas prácticas de ocultamiento en las que se asienta la aplicación sistemática de los conceptos mentales. El mero hecho de pensar, con independencia de su contenido, es lo que se quiere preservar, y este pensar está consustancialmente ligado a que pueda permanecer oculto³¹. Pero además, este sentido de privacidad explica por qué

²⁸ CAVELL, S., *op. cit.*, P. 257.

²⁹ Quien mejor ha abordado esta dimensión del «ocultamiento» en la delimitación del valor de la privacidad ha sido NAGEL, T. (2004) en su libro *Concealment and Exposure: And Other Essays*.

³⁰ Es también una diferencia esencial en el modo en que respondemos a las expresiones, si como mera evidencia o como algo así como declaraciones. Esto es crucial para distinguir lo que es un testimonio de una prueba o un indicio objetivo. Este es un tema con evidentes consecuencias jurídicas.

³¹ Este es un aspecto ineludible de lo que ha de entenderse por libertad de pensamiento, en tanto que esta remite no solo al contenido del pensamiento sino al pensar mismo. Este es

en ciertas ocasiones nos es exigible, quizá por razones morales, el que revelemos ciertos aspectos de nuestra vida mental. Exigencias sobre la expresión (o comunicación), como en el caso de una pareja, son inseparables de un derecho a preservar aspectos de nuestra vida (mental) en tanto no expresados. La estabilidad de ciertos valores es imposible en intercambios donde la presión por la pura *transparencia* se hace valer por encima de otras consideraciones. De ahí el cuidado que hemos de prestar a los mecanismos de presión social, a los requisitos morales o a las exigencias jurídicas por revelar muchos aspectos de la vida (mental) privada. Lo diré claramente: en nuestro experimento, los supuestos datos objetivos sobre la vida mental no son necesariamente tales, pues están ahí solo bajo la asunción de que han sido revelados siguiendo las pautas que subyacen a nuestras prácticas compartidas de atribución de estados mentales. Los tratamos no como datos sino como trozos de mente.

Acabo de mencionar la transparencia y es este otro de los asuntos que tradicionalmente se han asociado a la privacidad de lo mental. Por ello, conviene comentar, al menos brevemente, cómo se vería afectado el contraste transparencia/opacidad sobre el que han girado muchas de las discusiones sobre el concepto cartesiano de lo mental. Y merece la pena hacerlo porque aquí las asunciones tradicionales parecen haber revertido sus papeles³². La mente cartesiana es transparente a sí misma: uno no podría estar en tal o cual estado mental sin saber que uno lo tiene. Esa transparencia de la propia mente se dobla de la opacidad de las mentes de los demás, cuyo acceso siempre será indirecto, a partir de los cuerpos (a los que por otro lado también accedemos indirectamente si uno sigue a Descartes hasta el final).

Que haya un acuerdo en que no hay tal transparencia no hace menos difícil entender las relaciones entre transparencia y opacidad de lo mental. Es un hecho que los mecanismos cognitivos que subyacen a nuestra vida mental no nos son accesibles, menos aún transparentes, y que muchos aspectos de nuestra personalidad permanecen «inconscientes». También es un hecho, incontrovertido, que otros revelan aspectos clave de nuestra vida mental, nos devuelven una imagen a veces más adecuada de nosotros mismos. Una perspectiva externa, observacional, nos ayuda en muchas ocasiones en el conocimiento de nuestra vida, de nuestra experiencia y de los propios pensamientos. No debería sorprendernos el que mucha gente se vea tentada por todos los dispositivos de autorrastreo que ayudan a registrar datos objetivos fisiológicos

un tema discutido en buena parte de la literatura jurídica sobre el tema, como LIGTHART, S. (próxima publicación). TAMBIÉN LIGTHART, S.; BUBLITZ, C.; DOUGLAS, T.; FORSBERG, L. y MEYNEN, G. (2022).

³² Lo que quiero mostrar es que en el fondo muchas de las concepciones materialistas y centradas en el cerebro comparten asunciones cartesianas.

de todo tipo, o conductuales, en tanto que revelan factores esenciales para la propia autogestión. También están aquellos que confían en que estos medios de registro y de rastreo logren una mayor transparencia respecto a uno mismo y los demás. ¿Por qué no entonces con los registros de actividad cerebral? Los datos neurofisiológicos serían así una ventana que, en tanto fiable, abre con cierta transparencia a la vida psicológica de uno mismo y de los otros, de modo simétrico. Los dispositivos reforzarían la quizá exigible transparencia de lo mental, para evitar engaños quizá³³.

La opacidad de la mente quedaría disuelta si estos dispositivos se mostraran fiables y su difusión permitiera transformar nuestras prácticas en el modo que sugiere nuestro experimento mental; se disolvería la opacidad de las mentes de otros, pero también los residuos de opacidad en la mirada hacia uno mismo. Quizá no tenga sentido *desear* tal transparencia, por los riesgos que puede acarrear, pero ¿está en juego algo que sea consustancial a nuestra vida de seres con mente? Muchos piensan que no: como mucho se puede afirmar que los aspectos normativos solo dependen de una decisión puramente personal de considerar más o menos «deseable» dejar que se vean los propios pensamientos *de un modo objetivo y transparente*. Las razones son externas a la misma condición de existencia de lo mental y de los seres que se ven unos a otros como dotados de mente. Es una mera cuestión *técnica* y esta se evalúa solo en relación a lo que sea deseable y sobre lo que nos podamos poner de acuerdo, ahora a nivel colectivo. Las regulaciones son regulaciones de las tecnologías no en tanto que transforman profundamente nuestros modos de habitar el mundo o de ser sino en tanto que sus consecuencias sean o no deseables. Repito, es posible desear y aspirar a la transparencia de lo mental, y si las técnicas de leer cerebros e inscribir en ellos permiten lograrlo, ahora solo se requiere establecer límites a ese deseo cuando las consecuencias no son asumibles por los daños que provocan a uno mismo o a los otros.

Si tradujéramos el problema normativo simplemente a una cuestión sobre si cada uno de nosotros, o colectivamente, preferiríamos la transparencia mental que, presumiblemente, podrían permitir estas neurotecnologías, estaríamos lejos de dar cuenta de lo que normativamente

³³ Recuerdo una frase con la que Javier Sampedro acababa uno de sus artículos en torno al impacto de estas técnicas en *El País* hace ya un tiempo; en ese artículo, aceptaba que las cuestiones de privacidad y de transparencia (ante los demás) no es una preocupación que hayamos de asumir sin más, pues tal vez «solo les preocupe a los delincuentes y a los adúlteros» (Javier SAMPEDRO, «¿Seguro que tus pensamientos son privados?», *El País*, 23 de febrero de 2024: <https://elpais.com/opinion/2024-02-24/seguir-que-tus-pensamientos-son-privados.html>). Si esto mismo es base para ciertas exigencias de transparencia y para motivar requerimientos a «dejarse» leer el cerebro ante procesos penales (por ejemplo), quizá sea menos fácil de decidir. ¿Cuánta transparencia podemos exigir para una vida en común? El engaño puede ser sancionado moralmente, pero ¿hemos de garantizar por todos los medios que no se dé?

importa por lo que respecta a nuestra naturaleza de seres con mente. A lo largo de estas páginas, no he hecho sino insistir en que lo que sean nuestras mentes es inseparable de las prácticas e interacciones en que nos atribuimos estados específicos. Tienen razón los que dicen que las mentes no son transparentes a sí mismas. Yo diría que también tenemos razón los que defendemos que no se harán más transparentes con una actitud objetiva de registro neurofisiológico. El juego transparencia/opacidad es consustancial a las prácticas de interacción y de intercambio de perspectivas que es definitorio de la vida mental. Por expresarlo en un eslogan que no obvia, desde el inicio, la dimensión normativa, interna, de lo que significa compartir vida mental: nos importa especialmente no decirnos todo a nosotros mismos y no decir todo a otros. Este es otro modo de destacar lo que señalaba anteriormente en relación a la privacidad: refleja otro aspecto de la ineludible dimensión expresiva de lo mental. Necesitamos una mayor claridad sobre la *ontología* de lo mental sin dualismos, pero también sin reduccionismos que desvirtúen el modo en que lo normativo es un aspecto de lo que una mente es.

Se pone en juego lo mental, como algo distintivo a proteger más allá del cerebro en tanto otro órgano más y más allá de la propia integridad corporal, en el momento en que distintos aspectos de nuestra vida no son ya tratados como mera facticidad³⁴. Estos aspectos revelan la dimensión *subjektiva* que implica una mente, pero no como un dominio interno al que solo uno tiene acceso, o que es privado en un sentido de inaccesible o cerrado intrínsecamente a la inspección, sino en tanto que lo configuramos en nuestras prácticas de revelación (y ocultamiento) o de expresión, como hemos visto³⁵; pero, sobre todo, en tanto que en relación a ello exhibimos un cierto «privilegio» especial que se ha denominado autoridad en primera persona. Esta autoridad es crucial a la hora de dar cuenta de la asimetría entre las atribuciones en primera y en tercera persona de estados mentales. Si digo de mí mismo que siento dolor, a mi declaración acompaña una presunción de verdad que solo se ve socavada si uno puede mostrar que es engaño o disimulación, es decir, que no es sincera. Si digo de otro que siente dolor, tal afirmación no lleva aparejada ninguna presunción de verdad y puede ser disputada de muchos modos³⁶.

El escenario descrito en el experimento mental con el que inicié estas páginas tiene mucho que ver con nuestras prácticas ordinarias

³⁴ Es esta una idea que subyace a la caracterización sartreana de la conciencia en su clásico SARTRE (1943), *El ser y la nada*.

³⁵ Me permito mencionar de nuevo el artículo de J. C. Bublitz, en el que se analizan muchos de los supuestos que no se están teniendo en cuenta en las discusiones normativas sobre la protección de las mentes. BUBLITZ, J. C. (2024).

³⁶ Sería presuntuoso resumir brevemente los debates sobre la autoridad de primera persona. Una referencia fundacional es el clásico de DAVIDSON, D. (1984). La orientación que más me atrae en esta cuestión es la defendida por MORAN, R. (2001).

de reconocer autoridad de primera persona para cierto tipo de declaraciones. No es fácil explicar el fundamento de tal autoridad, más aún cuando hay argumentos bastante decisivos en contra de la explicación epistémica (de tipo cartesiano) de la autoridad, como si cada uno de nosotros tuviéramos un acceso privilegiado de tipo epistémico, un ojo interno o una capacidad especial de autoobservación³⁷. En las explicaciones epistémicas, la autoridad es la de un experto que exhibe su competencia en un dominio propio al cual solo él tiene acceso. Las alternativas, en ocasiones, han tendido a rebajar estas pretensiones y a ver tal autoridad (que se supone por defecto) como un mero título conferido. Es indudable, para mí, que la autoridad de primera persona se revela en los compromisos prácticos y expresivos que acompañan a nuestras declaraciones y que, como tal, se asienta sobre una práctica social compartida, en la que cada declaración reclama un reconocimiento, pero *desde la idea de que está en manos de uno dar forma al propio mundo mental* y sostenerlo³⁸. Y es esto mismo lo que se busca proteger, el que esté en nuestras manos rechazar ciertos rasgos que uno mismo (u otros) puede identificar como pertenecientes a la facticidad, y de los que uno no diría que le son propios. Si hay algo que nos debería preocupar de la situación creada por las neurotecnologías en el experimento de pensamiento es el hecho de que, con ellas, se podría socavar la presunción de autoridad respecto a los propios estados mentales que uno exhibe en primera persona. Recordemos que, en una de las lecturas, cabía la posibilidad de que el dispositivo nos diera información sobre lo que se supone que es nuestra vida mental *de un modo fiable*, quizá incluso con mayor fiabilidad de la que exhibimos en nuestras declaraciones ordinarias (pues autoridad por defecto no implica infalibilidad). Ante la fiabilidad de los dispositivos, ¿hacia dónde se inclinará la balanza a la hora de asignar autoridad? ¿En qué sentido puede verse afectada la autoridad *en primera persona*? ¿Podemos remplazarla por una autoridad independiente y «heredada» de métodos fiables de establecer lo que pensamos?

Considérese la siguiente situación: el dispositivo inscribe «deseo, chocolate». Cualquiera podría extraer la conclusión de que su estado mental sería el que uno identificaría mediante la atribución «S desea comer chocolate». Uno podría incluso ¡atribuírselo a sí mismo! Hay momentos en que podría suceder que uno «no hubiera reparado» en ello, en que lo desea, y ahora el dispositivo estaría en disposición, de modo fiable (se supone), de indicarle que de hecho es así y uno lo desea. Dos posibilidades están abiertas: o bien que mi consulta simplemente *confirme* lo que uno mismo aceptaría, o bien que no coincida con algo que uno mismo admitiría y que se produzca

³⁷ Un resumen de las críticas a las explicaciones epistémicas de la autoridad de primera persona se puede encontrar en LARMORE, C. (2004).

³⁸ Es esto lo que hace que la autoridad de primera persona sea una forma de la *libertad*.

una tensión entre la información del dispositivo, que uno vería ahora como pura facticidad —como algo que uno no reconocería como parte de sí—, pues no estaría dispuesto a declararlo. En cualquiera de estas situaciones no tenemos por qué asumir que se desdibuja nuestra autoridad de primera persona, en la medida en que nuestras prácticas de *reconocimiento* de tal autoridad no se dirigen primariamente a lo que se despliega en el dispositivo sino a lo que cada uno esté dispuesto a declarar como perteneciente a la propia mente. Es obvio, igualmente, que esta autoridad de primera persona exhibe una dimensión epistémica, pues es posible siempre que entre en colisión con información obtenida observacional e inferencialmente, en este caso la que proporciona el dispositivo. En pocas palabras, en nuestras prácticas psicológicas, lo que se dirime es si cierta externalidad es una manifestación o expresión que uno está dispuesto a suscribir como propia. La autoridad de primera persona es práctico-expresiva y epistémica a un tiempo, y en esa medida es la clave para apropiarnos de la propia experiencia y constituirnos como agentes. Es una autoridad por defecto, que arraiga en las propias prácticas de atribución psicológica en tanto que estas son expresión del *reconocimiento* de esa autoridad, expresión de ese aspecto de la vida mental que entendemos como *subjetivo*; esta es una autoridad revocable, que puede contrastar con información que uno mismo u otros pueden esgrimir tras procedimientos de examen³⁹. En nuestras prácticas, se entrelazan ambas perspectivas constantemente, de primera y de tercera persona, incluso cuando se trata de uno mismo. *Importan* los datos objetivos sobre la propia condición mental, pero es esencial que esté en nuestra mano escapar a la pura facticidad. La amenaza se conjura solo en la medida en que cada uno de nosotros no esté dispuesto a aceptar que lo que sea parte de la propia mente pueda ser completamente objetivado. Si es esa la asunción de escenarios en los que las neurociencias y las neurotecnologías registran y modifican la actividad cerebral, leen y modifican de un modo tan *directo* las mentes, entonces nuestras prácticas psicológicas se verían profundamente transformadas, y no es claro en qué medida cada sujeto, cada uno de nosotros, tendría ya algo que *decir* sobre sus experiencias y sobre sus actuaciones. Estarían ahí, como pura facticidad, no como parte de la vida consciente reconocible como tal. Socavar la autoridad de primera persona es un modo de afectar al significado de las propias experiencias y a nuestra capacidad de actuación.

A nadie se escapa la fascinante aventura que significa poder intervenir de este modo sobre las bases materiales de nuestra concien-

³⁹ No puedo entrar en ello, pero esta noción de «autoridad de primera persona», al menos entendida en estos términos, no puede segregarse de nuestra capacidad de autodeterminación. Por eso, creo que los autores que se han enfocado en proteger la autodeterminación —mediante un derecho específico— sean los que considero más encaminados en este ámbito.

cia; a nadie tampoco, los peligros y amenazas que surgen con ello. Muchos abogan por diseñar instrumentos (normativos) que protejan específicamente los cerebros de manipulaciones indebidas, básicamente por el hecho de que se podría ver afectada de un modo hasta ahora desconocido nuestra libertad cognitiva⁴⁰. Mis reflexiones no inciden, al menos de modo directo, en estos debates. Constituyen una llamada de atención ante modos de enmarcar los problemas quizá equivocados, excesivamente neurocéntricos, y que no dan cuenta de una dimensión *ontológica* de lo mental: este aspecto de lo que son las mentes no es identificable desde un enfoque que solo tenga en cuenta los realizadores neuronales de las mentes; solo se revela cuando uno atiende a las prácticas de lectura y manipulación de mentes a partir del dominio de cierto tipo de conceptos. No parece fácil abandonarse a una imagen de lo mental en la que nuestros estados —los que nos importan— se midan con varas conceptuales rígidas. Lo mental, en su dimensión expresiva, exhibe una cierta indeterminación e inseguridad, así se articula en nuestra vida⁴¹. ¿Estamos dispuestos a romper el frágil equilibrio entre la actitud participante y la actitud objetiva que marca las relaciones interpersonales? No hago aquí sino aplicar un cierto argumento strawsoniano sobre la compatibilidad o incompatibilidad del determinismo y la libertad⁴². El experimento de pensamiento con el que he abierto estas páginas simplemente pone el foco en la idea de que una reducción de las interacciones en las que desplegamos nuestros conceptos psicológicos a determinaciones objetivas de datos que fijan las condiciones mentales específicas de cada uno de nosotros compromete la vida humana que compartimos. Aunque sea teóricamente posible lo que se sugiere en el experimento, y quizá dentro de poco también técnicamente posible, me pregunto, como Strawson, si esto no es *prácticamente* inconcebible.

3. BIBLIOGRAFÍA

ALEGRE, S. (2017), «Rethinking freedom of thought for the 21st century». *European Human Rights Law Review*, 3, pp. 221-33.

⁴⁰ Por ejemplo, en el libro de FARAHANY (2023) antes citado, se postula la necesidad de un nuevo derecho a lo que ella denomina libertad cognitiva, «the right to self-determination over our brains and mental experiences», un derecho que incluye otros como la privacidad mental, la libertad de pensamiento y la auto-determinación, derechos que han de hacerse compatibles con una neurociencia y una neurotecnología que permiten rastrear y *hackear* (*tracking and hacking*) el cerebro. En términos jurídicos, el debate se ha cifrado en una decisión entre reinterpretar los derechos ya existentes o introducir nuevos derechos. Encuentro muy clarificador en relación a este tema el siguiente artículo de BUBLITZ, J.C. (2022).

⁴¹ Esta es la enseñanza de Wittgenstein sobre la naturaleza de lo mental, al menos como se recoge en sus *Últimos escritos de filosofía de la psicología* (WITTGENSTEIN, 2008).

⁴² Véase de nuevo STRAWSON (1963).

- AYER, A. J. (1963), «Privacy». En *The Concept of a Person*. Londres: Macmillan, pp. 52-81.
- BUBLITZ, J. C. (2024), «What an International Declaration on Neurotechnologies and Human Rights Could Look like: Ideas, Suggestions, Desiderata». *American Journal of Bioethics Neuroscience*, 15(2), pp. 96-112.
- BUBLITZ, J. C. (2022), «Novel Neurorights: From Nonsense to Substance». *Neuroethics* 15 (1), pp. 1-15.
- BUBLITZ, J. C. (2021), «Freedom of Thought as an International Human Right: Elements of a Theory of a Living Right». En BLITZ, M. J. y BUBLITZ, J. C. (eds.). *The Law and Ethics of Freedom of Thought, Volume 1. Neuroscience, Autonomy, and Individual Rights*. Londres: Palgrave Macmillan. pp. 49-101. https://doi.org/10.1007/978-3-030-84494-3_3.
- BUBLITZ, J. C. (2020), «The nascent right to psychological integrity and mental self-determination». En VON ARNAULD, A. et al. (eds.). *The Cambridge Handbook of New Human Rights: Recognition, Novelty, Rhetoric*. Cambridge: Cambridge University Press, pp. 387-403.
- BUBLITZ, J. C. (2015), «Cognitive Liberty or the International Human Right to Freedom of Thought». En CLAUSEN, J. y LEVY, N. (eds.). *Handbook of Neuroethics*. Dordrecht: Springer, pp. 1309-1333. https://doi.org/10.1007/978-94-007-4707-4_166.
- BUBLITZ, J. C. (2013), «My Mind Is Mine!? Cognitive Liberty as a Legal Concept». En HILDT, E. y FRANKE, A. (eds.). *Cognitive Enhancement: An Interdisciplinary Perspective*, vol. 1. Dordrecht: Springer, pp. 233-264. https://doi.org/10.1007/978-94-007-6253-4_19.
- BYRNE, R. y WHITEN, A. (eds.) (1988), *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. Oxford: Clarendon Press.
- CARRUTHERS, P. y SMITH, P. K. (eds.) (1996), *Theories of Theories of Mind*. Nueva York: Cambridge University Press.
- CASSINADRI, G. (2025), «The Right to Mental Integrity: Multidimensional, Multilayered and Extended». *Neuroethics*, 18 (16), pp. 1-21.
- CAVELL, S. (2002), «Knowing and Acknowledging». En *Must We Mean What We Say?: A Book of Essays*. Cambridge: Cambridge University Press, pp. 238-366.
- CAZETTES, F., REATO, D., AUGUSTO, E. et al. (2025), «Facial expressions in mice reveal latent cognitive variables and their neural correlates». *Nat Neurosci*, 28, pp. 2310-2318. <https://doi.org/10.1038/s41593-025-02071-5>.
- CHURCHLAND, P. M. (1981), «Eliminative Materialism and the Propositional Attitudes». *Journal of Philosophy*, 78(2), pp. 67-90.
- DAVIDSON, D. (1984), «First person authority». *Dialectica*, 38(2-3), pp. 101-112.
- DAVIES, M. y STONE, T. (eds.) (1995), *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell.
- FARAHANY, N. A. (2023), *The Battle for Your Brain. Defending the Right to Think Freely in the Age of Neurotechnology*, Nueva York: St. Martin's Press.
- FLORIDI, L. (2014), *The fourth revolution: how the infosphere is reshaping human reality*. Oxford: Oxford University Press.

- GOERING, S., KLEIN, E., *et al.* (2021), «Recommendations for Responsible Development and Application of Neurotechnologies». *Neuroethics*, 14(3), pp. 365-386.
- GOLDMAN, A.I. (1993), «The psychology of folk psychology». *Behavioral and Brain Sciences*, 16(1), pp. 15-28. <https://doi.org/10.1017/S0140525X00028648>.
- GOMILA, A. (2002), «La perspectiva de la segunda persona de la atribución intencional». *Azafra*, 4, pp. 123-138.
- IENCA, M. y ANDORNO, R. (2017), «Towards new human rights in the age of neuroscience and neurotechnology». *Life Sciences, Society and Policy*, 13(1), pp. 1-27.
- IENCA, M. (2021), «On neurorights». *Frontiers in Human Neuroscience*, 15:701258. <https://doi.org/10.3389/fnhum.2021.701258>.
- LARMORE, Ch. (2004), *Les pratiques du moi*. Paris: Presses Universitaires de France.
- LAVAZZA, A. (2018), «Freedom of Thought and Mental Integrity: The Moral Requirements for Any Neural Prosthesis». *Frontiers in Neuroscience*, 12:82. <https://doi.org/10.3389/fnins.2018.00082>.
- LAVAZZA, A. y GIORGI, R. (2023), «Philosophical foundation of the right to mental integrity in the age of neurotechnologies». *Neuroethics*, 16(1), pp. 1-13.
- LIGHTART, S. (próxima publicación). «Mental privacy as part of the human right to freedom of thought?». En BLITZ, M. y BUBLITZ, J.C. (eds.). *The law and ethics of freedom of thought, Volume 2. Cognitive Liberty and Privacy*. Londres: Palgrave Macmillan.
- LIGHTART, S. *et al.* (2023), «Minding Rights: Mapping Ethical and Legal Foundations of “Neurorights”». *Cambridge Quarterly of Healthcare Ethics*, 32(4), pp. 461-481.
- LIGHTART, S. *et al.* (2022), «Rethinking the Right to Freedom of Thought: A Multidisciplinary Analysis». *Human Rights Law Review*, 22(4), pp. 1-14.
- LIGHTART, S. *et al.* (2021), «Forensic brain-reading and mental privacy in European Human Rights Law: Foundations and challenges». *Neuroethics*, 14(2), pp. 191-203.
- MCCARTHY-JONES, S. (2019), «The autonomous mind: The right to freedom of thought in the twenty-first century». *Frontiers in Artificial Intelligence*, 2, pp. 1-7.
- MICHALOWSKI, S. (2020), «Critical reflections on the need for a right to mental self-determination». En von Arnould A. *et al.* (eds.). *The Cambridge Handbook of New Human Rights: Recognition, Novelty, Rhetoric*. Cambridge: Cambridge University Press, pp. 404-412.
- MORAN, R. (2001), *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.
- NAGEL, T. (2004), *Concealment and Exposure: And Other Essays*. Nueva York: Oxford University Press.
- NEFF, G. y NAFUS, D. (2016), *Self-tracking*, Cambridge (Mass.): The MIT Press.
- PÉREZ, D. y GOMILA, A. (2021), *Social Cognition and the Second Person in Human Interaction*. Londres - Nueva York: Routledge.
- SARTRE, J.P. (1943), *L'être et le néant*. Paris: Gallimard.

- SOMMAGGIO, P., MAZZOCCA, M., GEROLA, A. y FERRO, F. (2017), «Cognitive liberty. A first step towards a human neuro-rights declaration». *BioLaw Journal – Rivista di BioDiritto*, 3, pp. 27-45.
- STRAWSON, P.F. (1963), «Freedom and Resentment». En *Proceedings of the British Academy*, 48, pp. 187-211. (Traducción castellana en (1995), *Libertad y resentimiento y otros ensayos*. Barcelona: Paidós, pp. 37-68).
- TAYLOR, C. (1989), *The Sources of the Self. The Making of the Modern Identity*, Cambridge (Mass.): Harvard University Press.
- WAJNERMAN PAZ, A. (2022), «Is Your Neural Data Part of Your Mind? Exploring the Conceptual Basis of Mental Privacy». *Minds and Machines*, 32(2), pp. 395-415.
- WITTGENSTEIN, L. (1988), *Investigaciones Filosóficas*, Barcelona: Crítica.
- WITTGENSTEIN, L. (2008), *Últimos escritos de filosofía de la psicología*, vol. 1 y 2, Madrid: Tecnos.
- WRIGHT, C., SMITH, B. C. y MACDONALD, C. (eds.) (2000), *Knowing Our Own Minds*. Oxford: Clarendon Press.
- YUSTE, R. *et al.* (2017), «Four ethical priorities for neurotechnologies and AI». *Nature*, 551, pp. 159-163. <https://doi.org/10.1038/551159a>.