



RIDAA
Repositorio Institucional
Digital de Acceso Abierto de la
Universidad Nacional de Quilmes



**Universidad
Nacional
de Quilmes**

Cao, Stella Maris

Filosofía de las neuroimágenes funcionales



Esta obra está bajo una Licencia Creative Commons Argentina.
Atribución - No Comercial - Sin Obra Derivada 2.5
<https://creativecommons.org/licenses/by-nc-nd/2.5/ar/>

Documento descargado de RIDAA-UNQ Repositorio Institucional Digital de Acceso Abierto de la Universidad Nacional de Quilmes de la Universidad Nacional de Quilmes

Cita recomendada:

Cao, S. M. (2026). *Filosofía de las neuroimágenes funcionales. (Tesis de maestría)*. Universidad Nacional de Quilmes, Bernal, Argentina. Disponible en RIDAA-UNQ Repositorio Institucional Digital de Acceso Abierto de la Universidad Nacional de Quilmes <http://ridaa.unq.edu.ar/handle/20.500.11807/6052>

Puede encontrar éste y otros documentos en: <https://ridaa.unq.edu.ar>

Filosofía de las neuroimágenes funcionales

TESIS DE MAESTRÍA

Stella Maris Cao

stellamariscao@hotmail.com

Resumen

En las últimas décadas, el desarrollo de las técnicas de neuroimagen funcional y, en particular, de la resonancia magnética funcional (RMf), dio lugar a una expansión sin precedentes de discursos filosóficos, científicos, políticos y pseudocientíficos sobre la neurociencia y su capacidad para explicar la mente humana. La presencia creciente de estas tecnologías en la investigación, en el ámbito médico, en el social y en los medios de comunicación fue consolidando la idea de que es posible “visualizar los procesos cognitivos en acción”, otorgando a la neuroimagen un estatus privilegiado como fuente de conocimiento. Sin embargo, su interpretación plantea desafíos epistemológicos, metodológicos y éticos que requieren un análisis filosófico riguroso. Sostengo que las herramientas conceptuales propias de la filosofía analítica de la ciencia son adecuadas para ello.

Entre los principales obstáculos para una interpretación de las neuroimágenes, se encuentran la carga teórica de la observación, los problemas inherentes al diseño de investigación, los modelos inferenciales utilizados y la compleja relación entre neurociencia y psicología cognitiva.

Estas dificultades no solo conciernen al plano teórico: adquieren una dimensión práctica cuando se considera el impacto social de estas tecnologías. En el ámbito neuroético, se exploran las implicaciones médicas, sociales y políticas del uso de neuroimágenes, analizando el riesgo potencial de que su aplicación refuerce interpretaciones reduccionistas de la identidad y la autonomía humana. La creciente influencia de estas tecnologías en la toma de decisiones individuales y colectivas plantea interrogantes acerca de sus límites y confirma la necesidad de una reflexión ética y de una regulación normativa que garanticen su uso responsable.

En definitiva, este trabajo sostiene que la neuroimagen funcional no es un reflejo directo de la cognición, sino una construcción cuya interpretación debe abordarse con cautela. Desde una perspectiva filosófica, se problematiza su papel en la producción de conocimiento y su

impacto en la sociedad, cuestionando las narrativas deterministas y promoviendo un análisis crítico que además de evitar simplificaciones excesivas, asegure una evaluación rigurosa tanto de sus posibilidades como de sus limitaciones.

Palabras clave: Filosofía de la neurociencia; Neuroimagen funcional; Resonancia Magnética funcional (RMf); Ontología cognitiva; Epistemología; Neuroética.

Abstract

In recent decades, the development of functional neuroimaging techniques—particularly functional Magnetic Resonance Imaging (fMRI)—has generated an unprecedented expansion of philosophical, scientific, political, and pseudoscientific discourse surrounding neuroscience and its capacity to explain the human mind. The growing presence of these technologies across research, clinical practice, public debate, and mass media has strengthened the belief that it is possible to visualize cognitive processes in action, granting neuroimaging a privileged status as a source of knowledge. However, interpreting neuroimages presents epistemological, methodological, and ethical challenges that demand rigorous philosophical examination. This thesis argues that the conceptual tools offered by analytic philosophy of science are well suited to this task.

Major obstacles to the interpretation of neuroimaging include the theory-ladenness of observation, the methodological limitations of experimental design, the inferential models employed, and the complex relationship between neuroscience and cognitive psychology.

These challenges are not merely theoretical: they gain a practical dimension when considering the social impact of neuroimaging technologies. Within the field of neuroethics, the medical, social, and political implications of their use are examined, including the potential risk that such applications reinforce reductionist interpretations of human identity and autonomy. The increasing influence of neuroimaging on individual and collective decision-making raises questions regarding its limits and underscores the need for ethical reflection and regulatory frameworks that ensure responsible use.

Ultimately, this work maintains that functional neuroimaging is not a direct window into cognition, but rather a construction whose interpretation demands caution. From a philosophical perspective, its role in the production of knowledge and its social impact are critically examined, challenging deterministic narratives and promoting an analysis that—while avoiding excessive simplification—supports a rigorous evaluation of both its possibilities and its limits.

Keywords: Philosophy of Neuroscience; Functional Neuroimaging; fMRI; Cognitive Ontology; Epistemology; Neuroethics.



Filosofía de las neuroimágenes funcionales

Stella Maris Cao

Universidad Nacional de Quilmes

Trabajo final de Maestría en filosofía

Director: Dr. Sergio Barberis (CONICET-UNQ, FFyL-UBA, UDESA)

Codirectora: Dra. Lucía Federico (UNQ-UNTREF)

Buenos Aires, Argentina

Marzo de 2025

Agradecimientos

Cuando Sergio Barberis, mi director, me habló en nuestro primer encuentro de las transformaciones del espíritu según Nietzsche, algo se puso en marcha. Este trabajo final de Maestría es también el registro de ese devenir: las formas que el pensamiento, el afecto y el sentido fueron tomando a lo largo del camino.

Quiero agradecer profundamente a quienes acompañaron ese movimiento. A quienes estuvieron cerca, atentos, disponibles y dejaron una huella en este proceso de búsqueda.

A Sergio, que fue un espejo benévolo. Supo sostenerme en las inseguridades, reflejar mis búsquedas sin apurarlas, con una escucha atenta que acompañó incluso lo que todavía no encontraba palabras.

A Lucía Federico, mi codirectora, por su agudeza atenta, su mirada afinada que no deja pasar lo importante, su capacidad para señalar con precisión sin perder nunca la calidez. Por ayudarme a pensar mejor y a escribir con más cuidado.

A Emanuel López Muro, mi tutor, por su calidez y su humanidad. Por estar cuando el recorrido se volvía cuesta arriba, y acompañar con respeto, cercanía y compromiso genuino.

A mamá y a papá, que ya no están pero siguen siendo raíz. Su amor me dio fuerza cuando fui camello, coraje cuando quise ser león y confianza para soltar y volver a jugar como una niña.

A Eduardo, por enseñarme a jugar de grande. Por mostrarme que el pensamiento también puede ser lúdico y que en el juego habita una forma de libertad. Por hacer de la música un lenguaje compartido, capaz de decir incluso lo que no se puede pensar del todo.

A Raúl, por cultivar conmigo, casi cotidianamente, algunas de las más refinadas formas de la ignorancia. Por ensayar una metafísica sin garantías (un laberinto sin centro, un espejo opaco) y por hacer del no saber una práctica compartida y risueña.

A Paula, mi nueva compañera de juegos, con quien la risa, la ternura y la complicidad volvieron cotidiano el asombro.

A mis estudiantes, ese "público cautivo" que me permite reinventarme cada día. Que me enseña, sin saberlo, nuevas formas de mirar y de preguntar.

A la memoria del amor, en todas sus formas: el que sostiene, el que transforma, el que empuja, el que espera.

Y a quienes estuvieron cerca, con palabras, con silencios, con mates y tiempo compartido. Cada gesto fue parte de este tránsito, de esta transformación.

Este escrito, por muy humilde o imperfecto que sea, es una manera de dejar constancia: de que el pensamiento necesita afecto, de que el conocimiento se hace con otros, de que hay formas del amor que resisten el tiempo. Es apenas un gracias, un modo de quedarme un poco más, de seguir jugando con las palabras que alguna vez me tocaron y todavía me habitan.

Índice

Resumen.....	2
Abstract	3
1. <i>Introducción</i>	11
1.1. Definición del problema.....	11
1.2. Objetivos	12
1.2.1. Objetivo general.....	12
1.2.2. Objetivos específicos	12
1.3. Estado de la cuestión	12
1.4. Enfoque conceptual acerca del problema.....	14
1.5. Metodología	16
1.6. Plan de trabajo.....	16
2. <i>Pasado y presente de las neuroimágenes</i>	18
2.1. Apuntes históricos.....	18
2.1.1. Antecedentes	18
2.1.2. Dos teorías	20
2.1.3. Escuela de Alejandría	21
2.1.4. Galeno de Pérgamo	21
2.1.5. Edad Media	22
2.1.6. Renacimiento	23
2.1.7. René Descartes.....	25
2.1.8. La frenología.....	27
2.1.9. Paul Broca y Carl Wernicke	28
2.1.10. Las intuiciones de Brodmann.....	30
2.1.11. Técnicas de visualización	32
2.1.12. Nuevos desarrollos.....	34
2.1.13. La revolución en neuroimagen.....	35
2.2. ¿Cómo se generan las imágenes de resonancia magnética?.....	38
2.2.1. Resonancia magnética funcional (RMf)	42
2.2.2. El diseño de investigación	42
2.2.3. Desarrollo del experimento.....	44
2.2.4. Posprocesado.....	45

2.2.5.	Análisis estadístico.....	48
2.2.6.	Limitaciones.....	49
2.3.	Discusión y proyecciones.....	50
3.	<i>La RMf como fuente de evidencia</i>	53
3.1.	Regiones de interés.....	53
3.2.	El problema de las localizaciones	54
3.3.	Distancia inferencial real y aparente	59
3.3.1.	De la actividad neuronal a los datos en bruto	60
3.3.2.	De los datos en bruto a la neuroimagen	60
3.3.3.	De la neuroimagen a la interpretación de la conducta	61
3.4.	Lo que muestran las imágenes	62
3.4.1.	Neuroimagen y fotografía	62
3.4.2.	Comparaciones.....	66
3.5.	¿Qué es, entonces, una neuroimagen?.....	67
3.5.1.	Una neuroimagen es un mapa estadístico	67
3.5.2.	¿Qué tipo de evidencia nos brinda la neuroimagen?	68
3.6.	Las inferencias en neuroimágenes.....	70
3.6.1.	El teorema de Bayes.....	71
3.6.2.	Tipos de inferencias	74
3.6.3.	Inferencia directa	74
3.6.4.	Inferencia inversa.....	78
3.6.5.	Más allá de la inferencia	79
3.7.	Sesgos de publicación	81
3.8.	Discusión y proyecciones.....	84
4.	<i>Neuroimágenes y psicología cognitiva</i>	87
4.1.	Psicología cognitiva	87
4.1.1.	Las ciencias de la cognición	88
4.1.2.	El hexágono de Sloan.....	89
4.1.3.	Caracterización de la psicología cognitiva	91
4.2.	Sobre los constructos de la psicología cognitiva.....	93
4.2.2.	Modelos explicativos	96
4.2.3.	Un cerebro pluripotencial	100
4.3.	Ontología cognitiva	102
4.3.1.	Una expresión, más de tres significados	102

4.3.2.	Alcance y niveles de análisis	104
4.4.	El diseño experimental	108
4.4.1.	Confiabilidad y validez	109
4.4.2.	Validez ecológica.....	112
4.4.3.	Validez de constructo.....	114
4.5.	Discusión y proyecciones.....	115
5.	<i>El desafío de la neuroética</i>	121
5.1.	Contexto y delimitaciones del campo	121
5.2.	Neuroimágenes y conciencia.....	124
5.3.	Neuroimágenes y lectura de la mente	126
5.4.	Neuroimágenes y detección de mentiras.....	127
5.5.	Neuroimágenes y predicción de comportamientos	129
5.6.	Neuroimágenes y género.....	130
5.7.	Neuroimágenes y opinión pública.....	133
5.8.	Neuroimágenes e inteligencia artificial.....	136
5.9.	Neuroimágenes e identidad	138
5.10.	Discusión y proyecciones.....	141
6.	<i>Conclusiones</i>	142
7.	<i>Apéndice</i>	154
8.	<i>Referencias bibliográficas</i>	155

Ilustraciones y figuras

Figura 2-1.

Ilustración del siglo XV de Gregor Reisch

que muestra la concepción medieval de las funciones mentales. 22

Figura 2-2.

Imagen del libro VII de la obra De Humani Corporis Fabrica (1543).

Se muestra el cerebro de frente, con los hemisferios cerebrales abiertos, proporcionando una visión del cuerpo calloso y la apófisis crista galli. 24

Figura 2-3.

Leonardo da Vinci dibujó los ventrículos cerebrales después de extraer

modelos tridimensionales de cera inyectada en el cerebro (1504-1507). 25

Figura 2-4.

Imagen de la frenología que muestra los diferentes “órganos” de la conducta. 28

Figura 2-5.

El modelo de Wernicke. 29

Figura 2-6.

Las áreas de Brodmann 31

Figura 2-7.

Imagen perteneciente al primer atlas neuroanatómico fotográfico, Iconographie Photographique des Centres Nerveux (1873). Las fotografías (izquierda) del atlas se acompañaban con una litografía

(derecha) en la página contigua. 32

Figura 2-8.

Dibujo de Santiago Ramón y Cajal (1899) de diferentes tipos de neuronas del cerebelo de una paloma.

(A) célula de Purkinje, un ejemplo de neurona bipolar. (B) célula granulosa que es multipolar. 33

Figura 2-9.

Un ejemplo del aspecto de una placa radiográfica obtenida en 1919 utilizando el procedimiento de la neumoencefalografía..... 34

Figura 2-10.

Primera angiografía exitosa realizada por Egas Moniz..... 35

Figura 2-11.

Tomografía axial computarizada de cráneo del paciente. Presenta hematoma en proyección del núcleo lenticular izquierdo de 20 mm de diámetro mayor, rodeado de edema perilesional (flecha). 36

Figura 2-12.

En la imagen PET se visualiza el metabolismo de la glucosa en el cerebro a través de una molécula similar a esta que contiene un átomo de flúor radiactivo. Las áreas de color rojo representan las que más glucosa están captando. 37

Figura 2-13.

Arquitectura de las fibras de la materia blanca del cerebro medida a partir de imágenes de espectro de difusión (DSI). Se muestran las vías del cuerpo calloso y del tronco encefálico. Las fibras están codificadas por colores según su dirección: rojo = izquierda-derecha, verde = anteroposterior, azul = ascendente-descendente (RGB=XYZ). 38

Figura 2-14.

Cascada de relaciones neuronales, fisiológicas y biofísicas que conectan la actividad neuronal con los cambios de señal medidos utilizando iRMf. 41

Figura 2-15.

Ejemplos de transformaciones lineales. En cada figura, los puntos negros representan las ubicaciones de coordenadas originales y los puntos azules, las nuevas ubicaciones después de que se aplica la transformación. 45

Figura 2-16.

Ejemplo “casero” de filtro gaussiano. Obsérvese la eliminación del ruido pero también la pérdida de resolución espacial. 46

Figura 2-17.

Para hacer comparable las imágenes de un paciente con una base de imágenes de la población normal, es necesario “deformar” la superficie a un cerebro estandarizado.

La figura 8A muestra un corte transversal potenciado en T1 y la figura 8B muestra el mismo corte “normalizado” al estándar MNI 152 (Montreal Neurological Institute). 47

Figura 3-1.

Corriente causal y corriente funcional 60

Figura 4-1.

El “hexágono de Sloan” o hexágono cognitivo. 90

Figura 4-2.

Figura 4-3.

Traducción de dos esquemas presentados por Bickle para explicar su reduccionismo “ruthless”. La primera figura muestra el estándar de organización neural, con una vista de reducción “paso a paso” (flechas hacia abajo). La segunda imagen representa el reduccionismo implacable. Las flechas discontinuas representan niveles de intervención experimental; la flecha sólida, el nivel en el que se miden estas intervenciones. Según esta formulación, la psicología

“es un esfuerzo descriptivo que genera datos conductuales, en lugar de explicaciones”. 98

Figura 4-4.

El hexágono de Sloan “empobrecido” por la pérdida de la psicología cognitiva. 107

Figura 4-5.

El proceso de investigación según J. Sullivan. 108

Figura 5-1.

Los resultados del artículo de Adrian Owen de 2006, que muestran en rojo/amarillo las áreas del cerebro activas mientras se imagina jugando al tenis (izquierda) e imaginando navegar por la casa (derecha) en el paciente vegetativo (arriba) y los controles sanos (abajo). 125

Figura 5-2.

Imagen de la técnica de "Potenciales relacionados con eventos". 128

Figura 5-3.

Análisis de conexiones. Las conexiones intrahemisféricas se muestran en azul y las interhemisféricas se muestran en naranja. 131

Figura 5-4.

Algunos ejemplos de la difusión de la neurociencia en la prensa. 135

Tablas y cuadros

Tabla 2-1

<i>Teoría de los humores.....</i>	<i>19</i>
-----------------------------------	-----------

Tabla 2-2

<i>Valores T1 y T2 aproximados en el cerebro.</i>	<i>40</i>
--	-----------

Tabla 2-3

<i>Significación de los valores T1, T2 y T2*.....</i>	<i>41</i>
---	-----------

Tabla 2-4

<i>Métodos de análisis estadístico más frecuentemente utilizados en RMf.....</i>	<i>48</i>
--	-----------

Tabla 2-5

<i>Comparación de las técnicas de neuroimagen.</i>	<i>50</i>
---	-----------

Tabla 3-1

<i>Presupuestos teóricos y análisis de los datos</i>	<i>57</i>
--	-----------

Tabla 3-2

<i>Comparación de los modelos tradicionales y los bayesianos.....</i>	<i>73</i>
---	-----------

Tabla 3-3

<i>Esquema del experimento con el que Friston ilustra la sustracción cognitiva</i>	<i>76</i>
--	-----------

Tabla 4-1

<i>Esquema de la historia de las ciencias cognitivas, según Venturelli y Barberis (2023)</i>	<i>88</i>
--	-----------

Tabla 4-2

<i>Condiciones de la validez interna.....</i>	<i>116</i>
---	------------

Tabla 4-3

<i>Preguntas a las que la RMf puede responder según Kosslyn, 1999.</i>	<i>119</i>
---	------------

1. Introducción

1.1. Definición del problema

En las últimas décadas hemos asistido a una proliferación de diferentes discursos relativos a las neurociencias, con diversos niveles de científicidad puestos en juego. Dejando de lado las versiones de divulgación o aquellos estudios que presentan fallas burdas en el diseño de investigación, nos encontramos a menudo con ciertas conclusiones enunciadas apodícticamente y extrapoladas a “recetas” de intervención en distintos campos, entre ellos, por ejemplo, el educativo. Así, hoy se habla de neurodidáctica, neuroeducación o neuroaprendizaje.

Entre otros factores, esto se ha vuelto posible en virtud del fuerte impulso dado a las neurociencias a partir del desarrollo de las técnicas de neuroimágenes. Estas pueden ser tanto estructurales como funcionales y, entre estas últimas, se destaca la resonancia magnética funcional por contraste BOLD (acrónimo de *blood-oxygen-level-dependent* en inglés, es decir, “dependiente del nivel de oxígeno en la sangre”). La resonancia magnética funcional es una técnica de neuroimagen no invasiva que permite inferir la activación de grupos neuronales cuando se realizan distintas tareas cognitivas o conductuales.

Ahora bien, el horizonte planteado por las neuroimágenes se encuentra lleno de interrogantes. ¿En qué medida las imágenes obtenidas en una resonancia magnética funcional “representan” las áreas cerebrales implicadas en una determinada conducta o actividad cognitiva? Y más aún: ¿qué significa la activación de estos u otros grupos neuronales? ¿Es posible hablar de correlación, de evidencia, de causalidad? En resumidas cuentas, ¿qué relación puede establecerse entre la neuroimagen funcional y la conducta o actividad cognitiva que se pretende explorar?

Teniendo en cuenta, por otra parte, que toda disciplina académica es, al menos en parte, una construcción social, surge la pregunta por las implicaciones que pueden tener las inferencias sobre neuroimágenes funcionales en nuestra manera de percibir y de actuar en el mundo. Y es con base en ese interrogante que nos encontramos frente a nuevos problemas éticos: lo colegido a partir de las neuroimágenes no solo permitiría alcanzar una mayor comprensión de la mente humana, sino que podría convertirse en una poderosa herramienta para predecir la conducta, influir en ella e incluso controlarla (Farah, 2010).

1.2. Objetivos

1.2.1. Objetivo general

Elucidar los fundamentos epistemológicos, los principios experimentales y los desafíos neuroéticos de las técnicas de neuroimagen funcional (en particular, de la resonancia magnética funcional) aplicadas al estudio de la actividad cognitivo-conductual de los seres humanos.

1.2.2. Objetivos específicos

- Clasificar y discutir críticamente los problemas epistemológicos generales relativos a la utilización de técnicas de neuroimagen para identificar correlatos neuronales de procesos cognitivos-conductuales.
- Esclarecer las nociones de validez y confiabilidad experimental aplicadas a las investigaciones que se valen de las neuroimágenes funcionales.
- Analizar los desafíos neuroéticos asociados al uso de neuroimágenes funcionales, atendiendo a sus implicancias en contextos médicos, educativos, jurídicos y sociales, así como a su impacto potencial en la representación pública de la mente y en las configuraciones contemporáneas de la subjetividad.

1.3. Estado de la cuestión

Las neurociencias, como ya fue indicado, han tenido un fuerte impulso a partir del desarrollo de las técnicas de neuroimágenes, y de modo muy particular la resonancia magnética funcional. Pero, como bien señalan Nicolás Venturelli e Itatí Branca (2016, p. 178), “su estatus como evidencia no ha sido suficientemente discutido en la literatura filosófica”; además, es escasa la producción en castellano sobre este tema.

Para empezar, es necesario señalar que las limitaciones de la resonancia magnética funcional en términos de explicar o predecir una conducta no son solamente de orden técnico¹. Junto con ellas,

¹ Aquí resulta relevante mencionar el curioso artículo de Bennett et al. (2010) "Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument For Proper Multiple Comparisons Correction", publicado en el *Journal of Serendipitous and Unexpected Results* y ganador del Premio Ig Nobel, que

la filosofía de la neuroimagen propone como necesaria la reflexión acerca de una serie de tópicos, algunos de los cuales enunciaré a continuación, en una enumeración que no es exclusiva ni excluyente:

- a. Los presupuestos teóricos que orientan el diseño de un experimento científico dado (Venturelli y Branca, 2016); las técnicas de neuroimágenes resultan particularmente permeables a estos presupuestos, lo que puede llegar a cuestionar el estatus de evidencia de los resultados obtenidos.
- b. El diseño del experimento científico en sí mismo y el modo en que son obtenidos y procesados los datos (Unturbe, Lago, Alonso, 2007; Armony et al., 2012; Venturelli, 2021).
- c. La manera en la que se construye, a partir de estos datos, la imagen por RMf (Roskies, 2008; Klein, 2010). El producto final, la neuroimagen, guarda cierta similitud con “una fotografía del cerebro” y esto le otorga cierta “verosimilitud” que podría inducir al error de suponer que los procesos psicológicos devienen observables empíricos (Henson, 2005; Poldrack, 2018).
- d. Las condiciones de confiabilidad y validez del diseño experimental, de modo especial en cuanto a la posibilidad de explicar lo que sucede en contextos no experimentales (Poldrack, 2019; Button et al., 2013; Farah y Gillihan, 2013).

Tanto la confiabilidad como la validez de un experimento operan a la manera de restricciones normativas, pero en sentidos opuestos. Si, por una parte, la confiabilidad requiere la simplificación y reducción de las variables en juego, por la otra, la validez implica la inclusión de dimensiones complejas para que los resultados obtenidos sean significativos fuera del laboratorio experimental (Sullivan, 2009; Bickle, 2019).

- e. La solidez de las teorías propias de la psicología cognitiva que dan sustento a dichas investigaciones (Sullivan, 2017). En efecto, no existe una única teoría cognitiva aplicada a la investigación neurocientífica. Los debates actuales, orientados a lograr una ciencia de la cognición unificada, plantean sus construcciones a partir de la integración de explicaciones tanto psicológicas como neurocientíficas, con lo cual se exponen al riesgo de caer en una petición de principio.

más tarde llevó a la elaboración de un segundo artículo: Bennett, C. M. y Miller, M. B. (2010). “How reliable are the results from functional magnetic resonance imaging?” *Annals of the New York Academy of Sciences*, 1191, 133–155. <https://doi.org/10.1111/j.1749-6632.2010.05446.x>

- f. El tipo de inferencias que pueden realizarse a partir de los resultados obtenidos y algunas falacias asociadas (Poldrack, 2006; Klein, 2010b)².

Finalmente, la filosofía de la neuroimagen indaga también en las posibles implicaciones éticas que los resultados de las investigaciones pueden producir en los más variados ámbitos del universo social, de nuestra mirada del mundo y de nuestra comprensión del ser humano (Farah, 2012; Roskies, 2002). En efecto, las investigaciones en neuroética ya han abordado temas como el altruismo, la empatía, el uso de la realidad virtual (Roskies, 2021). Se han abierto nuevos campos de neurociencia social e incluso de neuroeconomía. Y no cabe duda de que avances mucho mayores tendrán lugar en los próximos años, con lo que la reflexión en neuroética, además de ser un desafío actual, tiende a convertirse en una urgencia futura.

1.4. Enfoque conceptual acerca del problema

El presente trabajo de maestría se encuadra conceptualmente dentro del campo de la filosofía de la neurociencia, en tanto subdisciplina de la filosofía de la ciencia.

A menudo se distingue la “filosofía de la neurociencia” de la “neurofilosofía”, en tanto esta última se refiere a la aplicación de conceptos neurocientíficos a las preguntas filosóficas tradicionales (Bickle, 2003, 2019).

Gold y Roskies (2008) se preguntan acerca de las motivaciones para emprender una investigación en filosofía de la neurociencia y señalan, entre otras, dos razones importantes:

En primer lugar, se apunta a **comprender la naturaleza de la ciencia en su conjunto**. A diferencia de otras disciplinas, según estos autores, “la neurociencia es una ciencia que consiste en una abundancia de datos y un gran número de modelos locales, pero ninguna teoría general articulada” (2008).

Dada la carencia de estas teorías generales, los datos que se obtienen a partir de las investigaciones en neurociencias suelen conformar lo que podríamos llamar “teorías de dominio restringido” (Gold y Roskies, 2008). La filosofía de la neurociencia, entonces, puede presentarse como desintegrada frente a la filosofía de otras disciplinas teóricamente más sólidas; no obstante, esa

² Solo a modo ilustrativo, podemos formularnos esta pregunta: si una investigación “demuestra” que un determinado movimiento provoca la activación de la zona conocida como “corteza motora”, ¿la activación de la corteza motora implica necesariamente que se esté realizando dicho movimiento? Este tipo de inferencias, conocidas como “inferencias inversas” son estudiadas, por ejemplo, en Poldrack (2006).

misma falta de un marco teórico universalmente aceptado puede darle “un margen mayor para contribuir al desarrollo conceptual de la neurociencia” (Gold y Roskies, 2008).

En segundo lugar, la filosofía de la neurociencia apunta a **aclarar los conceptos, métodos y principios básicos de esta disciplina**. Dentro de estos, se podrían considerar:

- El principio de las localizaciones funcionales, es decir, la asociación de diferentes funciones con regiones cerebrales específicas (Friston et al., 2006; Saxe, Brett, Kanwisher, 2006; Bunzl, Hanson, Poldrack, 2010).
- La metodología de obtención de datos, basada, de acuerdo con lo señalado en el punto anterior, en el entrelazamiento entre estructura y función. La pregunta central en esta área es la relación existente entre los datos obtenidos sobre la activación de las estructuras cerebrales y las teorías cognitivo-conductuales. Los constructos propios de dichas teorías son débiles, fragmentarios e incluso, de acuerdo con algunos autores como Piccinini y Craver (2011), aunque la psicología cognitiva y la neurociencia son consideradas disciplinas distintas, no constituyen campos explicativamente autónomos (cf. Sullivan, 2016).

La filosofía de la neurociencia plantea también interrogantes que guardan relación con otras diversas ramas de la filosofía, desde la metafísica a la ética.

En este último campo, el de la neuroética, surge una larga lista de desafíos vinculados con la filosofía de la neurociencia. Algunos de estos desafíos pueden ser de naturaleza práctica y actual: por ejemplo, la neuroeducación, la neuroeconomía o el neuromarketing y sus implicaciones individuales y sociales. Otros, en cambio, revisten un valor histórico y filosófico, en la medida en que empiezan a revisarse nociones clásicas dentro de la reflexión filosófica, como el concepto de “persona”, nuestra autocomprensión como agentes morales e incluso como seres espirituales (Farah, 2005).

Finalmente, dirigiendo la mirada al futuro, se vislumbra la necesidad de una profunda reflexión ética vinculada a la aplicación de las técnicas de neuroimagen funcional. Su empleo en contextos clínico-psicológicos o psiquiátricos, en la detección de mentiras o en la predicción de la conducta humana podría llegar a alcanzar dimensiones incluso distópicas (Roskies, 2015, 2015b; Piccinini, 2015).

La neuroimagen es un campo en rápido movimiento y evolución [...]. Sin duda, a medida que las técnicas de imagen se vuelvan más potentes y nuestro conocimiento sobre la organización y la función del cerebro se articule mejor, surgirán problemas éticos en torno a realidades prácticas que ahora parecen solo hipotéticas, y tal vez surjan problemas

completamente nuevos. [...] La neuroimagen plantea desafíos significativos, desafíos que deben abordarse mediante un análisis técnicamente informado, filosóficamente sofisticado y de mente clara. (Roskies, 2015, p. 662)

1.5. Metodología

Utilizaré los métodos propios de análisis lógico-conceptual. En tal sentido, luego de la revisión de los trabajos científicos y metacientíficos más importantes sobre los temas abordados, se realizará el análisis crítico y comparativo de las posiciones sostenidas por los distintos autores. Se buscarán convergencias y divergencias entre ellas y se procederá a la argumentación rigurosa en favor de algunas de las tesis planteadas.

1.6. Plan de trabajo

En el **segundo capítulo**, haré un recorrido por la historia de las representaciones estructurales y funcionales del cerebro a lo largo del tiempo, hasta llegar al surgimiento de las técnicas de neuroimagen. Presentaré los orígenes de estas técnicas hasta llegar a la que concentrará nuestra atención: la resonancia magnética funcional. Describiré allí brevemente cómo se desarrolla la técnica de RMf en la investigación científica, para poner en evidencia algunos problemas fundamentales de orden técnico a la hora de realizar un diseño de investigación.

En el **tercer capítulo**, analizaré el carácter de las imágenes producidas por RMf en tanto fuente de evidencia. Reflexionaré sobre la pregnancia de la neuroimagen contrastándola con las fotografías y veremos distintas posturas sobre los tipos de inferencias. Presentaré de manera intuitiva el teorema de Bayes (fundamental para entender estos procesos) y mencionaré los sesgos de publicación.

En el **cuarto capítulo**, reflexionaré sobre el estatuto de la psicología cognitiva y la viabilidad de que sus constructos sirvan de base a las investigaciones neurocientíficas. A partir de allí, desarrollaré el concepto de “ontología cognitiva”. Finalmente, revisaré el diseño experimental a la luz de los criterios de confiabilidad y validez, así como la relación entre ambos.

El **quinto capítulo** estará dedicado a la neuroética y allí me propongo plantear el impacto que la aplicación de técnicas de neuroimagen puede producir en algunos ámbitos socioculturales.

Como estrategia para favorecer la cohesión interna del texto y propiciar una lectura argumentativa de los contenidos, al final de cada uno de estos capítulos se incluye una breve sección titulada **“Discusión y proyecciones”**.

En ella se sintetizan los principales hallazgos del capítulo, se esboza una toma de posición argumentativa y se anticipan los ejes que serán retomados en el capítulo siguiente. Esta estructura busca evitar una exposición meramente acumulativa y contribuir a una mayor articulación entre las distintas dimensiones del problema abordado.

1.6.1.1. Siglas utilizadas:

RMf: resonancia magnética funcional

iRMf: imagen por resonancia magnética funcional

1.6.1.2. Traducciones:

En las páginas que siguen y salvo que esté expresamente indicado de otro modo, las traducciones del inglés y del italiano son propias.

2. Pasado y presente de las neuroimágenes

Me propongo partir del desarrollo histórico y técnico de la neuroimagen funcional, convencida de que su relevancia epistemológica no puede entenderse sin revisar las ideas que la anteceden. Por eso, en este capítulo, se revisará la evolución de las representaciones estructurales y funcionales del cerebro a lo largo del tiempo, hasta el surgimiento de la resonancia magnética funcional (RMf). Llegado este punto, se describirá esta técnica y las características del diseño de investigación y se señalarán algunos problemas de orden técnico que pueden surgir.

Este recorrido permitirá identificar los supuestos teóricos y metodológicos que, a lo largo de la historia y hasta hoy, acompañaron la construcción del conocimiento sobre la actividad cognitivo-conductual humana.

2.1. Apuntes históricos

Parece casi imposible comenzar un trabajo sobre las neuroimágenes funcionales sin aludir, al menos brevemente, a los hitos científicos y filosóficos que, de alguna manera, sirven de marco histórico a la investigación, mostrando cómo fueron modificándose con el tiempo las representaciones tanto estructurales como funcionales del cerebro.

2.1.1. Antecedentes

Si bien los primeros registros escritos de la descripción de la anatomía cerebral se remontan al antiguo Egipto, con el llamado “Papiro de Edwin Smith” —escrito en torno al año 1600 a.C. y que, según Martín-Araguz et al (2002, p. 187) sería copia de un tratado anterior que dataría del 3600 a. C.—, no existen testimonios gráficos pertenecientes a esta época (Martín-Araguz et al, 2002).

Siglos más tarde, Alcmeón de Crotona (siglo VI a.C.) fue un filósofo médico, contemporáneo de los pitagóricos y considerado el primero que logró establecer la relación anatómica y fisiológica entre el cerebro y los nervios de los sentidos (Outes y Orlando, 2008).

A partir de sus estudios empíricos, Alcmeón infirió que los órganos sensoriales están conectados al cerebro, y el cerebro a los músculos, a través de canales (*poroi*) y cuerdas

(*neura, tonoi*), y fue uno de los primeros defensores de la tesis de que transportan varios "espíritus" (*naturalis, zooticon, psychicon*). (Barberis, Wright, 2022, p. 1).

Alcmeón fue también un propulsor de la teoría de los humores, que más tarde sería sistematizada por Hipócrates y permanecería vigente por muchos siglos.

Hipócrates se valía de un sano escepticismo científico al plantear que las enfermedades no eran de origen divino. En el caso particular de la epilepsia, considerada en el momento como “la enfermedad sagrada”, afirmaba:

A propósito de la llamada enfermedad sagrada he aquí lo que ocurre: me parece que no es en modo alguno más divina ni más sagrada que las demás enfermedades, sino que tiene una causa natural. Pero los hombres creyeron que su causa era divina o por inexperiencia o por el carácter maravilloso de la dolencia, que no se parece en nada a las otras enfermedades. Y si la imposibilidad de conocer lo divino confirma su punto de vista, la banalidad del sistema de curación que adoptan lo contradice, dado que la tratan por medio de purificaciones y encantamientos. (Hipócrates, “Sobre la Enfermedad Sagrada”, en Alsina, 1970)

Ahora bien, si las enfermedades no eran producidas por los dioses sino que se asociaban a causas naturales, la teoría de los cuatro humores se constituiría en su modelo explicativo.

En la Grecia clásica se afirmaba que todo estaba formado por cuatro elementos que se combinaban en distintas proporciones: tierra, fuego, aire y agua. La escuela hipocrática vinculó cada elemento a un fluido o humor corporal, según la siguiente tabla³:

Tabla 2-1

Teoría de los humores

Elemento	Humor	Órgano	Temperamento
Aire	Sangre	Hígado	Sanguíneo/optimista
Fuego	Bilis amarilla	Vesícula biliar	Colérico
Tierra	Bilis negra	Bazo	Melancólico
Agua	Flema	Cerebro / Pulmones	Flemático

³ A título meramente informativo hemos colocado también en el cuadro el “temperamento” que fue asociado a estos humores. Por lo demás, estas asociaciones pueden resultar sorprendentes, pero no son caprichosas ni arbitrarias. La ictericia, por ejemplo, caracterizada por la coloración amarillenta de piel y ojos, estaba asociada a un aumento de la bilis amarilla. En cuanto a la bilis negra, la etimología de la palabra “melancolía” proviene justamente de esa expresión: [melan(o)- μέλαν/μέλαινα gr. 'negro' + khol(ē) χολή gr. 'bilis' + -ía gr. 'cualidad'].

El paso de la salud (*eukrasia*) a la enfermedad (*dyskrasia*) estaría dado, entonces, por un desequilibrio entre los diferentes humores corporales, lo que muchos siglos más tarde se llamaría “pérdida de la homeostasis”. (Barberis, Wright, 2022; Karenberg, 2015; González Álvarez, 2012).

2.1.2. Dos teorías

Desde la antigüedad y aún en nuestros días han existido dos posturas contrapuestas sobre la “sede del alma”: el cardiocentrismo (que establece al corazón como el centro de la vida mental) y el cerebrocentrismo o encéfalocentrismo (que sostiene que el asiento del alma es el cerebro).

El corazón mítico ha sido representado y sigue haciendo como sede de las emociones, de las intuiciones, los impulsos, las ideas y las sensaciones físicas del pecho y del cuerpo que asociamos con el amor (Baeza, 2001). Esta visión tiene varios milenios y no ha desaparecido del todo. La convicción de que las funciones psíquicas residen en el cerebro es una de las ideas que más han demorado en aceptarse, levantando polémicas y teniendo que vencer obstáculos epistemológicos que provienen de distintos puntos de vista como el científico e ideologías sociales, culturales y religiosas. (Salazar, 2016, p. 21)

La concepción aristotélica, por ejemplo, era cardiocéntrica:

Se puede afirmar que, en líneas generales, Aristóteles adopta una concepción "termodinámica" del funcionamiento del organismo. El ser vivo consta de un polo caliente, el corazón, centro también de la sensación, y dos sistemas de refrigeración que mantienen el equilibrio térmico: la respiración y la acción refrigerante del cerebro. (Alfredo Marcos en Aristóteles, 2010, *Introducción*, p. 48)

Dice Aristóteles en el capítulo 7 de su obra *De Partibus Animalium*: “La frialdad del cerebro se manifiesta al tacto y, además, es la menos sanguínea de todas las partes húmedas del cuerpo (de hecho, ni siquiera contiene sangre en sí mismo)” [...] “Así pues, **el cerebro regula el calor y la ebullición del corazón**”.

Algunos autores contemporáneos asocian metafóricamente la posición del Estagirita con la función moduladora o inhibitoria del cerebro:

Podría decirse que, en un sentido muy elemental, pero también esencial y originario, los *Tratados biológicos* contienen la primera visión (*theoria*) del sistema nervioso como órgano (en su sentido etimológico, instrumento) dirigido al control (inhibitorio) de la acción interna y externa del organismo. Podría decirse también, con una metáfora artística,

como las que abundan en los *Tratados*, que la noción de cerebro en Aristóteles se inscribe en un cardiocentrismo “bien temperado”. (Rábano, 2019, p. 143)

En oposición, Galeno sostenía una posición cerebrocéntrica, según la cual:

[...] los nervios estaban huecos y a través de ellos viajaban los espíritus animales que movían los músculos o producían las sensaciones. Estos espíritus se originaban en la sustancia del cerebro, pero se almacenaban en los ventrículos, o cavidades cerebrales, hasta que se necesitaban. (González Álvarez, 2012, p. 16)⁴

2.1.3. Escuela de Alejandría

En la escuela de Alejandría (siglo III a. C.) sus máximos representantes, Herófilo (325-280 a. C.) y Erasítrato de Cos (304-250 a. C.) fueron los primeros médicos griegos en realizar disecciones sistemáticas de cadáveres humanos. Se destacan en Herófilo sus trabajos de anatomía cerebral y en Erasítrato su estudio de las circunvoluciones cerebrales (García de Lucas, 2020). Este último llegó incluso a anticipar la circulación sanguínea, descubierta por William Harvey recién en el siglo XVII. Pero quizá su mayor aporte fue su giro decisivo hacia métodos empíricos que permitió entender al cerebro como un “órgano de gobierno” (Barberis y Wright, 2022, p. 2) y facilitó el triunfo del cerebrocentrismo sobre las teorías cardiocentristas.

2.1.4. Galeno de Pérgamo

Poco a poco la disección cadavérica fue cayendo en el olvido, dado que la sociedad grecorromana consideraba “el examen de los cadáveres como algo vil y repugnante” (Vargas Sanabria, 2014). En efecto, los tabúes al respecto hicieron que Galeno de Pérgamo (129-199 d. C.) debiera realizar disecciones animales para sus estudios de anatomía (García de Lucas, 2020).

A pesar de esta limitación, es significativo el aporte de Galeno al conocimiento de sistema nervioso: entre otros hallazgos, identificó dos de las tres membranas menínges (piamadre y duramadre); estudió los nervios motores y sensoriales (a los que supuestamente distinguía por su dureza, dado que estos últimos debían ser más maleables y alterables por la influencia de los

⁴ El concepto de “espíritus animales” era muy anterior a la época de Galeno: ya estaba presente en los tiempos de los médicos alejandrinos del s. III a.C. Estos espíritus:

Se trata de los *pneuma* psíquicos, derivados de los *pneuma* vitales, en el sentido del término griego que significa aire, soplo, espíritu, hálito. Arranca de la creencia alejandrina de que la vida está asociada a un sutil vapor o *pneuma* relacionado con la respiración. El aire entra en los pulmones a través de la tráquea y se transforma en neumonía vital. (González Álvarez, 2012, p. 16)

estímulos del exterior) (Barberis, Wright, 2022, p. 3). En otro orden, Galeno sostuvo la teoría humoral, cerebrocéntrica y ventricular.

2.1.5. Edad Media

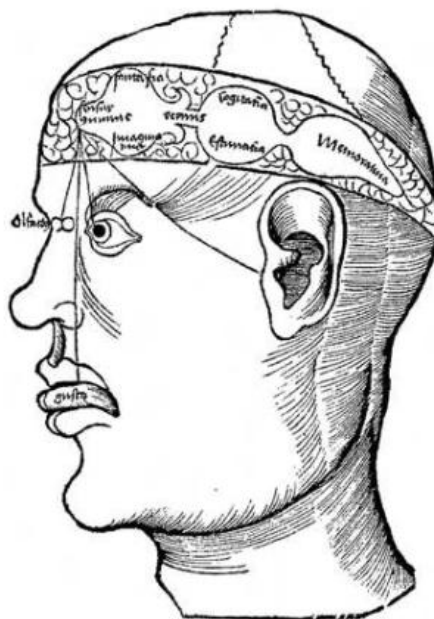
Tanto en Oriente Medio como en Europa, la creencia de la localización ventricular (ventrículos distintos asociados a diferentes funciones), que fuera propuesta por Galeno, se convirtió en una especie de dogma que tuvo vigencia a lo largo de toda la Edad Media.

La idea general era la siguiente:

- a) La percepción está asociada a un ventrículo delantero, una inexistente cavidad frontal del cerebro, creencia basada posiblemente en la proximidad de los órganos sensoriales. Junto con estos sentidos externos, se creía la existencia de “sentidos internos” que formaban *un senso comune*, donde residía el alma humana.
- b) En el ventrículo del medio se ubicaba la razón o cognición (específicamente humana) y el juicio o la estimación, facultad innata y fija compartida con los animales.
- c) La memoria se ubicaría en el ventrículo trasero. Es decir, una vez que se han realizado las estimaciones y cogniciones, sus resultados se conservan para el futuro en esa cavidad posterior.

Figura 2-1

Ilustración del siglo XV de Gregor Reisch que muestra la concepción medieval de las funciones mentales.



Nota: Tomado de González A., 2012.

Solo un pequeño matiz diferenciaba la doctrina galénica de las creencias medievales. Si Galeno creía que los espíritus animales se originaban en la materia sólida del cerebro, los Padres de la iglesia creían que el tejido cerebral era demasiado “sucio” para oficiar de intermediario entre el cuerpo y el alma. Por ello, se sostenía que los espíritus animales se generaban en las cavidades huecas de los ventrículos (González Álvarez, 2012, p. 17).

Hacia los siglos XI y XII, en el marco de la llamada *alta Edad Media*, comienzan a soplar vientos de cambio. La fundación de la primera universidad en Bolonia en 1088 —y con ella, el advenimiento del sistema universitario— dio lugar a cierta institucionalización de los saberes. En torno a 1260, Bolonia incorpora una escuela de medicina, y empieza a distinguirse más claramente entre los cirujanos barberos vocacionales y los profesores de medicina con formación universitaria (Barberis y Wright, 2022, p. 3). Este proceso marca el inicio de una transformación en el modo de producción y transmisión del conocimiento anatómico.

Las primeras concepciones sobre el cerebro, entonces, oscilaron entre enfoques cardiocentristas y cerebrocentristas, reflejando así el modo en que las explicaciones científicas se entrelazaban con creencias filosóficas y religiosas propias de la época, como la noción del corazón como “asiento del alma”. A pesar del avance en el conocimiento anatómico producido en la Escuela de Alejandría, el dogma ventricular medieval consolidó la visión de un cerebro estructurado en compartimentos funcionales. Este marco teórico perduró durante siglos y sentó las bases para la posterior **localización funcional de la cognición**, una idea que resurgiría con fuerza en el siglo XIX con la frenología y los estudios de Broca y Wernicke. Hoy en día, muchas de las investigaciones contemporáneas con neuroimágenes funcionales pueden leerse como una continuación —si bien epistemológicamente más compleja— de esta aspiración localizadora.

2.1.6. Renacimiento

En 1543, Andrés Vesalio (1514-1564) publica su obra magna *De Humani Corporis Fabrica*⁵, que reexamina en su libro VII la estructura del cerebro.

⁵ El libro puede ser consultado y descargado de <https://ucm.on.worldcat.org/discovery>

Figura 2-2.

Imagen del libro VII de la obra De Humani Corporis Fabrica (1543). Se muestra el cerebro de frente, con los hemisferios cerebrales abiertos, proporcionando una visión del cuerpo calloso y la apófisis crista galli.



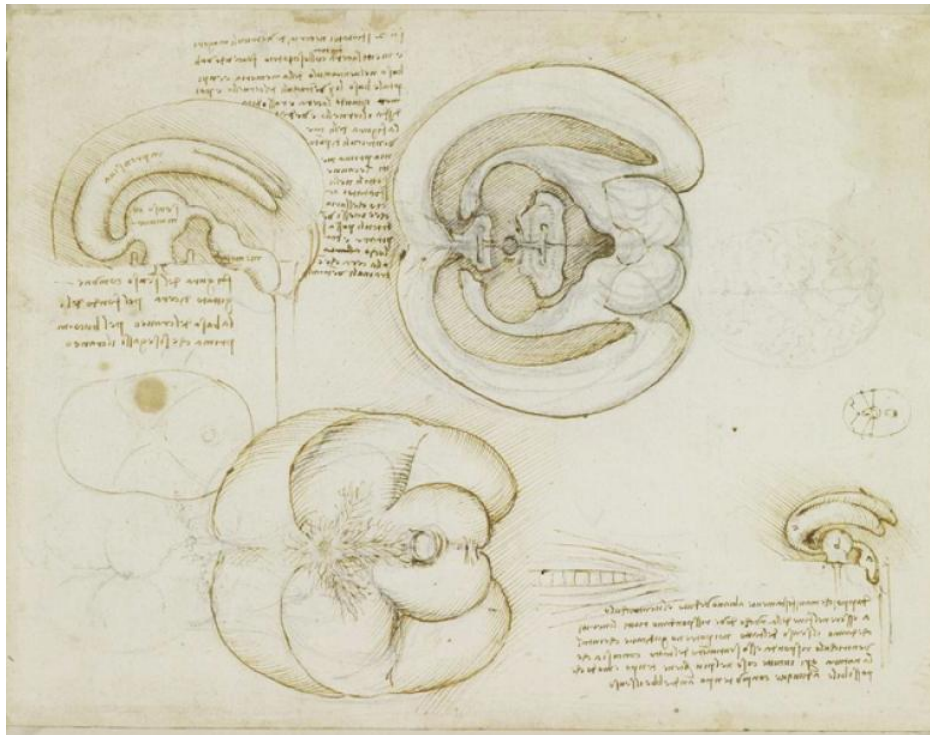
Nota: tomado de <https://www.medigraphic.com/pdfs/actmed/am-2015/am153m.pdf>

Por esta misma época, el cerebro comienza a representarse no solo por el dibujo, sino con la construcción de modelos en cera. Pionero en esta técnica fue el polímata Leonardo da Vinci, quien estudió los ventrículos cerebrales inyectando cera en el cerebro de los cadáveres y obteniendo a partir de allí figuras tridimensionales (García de Lucas, 2020).

No obstante, ni Leonardo Da Vinci ni Andrés Vesalio pusieron en duda la función del cerebro impulsado por los “espíritus animales”.

Figura 2-3.

Leonardo da Vinci dibujó los ventrículos cerebrales después de extraer modelos tridimensionales de cera inyectada en el cerebro (1504-1507).



Nota: Imagen tomada de la Royal Collection Trust. <https://www.rct.uk/>

2.1.7. René Descartes

Tampoco Descartes cuestiona los ancestrales espíritus animales, pero intentará explicar su funcionamiento de acuerdo con el conocimiento tecnológico de la época (el paradigma de referencia era el mecánico).

Para el filósofo francés, también los espíritus animales están encerrados en los ventrículos cerebrales pero, al contrario que en las explicaciones medievales, ahora no actúan de forma misteriosa, sino que siguen los principios mecánicos de la hidráulica a través de los nervios huecos. [...]

Es la glándula pineal la que crea los espíritus animales a partir de la sangre y estos se van acumulando en la cavidad del ventrículo que sirve de reservorio. A partir de aquí, la forma de actuar de estos espíritus responde simplemente a las leyes de la hidráulica. (González Álvarez, 2012, pp. 30-31).

Si bien el “dualismo cartesiano” suele invocarse, cargado de connotaciones negativas, como un obstáculo en los debates contemporáneos sobre la mente, es necesario reconocer que la propuesta de Descartes, en el contexto de su tiempo, constituyó una estrategia epistemológica que permitió desvincular el estudio del cuerpo de las restricciones teológicas. Su mecanicismo, *lejos de ser un error*, sentó las bases para el desarrollo de la fisiología moderna.

La elección del término “error” en el párrafo anterior es deliberada y alude al libro de António Damásio, *El error de Descartes* (1994): un título por demás sugerente pero que supone, a mi juicio, una lectura anacrónica del proyecto cartesiano. En contraste con el monismo emergentista postulado por Damásio, en este trabajo defiende la idea de que las teorías sobre **mente y cuerpo son, al menos parcialmente, inconmensurables**: la pretensión de explicar uno en los términos del otro necesariamente implica algún tipo de pérdida conceptual. Considerar uno como fundamento exclusivo del otro implica incurrir en un error o en una reducción categorial, esto es, aplicar al dominio de lo mental las categorías ontológicas o explicativas del dominio biológico⁶. En el marco del racionalismo del siglo XVII, la distinción cartesiana de sustancias permitió entender al cuerpo según principios mecánicos, mientras la conciencia era irreducible al mundo de lo físico. De alguna manera, el dualismo tuvo un valor emancipador, en tanto permitió el estudio científico del cuerpo sin que esto constituyera una “herejía”, mientras se reservaba para el alma una esfera inmaterial.

Así, el dualismo cartesiano constituyó para los tiempos posteriores una suerte de **matriz conceptual** que organizó diversas —y hasta antagónicas— formas de pensar la experiencia, el conocimiento y la subjetividad. Entre ellas:

- Con la separación radical entre sujeto y objeto, Descartes funda un modelo donde conocer comporta representar lo otro desde una posición exterior, “clara y distinta”. El sujeto se ubica como espectador racional del mundo, lo cual contribuyó a la consolidación del concepto moderno de ciencia, basado en la observación, medición y control de objetos separados del observador.
- La lectura reduccionista que se consolidó a partir del avance de las ideas positivistas, de la psicofisiología y, más recientemente, de la neurociencia. Lo que en Descartes era una

⁶ Este problema no es exclusivo del dualismo cartesiano, sino que se manifiesta también en la evolución de la neurociencia. Como señala Barberis (2024), la transición desde la afasiología clásica de Wernicke a la neurología del comportamiento de Geschwind es un caso de sustitución con inconmensurabilidad parcial, en el que algunos conceptos se mantuvieron, pero otros cambiaron de significado o fueron redefinidos en un marco teórico distinto. De manera similar, la disparidad entre teorías dualistas y monistas en filosofía de la mente refleja el problema más amplio de la discontinuidad entre los lenguajes explicativos de las diferentes disciplinas.

distinción ontológica que permitía organizar el saber terminó convertido en uno de los más decisivos soportes del materialismo.

Esta matriz conceptual dualista dejó una profunda huella en la ciencia moderna y contemporánea, tanto en aquellas teorías que retoman de forma más o menos explícita la distinción entre mente y cuerpo como en las que se constituyen por oposición a ella.

2.1.8. La frenología

El sistema fue creado por Franz Joseph Gall, médico alemán nacido en 1758. Su obra más importante, *Anatomie et physiologie du système nerveux en général et du cerveau en particulier, avec des observations sur la possibilité de reconnaître plusieurs dispositions intellectuelles et morales de l'homme et des animaux par la configuration de leurs têtes* (Anatomía y fisiología del sistema nervioso en general y del cerebro en particular, con observaciones sobre la posibilidad de reconocer diversas disposiciones intelectuales y morales del hombre y de los animales por la configuración de sus cabezas) en cuatro tomos, se publicó entre 1810 y 1819, año en que Gall se naturalizó francés. (Domènech, 1977, p. 12)

De acuerdo con Domènech, se basa en los siguientes postulados:

- a. El cerebro no debía ser considerado como un órgano único, sino como una víscera en la que distintas regiones correspondían a una función determinada (doctrina de las localizaciones cerebrales). Gall estableció la existencia de 27 órganos.
- b. El estado de las funciones estaba en relación directa con el tamaño del órgano en el que se asentaba.
- c. La forma de cada una de las regiones del cerebro influía de modo directo en la forma de la cubierta ósea.
- d. Así era posible conocer el estado de cada órgano mediante el estudio externo del cráneo o craneoscopía. (Domènech, 1977, p. 10)

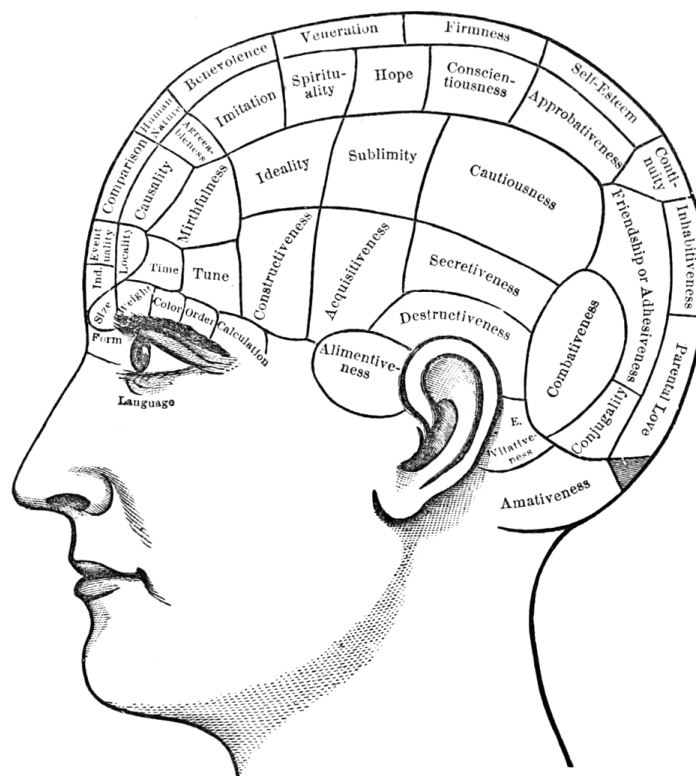
Entre otros, Gall describió un órgano relacionado con el instinto de agresión o *Wurgsinn*, llamado en francés *organe du meurtre* (órgano del asesinato). Si bien puede ser considerado un antecedente de las ideas de Lombroso sobre el determinismo criminológico, su intención era mucho más amplia, en tanto intentaba construir una doctrina completa para explicar la conducta humana como dependiente de factores previos, vinculados con el cerebro. En otras palabras: postulaba el determinismo en la conducta. Además, se observa cómo estamos en los inicios de la doctrina de las localizaciones cerebrales y también de una teoría organicista de la actividad mental.

Finalmente, debemos tener presente la importancia que tuvo en Francia el debate sobre las localizaciones cerebrales y, en particular, la facultad del habla, con los aportes de Paul Broca:

A pesar de que éste en principio no se mostró adversario de la frenología, el hecho es que sus aportaciones, y la escuela que abrió, marcaron el camino cierto para el conocimiento de la actividad del sistema nervioso central, en lo que toca a las localizaciones cerebrales. (Domènech, 1977, p. 31).

Figura 2-4.

Imagen de la frenología que muestra los diferentes “órganos” de la conducta.



NUMBERING AND DEFINITION OF THE ORGANS.

- | | |
|---|--|
| 1. AMATIVENESS, Love between the sexes. | 20. CONSTRUCTIVENESS, Mechanical ingenuity. |
| 2. CONJUGALITY, Matrimony—love of one. [etc. | 21. IDEALITY, Refinement—taste—purity. |
| 3. PARENTAL LOVE, Regard for offspring, pets, | 22. IMITATION, Copying—patterning. |
| 4. FRIENDSHIP, Adhesiveness—sociability. | 23. MIRTHFULNESS, Jocoseness—wit—fun. |
| 5. INHABITIVENESS, Love of home. | 24. INDIVIDUALITY, Observation. |
| 6. CONTINUITY, One thing at a time. | 25. FORM, Recollection of shape. |
| 7. VITATIVENESS, Love of life. | 26. SIZE, Measuring by the eye. |
| 8. COMBATIVENESS, Resistance—defense. | 27. WEIGHT, Balancing—climbing. |
| 9. DESTRUCTIVENESS, Executiveness—force. | 28. COLOR, Judgment of colors. |
| 10. ALIMENTIVENESS, Appetite—hunger. | 29. ORDER, Method—system—arrangement. |
| 11. ACQUISITIVENESS, Accumulation. | 30. CALCULATION, Mental arithmetic. |
| 12. SECRETIVENESS, Policy—management. | 31. LOCALITY, Recollection of places. |
| 13. CAUTIOUSNESS, Prudence—provision. | 32. EVENTUALITY, Memory of facts. |
| 14. APPROBATIVENESS, Ambition—display. | 33. TIME, Cognizance of duration. |
| 15. SELF-ESTEEM, Self-respect—dignity. | 34. TUNE, Sense of harmony and melody. |
| 16. FIRMNESS, Decision—perseverance. | 35. LANGUAGE, Expression of ideas. |
| 17. CONSCIENTIOUSNESS, Justice—equity. | 36. CAUSALITY, Applying causes to effect. [tion. |
| 18. HOPE, Expectation—enterprise. | 37. COMPARISON, Inductive reasoning—illustra- |
| 19. SPIRITUALITY, Intuition—faith—credulity. | C. HUMAN NATURE, Perception of motives. |
| 20. VENERATION, Devotion—respect. | D. AGREEABLENESS, Pleasantness—suavity. |
| 21. BENEVOLENCE, Kindness—goodness. | |

Nota: Tomado de Sizer, N. y Drayton, H. S. (1892). *Heads and Faces, and how to Study Them: A Manual of Phrenology and Physiognomy for the People*. Fowler & Wells, 203.

2.1.9. Paul Broca y Carl Wernicke

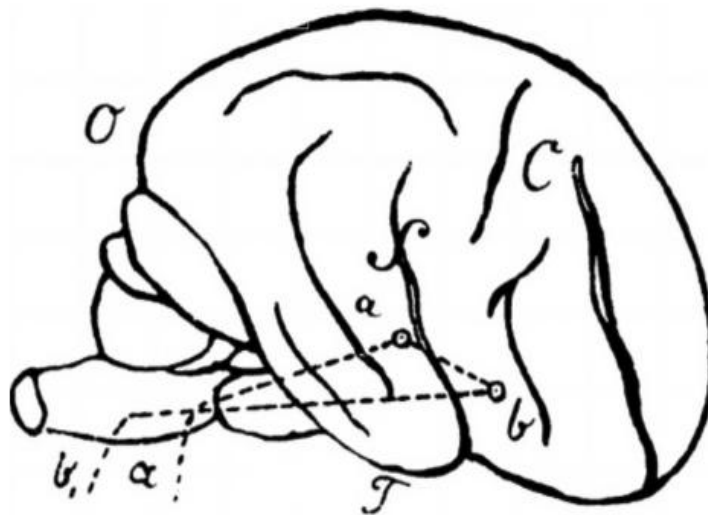
El neurólogo francés Pierre-Paul Broca (1824-1880) presentó en 1861 el caso de Monsieur Leborgne, un paciente epiléptico con una parálisis hemilateral y un singular trastorno del habla.

Se le conocía como “Monsieur Tan”, o “Tan-Tan”. Parecía comprender el lenguaje sin dificultad y podía mover bien los músculos de la boca y la lengua, pero sólo conseguía pronunciar algo parecido a la sílaba “tan”, de ahí su apodo. Cuando Leborgne muere una semana después de su ingreso, Broca le practica la autopsia y encuentra una lesión importante en su cerebro, en el lóbulo frontal del hemisferio izquierdo (González Álvarez, 2012, pp. 64-75).

En la década siguiente, un neurólogo alemán de origen polaco, Carl Wernicke (1848-1904), describió un nuevo tipo de trastorno del lenguaje, resultante de lesiones en el lóbulo temporal izquierdo. Se trataba de la afasia sensorial o de Wernicke, que presenta dificultades de comprensión (González Álvarez, 2012, p. 78).

Figura 2-5.

El modelo de Wernicke.



Nota: tomado de Barberis y Wright, 2020.

Podemos observar que la frenología, si bien fue prontamente desacreditada, constituyó el primer intento sistemático de “mapear” la relación entre cerebro y comportamiento. Por su parte, las investigaciones posteriores de Broca y Wernicke contribuyeron a validar la idea de que ciertas funciones específicas podían asociarse con regiones cerebrales concretas, impulsando el enfoque localizacionista. Este principio orientó el desarrollo de la neurociencia moderna y aún hoy es central en las técnicas de neuroimagen funcional.

Este énfasis en la localización ha sido objeto de críticas contemporáneas. William Uttal (2001) describió la interpretación de neuroimágenes como una “nueva frenología”, advirtiendo que la confianza excesiva en la asignación estricta de funciones cognitivas a regiones específicas del

cerebro, junto con la confusión entre correlación y causalidad y otros problemas metodológicos, comprometen severamente la validez de los resultados.

2.1.10. Las intuiciones de Brodmann

Korbinian Brodmann (1868-1918) fue un neurólogo alemán que describió la corteza cerebral en función de 52 regiones distintas de acuerdo con su **citoarquitectura**. Si bien su trabajo se funda en el estudio de las características estructurales del cerebro humano y otros cerebros animales, con el paso del tiempo sirvió de base para los modelos teóricos basados en el localizacionismo funcional.

En el prólogo de su libro, Brodmann señala que se le había encomendado la tarea de

[...] realizar un análisis topográfico de la corteza cerebral del ser humano según su estructura celular, para lograr una división topográfico-localizatoria detallada de la misma, que tuviera aplicación clínica y también ofreciera una imagen de la estructura histológica normal, para su aplicación en la patología. (Brodmann en Fontana, 2010, S2)

Brodmann analiza datos cuantitativos sobre la corteza humana, comparándola a su vez con la de primates y no primates. Estudia también la variabilidad en la topografía cortical, cuya diferenciación celular expresaría una especificación histológica y una división funcional del trabajo.

El polimorfismo de las células, su diferenciación histogenética asincrónica, la estricta separación regional de ciertos tipos celulares y la aparición regular de tipos celulares similares (homólogos) en posiciones idénticas en la superficie cortical de todos los mamíferos, justifican la visión de que en la corteza cerebral ha ocurrido una amplia especialización de funciones entre los elementos celulares; en otras palabras, que la especificidad funcional está dividida entre células de diferente morfología individual, diferente localización y diferente capacidad (Brodmann, Garey, 2005, p. 251).

Brodmann ofrece algunas reflexiones interesantes sobre la relación entre estructura y función:

- No es posible deducir la importancia psicológica a partir de la mayor masa del grupo, “sino por la organización interna (química, dinámica o aún histológica, pero inaccesible a las demostraciones microscópicas del momento)” (Fontana 2010, S13).
- Si bien confirma la presencia de una localización regional estricta de ciertas funciones, sugiere que las funciones complejas se corresponden con un complejo de regiones (Fontana 2010, S13).

- En cuanto a los procesos psíquicos superiores, señala que requieren el entrelazamiento e interacción de procesos elementales.

En verdad, existe sólo un centro psíquico y este es el cerebro, como un todo, con todos sus órganos. No se debería decir con esto que para los procesos psíquicos superiores todos los órganos particulares colaboran de igual manera, sino que, para cada caso concreto, los “focos funcionales” a disposición entran en acción en distinto número, grado y forma de conexión (Fontana 2010, S14).

- Sobre la superposición de funciones:

En aquellas áreas con límites poco netos, hay superposición de campos y, se podría suponer, también de funciones. Se entiende con esto que existen, para diferentes zonas del organismo, capacidades corticales subordinadas que se superponen parcialmente en la corteza. Se trata de territorios que pueden también abastecer otras funciones (*principio de la representación múltiple de territorios corticales*). Estas zonas serían quizá responsables de la restitución de funciones perdidas por procesos patológicos incluso luego de años de instalado el déficit, por el reemplazo de las zonas lesionadas por otras relacionadas (*principio de la suplencia funcional*). Esta restitución suele ser debida a la ejercitación (Fontana 2010, S14).

Figura 2-6.

Las áreas de Brodmann



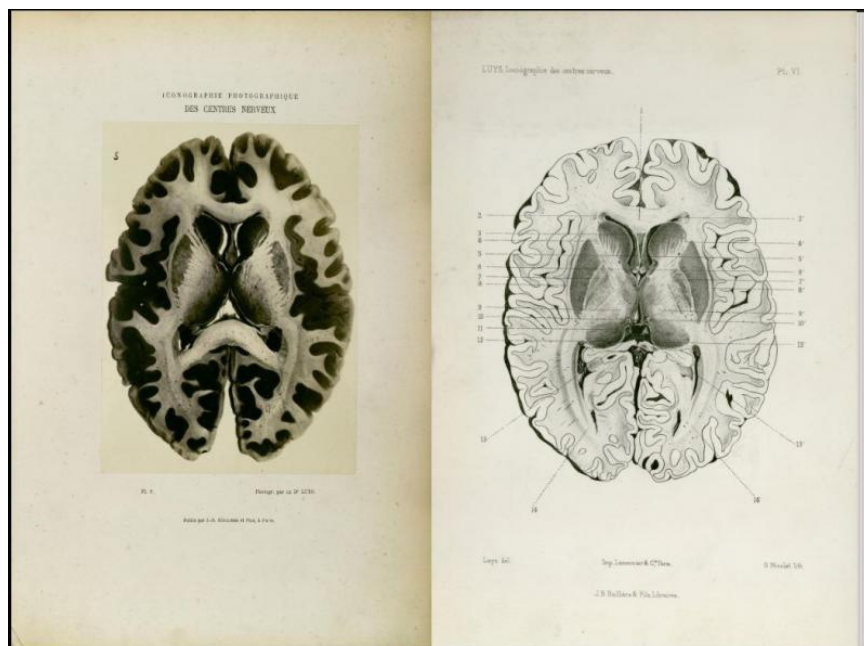
Nota: tomado de Brodmann, 1909, p. 131.

2.1.11. Técnicas de visualización

Tras la invención del daguerrotipo en 1839, por Louis Daguerre (1787–1851), poco a poco la fotografía se fue convirtiendo en una poderosa herramienta: es así como, en 1873, el francés Jules Bernard Luys (1828–1897) publicó el primer atlas neuroanatómico fotográfico: *Iconographie Photographique des Centres Nerveux*⁷.

Figura 2-7.

Imagen perteneciente al primer atlas neuroanatómico fotográfico, *Iconographie Photographique des Centres Nerveux* (1873). Las fotografías (izquierda) del atlas se acompañaban con una litografía (derecha) en la página contigua.



Nota: tomado de *Iconographie Photographique des Centres Nerveux* (1873)

La incorporación de la fotografía en el ámbito de las neurociencias supuso un cambio significativo en la representación visual del cerebro, pero no resolvió la cuestión de la interpretación de las imágenes. Este problema se amplificaría más tarde con la llegada de la neuroimagen moderna, que, aun teniendo a menudo una apariencia casi “fotográfica” (como veremos en los siguientes capítulos) se funda en mediciones basadas en supuestos tanto teóricos como metodológicos.

En paralelo con el desarrollo de la fotografía, comenzaron a desarrollarse técnicas de visualización más precisas. Me refiero, en particular, al método de tinción con sales de plata desarrollado por Camillo Golgi. Esta técnica permitía observar con gran detalle las estructuras del sistema nervioso,

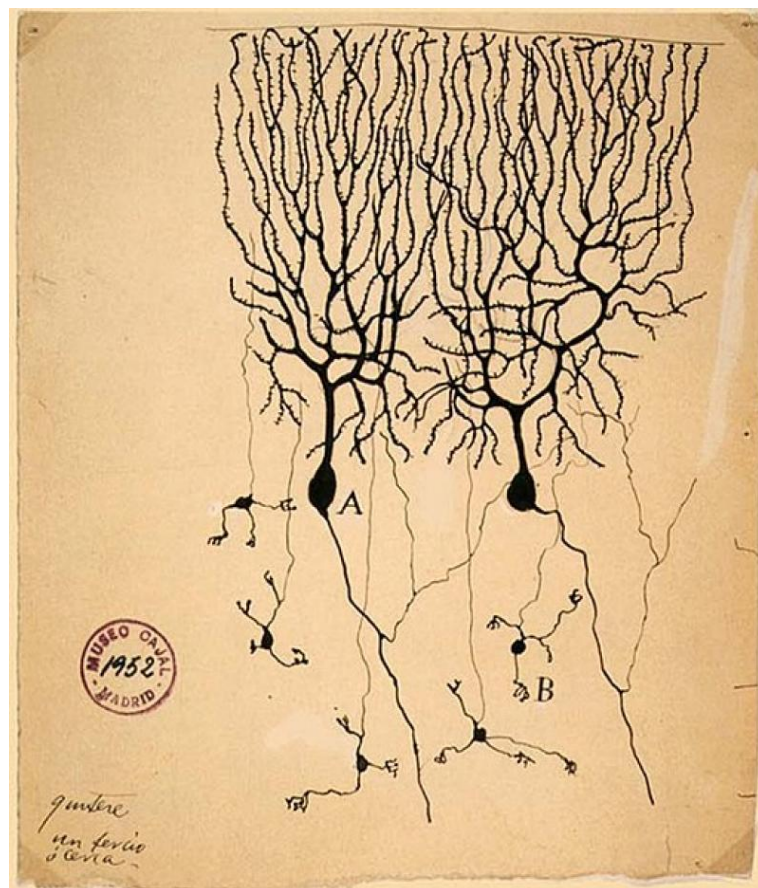
⁷ Puede ser descargado de <https://patrimoine.sorbonne-universite.fr/fonds/>

aunque Golgi interpretaba que las neuronas formaban una red continua, en una concepción conocida como **teoría reticular**.

Poco más de una década después, Santiago Ramón y Cajal (1852–1934) adoptó y perfeccionó el método de Golgi, utilizándolo para llegar a una conclusión radicalmente distinta: formuló así la “teoría de la neurona”, según la cual el sistema nervioso está compuesto por células discretas y separadas. Esta interpretación, sustentada en evidencia empírica y plasmada en sus detallados dibujos, se impondría finalmente como el paradigma dominante de la neurociencia moderna. Paradójicamente, en 1906, ambos científicos compartieron el Premio Nobel, a pesar de sus posturas enfrentadas.

Figura 2-8.

Dibujo de Santiago Ramón y Cajal (1899) de diferentes tipos de neuronas del cerebelo de una paloma. (A) célula de Purkinje, un ejemplo de neurona bipolar. (B) célula granulosa que es multipolar.



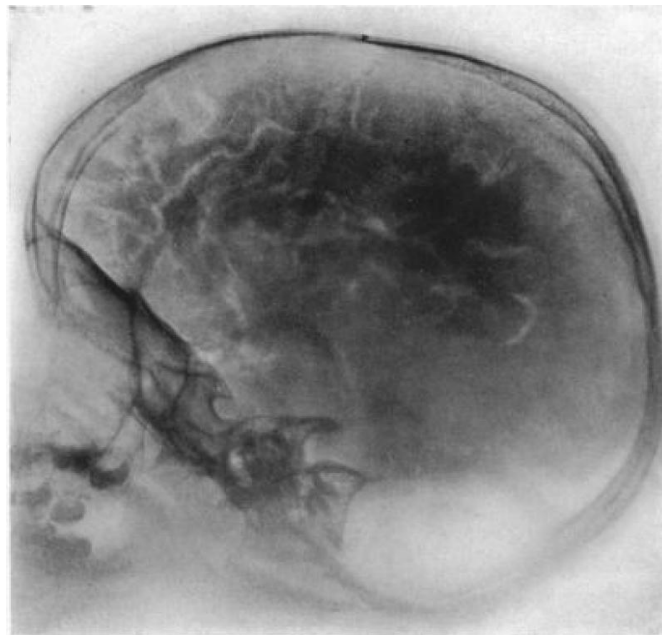
Nota: tomado de https://www.wikiwand.com/es/articles/C%C3%A9lulas_de_Purkinje

2.1.12. Nuevos desarrollos

El descubrimiento de los rayos X por Wilhelm Roentgen (1845–1923) en diciembre de 1895 tuvo fuerte impacto en medicina, pero no tanto en neurología, dado que el cerebro no podía ser visualizado con esta técnica. En 1918, Walter Dandy (1886–1946), neurocirujano del hospital John Hopkins, desarrolló la ventriculografía: a través de esta técnica se introducía aire en los ventrículos laterales del cerebro para mejorar el contraste del sistema ventricular. Tiempo después, en 1919, introdujo una variante de esta técnica conocida como neumoencefalografía (García de Lucas, 2020). Pero ambos procedimientos tuvieron un uso muy limitado, dados los riesgos inherentes a su aplicación.

Figura 2-9.

Un ejemplo del aspecto de una placa radiográfica obtenida en 1919 utilizando el procedimiento de la neumoencefalografía.

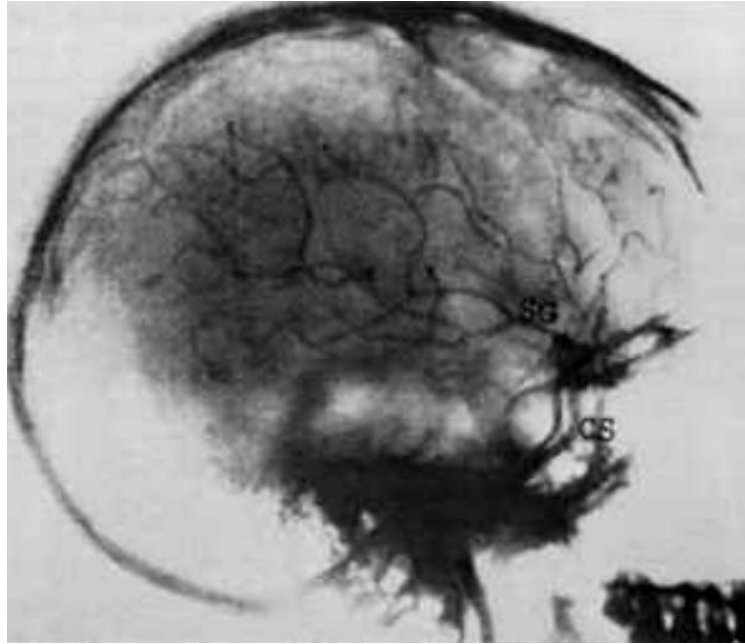


Nota: tomado de <https://radiopaedia.org/cases/pneumoencephalography>

El siguiente hito importante en el campo de la neuroimagen fue el desarrollo de la angiografía cerebral por el neurólogo portugués Egas Moniz (1874–1955) en 1927. Este procedimiento consistía en la inyección intravenosa de un compuesto opaco a los rayos X para visualizar los vasos sanguíneos del cerebro. Sin embargo, al igual que las dos técnicas mencionadas anteriormente, su uso se vio limitado debido al riesgo que representaba para los pacientes.

Figura 2-10.

Primera angiografía exitosa realizada por Egas Moniz.



Nota: imagen tomada de <https://www.redalyc.org/pdf/3825/382538480002.pdf>

Junto con los aspectos anatómicos, el cerebro empezó a estudiarse a través de la obtención de imágenes o registros de su actividad eléctrica o metabólica. En particular, el registro de la actividad eléctrica del cerebro, conocido como EEG (electroencefalograma), se practicó desde la década de 1930 (Beaumont, 2008, p. 264).

2.1.13. La revolución en neuroimagen

En 1971 y a partir de los trabajos de Allan McLeod Cormack, el ingeniero Godfrey Hounsfield (1919-2004) junto con el radiólogo Jamie Ambrose (1923-2006) lograron generar por primera vez una imagen del cerebro a través de la tomografía computarizada (CT por sus siglas en inglés) (García de Lucas, 2020), abriendo así las puertas al desarrollo de la neurocirugía y de la neuroimagen moderna. Godfrey Hounsfield y Allan Cormack obtuvieron el Premio Nobel en 1979. La técnica emplea rayos X dirigidos desde un gran número de orientaciones, que, mediante métodos computacionales, permite reconstruir “cortes” a través del cuerpo (Beaumont, 2008, pp. 280-281).

Figura 2-11.

Tomografía axial computarizada de cráneo del paciente. Presenta hematoma en proyección del núcleo lenticular izquierdo de 20 mm de diámetro mayor, rodeado de edema perilesional (flecha).



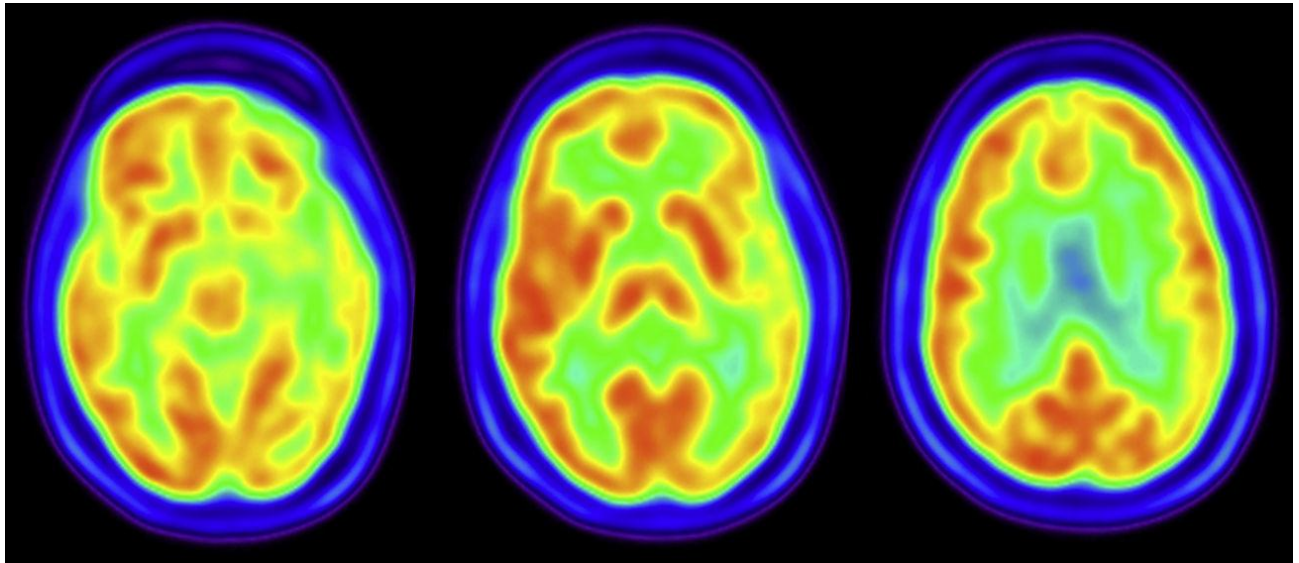
Nota: tomado de Cogollo G, Marysabel, Julio N, Luis Carlos, & Alvarado C, Daniel Enrique. (2016). Hemorragia intracraneal asociada a preeclampsia severa. *Revista chilena de obstetricia y ginecología*, 81(6), 511-514. <https://dx.doi.org/10.4067/S0717-75262016000600010>

A partir de ese momento, los avances en el campo de la neuroimagen comenzaron a multiplicarse, dando lugar al desarrollo de diversas tecnologías orientadas al estudio estructural y funcional del cerebro. Entre ellas, la resonancia magnética —en sus modalidades estructural y funcional— se ha consolidado como una de las técnicas más utilizadas, motivo por el cual le dedicaré un apartado específico en este trabajo.

En 1975, Michael Phelps y Edward Hoffman desarrollaron el primer escáner de tomografía por emisión de positrones (PET por sus siglas en inglés). Mediante esta técnica, es posible estudiar el metabolismo cerebral a través de la inyección de moléculas radiactivas.

Figura 2-12.

En la imagen PET se visualiza el metabolismo de la glucosa en el cerebro a través de una molécula similar a esta que contiene un átomo de flúor radiactivo. Las áreas de color rojo representan las que más glucosa están captando.



Nota: tomado de Matías-Guiu, J. A., García-Ramos, R., Cabrera-Martín, M. N., Moreno-Ramos, T. (2015). Afasia progresiva logopénica asociada a enfermedad de Parkinson idiopática [Logopenic progressive aphasia associated with idiopathic Parkinson's disease]. *Neurología*, 30(8), 521–524. <https://doi.org/10.1016/j.nrl.2014.01.007>

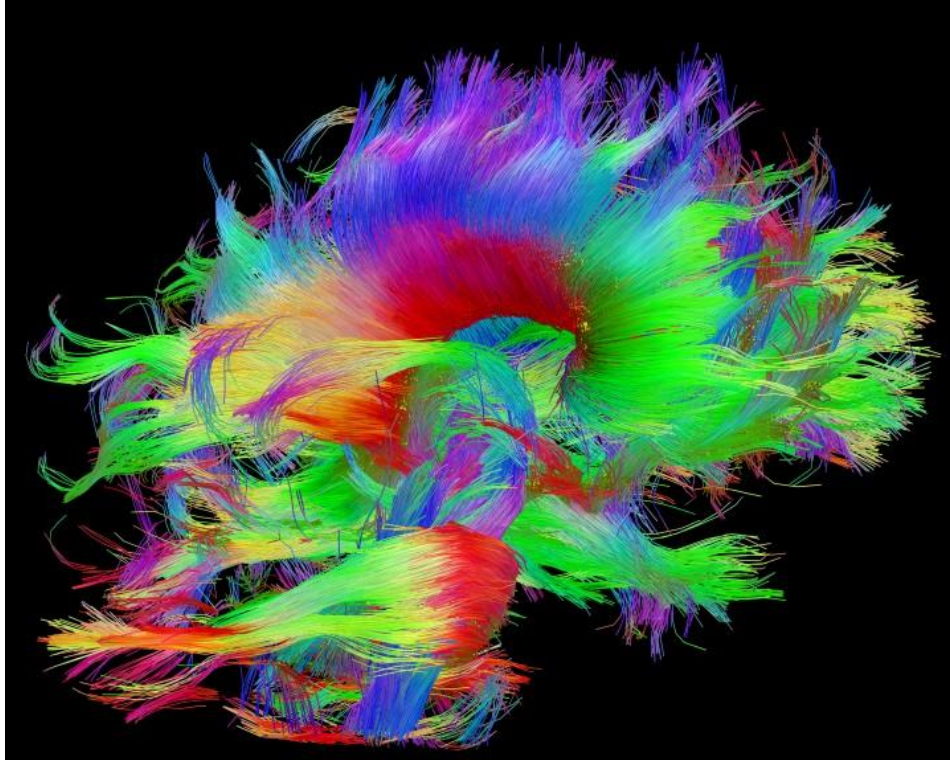
Finalmente, una de las últimas técnicas es la de neuroimágenes por tensor de difusión (ITD), en la cual el desplazamiento de las moléculas de agua por las fibras nerviosas revela el proceso de comunicación neuronal, que se muestra a través de cartografías de redes neuronales (muchas de ellas de gran belleza) conocidas como “conectomas”⁸. “No obstante, las imágenes que se obtienen se basan en proyecciones matemáticas y estadísticas. En otras palabras, no reproducen las comunicaciones reales” (Bereither, 2014, p. 44).

De acuerdo con Haueis y Slaby (2017), pensar los conectomas como “descripciones completas” es erróneo porque se trata de “objetos constitutivamente epistémicos”, lo que significa que su estructura y significado dependen del conjunto particular de factores de objeto e investigación. Por esto, optan por pensar la conectómica como “la práctica de rastrear indexicalmente propiedades de conexión parcialmente desconocidas bajo supuestos de modelos para crear diagramas idealizados y revisables de la conectividad del cerebro” (2017, p. 16). En otras palabras, la conectómica se orienta a identificar y catalogar las conexiones neuronales en un sistema organizado, a partir de modelos teóricos o computacionales, para plasmar los resultados en esquemas que no son representaciones de la realidad, sino aproximaciones idealizadas y susceptibles de revisión.

⁸ Para ampliar la noción de conectómica, sugerimos visitar el sitio web www.humanconnectomeproject.org

Figura 2-13.

Arquitectura de las fibras de la materia blanca del cerebro medida a partir de imágenes de espectro de difusión (DSI). Se muestran las vías del cuerpo calloso y del tronco encefálico. Las fibras están codificadas por colores según su dirección: rojo = izquierda-derecha, verde = anteroposterior, azul = ascendente-descendente (RGB=XYZ).



Nota: tomado de www.humanconnectomeproject.org

2.2. ¿Cómo se generan las imágenes de resonancia magnética?

Para esta primera parte de la explicación acerca de las generalidades de la técnica, realizaré una apretada síntesis de lo desarrollado por Armony, Trejo-Martínez y Hernández (2012).

Los átomos poseen un núcleo conformado, en su mayoría, por dos tipos de partículas elementales: neutrones (de carga eléctrica nula) y protones (electropositivos). Además de masa y carga eléctrica, los núcleos atómicos pueden poseer (en el caso de que tengan un número impar de protones o neutrones) un momento angular llamado *spin*. Los núcleos con *spin* también poseen propiedades magnéticas, a través de lo que se conoce como **momento magnético**.

En ausencia de campos externos, los *spin* de un conjunto de átomos de hidrógeno apuntarán en direcciones aleatorias, por lo que la suma vectorial total será cero y el momento magnético total será nulo. Sin embargo, si colocamos la muestra en un campo magnético externo, los *spin* de los núcleos tenderán a alinearse en la dirección principal del campo, en la misma dirección que este (paralelo) o en dirección opuesta (antiparalelo). Dado que el *spin* requiere menos energía para

mantenerse en posición paralela que antiparalela, habrá un exceso de núcleos en posición paralela, por lo que no llegarán a cancelarse con los que se alinean en la dirección opuesta, resultando en **un momento magnético neto distinto de cero**, en dirección paralela al campo magnético exterior.

Relajación longitudinal o T1

Si absorbe la cantidad de energía adecuada, un *spin* en estado paralelo puede pasar al estado antiparalelo (alta energía). En la resonancia magnética, dicha energía se genera utilizando un pulso de radiofrecuencia que oscila a la frecuencia conocida como **frecuencia de Larmor**⁹. Si se emiten suficientes ondas electromagnéticas, la proporción de *spin* en estado antiparalelo pasará a ser igual al de aquellos en estado paralelo, con lo cual el momento magnético total será cero. Una vez que se apaga el pulso electromagnético, los *spin* vuelven a su estado original, recuperándose así la magnetización longitudinal (paralela al campo externo). El tiempo de este proceso, denominado **relajación longitudinal**, está asociado con una constante de tiempo **T1**, que depende directamente de las propiedades magnéticas del medio en el que se encuentran los núcleos y, por consiguiente, permite identificarlo. En particular, en el cerebro se pueden identificar tres tiempos T1, correspondientes, en forma creciente, a la sustancia blanca, la materia gris y el líquido cefalorraquídeo. Mediante la RM se puede entonces crear un “mapa” de constantes de tiempo T1 que nos permite obtener una imagen anatómica del cerebro.

Relajación transversal o T2

En presencia de un campo magnético externo, los *spin* no se alinean perfectamente en la dirección externa a este, sino que rotan a su alrededor con un **movimiento de precesión**¹⁰ de frecuencia

⁹ La frecuencia de Larmor está dada por la ecuación:

$$\omega_0 = \gamma B_0$$

donde ω_0 es la frecuencia de precesión de Larmor (en radianes por segundo), γ es la razón giromagnética (una constante que depende del tipo de núcleo) y B_0 es la intensidad del campo magnético externo aplicado (en teslas, T). Se infiere de aquí que, a mayor intensidad del campo magnético, la frecuencia de Larmor aumenta, lo que exige una mayor precisión en la transmisión y recepción de señales en iRMf. Además, un campo magnético más alto permite una mayor resolución espacial y temporal en la detección de la señal. Sin embargo, esto también conlleva una mayor sensibilidad a artefactos y variaciones de susceptibilidad magnética, lo que puede generar distorsiones en la imagen. Por ello, la elección del campo magnético óptimo en iRMf depende del balance entre mejor resolución y minimización de artefactos.

Fuente: Pebet, N. (2004). Resonancia nuclear magnética. *XIII Seminario de Ingeniería Biomédica*, 1-5.

¹⁰ El movimiento de precesión es el movimiento asociado con el cambio de dirección en el espacio que experimenta el eje instantáneo de rotación de un cuerpo.

igual a la de Larmor; este movimiento hace que, además de un momento magnético longitudinal, cada *spin* tenga un momento magnético transversal (que varía en el tiempo) en el plano perpendicular. En general, el momento magnético transversal neto es nulo debido a que la rotación de los distintos *spin* está fuera de fase y la suma vectorial es cero. Sin embargo, el pulso de radiofrecuencia que invierte los *spin* también hace que su rotación transversal entre en fase, generándose un momento magnético neto en este plano. Una vez que se apaga el pulso y debido a las interacciones entre los núcleos, el momento magnético transversal desaparece. Este proceso de **relajación transversal** posee como tiempo característico la constante **T2**, que también depende del medio en el que se encuentran los núcleos. Entonces, utilizando una secuencia de adquisición de imágenes sensible a la relajación T2 se obtienen imágenes anatómicas que complementan aquellas obtenidas por las secuencias T1.

Las diferencias en las propiedades magnéticas en una región dada provocarán que los *spin* roten con diferentes frecuencias de Larmor, acelerando la pérdida de fase entre los *spin* y, por lo tanto, reduciendo el tiempo de relajación. El efecto combinado de la interacción entre *spin* y la inhomogeneidad magnética local en la relajación transversal está dado por una constante de tiempo **T2***. Este factor es esencial para la adquisición de imágenes de resonancia magnética funcional.

Tabla 2-2

Valores T1 y T2 aproximados en el cerebro.

	T1 (MS)	T2 (MS)
Materia blanca	510	67
Materia gris	760	77
Edema	900	126
Líquido cefalorraquídeo	2650	180

Nota: tomado de Maestú-Unturbe, 2007, p. 33

Cabe subrayar aquí la diferencia entre T2 y T2*. T2 hace referencia al tiempo que tarda la señal en desaparecer luego de que se haya aplicado un pulso de radiofrecuencia. Las imágenes ponderadas en T2 se usan para identificar estructuras que contienen agua y permiten diferenciar

Un ejemplo de precesión lo tenemos en el movimiento que realiza un trompo en rotación: cuando el eje de rotación no es vertical, el trompo realiza un “cabeceo” similar al movimiento de precesión.

Para introducirnos de manera informal a este sugerimos el siguiente video sobre el giroscopio y sus movimientos https://youtu.be/cquvA_IpEsA?si=plEfpCPbDIJq-TFb

tejidos blandos. $T2^*$ se relaciona también con el tiempo de desaparición de la señal, pero está influido por otros factores, como los gradientes de campo magnético y los cambios en la susceptibilidad magnética. Las imágenes ponderadas en $T2^*$ son las comúnmente usadas en la resonancia magnética funcional.

Es posible sintetizar lo anterior en la siguiente tabla (Westbrook y Kaut, 1988, pp. 19-22):

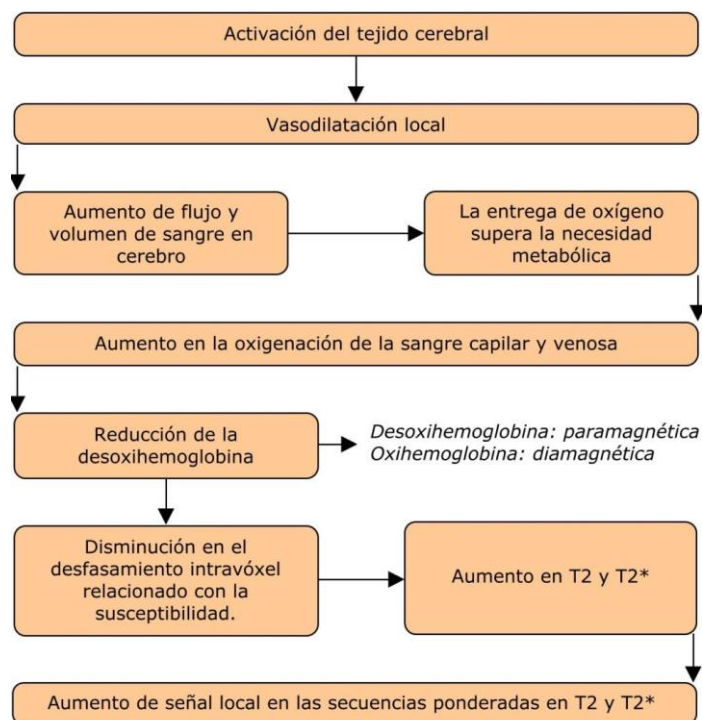
Tabla 2-3

Significación de los valores $T1$, $T2$ y $T2^$*

Relajación longitudinal $T1$	Evaluación de la anatomía normal y patológica. Resalta tejidos con alto contenido de lípidos o proteínas.
Relajación transversal $T2$	Evaluación de la anatomía normal y patológica. Resalta tejidos con alto contenido de agua.
Relajación transversal $T2^*$	Refleja la demanda metabólica de oxígeno de las áreas cerebrales activas. Se utiliza para crear mapas de la actividad cerebral funcional.

Figura 2-14.

Cascada de relaciones neuronales, fisiológicas y biofísicas que conectan la actividad neuronal con los cambios de señal medidos utilizando iRMf.



Nota: elaboración propia a partir de Hillary y DeLuca, 2007.

2.2.1. Resonancia magnética funcional (RMf)

Las imágenes obtenidas por RMf se basan en la premisa que las neuronas involucradas en un proceso mental requieren de una mayor cantidad de energía, que obtienen a través de la hemoglobina. Se incrementa la demanda local de oxígeno, lo que causa un incremento gradual de sangre oxigenada (oxihemoglobina) en esa región y una consecuente disminución de la concentración de desoxihemoglobina (sangre sin oxígeno). La hemoglobina oxigenada es ligeramente diamagnética, es decir, presenta una débil repulsión ante un campo magnético, mientras que la hemoglobina sin oxígeno es paramagnética, o sea que posee una susceptibilidad magnética positiva y por ello es atraída hacia un campo magnético.

Si comparamos la señal (proporcional al tiempo de relajación $T2^*$) en una misma región bajo dos condiciones, una de reposo y otra en la que está activa, encontraremos distintos valores. La diferencia puede tomarse como un índice indirecto del grado de actividad neuronal en esa área. Este tipo de señal, llamado contraste BOLD (del inglés *Blood Oxygen Level Dependent* o, en español, dependiente del nivel de oxígeno en la sangre) fue descrito por Ogawa y colaboradores en 1990 y es el más utilizado en la adquisición de imágenes de RMf.

2.2.2. El diseño de investigación

Ahora bien, ¿a qué condiciones debe responder un diseño de investigación?

En líneas generales, hay cuatro variables a ser controladas (Maestú-Unturbe et al., 2007, p. 41):

1. Comparaciones experimentales: ¿qué vamos a medir?
2. Propiedades de los estímulos: ¿qué se presenta?
3. Tiempos de presentación: ¿cuándo se presenta?
4. Instrucciones para el sujeto: ¿qué tienen que hacer los sujetos?

El diseño experimental más básico y, con todo, el que es todavía más utilizado en la resonancia magnética funcional es el método de sustracción cognitiva, basado en la hipótesis de la “inserción pura”. (Maestú-Unturbe et al., 2007, p. 41; Poldrack, 2010, pp. 147-148).

- a. **Principio de sustracción cognitiva:** implica comparar dos eventos que difieren en una única característica. Por ejemplo, consideremos los tiempos de reacción t_1 y t_2 de dos condiciones experimentales:

C1. El sujeto debe realizar una acción a cada vez que se enciende una luz;

C2. El sujeto debe realizar la misma acción cada vez que se enciende una luz roja.

Por el principio de sustracción cognitiva, se entiende que la diferencia en los tiempos de reacción

$$t_d = \Delta t_2 - t_1$$

representaría, en el ejemplo mencionado, el tiempo empleado para la discriminación del color.

- b. **Principio de inserción pura:** es el supuesto según el cual se puede sumar un proceso cognitivo a otro conjunto de procesos sin afectarlos (Maestú-Unturbe, 2007, p. 42; Venturelli y Branca, 2016, p. 186). Se presume que los procesos inicialmente implicados no se modificaron y los nuevos pueden ser estudiados de manera independiente.

Históricamente, estos principios de inserción pura y de sustracción cognitiva han sido ampliamente utilizados en los paradigmas de la psicología clásica¹¹. Pero cabe señalar que se trata de supuestos metodológicos propios del diseño experimental en neurociencia cognitiva. Se los utiliza por comodidad metodológica, porque permiten la operacionalización experimental de procesos mentales complejos, pero constituyen idealizaciones cuya validez ha sido y es ampliamente discutida. Como advierte Poldrack (2010), el principio de inserción pura presupone una independencia entre los procesos cognitivos que rara vez se corrobora empíricamente. Además, los patrones de activación cerebral no responden a una lógica lineal o aditiva, lo que permite poner en cuestión las inferencias que se realicen a partir de estos principios.

Una lectura crítica del diseño experimental obliga a interrogar no solo la confiabilidad metodológica de estas técnicas, sino también los supuestos teóricos que las sustentan. Junto con los principios de inserción pura y de sustracción cognitiva, otros supuestos que abordaré más adelante en este trabajo son la presunción de estabilidad y especificidad de los constructos cognitivos, la inferencia causal realizada a partir de correlaciones y la suposición de homogeneidad interindividual en la organización cerebral.

Establecimiento de la condición de control

El investigador ha de seleccionar aquella tarea que elimine todo aquello que no sea componentes de interés. Puede tratarse de un período de reposo, pero en el caso de que se trate de una investigación de tipo cognitivo, el cerebro sigue manteniéndose activo. En ocasiones se sugiere la realización de una tarea distractora, por ejemplo, el repaso de material aprendido en los estudios de memoria.

¹¹ Para profundizar en el tema de la sustracción cognitiva: Friston, K., Price, C. J., Fletcher, P. C., Moore, C., Frackowiak, R. S. J. y Dolan, R. J. (1996) The trouble with cognitive subtraction. *Neuroimage*. 4:97-104.

Cuando la activación cerebral es mayor en la condición de control que en las condiciones experimentales, suele interpretarse como una desactivación.

Tipos de diseño:

Básicamente, los tipos de diseño utilizados en la RMf son:

- Diseño de bloques: en él, se alternan bloques de reposo con bloques en los que se realiza la tarea de interés. Está basado en el principio de la inserción pura.
- Diseño ligado a eventos (*ER-event related*): en este caso, se registran las respuestas ante eventos discretos y únicos, en lugar de registrar la integración de señales durante una cierta cantidad de tiempo.

2.2.3. Desarrollo del experimento

El experimento implica llevar adelante la siguiente secuencia de pasos:

- a. Preparación del sujeto para el estudio.
- b. Presentación de estímulos:
 - Visuales: permiten al experimentador diseñar gran cantidad de experimentos cognitivos (lenguaje, memoria, atención, percepción, etc.).
 - Auditivos: El factor que más complica la presentación de sonidos es el ruido de fondo presente en la RM. El sonido más intenso que genera la bobina de gradientes puede alcanzar entre 103 y 134 dB. Este sonido genera una actividad cerebral no relacionada con los estímulos de interés¹². Por esta razón, se debe utilizar un sistema de protección que reduzca al mínimo la intensidad del sonido ambiental (tapones de cera, cascos con un sistema aislante, etc.).
 - Táctiles, olfativos, gustativos¹³, dolorosos y propioceptivos: no son muy frecuentes.
- c. Registros conductuales y fisiológicos, si corresponde al diseño experimental.

¹² Al respecto se puede consultar, por ejemplo: Bandettini, P. A., Jesmanowicz, A., Kylen, J. V., Birm, R. M., Hyde, J. S. Functional MRI of brain activation induced by scanner acoustic noise. *Magn Reson Med*. 1998;39:410-6.

¹³ En Maestú-Unturbe et al., 2007, p. 54 (así como en un número considerable de estudios) dice “gestatorios” en lugar de “gustatorios” o “gustativos”, pero entendemos que se trata de un error tipográfico.

2.2.4. Posprocesado

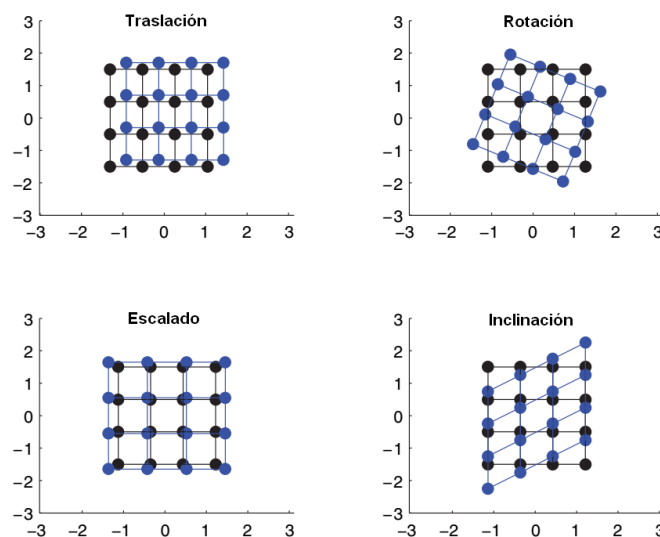
Luego de la aplicación de la técnica, son necesarios varios procesos antes de poder acceder a la neuroimagen tal y como viene a nuestra mente. Los más importantes son los siguientes:

2.2.4.1. Registro

Para registrar los aspectos fisiológicos, se hacen necesarias numerosas adquisiciones en las que es prácticamente imposible conseguir que el sujeto permanezca inmóvil. Por ello, el registro consiste en una realineación geométrica de las imágenes mediante transformaciones rígidas afines¹⁴.

Figura 2-15.

Ejemplos de transformaciones lineales. En cada figura, los puntos negros representan las ubicaciones de coordenadas originales y los puntos azules, las nuevas ubicaciones después de que se aplica la transformación.



Nota: traducción de Poldrack, Mumford, Nichols (2016, p. 19).

2.2.4.2. Corrección de tiempo de corte (slice timing correction)

En la práctica, cada corte de un volumen es adquirido en instantes de tiempo diferentes aunque sucesivos, por lo que no se corresponde con el mismo instante en la ejecución de la tarea por parte del sujeto. Por medio de interpolaciones matemáticas, más o menos sofisticadas, se puede regenerar el estudio de forma que los cortes de un mismo volumen sean muy similares a una adquisición ideal en la que todo un volumen se adquiriese en el mismo instante de tiempo.

¹⁴ Una transformación afín es una función invertible del plano que preserva líneas rectas y preserva paralelas. Por otro lado, una transformación rígida preserva longitudes.

2.2.4.3. *Filtrado paso bajo*

Las tareas suelen ser repetitivas con objeto de adquirir varias muestras en el mismo estado del sujeto (ya sea en reposo o activación). Esto significa que la actividad ha de aparecer y desaparecer con una frecuencia cercana a la frecuencia de repetición de los ciclos de activación-reposo. A través del filtrado paso bajo, se eliminan componentes en la señal vinculados con ruido de alta frecuencia y algunos artefactos.

2.2.4.4. *Eliminación de tendencias lineales (linear trend removal)*

El estado basal puede aumentar o disminuir a lo largo de la tarea. Esto hace que podamos tener diferencias significativas entre la amplitud de la señal al principio de la tarea con respecto al final de esta. Estas tendencias aumentan las proporciones de falsos positivos y negativos dependiendo también de la forma en que se realice la estadística. Conviene tratar de detectar estas tendencias para su eliminación previamente a cualquier tipo de análisis estadístico.

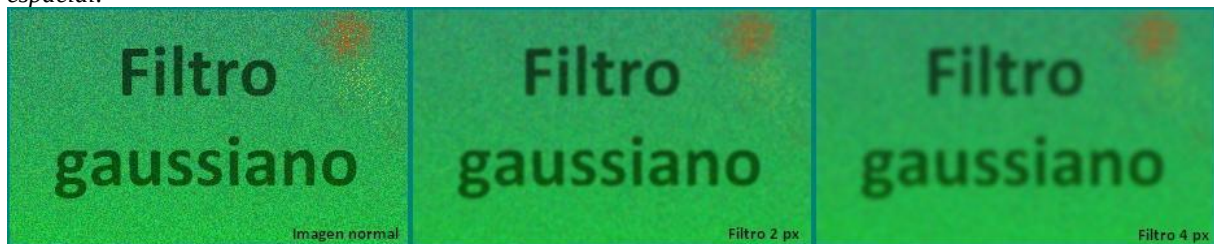
2.2.4.5. *Suavizado o smoothing*

Es habitual la utilización de un filtrado gaussiano para mejorar la relación señal/ruido.

En esencia, esto implica una homogeneización de las imágenes, lo que genera una pérdida de resolución espacial (como lo sería, por ejemplo, ver una proyección fuera de foco). Sin embargo, este paso es necesario a fin de eliminar las diferencias residuales en la anatomía de los sujetos (es decir, lo que no pudo ser ajustado satisfactoriamente durante la normalización), así como para aumentar la relación señal-ruido (Harmony et al., 2012).

Figura 2-16.

Ejemplo “casero” de filtro gaussiano. Obsérvese la eliminación del ruido pero también la pérdida de resolución espacial.



Nota: elaboración propia.

2.2.4.6. *Normalización*

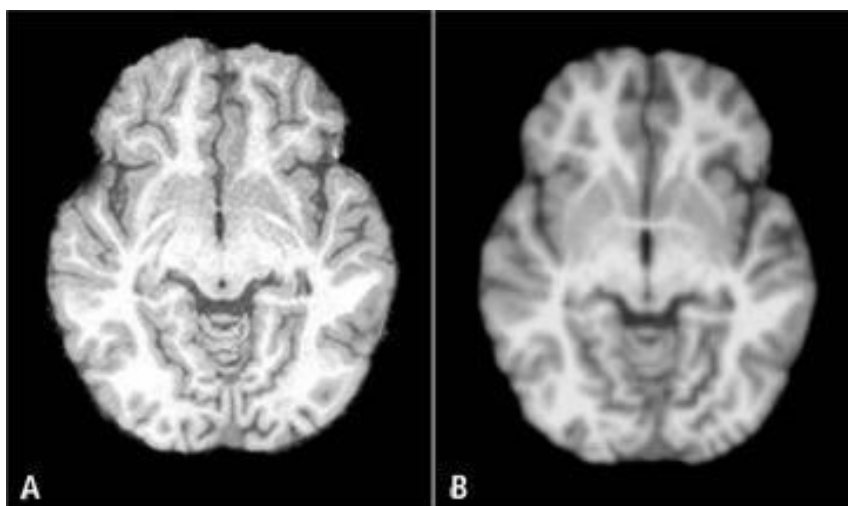
Este es el punto más delicado del posproceso y es necesario solo cuando se requieren comparaciones.

Suelen emplearse dos métodos de normalización: el ajuste a la rejilla de **Talairach**, o bien al cerebro normalizado del **Instituto Neurológico de Montreal** (MNI, *Montreal Neurology*

Institute). Talairach, al ser más antiguo, es el más referenciado en la literatura médica; sin embargo, dicho atlas se realizó sobre un cerebro de un determinado sexo, por lo que dista bastante de lo que se entiende por un cerebro estándar o promedio; además, no contempla totalmente las diferencias anatómicas individuales. El cerebro estereotáctico del MNI tiene la ventaja de ser el promedio de numerosos casos adquiridos, tanto con resonancia como con PET. La normalización, de por sí compleja, puede serlo aún más si los sujetos del estudio tienen malformaciones o han sido intervenidos quirúrgicamente dando lugar a pérdidas de masa encefálica, como puede ocurrir en estudios funcionales de pacientes que han sufrido daño cerebral. En estos casos, la normalización se convierte en un proceso casi manual y, por tanto, muy tedioso.

Figura 2-17.

Para hacer comparable las imágenes de un paciente con una base de imágenes de la población normal, es necesario “deformar” la superficie a un cerebro estandarizado. La figura 8A muestra un corte transversal potenciado en T1 y la figura 8B muestra el mismo corte “normalizado” al estándar MNI 152 (Montreal Neurological Institute).



Nota: tomado de Rojas, C. G., Gálvez, M. M., Cordovez, M. J., Pardo, B. C., Cisternas, E. J. y Asahi, K. T. (2010). Imágenes Neuro-Oftalmológicas. Revista Médica Clínica Las Condes, 21(6), 997-1002.

La etapa de normalización en el procesamiento de imágenes cerebrales funcionales pone en evidencia uno de los supuestos centrales de estas técnicas: la **presunción de homogeneidad interindividual**. Para poder realizar comparaciones grupales, se ajustan los datos de distintos sujetos a un modelo estándar. Pero esto implica suponer que las diferencias individuales son irrelevantes o que pueden ser corregidas, lo que resulta cuestionable tanto desde el punto de vista técnico como conceptual. Como advierten diversos autores, esta práctica puede oscurecer variaciones funcionales relevantes, especialmente en poblaciones no normativas (Uttal, 2001; Friston, 2007).

2.2.5. Análisis estadístico

Después del posproceso, el problema estadístico consiste en localizar, en presencia de ruido, las regiones donde hay activación.

Es necesario saber en qué zonas del cerebro aparece una diferencia significativa entre la condición de activación y la condición de control. Estadísticamente requiere una prueba que puede ser implementado de diversos modos: correlación, t-Student, ANOVA (análisis de la varianza), modelo general lineal (GLM, *general linear model*).¹⁵

Tabla 2-4

Métodos de análisis estadístico más frecuentemente utilizados en RMf.

Modelo	Descripción	Ventajas	Limitaciones	Métodos Estadísticos Comunes	Frecuencia de Aplicación
Modelo General Lineal (GLM)	Utiliza un modelo lineal para analizar la relación entre la señal BOLD y las condiciones experimentales. Estima los efectos de cada condición.	Flexible, permite el análisis de múltiples condiciones, covariables e interacciones.	Supone que los datos tienen una distribución normal y requiere corrección por comparaciones múltiples.	t-Student, ANOVA, F-Tests	★★★★★ Muy frecuente: es el enfoque estándar para el análisis de iRMf.
Prueba t de Student	Compara las medias de dos condiciones para evaluar si hay diferencias significativas entre ellas.	Simple y directa para comparar dos condiciones.	No maneja más de dos condiciones, no tiene en cuenta factores adicionales o interacciones.	Prueba t (para dos condiciones)	★★★★★ frecuente: Usado en estudios simples con dos condiciones.
ANOVA (Análisis de varianza)	Compara las medias entre tres o más grupos o condiciones experimentales.	Eficaz para comparar múltiples condiciones, evalúa efectos principales e interacciones.	No identifica qué condiciones son significativamente diferentes sin pruebas post-hoc.	ANOVA de un factor, ANOVA de dos factores	★★★★★ frecuente: Usado en estudios con múltiples condiciones.
Modelos de regresión lineal	Modela la relación entre la señal BOLD y las condiciones experimentales mediante una regresión lineal.	Flexible, permite análisis más detallados e incluye covariables.	Supone linealidad entre las variables y normalidad en los errores.	Regresión lineal simple/múltiple	★★★★★ moderada: Usado en estudios con covariables y análisis detallados.

Nota: elaboración propia basada en Faro, S. H., Mohamed, F. B. (eds.). (2010). *BOLD fMRI: A guide to functional imaging for neuroscientists*. Springer Science & Business Media.

¹⁵ Estos son los modelos estadísticos clásicos utilizados para la interpretación de las imágenes por RMf.

2.2.6. Limitaciones

Teniendo en cuenta lo visto, se puede concluir que la aplicación de la resonancia magnética funcional de por sí presenta limitaciones; algunas son solo técnicas y otras permeables a distintos tipos de sesgos teóricos. Entre las más importantes, según Venturelli y Branca (2016, pp. 190-194), podríamos mencionar:

- a. La determinación del tamaño de la muestra (a menudo exigua) y el control sobre la tarea experimental.
- b. El carácter indirecto del registro de la herramienta, que implica la presencia de un hiato temporal entre la actividad hemodinámica y la activación/desactivación.
- c. La posible variabilidad estructural intersujeto.
- d. La presencia de fuentes de ruido o interferencias (entre ellos, el ruido mismo del escáner).
- e. La técnica que se aplique será diferente según el objetivo que se desee alcanzar; si se establece un umbral estadístico muy estricto, solo se detectarán las activaciones más fuertes, con la posibilidad de que se presenten errores de tipo II (falsos negativos). En cambio, si el umbral es bajo, aumenta la posibilidad de errores tipo I (falsos positivos).
- f. El uso de diferentes diseños factoriales, herramientas automatizadas, modelizaciones matemáticas (como secciones de Poincaré¹⁶).

Por otra parte, aunque en ocasiones se afirma que la resonancia magnética funcional mide la actividad neuronal “en tiempo real”¹⁷, esta afirmación no es correcta. Además de ser una medida indirecta, basada en cambios hemodinámicos, la RMf tiene un desfase de algunos segundos. Nos parece relevante, entonces, hacer una comparación entre las técnicas más utilizadas y su resolución espacial y temporal.

¹⁶ Las secciones de Poincaré es un tipo de modelización matemática que permite la elaboración de mapeos que heredan muchas de las propiedades del sistema original pero con una dimensión menor, lo que facilita el análisis. Cf. Núñez-Yepetz, H. N. y Salas-Brito, A. L. (2013). Poincaré, la mecánica clásica y el teorema de la recurrencia. *Revista mexicana de física E*, 59(2), 91-100.

¹⁷ Esta expresión aparece, entre otros sitios web, en:

<https://www.apollohospitals.com/es/diagnostics-investigations/functional-mri>,

<https://resonancia-magnetica.com/anatomias2/partes2/funcional/>,

<https://radiologiaclub.com/2016/05/10/utilidad-de-la-resonancia-magnetica-en-el-diagnostico-de-alzheimer/>

Tabla 2-5*Comparación de las técnicas de neuroimagen.*

Técnica	Resolución temporal	Resolución espacial	Qué mide	Invasividad
Resonancia magnética funcional (RMf)	1–3 segundos (por volumen); efecto BOLD con desfase	Alta (~1–3 mm); buena localización regional	Cambios en oxigenación sanguínea (actividad metabólica)	No invasiva
Electroencefalograma (EEG)	Milisegundos; excelente para eventos rápidos	Baja (~centímetros); difícil localizar con precisión	Actividad eléctrica cortical	No invasiva
Magnetoencefalografía (MEG)	Milisegundos; buena resolución temporal y espacial	Alta (~mm); mejor que EEG	Campos magnéticos generados por la actividad neuronal	No invasiva
Tomografía por emisión de positrones (PET)	Minutos; muy baja resolución temporal	Media (~4–5 mm); limitada pero útil en ciertos casos	Metabolismo cerebral mediante trazadores radioactivos	Moderadamente invasiva (requiere inyección de trazador)

Nota: adaptado de Huettel, Song y McCarthy (2014); Gazzaniga, Ivry y Mangun (2019).

2.3. Discusión y proyecciones

El recorrido histórico realizado en la primera parte de este capítulo nos mostró que el surgimiento de la RMf se sitúa en una tradición de pensamiento que, desde sus mismos inicios, produjo diferentes modos de “representar el cerebro” condicionados por las ideas filosóficas, científicas y sociales propias de cada época. Desde la frenología hasta la resonancia magnética funcional, la búsqueda de correlaciones entre estructuras cerebrales y funciones cognitivas ha guiado la investigación en neurociencia, consolidando ciertos paradigmas que persisten hasta hoy.

Este recorrido histórico ilustra que la observación científica nunca es completamente neutral. Como señala Hanson (1961), los datos están siempre mediados por las teorías que los organizan

(*theory-loaded*), lo que se conoce como **carga teórica de la observación**¹⁸. Díez y Moulines (2017) afirmaron también que “toda observación, o mejor dicho todo informe observacional, supone una interpretación de los datos de los sentidos, y una interpretación no es más que una conceptualización teórica, sea explícita o implícita” (pp. 301–302). En este sentido, cabe una distinción adicional: no solo el ojo del observador, sino también los instrumentos con los que se observa y registra, están atravesados por supuestos teóricos que condicionan qué se considera observable, qué se registra y cómo se interpreta.

Esta problematización de la observación como instancia neutra cobra especial relevancia en el campo de las neurociencias cognitivas. Como sostienen Barberis, Roffé y Ginnobili (2025), incluso cuando se parte de la descripción de una estructura dentro de un sistema, su interpretación funcional implica un proceso de enriquecimiento conceptual guiado por marcos teóricos. Esto supone que el fenómeno a explicar (*explanandum*) y aquello que lo explica (*explanans*) no se definen de forma fija ni neutral, sino que dependen del marco teórico desde el cual se organiza la explicación.

En otras palabras, identificar una estructura no explica automáticamente la función asociada. Lo que aparece como activación significativa es, en realidad, el resultado de una serie de decisiones de distinto tipo:

- Conceptuales: qué se investiga, qué se considera función relevante, qué marcos teóricos se adoptan.
- Experimentales: el diseño de la tarea, los estímulos, las condiciones de control.
- Estadísticas: tipo de análisis, umbral de significancia, tratamiento de los datos.

La conciencia sobre la carga teórica que atraviesa tanto la observación como el instrumental nos invita también a revisar críticamente los saberes que se fueron construyendo a partir de esas observaciones. A lo largo de la historia, prácticas científicas que en su momento fueron consideradas legítimas —como la frenología— terminaron siendo descartadas al evidenciarse los supuestos teóricos erróneos que las sustentaban. Esto abre interrogantes sobre nuestras propias herramientas actuales. ¿Podría ocurrir algo similar con las modernas técnicas de neuroimagen? ¿Hasta qué punto sus inferencias se sostendrán filosófica y metodológicamente en el tiempo?

La llegada de la resonancia magnética funcional marcó un punto de inflexión en la investigación del sistema nervioso, ya que, por su capacidad para evidenciar patrones de activación con alta

¹⁸ “De hecho, lo que llamamos ‘causas’ está cargado de teoría de principio a fin. No son simples enlaces tangibles en la cadena de la experiencia sensorial, sino más bien detalles dentro de un patrón intrincado de conceptos.” (Hanson, 1961, p. 54)

resolución espacial, pareció ofrecer una solución definitiva al problema de la localización funcional. Sin embargo, su interpretación sigue dependiendo de modelos estadísticos y conceptuales que, lejos de proporcionar una imagen objetiva del cerebro, reflejan decisiones teóricas previas. Entre ellas, hemos examinado y cuestionado en este capítulo los principios de sustracción cognitiva y el de inserción pura y la presunción de homogeneidad interindividual. Abordamos también la complejidad del diseño experimental desde una dimensión predominantemente técnica, desde el diseño inicial al análisis estadístico.

A partir de este análisis, en el siguiente capítulo examinaremos en qué medida las imágenes de RMf pueden considerarse fuentes confiables de evidencia y cómo los métodos de inferencia impactan en su validez epistemológica.

3. La RMf como fuente de evidencia

En el capítulo anterior se realizó una revisión de la evolución histórica de las técnicas de neuroimagen y cómo su desarrollo estuvo guiado por concepciones previas sobre la relación entre cerebro y cognición. En este capítulo, se explorará el estatus epistemológico de la resonancia magnética funcional (RMf), analizando sus alcances y limitaciones como fuente de evidencia en neurociencia cognitiva. Se pondrá el foco en los problemas metodológicos asociados con la generación, procesamiento e interpretación de neuroimágenes, así como en las dificultades inferenciales que plantea su uso en la investigación psicológica y cognitiva. También se presentará de manera intuitiva el teorema de Bayes y, finalmente, se abordará el tema de los sesgos de publicación.

3.1. Regiones de interés

Una región de interés (ROI) es una zona específica del cerebro que se selecciona para un análisis. En lugar de examinar de manera individual la activación en cada vóxel¹⁹, al realizar un análisis de ROI se observa la señal de modo conjunto dentro de la región delimitada.

Una ROI puede ser tanto estructural como funcional. Las primeras se basan en la anatomía cerebral y suelen definirse a partir de atlas anatómicos. Las ROI funcionales, en cambio, se basan en patrones de actividad relacionados con tareas cognitivas o con comportamientos específicos (Poldrack, 2007).

Si bien las ROI constituyen una herramienta poderosa para el análisis, su uso no está exento de potenciales problemas. Entre ellos:

1. **Sesgo hacia la hipótesis nula** (Poldrack, Mumford y Nichols, 2011; Poldrack, 2007; Poldrack y Mumford, 2010): cuando las ROI anatómicas son grandes, los vóxeles realmente activos pueden constituir solo una pequeña parte de la región. Por ello, al promediar la señal para toda la ROI, la activación de los vóxeles puede diluirse entre los no activos.

¹⁹ El término “vóxel” es acrónimo de “volumetric pixel” y es la unidad cúbica que compone un objeto tridimensional. Es el equivalente de un píxel en un objeto 2D. El tamaño exacto de un vóxel varía según la tecnología. Los vóxeles de resonancia magnética funcional representan típicamente un volumen de 27 mm³ en un cuboide equilátero.

Fuente: https://hmong.es/wiki/Statistical_parametric_mapping

2. **Suposición de homogeneidad** (Poldrack y Mumford, 2010): el uso de ROI puede dar lugar al supuesto erróneo de que todos los vóxeles dentro de la ROI son funcionalmente homogéneos; esto puede afectar la interpretación de los resultados.
3. Las ROI pueden ser útiles a la hora de estudiar la **selectividad funcional**, esto es, la respuesta específica de una región a un tipo particular de estímulo o tarea. Pero no permiten realizar inferencias acerca de la **especialización funcional**, es decir, si la región es la única o la principal responsable de una función cognitiva particular (Friston, Rotshtein, Geng, Sterzer y Henson, 2006).
4. **Dependencia del contexto**: las definiciones operacionales de las ROI funcionales presuponen la invariancia del contexto; no obstante, la respuesta de una región puede variar según el contexto experimental (Friston, Rotshtein, Geng, Sterzer y Henson, 2006).

A estas problemáticas podríamos agregar una más, que usualmente es conocida como “correlaciones vudú”, a partir de un artículo publicado por Vul et al. (2009) en el que describían un problema de **circularidad** observado en el análisis de datos de iRMf. Este problema se presenta en cualquier estudio que pretenda correlacionar una medida externa con la actividad cerebral y se basa en la utilización de los mismos datos para definir una región de interés (o ROI) y para analizar la actividad dentro de ellas. Si se selecciona un conjunto de vóxeles a partir de su relación con una variable externa, se produce un sesgo en el análisis que hace muy probable encontrar una correlación significativa, dado que se aumenta el tamaño del efecto, lo que puede conducir a conclusiones erróneas. En esencia, se está correlacionando la variable consigo misma.

Las regiones de interés (ROI), entonces, permiten acotar el análisis de datos en neuroimagen y facilitar la interpretación de resultados. Sin embargo, su uso no está exento de problemas metodológicos, como el sesgo hacia la hipótesis nula, la suposición de homogeneidad funcional dentro de una región y el riesgo de correlaciones espurias. Estos factores muestran que, lejos de ser una representación objetiva de la actividad cerebral, las neuroimágenes son producto de decisiones metodológicas y teóricas que influyen en la construcción del conocimiento neurocientífico.

3.2. El problema de las localizaciones

Reflexionar sobre las dificultades inherentes a la investigación mediante ROI nos introduce a un problema más amplio: el de las concepciones localizacionistas en neuroimágenes.

En efecto, desde la creación de la RMf, se fueron defendiendo ideas **modulares** y **localizacionistas** sobre acerca de la actividad del cerebro: “esto es, distintas áreas del cerebro instancian operaciones precisas y se relacionan entre sí de una manera discreta para dar lugar a una actividad cognitiva como puede ser la visión, el lenguaje, la percepción, entre otras” (Venturelli, Branca, 2016, 179). Otros enfoques, conocidos como **dinamicistas**, se orientan a capturar patrones de dinámicas neuronales, entendiendo al cerebro como un sistema complejo. La tendencia localizacionista en neurociencias ha sido estudiada por Büchel y Friston (2000) entre otros. Tressoldi, Sella, Colheart y Umiltà (2012), por su parte, realizaron un metaanálisis²⁰ de estudios entre 2007 y 2011 en revistas de reconocimiento por parte de la comunidad científica.

Intentaron determinar en qué medida los artículos analizados tenían como objetivo: (1) la localización de procesos cognitivos o (2) poner a prueba una teoría sobre los procesos cognitivos. Los resultados mostraron que en el 89% de los casos se había seguido el primer objetivo, es decir, la localización. (Venturelli, Branca, 2016, p. 184)²¹

Las posiciones localizacionistas clásicas fueron modificándose: al presente, la tendencia dominante es la de “individualizar redes operativas que pueden estar distribuidas en el cerebro pero que son responsables de operaciones cognitivas específicas y cuya contribución en el marco de la actividad psicológica global puede aislarse” (Venturelli, Branca, 2016, p. 184): esto se asocia al modularismo en el campo de las neurociencias. A diferencia de la posición elaborada por Fodor en torno de la particular organización de nuestra arquitectura cognitiva y centrada en una noción muy elaborada de “módulo cognitivo”, los abordajes modularistas que nos ocupan no son tan afines al concepto fodoriano de “módulo”, sino que se basan en una noción más inespecífica.

Para las neurociencias se suele aceptar la definición de Saul Sternberg, de acuerdo con la cual “Dos (sub)procesos A y B de un proceso complejo (mental o neuronal) son módulos si y solo si cada uno puede cambiarse independientemente del otro” (2011, p. 159)

Este criterio de modularidad parece ser mucho más débil que el conjunto de propiedades del módulo sugerido por Fodor (1983), según el cual los módulos suelen ser innatos, encapsulados en información, específicos del dominio, “cableados”, autónomos y rápidos. Sin embargo, la

²⁰ El metaanálisis es una de las principales metodologías aplicadas en los artículos de revisión y puede ser descrito como “una herramienta estadística para estimar la media y la varianza de los efectos poblacionales subyacentes a partir de una colección de estudios empíricos que abordan aparentemente la misma pregunta de investigación” (Field, Gillett, 2010). El proceso de realización de un metaanálisis incluye la selección de artículos, la determinación de criterios de inclusión, el cálculo de tamaños de efecto, el análisis real y la estimación de efectos del sesgo de publicación.

²¹ Tengamos en cuenta, en este sentido, los trabajos de Poldrack (2006) y Klein (2010), sobre las dificultades asociadas a la “inferencia inversa” que representaría el segundo de los objetivos planteados en la cita.

especificidad de dominio parece implicar una modificabilidad separada (Sternberg, 2011, p. 159, nota al pie 4). Venturelli y Branca (2016, p. 185) afirman que en neuropsicología “opera un criterio de ‘influencia diferencial’ por el cual no es necesario que dos sistemas sean modificados de forma independiente para que puedan ser considerados módulos, sino que con sólo registrar cambios con una diferencia significativa es suficiente”.

La asunción fuerte aquí es que una operación cognitiva delimitada se corresponderá con la actividad del mismo sistema neuronal —esté asociado a una región o área específica o se encuentre más bien distribuido en el cerebro (ver, por ejemplo, Friston y Price, 2011)— tanto en el caso de que esté acompañada por otras operaciones diferentes como en el caso de que no lo esté. (Venturelli y Branca, 2016, p. 186)

Cabe señalar que tanto el localizacionismo como el modularismo plasman la noción de **sustracción cognitiva** que analizaremos más adelante.

Adele Abrahamsen y William Bechtel (2012) contrastan los puntos de vista modularistas y localizacionistas (a la que denominan “**concepciones reactivas**”) con las **concepciones endógenamente activas del cerebro**. Señalan un movimiento reciente hacia estas últimas y destacan el tratamiento de la variabilidad como objeto de estudio en sí mismo. “El desafío es detectar y analizar patrones de [la actividad endógena del cerebro] y develar sus orígenes y funciones” (Abrahamsen y Bechtel, 2012, p. 344; traducción de Venturelli y Branca). Esto implica “el relativo abandono de la idea de la neurona individual como unidad fundamental de análisis para la transmisión de información” (Venturelli, Branca, 2016, p. 188).

En particular, Nicolás Venturelli (2015) describe el enfoque dinamicista en neurociencia cognitiva, centrándose en la conceptualización del cerebro como un sistema complejo y dinámico. Según este autor, los tres principios clave que subyacen a esta visión del cerebro dinámico son:

1. *La búsqueda de patrones espaciotemporales a gran escala.* El enfoque dinamicista se aleja del estudio de neuronas individuales como unidades básicas de análisis, para enfocar la búsqueda de patrones de actividad neuronal a gran escala que involucran a grupos de neuronas en diferentes áreas del cerebro. Se busca identificar cómo estos patrones cambian a lo largo del tiempo y cómo se relacionan con el desempeño cognitivo.
2. *La no-linealidad del cerebro.* El cerebro es un sistema altamente no lineal²², donde el comportamiento del sistema no puede ser simplemente la suma del comportamiento de sus partes. Esto trae aparejado que pequeñas variaciones en las condiciones iniciales pueden

²² En el capítulo anterior, realizamos una crítica a los principios de inserción pura y de sustracción cognitiva a partir de la no linealidad del cerebro.

tener grandes consecuencias en la actividad cerebral y que la interacción entre las diferentes áreas del cerebro es crucial para comprender su funcionamiento.

3. *La autoorganización de los procesos cerebrales.* Se plantea que el cerebro es un sistema autoorganizado en el que los patrones de actividad neuronal emergen de manera espontánea a partir de la interacción de sus componentes. Esta capacidad de autoorganización permite al cerebro adaptarse a nuevas situaciones y generar respuestas complejas a partir de la interacción de elementos simples (Venturelli, 2015, pp. 22-23).

Cada enfoque implica diferentes presupuestos teóricos y diversos abordajes para el análisis de los datos. Presentaré estos contrastes en una tabla (a partir de Venturelli, Branca, 2016, pp. 194-198):

Tabla 3-1

Presupuestos teóricos y análisis de los datos

	Enfoque localizacionista/modularista	Enfoque dinamicista
Priorización de aspectos espaciales o temporales	Privilegia los aspectos espaciales (qué áreas/vóxeles sobrepasan cierto umbral de activación de manera confiable, en asociación a una actividad).	Privilegia los cambios de estado. Se pierde en algún grado la precisión espacial. Suelen incluir una modelización matemática que permite inferir el tiempo real de activación de las neuronas a partir de la consideración de los tiempos de fluctuación en la respuesta hemodinámica.
Focalización sobre una región o la identificación de patrones generales de activación en el cerebro	Una estrategia común es utilizar localizadores funcionales como lo son las regiones funcionales de interés (fROI). Esto consiste en predefinir un área de interés y registrar los cambios y variaciones solo en esa área; los investigadores reducen la anatomía funcional a una sola región del cerebro que se define operacionalmente por el localizador funcional.	En estos casos la estrategia de análisis, en lugar de restringir la información considerada, ha consistido en tomar en cuenta todos los vóxeles e introducir algunos parámetros que permitan identificar patrones dinámicos y correlaciones de gran alcance en la actividad neural (por ejemplo, la propiedad de autoorganización o metaestabilidad). En esta dirección, una técnica que ha sido utilizada es la de “grano grueso” que consiste en aumentar mediante fórmulas de física estadística el tamaño de los vóxeles (aspecto espacial) y las unidades de tiempo en que se producen las activaciones (aspecto temporal).
Tipo de análisis matemático que se realiza sobre los datos	Análisis estandarizados y relativamente automatizados.	Mayor grado de libertad para el análisis de los datos. En estos casos, los datos son manipulados matemática o computacionalmente de modo que permitan inferir determinados patrones dinámicos de actividad.

El enfoque localizacionista ha sido dominante en la interpretación de neuroimágenes a lo largo de la historia, pero es dable cuestionar su énfasis en la asignación de funciones a regiones específicas

del cerebro. La evidencia creciente sobre la conectividad funcional y, en particular, sobre la existencia de redes funcionales dinámicas²³ permite pensar que la cognición no puede reducirse a la activación aislada de una ROI dada.

Por otra parte, entender al cerebro como un sistema autoorganizado nos permite modificar radicalmente nuestra conceptualización acerca de la actividad neuronal y su relación con la cognición. De acuerdo con esta perspectiva, en lugar de explicar los procesos mentales como la activación de **regiones o módulos discretos** —en la línea de los principios de sustracción e inserción pura— la autoorganización postula que la actividad cognitiva surge a partir de procesos no lineales, distribuidos y dependientes del contexto. Se libera, así, de las concepciones lineales y aditivas vistas en el capítulo anterior y abre las puertas a enfoques más integrativos y probabilísticos.

Asumir la autoorganización como principio epistemológico es un cambio en la lógica de la explicación en neurociencias. En lugar de buscar isomorfismos o correspondencias uno a uno entre función y localización, se busca comprender el surgimiento de determinadas configuraciones funcionales transitorias. Así, no solo se problematizan los supuestos metodológicos vigentes, sino que se desafía la idea misma de evidencia en el sentido de “prueba de localización” para dar lugar a una lógica emergente y distribuida del funcionamiento mental.

A manera de ejemplo, en un estudio realizado por Haxby et al. (2001), se analizó el procesamiento del reconocimiento de rostros y objetos mediante técnicas de resonancia magnética funcional. Esta función se atribuía tradicionalmente a áreas especializadas, como el área fusiforme de la cara (*fusiform face area*, FFA), en línea con un enfoque localizacionista.

Los resultados de Haxby y sus colegas ofrecen una perspectiva diferente, más cercana a un modelo de cerebro autoorganizado, con un funcionamiento de lógica emergente y distribuida. Los autores encontraron que las representaciones se distribuyen ampliamente en la corteza temporal ventral y que las regiones que responden fuertemente a una categoría (por ejemplo, rostros) pueden también contener información sobre otras categorías (como herramientas o casas). Esto sugiere, entonces, que las funciones no están rígidamente “localizadas” en áreas específicas, sino que emergen de

²³ Las redes funcionales dinámicas refieren a patrones de activación simultánea entre distintas regiones cerebrales que se conectan de forma flexible según la tarea, el contexto o el estado interno del sujeto. A diferencia del enfoque localizacionista, que asocia funciones mentales a áreas cerebrales específicas, este modelo propone que procesos como la memoria, la atención o el lenguaje surgen de la interacción entre múltiples regiones que se reorganizan en tiempo real. Un mismo grupo de áreas puede participar en distintas funciones, dependiendo de cómo se conecten en cada momento (Pessoa, 2014).

patrones de actividad distribuidos y superpuestos, que son el resultado de cómo las poblaciones neuronales se organizan en respuesta a las regularidades del entorno.

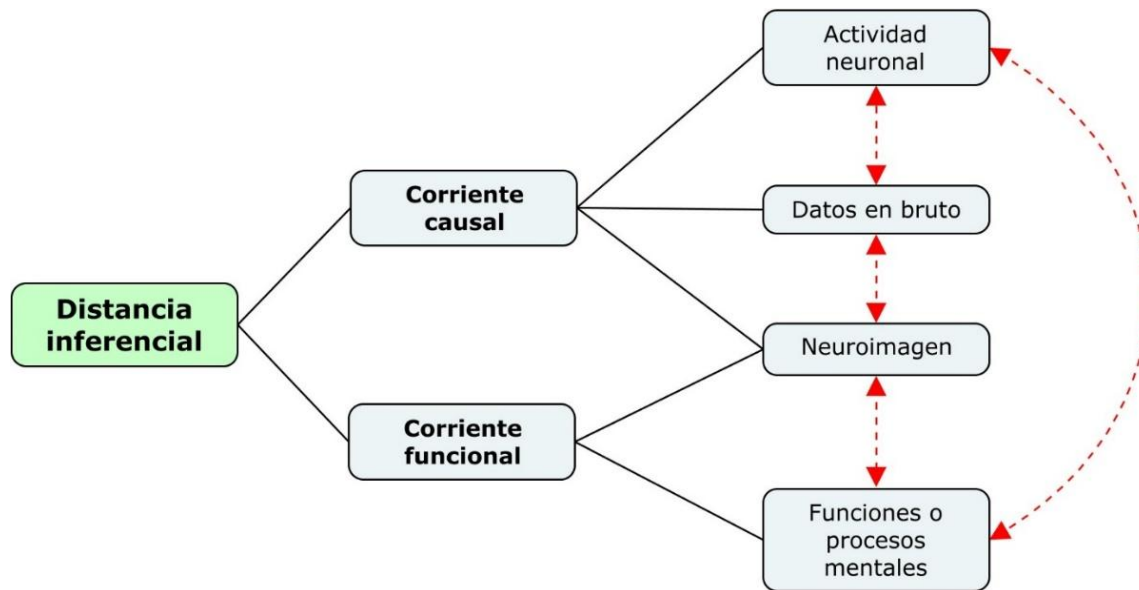
3.3. Distancia inferencial real y aparente

Adina Roskies (2010) hace referencia a los efectos epistémicos de las neuroimágenes en los no expertos, aludiendo a lo que llama “distancia inferencial real” y “distancia inferencial aparente”. La primera se refiere a las inferencias explícitamente empleadas en una práctica científica, mientras que “distancia inferencial aparente” (o percibida) indica una medida más subjetiva, relativa a la confianza que las personas depositan en una conclusión basada en la evidencia. Y señala que nuestras conclusiones “no pueden ser más seguras que los presupuestos en los que descansan” (2010, p. 197).

Las neuroimágenes cognitivas procuran medir la actividad neural que subyace a procesos cognitivos o comportamentales. Los observables, entonces, serían tanto las conductas de los sujetos como las señales de RM obtenidas, por lo cual se trataría de inferir los “elementos de conducta” por una parte y las “funciones elementales” del cerebro o patrones de actividad neural que provocan las señales de RM, por otra. Sugiere entonces Roskies que es posible entender el proceso a partir de dos instancias diferenciadas, a las que denomina la “corriente causal” y la “corriente funcional”.

- La primera o “corriente causal” estaría vinculada a la relación entre la actividad neural y la iRMf; por consiguiente, podría ser a su vez descompuesta en dos relaciones: la existente entre la actividad neural y los datos en bruto obtenidos, por un lado, y entre estos datos y la neuroimagen construida, por el otro.
- La segunda, llamada “corriente funcional” estaría asociada a la relación entre la iRMf y los elementos cognitivo-conductuales. Implica la comprensión del diseño experimental y sus supuestos subyacentes y según Roskies, es la más problemática en términos inferenciales.

Representaré sucintamente lo mencionado en la siguiente figura, en la cual las flechas dobles indican aquellas relaciones que se deben indagar.

Figura 3-1.*Corriente causal y corriente funcional**Nota: elaboración propia.*

3.3.1. De la actividad neuronal a los datos en bruto

En los capítulos anteriores hemos visto que, contra lo usualmente supuesto en ámbitos no especializados, la resonancia magnética no mide las señales eléctricas generadas en y por las neuronas, sino el nivel de oxigenación sanguínea, y lo hace, además, de manera indirecta.

Señala Roskies (2010, p. 200):

Nuestra confianza en nuestra capacidad para relacionar los cambios en la señal de iRMf con la actividad neuronal depende de la generalización empírica y la corroboración por otros métodos que miden más directamente las demandas metabólicas y el flujo sanguíneo. Es relevante recordar que **la actividad sináptica reflejada indirectamente en las neuroimágenes puede ser tanto excitatoria como inhibitoria** y la RMf no nos permite distinguir una de otra. Además, la respuesta vascular ocurre con un retraso de unos segundos y se despliega en un tiempo mayor que el de la actividad que pretende reflejar, si bien los estudios muestran que estas variaciones son relativamente estereotípicas.

3.3.2. De los datos en bruto a la neuroimagen

Como también vimos en el capítulo anterior, para obtener cierto grado de significatividad y minimizar los efectos del “ruido” aleatorio, los resultados obtenidos deben ser sometidos a

diversos posprocesos que alteran la distancia inferencial, porque introduce factores adicionales que deben tenerse en cuenta a la hora de interpretar las imágenes.

La construcción de la imagen se funda en ciertas decisiones convencionales y está marcada por la multivocidad: “La misma imagen visual podría representar cualquiera de un grupo de conjuntos de datos y el mismo conjunto de datos podría usarse para generar una variedad de imágenes” (Roskies, 2010, p. 202).

3.3.3. De la neuroimagen a la interpretación de la conducta

Cualquier forma de interpretación de las neuroimágenes **implica reducción de datos**. Y a menudo es en función de los presupuestos de partida que se asigna significatividad a los datos; los localizacionistas fuertes, por ejemplo, pueden considerar un umbral de datos más bajo que aquellos que se enfocan en el análisis de patrones.

“En última instancia”, señala Roskies, “cuál es la mejor opción dependerá de una teoría de la función cerebral aún no articulada” (2010, p. 204). Y finalmente, destaca (p. 207):

De las dos corrientes discutidas, la funcional y la causal, la corriente funcional es la más problemática en términos inferenciales. La corriente causal implica suposiciones para las cuales nuestra evidencia es incompleta o no concluyente en muchos lugares, pero esto puede ser fácilmente remediado con trabajos ulteriores y es probable que nuestras suposiciones actuales estén cerca de ser correctas.

Los supuestos que se invocan en la corriente funcional son más globales y difíciles de verificar de forma independiente y en algún nivel involucran todo el marco conceptual en el que se basan las neuroimágenes. Especialmente en ausencia de una teoría general de la función cognitiva, o de una corroboración fisiológica cuidadosa en los niveles inferiores, las conclusiones sobre la función deben considerarse provisionales. Los especialistas en neuroimágenes, como los de cualquier otra ciencia, tienden a trabajar dentro de los paradigmas establecidos para su campo, pero es esencial reconocer que nuestra ontología funcional actual se encuentra en un terreno epistémicamente inestable.

El análisis de la RMf como fuente de evidencia revela una discrepancia entre la confianza que se deposita en estas imágenes y la complejidad del proceso inferencial que las sustenta. Como señala Roskies, la distancia inferencial real entre actividad neuronal y neuroimagen es mucho mayor de lo que se percibe intuitivamente. Esto pone en cuestión el valor epistémico de la RMf como herramienta de investigación y nos obliga a examinar con mayor detenimiento las inferencias que extraemos de sus datos.

3.4. Lo que muestran las imágenes

Nuestra lengua abunda en frases que resaltan la primacía de lo visual. La imagen persuade, convence, muestra y aparentemente “demuestra”. Parecería tener cierto poder veritativo, en tanto fenomenológicamente se nos presenta como una puerta de acceso directo al mundo, aparentemente sin mediaciones.

Esta lectura se apoya en una tradición profunda del pensamiento occidental, donde **ver es conocer**: desde Platón hasta Aristóteles, la visión fue considerada el sentido más elevado y confiable para acceder al mundo. Aristóteles afirmaba en la *Metafísica* que “de todos los sentidos, preferimos la vista porque nos hace conocer más” (980a21–27), mientras que Platón, en la alegoría de la caverna (*República*, 518c), proponía una ascensión del alma hacia la luz como condición del verdadero saber.

Sin embargo, la información que llega a los órganos de los sentidos (es decir, la sensación) es sometida a una compleja serie de procesos neurofisiológicos, entre los cuales se incluye su contraste con experiencias previas —proceso comúnmente denominado apercepción—, que da lugar a la percepción. Esta no refleja pasivamente el mundo, sino que **construye** el objeto, lo interpreta, lo **re-presenta** a nuestra intelección.

3.4.1. Neuroimagen y fotografía

Como se mencionó en el punto anterior, las imágenes, en general, “nos invitan a creer”. Se les supone una proximidad inferencial, por lo que suelen cumplir una función probatoria importante entre nosotros.

Colin Klein (2010, p. 187) señala, además, que el estatus evidencial de las neuroimágenes puede parecer no problemático, sobre todo para el lego, como si fueran las imágenes producidas por el mismo cerebro en acción. Pero no solo para el lego; en el mismo artículo, agrega:

Algunos autores describen a la iRMf como siendo ‘para la ciencia cerebral lo que el telescopio de Galileo fue para la astronomía’ (Jacobs 2009, p. 121); otros hablan de la apertura de la caja negra de la mente para la observación directa (Camerer et al., 2005).

A este respecto, Roskies (2008) afirma:

Superficialmente, las neuroimágenes son como fotografías. Son generadas por máquinas que muchos consideran como cámaras cerebrales, que producen representaciones sobre eventos en la mente y el cerebro. Estas representaciones son imágenes, típicamente de

colores brillantes, que parecen representar un objeto real (un cerebro o corte cerebral) y algunas de sus características visuales; los contornos de colores delimitan regiones de actividad neuronal. Por lo tanto, mi opinión es que las personas tienden a apreciar las neuroimágenes como si fueran fotografías de la actividad cerebral. Es plausible que nuestra familiaridad con la fotografía y con el uso de imágenes fotográficas en ciencia, y la forma automática en la que interpretamos las fotografías, nos dispone a impartir a otras imágenes que guardan un parecido superficial con las fotografías características probatorias similares.

Para establecer las diferencias entre unas y otras, Roskies aborda el análisis de tres características usualmente asignadas a la fotografía:

- Mimesis
- Dependencia contrafáctica
- Independencia teórica

Indagaré un poco en cada uno de estos conceptos.

3.4.1.1. *Mimesis*

El concepto de *mīmēsis* (μίμησις) data de la Antigua Grecia, aunque fue Platón quien le dio lo que podríamos considerar cierta precisión conceptual. Podríamos asociar la mimesis con “imitación”, “representación” e incluso “emulación”.

Según Kulvicky (en Roskies, 2008) “una representación es mimética si instancia una propiedad representada por medio de la misma propiedad”. Las fotografías pueden considerarse miméticas en varios aspectos, ya que muchas de las propiedades visuales de los objetos y escenas que se fotografían aparecen representadas en ellas.

3.4.1.2. *Dependencia contrafáctica*

El concepto de dependencia contrafáctica fue introducido en 1973 por David Lewis en su artículo “Causation”.

Señala Lewis que Hume elabora al menos dos definiciones de “causación”. En la primera, que aún dominaría el terreno de la filosofía, una sucesión causal es una sucesión que instancia una regularidad. Postula entonces que las definiciones regularistas de causa podrían ser formuladas aproximadamente en estos términos (p. 556):

Supongamos que **C** es la proposición de que *c* existe (u ocurre) y supongamos que **E** es la proposición de que *e* existe. Entonces, *c* causa *e*, de acuerdo con un análisis regularista típico, si y solo si (1) **C** y **E** son verdaderas; y (2) para algún conjunto no vacío **L** de proposiciones de ley verdaderas y algún conjunto **H** de proposiciones de hechos particulares verdaderas, **L** y **H** implican conjuntamente **C** → **E**, si bien **L** y **H** conjuntamente no implican **E**, y **H** solo no implica **C** → **E**.

Señala a continuación que, si bien **C** puede **ser** causa de **E**, también podría ser **efecto** de **E**, un **epifenómeno en la historia causal** de **E**, o una **causa potencial** de **E** (esto es, no causó **E** pero lo habría hecho en ausencia de lo que realmente causó **E**).

Para sortear esta dificultad acude a la definición del contrafáctico (pp. 373-374) recurriendo a la noción de “mundos posibles”²⁴, de la siguiente manera:

Dado cualquier par de proposiciones **A** y **C**, tenemos su contrafáctico **A** □→ **C**: la proposición según la cual, si **A** fuera verdadera, entonces **C** también sería verdadera. La operación □→ es definida por una regla de verdad, como sigue. **A** □→ **C** es verdadera (en un mundo *m*) si y solo si o bien (1) no hay mundos-**A** posibles (en cuyo caso **A** □→ **C** es vacua), o (2) algún mundo-**A** donde se da **C** es más cercano (a *m*) que cualquier mundo-**A** en el que **C** no se da. En otras palabras, un contrafáctico es verdadero de modo no vacuo si y solo si se requiere una divergencia menor respecto de la actualidad para hacer verdadero el consecuente junto con el antecedente de la que se requiere para hacer verdadero al antecedente sin el consecuente.

Y finalmente vuelve a Hume²⁵ para sintetizar (p. 375):

[...] si *c* y *e* son eventos actuales [...] **e depende causalmente de c si y solo si, si c no hubiera ocurrido, e nunca habría existido**. Tomo a la segunda definición de Hume como mi definición, no de la causación misma, sino de la dependencia causal entre eventos actuales.

²⁴ El concepto de “mundos posibles” es indudablemente ambiguo y fue utilizado de modos diversos por los filósofos del siglo XX. Lewis se inscribe como un representante del “concretismo” (Menzel, 2017).

²⁵ Me refiero aquí a las dos definiciones que brinda Hume en *Investigación sobre el conocimiento humano*, sección VII, p. 101:

Podemos, pues, definir una causa como un objeto seguido de otro, cuando todos los objetos similares al primero son seguidos por objetos similares al segundo. O, en otras palabras, **el segundo objeto nunca ha existido sin que el primer objeto no se hubiera dado**.

Sea $A_1, A_2 \dots$ es una familia de proposiciones posibles no conjuntas; sea $C_1, C_2 \dots$ otra familia (de igual dimensión). Entonces si todos los contrafácticos $A_1 \square \rightarrow C_1, A_2 \square \rightarrow C_2 \dots$ entre las proposiciones correspondientes en las dos familias son verdaderos, podemos decir que C **depende contrafácticamente** de A . (Lewis, 1973, p. 561)

La dependencia causal entre eventos actuales [*actual*] implica entonces causación. Pero si bien la causación debe ser transitiva, la dependencia causal puede no serlo. Entonces, según este autor, puede haber causación sin dependencia causal.²⁶

En el caso que nos ocupa, la dependencia contrafáctica se refiere al hecho de que el contenido de las fotografías depende sistemáticamente de la escena fotografiada. Si la escena fuera diferente, la fotografía sería diferente. Esta dependencia contrafáctica está asegurada por los métodos mecánicos de la fotografía. El valor epistémico de las fotografías está relacionado con esta dependencia contrafáctica.

Trasladado al contexto de la neuroimagen, la dependencia contrafáctica exigiría que, si la actividad neuronal no hubiera ocurrido, la imagen de activación no debería existir. Pero, debido entre otros factores al ruido estadístico y al posprocesado, esta garantía no siempre se cumple en la RMf con la misma certeza que en la fotografía tradicional.

3.4.1.3. *Independencia teórica*

Roskies (2010, p. 209) explica de este modo la independencia teórica:

Nuestras capacidades intuitivas para interpretar fotografías obvian la necesidad de teorías explícitas para interpretarlas. En este sentido, las fotografías son independientes de la teoría. Además, las fotografías tienen una característica que Walton ha denominado independencia de creencias (Cohen y Meskin, 2004; Walton, 1984). Las fotografías son independientes de las creencias, ya que las creencias del fotógrafo no influyen en el

²⁶ Cabe mencionar aquí un libro de Thomas Kroedel (2020), *Mental Causation. Counterfactual Theory*, que se basa justamente en Lewis para intentar fundamentar una teoría de la causación mental que pretende ser superadora tanto del dualismo como del monismo fiscalista, a partir del desarrollo del **condicional contrafáctico**. En este texto, señala que nuestras mentes producen efectos físicos (por ejemplo, el movimiento de nuestros cuerpos). Y mediante el condicional contrafáctico, defiende una teoría de causación mental en término de producción de diferencias. Si nuestras mentes hubieran sido diferentes, el mundo físico habría sido diferente; por consiguiente, la mente causa eventos en el mundo físico (Kroedel, 2020, ix).

El tema del condicional ha dado lugar a numerosas reflexiones, tanto de lógicos como, incluso, de investigadores de la inteligencia como Jean Piaget. Excede a los límites de este trabajo abordar este tema apasionante pero, al mismo tiempo, sumamente complejo.

contenido de la fotografía al interrumpir o desviar el proceso causal por el cual transmite información sobre su temática; tampoco es necesario invocar las creencias del fotógrafo para interpretar la fotografía. Ambas cualidades las subsumo bajo la rúbrica de independencia teórica.

En cambio, en el caso de la iRMf, la interpretación de la señal depende de la teoría: ciertas relaciones teóricas entre la neuroimagen y la actividad neuronal son necesarias para justificar la significatividad. “Los límites de nuestra interpretación están regidos por la naturaleza de estas relaciones” (Roskies, 2010, p. 210).

3.4.2. Comparaciones

Aunque las coloridas representaciones obtenidas mediante resonancia magnética funcional pueden parecer “fotografías” del cerebro en acción, su valor ontológico (i.e., su estatus de realidad) es profundamente distinto. A diferencia de la fotografía, **las neuroimágenes no son miméticas** en relación con la función cognitivo-conductual, ya que no existe una correspondencia directa entre la señal registrada y la actividad mental representada. Además, la relación entre la señal de RMf y la actividad neuronal es problemática: distintos patrones de actividad cerebral pueden producir la misma señal, lo que implica que **no se cumple el criterio de dependencia contrafáctica** (Roskies, 2008).

Otro aspecto crucial es que la interpretación de las neuroimágenes **depende fuertemente de modelos teóricos**. Mientras que una fotografía captura patrones de luz de manera directa, las imágenes de activación cerebral son construcciones estadísticas inferidas a partir de datos brutos. Asumir que estas imágenes constituyen evidencia autosuficiente es un error metodológico, ya que su significado solo puede comprenderse dentro de un marco teórico determinado.

Un estudio realizado por McCabe y Castel (2008) demostró que la presencia de neuroimágenes puede influir en la percepción de credibilidad científica. En una serie de experimentos, se presentaron artículos sobre neurociencia cognitiva a diferentes grupos de participantes. Mientras que algunos de estos artículos incluían gráficos o diagramas, otros mostraban imágenes cerebrales obtenidas por resonancia magnética funcional (RMf). Los resultados indicaron que los artículos acompañados por neuroimágenes fueron evaluados como más creíbles, incluso cuando el contenido científico era idéntico. Este hallazgo sugiere que el atractivo visual de las imágenes cerebrales puede llevar a una aceptación acrítica de los resultados, reforzando la tendencia a interpretar las neuroimágenes como evidencia autosuficiente.

Para comprender plenamente su significado, es necesario preguntarse qué tipo de entidad representa realmente una neuroimagen y bajo qué condiciones puede considerarse evidencia

válida. Estas cuestiones, tanto en su dimensión ontológica como epistemológica, serán exploradas a continuación.

3.5. ¿Qué es, entonces, una neuroimagen?

Si la neuroimagen no es similar a una fotografía, pero tampoco es un dibujo... ¿Qué es? Podemos afirmar que **es la representación gráfica de un mapa estadístico**.

3.5.1. Una neuroimagen es un mapa estadístico

En la opinión de Colin Klein (2010b, 267), el estatus de la iRMf como medida indirecta de la actividad cerebral es, en realidad, irrelevante. Lo que se debe tener en cuenta es que las neuroimágenes muestran mapas de regiones donde hubo una diferencia **estadísticamente significativa**.

Para establecer esta diferencia en la resonancia magnética funcional, la señal BOLD en cada subregión (o vóxel) es sometida a un NHST²⁷ (prueba de la hipótesis nula):

En primer lugar, se calcula la probabilidad de que la hipótesis nula sea cierta: es decir, una condición experimental no tuvo un efecto real sobre la señal de RM observada, por lo que la actividad neuronal en una región permaneció sin cambios mientras el sujeto realizaba diferentes tareas cognitivas. Este valor, que representa la probabilidad de observar datos condicionados a la hipótesis nula, se denomina valor p. En el segundo paso, se compara el valor p con un nivel de significancia predeterminado α . Si el valor p es inferior a α , los datos son *estadísticamente significativos*. Un resultado significativo es uno que sería poco probable que se observara si la hipótesis nula fuera cierta; con un α de 0.01, por ejemplo, uno podría esperar ver datos significativos en solo alrededor de 1 de 100 observaciones en regiones donde la hipótesis nula fuera cierta. (Klein, 2010b, p. 267)

²⁷ De acuerdo con Nickerson (2000, p. 242) el concepto de NHST o “Prueba de significancia estadística de la hipótesis nula” puede tener distintas connotaciones pero, en investigación psicológica, representa la hipótesis de que no existe ninguna diferencia entre dos conjuntos de datos con respecto a algún parámetro, generalmente sus medias, o de que no se produce ningún efecto de una manipulación experimental en la variable dependiente de interés.

La hipótesis nula establece que el grupo experimental y el grupo de control no son diferentes con respecto a una propiedad específica de interés y que cualquier diferencia encontrada entre sus medias se debe a la fluctuación del muestreo. (Carver, 1978, p. 381)

Los resultados de estos cálculos de probabilidades se representan como mapas paramétricos estadísticos (SPM). Los SPM, entonces, muestran aquellos vóxeles en los que el valor p para esa región menor o igual que el nivel de significancia α . El método empleado es univariante, es decir que cada vóxel es relacionado de manera independiente con el estímulo aplicado.

Por lo general, se usa un rango de colores para representar la magnitud del valor p calculado para una región, con colores más brillantes que indican valores p más bajos. **La neuroimagen, entonces, es la representación gráfica de los resultados de miles de pruebas simultáneas de significancia, mostrando las áreas en las que nuestros datos nos permiten rechazar la hipótesis nula.**

3.5.2. ¿Qué tipo de evidencia nos brinda la neuroimagen?

Klein llega a afirmar que “las neuroimágenes no proveen ni siquiera un sustento débil para las hipótesis de tipo funcional” (2010b, p. 268) y basa su afirmación en lo que llama “el argumento escéptico” (en la cita que sigue, “D” corresponde a la noción de evidencia).

El escepticismo acerca de las neuroimágenes equivale a la proposición de que la relación de probabilidad de una hipótesis funcional a su nulidad siempre es muy baja cuando tratamos un SPM como datos. La forma precisa de este escepticismo depende de cuál de las tres formas de construir D que elegimos. En primer lugar, D podría ser el hecho de que hubo una *mayor actividad*: es decir, el hecho de que hubo más actividad cerebral en una condición que en otra. En segundo lugar, D podría ser el hecho de que hubo una *diferencia estadísticamente significativa en la actividad*: no solo diferencia, sino diferencia estadísticamente detectable. En tercer lugar, D podría asociarse con el curso temporal real de algunos datos estadísticamente significativos: no solo el hecho de la significancia, en otras palabras, sino el hecho de que la *diferencia era significativa y tomaba tal o cual forma*. Cada una de las tres formas de leer D hace que las neuroimágenes sean problemáticas como evidencia. (2010b, pp. 268-269)

En cada uno de estos casos, los argumentos de Klein girarán en torno a la noción de que los cerebros son sistemas *causalmente densos* (2010b, p. 269) y que, por consiguiente, existen razones para creer que cualquier tarea tendrá efectos generalizados en el cerebro. Estos efectos (activaciones) pueden aparecer como funcionalmente no significativos, pero esto dependerá de cómo se elija el nivel de significancia denominado α . “La justificación teórica para elegir un nivel α depende de la conveniencia de reducir los falsos positivos.” (Klein, 2010b, p. 270)

Empero:

Si todo en el cerebro está débilmente conectado con todo lo demás, entonces se debe esperar que cada tarea resulte en alguna diferencia en la activación neuronal. Esto significa que la hipótesis nula H_0 siempre es estrictamente falsa. Pero si no hay nulos verdaderos, entonces es trivialmente imposible tener un falso positivo, sin importar qué α elija. Entonces, si los cerebros son causalmente densos, cualquier α servirá, y la elección de uno es arbitraria. (Klein, 2010b, p. 270)

La raíz de la posición escéptica radica en que el NHST es controvertido: no brinda información relevante para probar hipótesis en sistemas en los que las hipótesis nulas suelen ser falsas. Y las hipótesis nulas rara vez son ciertas en los sistemas causalmente densos. Como señala, no sin cierta ironía, Robert Abelson (1997, p. 12), “Actuamos tontamente cuando celebramos los resultados con $p = 0.05$, pero nos revolcamos en la autocompasión cuando $p = 0.07$ ”.²⁸

Decir que algún área del cerebro A contribuye a la ejecución de E no es simplemente decir que hace algo cuando se ejecuta E. En cambio, es una afirmación acerca de lo que hace A, es decir, que hace algo particular que contribuye al desempeño de E. Es perfectamente posible que A haga algo que haga una pequeña pero necesaria contribución a E, o que A sea extremadamente activa pero no tenga efecto sobre E. (Klein, 2010b, p. 274)

No obstante, Klein concluye su artículo señalando que el escepticismo sobre la neuroimagen funcional no debe llevarnos a un total rechazo de la técnica: no solo porque la experimentación hoy emplea análisis estadísticos más sofisticados, sino porque la información (sobre todo anatómica neural) puede ayudar a restringir las implicaciones de las hipótesis funcionales para así permitir pruebas más detalladas. “Encontrar la significación estadística nunca es suficiente para

²⁸ Esta frase de R. Abelson ilustra un problema fundamental con las pruebas de significancia de hipótesis nula (NHST). El problema radica en la obsesión con un umbral arbitrario de $p = 0,05$ para determinar la “significancia” de un resultado. Esto conduce a una interpretación dicotómica y simplista de los datos, dado que se crea la ilusión de que los resultados con $p < 0,05$ son “verdaderos” o “importantes”, mientras que los resultados con $p > 0,05$ son “falsos” o “irrelevantes”. Esta “fijación” con el umbral de $p = 0,05$ puede llevar a conclusiones engañosas (como ignorar un resultado con $p = 0,051$ mientras se celebra otro con $p = 0,049$ como “un gran descubrimiento”, a pesar de que la diferencia práctica entre ambos puede ser mínima. En lugar de aferrarse a un umbral arbitrario, Abelson aboga por un enfoque más holístico que no solo tenga en cuenta la significancia estadística. En efecto, en su libro *Statistics as Principled Arguments* (1995) acuña el acrónimo MAGIC para establecer las condiciones de la “fuerza persuasiva” de una investigación: magnitud (o tamaño del efecto), articulación (claridad y precisión con las que se presentan los resultados), generalidad (aplicabilidad de los hallazgos a contextos más amplios), interés (potencial de cambiar lo que se cree sobre un tema importante) y credibilidad (confianza en los resultados y en la metodología utilizada). (Cf. Abelson, 1995, p. 11-13)

confirmar una hipótesis, pero proporciona cierta garantía para tomar los datos en serio y realizar análisis adicionales sobre ellos”. (Klein, 2010b, p. 275)

El análisis estadístico de la RMf nos obliga a replantearnos una vez más su estatus como fuente de evidencia. Más que capturar directamente la actividad cognitiva, las neuroimágenes representan **probabilidades de activación en función de modelos predefinidos**. Esta perspectiva introduce una pregunta fundamental: ¿hasta qué punto la neuroimagen puede validar teorías psicológicas o cognitivas sin caer en la falacia de inferencia inversa? Esta cuestión será central en los próximos capítulos, donde examinaremos el vínculo entre neuroimagen y psicología cognitiva.

3.6. Las inferencias en neuroimágenes

Si llevamos al extremo la afirmación "**correlación no implica causalidad**", podríamos caer en una postura escéptica que impediría cualquier inferencia científica basada en regularidades empíricas. Sin embargo, en la práctica, la correlación es un punto de partida fundamental en la investigación, siempre que se inserte dentro de un **marco explicativo** que justifique su interpretación.

Por ejemplo, en epidemiología, la relación entre fumar y el cáncer de pulmón fue inicialmente observada como una correlación estadística. Sin embargo, su interpretación como una relación causal requirió la integración de estudios experimentales, modelos biológicos y la identificación de mecanismos moleculares dentro de un marco teórico que proporcionara una explicación coherente de dicha relación (Bradford Hill, 1965).

Este problema se manifiesta de manera similar en materia de neuroimágenes. Encontrar que una región cerebral se activa durante una tarea cognitiva no implica automáticamente que dicha región sea la causa del proceso mental en cuestión.

Hecha esta aclaración, me propongo introducir el tema de los diversos tipos de inferencia en materia de neuroimágenes. Para ello, se hace necesario un primer acercamiento a las nociones básicas de la estadística bayesiana.

3.6.1. El teorema de Bayes

Es una fórmula matemática simple utilizada para calcular probabilidades condicionales, es decir, la probabilidad de que tenga lugar un evento X dado un evento Y. (Joyce, 2021; Rangel, 2015; Poldrack, 2006). En otras palabras, si tenemos alguna hipótesis sobre el mundo, posiblemente basada en información previa, la regla de Bayes nos indica cómo reevaluar la probabilidad de esta hipótesis a la luz de nueva evidencia (Bower y Davis, 2012). Es decir, nos permite actualizar las creencias o probabilidades a la luz de nueva evidencia; nos ayuda a ajustar nuestras ideas sobre algo en función de nueva información.²⁹

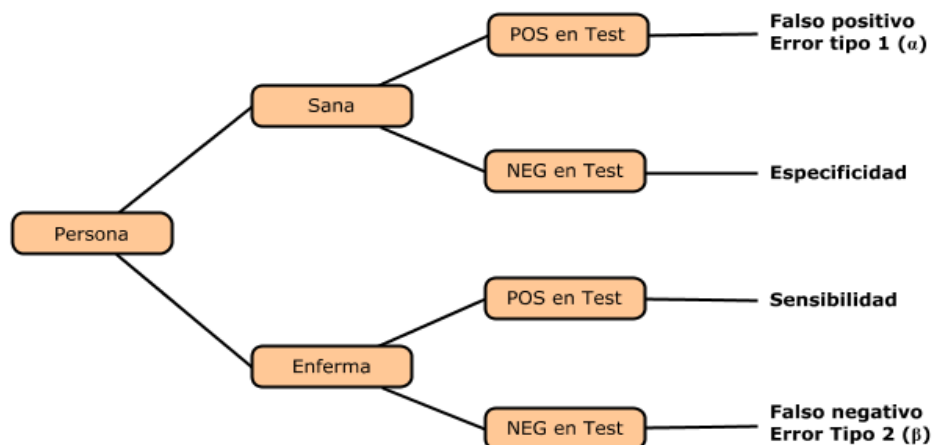
²⁹ Para ilustrar con un ejemplo la potencia del Teorema de Bayes, consideremos los datos siguientes:

Según los estudios epidemiológicos actuales (tomados de *ALS News Today*) la esclerosis lateral amiotrófica (ELA) tiene una prevalencia de aproximadamente 4,5 por cada 100.000 personas, lo cual se traduce a 0,000045.

Imaginemos un test para detectar la enfermedad con un 95% de confiabilidad. Esto significa:

- Sensibilidad del test ($P(\text{positivo}|\text{ELA})$): 0,95 (95%), que representa la capacidad del test para identificar correctamente a los individuos con ELA.
- Especificidad del test ($P(\text{negativo}|\text{no ELA})$): 0,95 (95%), que indica la capacidad del test para identificar correctamente a los individuos que no tienen ELA. La probabilidad de un falso positivo ($P(\text{positivo}|\text{no ELA})$) sería entonces 0,05 (5%).

Para una persona que se somete a dicho test, existen cuatro posibilidades:



Aplicando el teorema de Bayes, encontramos que la posibilidad de llegar a padecer ELA habiendo obtenido un resultado positivo es:

- $P(\text{ELA}|\text{positivo}) \approx 0.0008549$

Este valor obtenido significa que, si una persona obtiene un resultado positivo en el test, la probabilidad de que realmente tenga ELA es de aproximadamente 0,08549 %, o 0,0008549.

Esto equivale a aproximadamente 1 en 1170. Este resultado ilustra cómo, a pesar de la alta confiabilidad del test, la baja prevalencia de la enfermedad resulta en una alta tasa de falsos positivos.

Cabe señalar también aquí una derivación especulativa pero influyente del teorema: la noción de “cerebro bayesiano”, según la cual el sistema nervioso opera como un mecanismo de inferencia probabilística que actualiza hipótesis internas a partir de los errores de predicción entre lo esperado y lo percibido. Este modelo sugiere que la cognición

La probabilidad de A dado B se define como:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Ahora bien, tendemos a asumir, erróneamente, que $P(A|B)$ es igual o casi igual a $P(B|A)$ y esto es conocido como “**falacia de la probabilidad condicional**”.

Para calcular $P(A|B)$ conociendo $P(B|A)$ se utiliza el teorema de Bayes, que suele tomar algunas de las formas siguientes:

$$P(A|B) = P(B|A) \frac{P(A)}{P(B)}$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|-A)P(-A)}$$

Supongamos, para ilustrar esta fórmula, que estamos investigando la respuesta del cerebro a un estímulo visual específico y que nos valemos de la RMf para medir la activación en una región particular del cerebro. Contamos, además, con la información siguiente:

- **Probabilidad previa (P(A))**: porcentaje de probabilidad de que la región se active en respuesta al estímulo, a partir de estudios previos.
- **Sensibilidad del test (P(B|A))**: probabilidad de la RMf de detectar correctamente la activación si está presente.
- **Especificidad del test (P(B|-A))**: probabilidad de que el estudio arroje un resultado positivo falso si no hay activación.

Si obtenemos un resultado positivo en la iRMf, aplicamos el teorema de Bayes para calcular la probabilidad posterior de que realmente haya activación cerebral. El resultado obtenido

humana generaría constantemente conjeturas sobre el entorno y las contrastaría con las entradas sensoriales. La discrepancia entre ambos procesos activaría ajustes para reducir el error predictivo. Algunos autores advierten sobre la escasa falsabilidad de este modelo y también sobre la insuficiencia de evidencia empírica que respalde que las neuronas operan *literalmente* bajo cálculos bayesianos. Por ello, se propone considerarlo más bien como un marco heurístico para construir modelos de cognición y no como una descripción directa del funcionamiento cerebral (Ongaro y Kaptchuk, 2019; Bowers y Davis, 2012; Tenenbaum et al., 2011).

proporciona una probabilidad actualizada basada en los datos observados y las probabilidades previas, mejorando así la interpretación de los resultados.

En el capítulo anterior se consideraron los distintos modelos “clásicos” de análisis estadístico de las neuroimágenes. El teorema de Bayes nos ofrece un nuevo conjunto de modelos; en el cuadro siguiente consideraremos las principales diferencias.

Tabla 3-2

Comparación de los modelos tradicionales y los bayesianos

Aspecto	Modelos tradicionales	Modelos bayesianos
Enfoque	Se basan en pruebas de hipótesis y en la idea de rechazar la hipótesis nula.	Permiten incorporar información previa sobre los parámetros del modelo.
Inferencia	Se realiza a partir del valor p , que indica la probabilidad de observar los datos si la hipótesis nula fuera cierta.	La inferencia se basa en la distribución posterior, que combina los datos con la información previa.
Problema de comparaciones múltiples	Sufren de este problema: aumenta la probabilidad de obtener un falso positivo cuando se realizan múltiples pruebas.	No sufren de este problema, ya que la inferencia en un vóxel no depende de la inferencia en otros.
Incorporación de información previa	Limitada en su capacidad para incorporar información previa sobre los parámetros del modelo.	Permiten incorporar de manera explícita la información previa mediante la distribución previa.
Flexibilidad	Pueden ser limitados, especialmente si asumen linealidad y no son generalizables a diferentes poblaciones.	Ofrecen una mayor flexibilidad en la especificación del modelo y pueden manejar relaciones no lineales.
Validez aparente	Se centra en probar si la diferencia es significativa con respecto a cero.	Se centra en la probabilidad de activación dado un tamaño específico, lo que mejora la validez aparente.
Ventajas	Simplicidad, menos intensivo computacionalmente.	Incorporación de información previa, mayor flexibilidad, evita el problema de comparaciones múltiples.
Desventajas	Puede no ser generalizable a diferentes poblaciones y no manejar relaciones no lineales. Además, puede tener un mayor riesgo de falsos positivos debido a las comparaciones múltiples.	Más intensivo computacionalmente y requiere mayor capacidad para estimar distribuciones posteriores.

Nota: elaboración propia a partir de Faro (2010) y Filippi (2016).

3.6.1.1. *Falacia de la probabilidad condicional*

Señalé antes que confundir $P(A|B)$ con $P(B|A)$ da lugar a la “falacia de la probabilidad condicional”. En lo que hace a nuestro tema, esta se produce cuando se malinterpreta la relación entre la activación cerebral y los estados experimentales.

Supongamos, por ejemplo, que se observa la activación de una región cerebral específica cuando los sujetos realizan una tarea cognitiva particular (A). A partir de allí, los investigadores podrían

inferir que esta activación es evidencia de que la región cerebral es responsable de la tarea (B). La probabilidad de que una tarea active una región cerebral específica ($P(A|B)$) no es necesariamente igual a la probabilidad de que la región cerebral específica se active durante una tarea dada ($P(B|A)$). En otras palabras, que una región cerebral se active durante una tarea no implica que esa región sea la causa directa de la tarea.

3.6.2. Tipos de inferencias

Las neuroimágenes pueden ser usadas en dos sentidos:

1. Dada una teoría cognitiva, la neuroimagen puede ser empleada para inferir funciones cerebrales: esto sería lo que Poldrack denomina “**inferencia directa**” (*forward inference*) (Poldrack 2006, p. 59; Klein, 2010, p. 188).

Dado un proceso cognitivo C1 y una región cerebral R1 que puede estar activada (+) o desactivada (-), la inferencia cobraría esta forma lógica:

$C1 \Rightarrow + R1$

2. El segundo uso, más controversial, consiste en el empleo de la RMf para confirmar o desconfirmar las teorías cognitivas en sí mismas. Esto suele conocerse como **inferencia inversa** (*reverse inference*) (Poldrack 2006, p. 59; Klein, 2010, p. 188): a partir de la presencia de una activación cerebral se infiere la participación de una función cognitiva.

$+R1 \Rightarrow C1$

Desarrollaré ahora el primero de los usos, el referido a la inferencia directa.

3.6.3. Inferencia directa

3.6.3.1. Inferencia directa y principio de inserción pura:

Colin Klein (2010, p. 188) plantea que el primero de los usos de la neuroimagen no necesariamente permite inferir la organización del cerebro. Se vale para ello de un ejemplo tomado de Friston et al. (1996, p. 99), cuyo diseño experimental nos interesa profundizar un poco más.

Consideremos el problema de diseñar un experimento para identificar áreas del cerebro activadas selectivamente por recuperación fonológica durante la denominación de objetos. Los procesos cognitivos involucrados en esta tarea incluyen el análisis visual, el reconocimiento de objetos, la recuperación fonológica y el habla. Supongamos que no nos

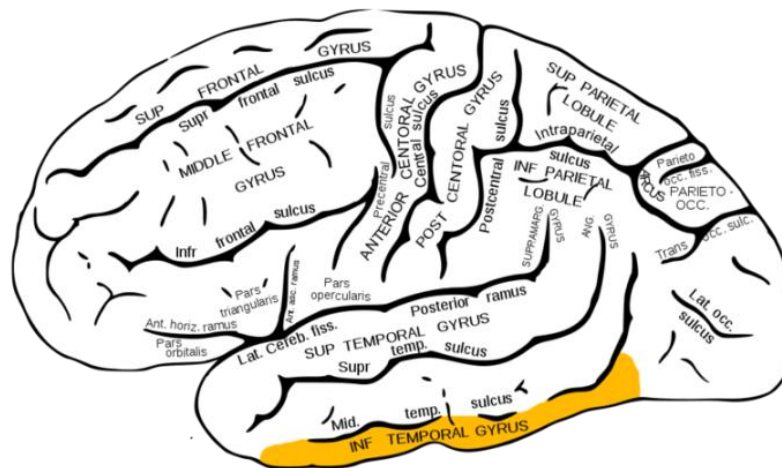
interesan los aspectos sensoriomotores de la tarea, sino que deseamos probar la hipótesis de que las regiones inferotemporales están involucradas tanto en el reconocimiento de objetos como en la recuperación fonológica. En este caso queremos una serie de tareas que, en restas sucesivas, aíslen estos dos componentes cognitivos (es decir, tres tareas). Las tareas utilizadas fueron:

- (A) decir "sí" cuando se presenta una forma coloreada (análisis visual y habla),
- (B) decir "sí" cuando se le presenta un objeto coloreado (análisis visual, habla y reconocimiento de objeto) y
- (C) nombrar un objeto de color presentado visualmente (análisis visual, habla, reconocimiento de objetos y recuperación fonológica).

La sustracción de la tarea A de la tarea B debe identificar las regiones del cerebro asociadas con el reconocimiento de objetos, y la sustracción de la tarea B de la tarea C debe identificar las regiones implicadas en la recuperación fonológica. La hipótesis aquí es que ambas sustracciones deberían activar las regiones inferotemporales izquierdas (entre otras regiones). (Friston et al., 1996, p. 99)

Figura 3.2.

Región inferotemporal.



Nota: modificado a partir de Gray, H. (1918). *Anatomy of the human body*. 20a. ed.

Los resultados arrojaron que la región inferotemporal fue más activa cuando los sujetos tenían que detectar objetos en lugar de formas, **pero no se activó más cuando los sujetos tenían que nombrar objetos además de detectarlos**³⁰.

Se podría inferir a partir de allí que la región inferotemporal estudiada es relevante para el reconocimiento de objetos, pero no para su denominación. Pero la investigación propuso el agregado de una nueva tarea:

D) Indicar el color de una forma presentada (análisis visual, habla y recuperación fonológica)

La sustracción de la tarea A de la tarea B da la activación debida al reconocimiento de objetos en ausencia de recuperación fonológica y la sustracción de la tarea D de la C da la activación dependiente del reconocimiento en el contexto de la recuperación fonológica. La inserción pura requiere que estas activaciones sean las mismas y este no fue el caso: la activación en el contexto de la recuperación fonológica fue mucho mayor que sin recuperación fonológica.

Para mayor claridad, representaré lo señalado en un cuadro según lo que se esperaría por el principio de sustracción cognitiva:

Tabla 3-3

Esquema del experimento con el que Friston et al. ilustran la sustracción cognitiva

A-B = reconocimiento de objetos		
D-A=recuperación fonológica	Tarea A	Tarea B
	análisis visual habla	análisis visual habla reconocimiento de objetos
	Tarea D	Tarea C
	análisis visual habla recuperación fonológica	análisis visual habla reconocimiento de objetos recuperación fonológica
		C-B=recuperación fonológica

Como vemos, la sustracción de D-A y la de C-B deberían corresponder, ambas, a la recuperación fonológica y, por consiguiente, dar resultados similares, lo que no ocurrió (el valor de activación

³⁰ Klein, 2010, p. 188. El original en inglés dice “but was not further activated when subjects had to additionally name shapes along with detecting them”. Según el experimento realizado por Friston et al., consideramos que colocar “objetos” en lugar de “formas” resulta más adecuado.

de D-A fue muy superior al de C-B). Se demuestra así cómo se produce una violación del principio de inserción pura (y por ende al de sustracción cognitiva).

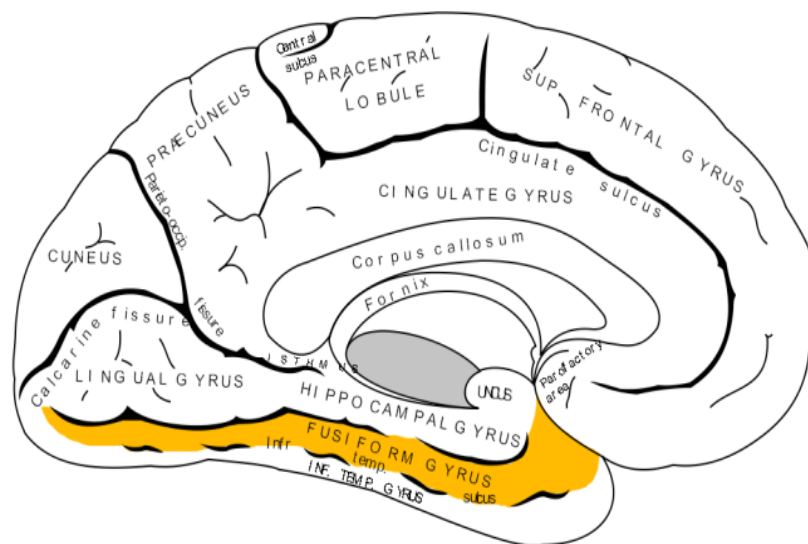
3.6.3.2. Inferencia y ontología cognitiva

Al respecto, señala Klein (2010, 190):

Muchos estudios de NI son relativamente inespecíficos, en el sentido de que diferentes experimentos atribuyen diferentes funciones a la misma región del cerebro. Price y Friston señalan que una región del giro fusiforme lateral posterior izquierdo (PLF) se ha asociado con una serie de funciones: procesar formas de palabras visuales, detectar los atributos visuales de los animales, nombrar colores, decodificar Braille, imaginar objetos y tomar decisiones de acción sobre objetos familiares (2005, p. 265–7). ¿Qué función, entonces, debemos atribuir a PLF? Decir que la función es nombrar colores o animales es engañoso: hace más que eso. Decir que hace todas estas cosas no es esclarecedor: una buena teoría debería unificar nuestras observaciones, no solo enumerarlas.

Figura 3-3.

Giro fusiforme lateral.



Nota: modificado a partir de Gray, H. (1918). *Anatomy of the human body*. 20a. ed.

Cada una de estas funciones podría considerarse diferentes desempeños de una función más general, que las engloba. Es la conclusión a la que arriban Friston y Price en relación con el PLF: “En un nivel de descripción, el área tiene múltiples funciones que están determinadas por las

regiones junto con las cuales se activa, mientras que en otro nivel, se le puede asignar una única función: integrar las señales sensoriales con la salida motora” (2005, p. 268). Indican así también la necesidad de una “ontología cognitiva” que permita interpretar los hallazgos de las NI en el nivel adecuado de abstracción³¹.

3.6.4. Inferencia inversa

De acuerdo con Poldrack (2006, p. 59), en algunos estudios la inferencia inversa ocupa un lugar central. Refiere, por ejemplo, una investigación que involucra transacciones económicas entre dos sujetos, con la posibilidad de que uno de ellos traicione la confianza del otro y de que el primero le imponga algún tipo de castigo. En la conclusión, se señala:

Hipotetizamos que el castigo altruista proporciona alivio o satisfacción al castigador y activa, por lo tanto, regiones cerebrales relacionadas con la recompensa. Nuestro diseño genera cinco contrastes en los que se puede probar esta hipótesis, y el estriado dorsal anterior se activa en los cinco contrastes, lo que sugiere que el caudado juega un papel decisivo en el castigo altruista. La activación del caudado es particularmente interesante porque esta región del cerebro ha estado implicada en la toma de decisiones o acciones motivadas por recompensas anticipadas. (De Quervain, 2004, p. 1268).

Analicemos lo indicado aquí.

1. Se quiere probar que el castigo (C) proporciona satisfacción o sensación de recompensa (R).
2. Se sabe, por investigaciones anteriores, que la sensación de recompensa provoca la activación del núcleo caudado (NC).

Podríamos expresar de forma simbólica la inferencia realizada:

$R \Rightarrow + NC$
$C \Rightarrow + NC$
$\therefore C \Rightarrow R$

³¹ En el siguiente capítulo enfocaremos el complejo tema de esta “ontología cognitiva”.

Queda claro que se trata de una inferencia inválida, por afirmación del consecuente (Poldrack, 2006, p. 60). Solo podría ser válida si se hubiera verificado que $R \Leftrightarrow + NC$, es decir, que el núcleo caudado se activa **si y solo si** estamos en presencia de una sensación de recompensa.

Esta fuerte premisa “es equivalente a la afirmación de que la relación entre el razonamiento y R es un invariante psicofísico (Cacioppo et al. 2007, 13-14); en términos técnicos, que existe un mapeo uno a uno independiente del contexto entre la función cognitiva y la estructura cerebral” (Klein, 2010, p. 192)

Por ello, como argumenta Klein, para algunos autores la inferencia inversa es siempre inválida. Entre ellos, para Fodor, quien considera “sobre bases filosóficas, que la ciencia cognitiva es metodológicamente independiente de la neurociencia (Fodor, 1997). Esta posición, conocida como *fisicalismo no reduccionista*³², sigue siendo popular entre los filósofos de la mente”. (Klein, 2010, p. 191).

En una línea con puntos de contacto con lo anterior, “Harley argumenta que los datos del cerebro son irrelevantes para la psicología cognitiva hasta que tengamos una teoría psicológica completa³³— momento en el cual NI [la neuroimagen funcional] no tendría nada de interés para contribuir a la psicología.” (Klein, 2010, p. 191)

En cambio, Florian Hutzler (2014) adopta una postura matizada, más optimista, con respecto a la inferencia inversa, argumentando que, si bien presenta desafíos metodológicos, no debe considerarse intrínsecamente inválida y recomienda:

- La integración de probabilidades previas mediante un abordaje bayesiano, que incorpore conocimientos previos sobre la relación entre regiones cerebrales y funciones cognitivas (Hutzler, 2014, p. 1065).
- La utilización de datos convergentes de múltiples estudios, minimizando los riesgos asociados con la multifuncionalidad de las regiones cerebrales.

3.6.5. Más allá de la inferencia

Hemos visto que tanto la inferencia directa (de la teoría cognitiva a la neuroimagen) como la inversa (de la neuroimagen a la teoría cognitiva) tienen dimensiones problemáticas. Esto no significa que se instale alguna forma de escepticismo radical con respecto a las neuroimágenes.

³² El fisicalismo no reduccionista o no eliminativo sostiene que todo lo que existe es, en última instancia, físico, pero no todo puede ser reducido de manera completa a términos físicos sin perder información importante o conceptos valiosos.

³³ En el siguiente capítulo argumentaré sobre esta temática.

En este contexto, Klein plantea la existencia de otros posibles usos de la evidencia provista por las neuroimágenes. Los mencionaremos brevemente.

a) Problemas de consistencia

La evidencia en NI puede ser contrastada con las predicciones de teorías en competencia, con miras a favorecer una u otra. Pero señala Klein que estas pruebas de consistencia constituyen “una visión falsacionista de la comprobación de teorías (Popper, 1959)” (2010, p. 193). Y agrega:

En una explicación falsacionista, las teorías consistentes con los datos permanecen en juego, mientras que las teorías que son inconsistentes con los datos deben ser abandonadas. Sin embargo, el falsacionismo no permite decidir entre teorías no falsificadas. Y como los filósofos han señalado durante mucho tiempo, la mayoría de las teorías pueden hacerse consistentes con evidencia aparentemente recalcitrante simplemente adoptando hipótesis auxiliares apropiadas (Klee 1996, Capítulo 4; Harley 2004b, p. 50).

b) Explicaciones probabilísticas

Señala Klein que otra aproximación a los datos de NI es la probabilística: en un marco bayesiano probabilístico, los datos D de NI respaldan una hipótesis H en caso de que aumente la probabilidad de que esa hipótesis sea cierta (Klein, 2010, 194).

Entre otras, señala como ventaja de este tipo de explicaciones:

- Son informativas, es decir que dadas dos hipótesis en juego, inclinan la balanza a favor de alguna de ellas.
- Permite una versión “débil” de la inferencia inversa: es decir, dada una activación cerebral, es posible inferir la probabilidad de que ocurra cierto proceso cognitivo. No obstante,

Poldrack utilizó estudios existentes para estimar el factor de Bayes de las inferencias inversas sobre el procesamiento del lenguaje. Encontró que la inferencia inversa podría proporcionar un 'aumento positivo pero relativamente débil en la confianza' en las hipótesis cognitivas (2006, p. 62). (Klein, 2010, p. 194)

c) Información exploratoria

El análisis exploratorio de datos, por el contrario, simplemente busca patrones en los datos y sugiere interpretaciones de ellos (Tukey, 1977; Good, 1983). En la literatura de imágenes, el análisis exploratorio de datos a veces se denomina análisis basado en datos (en oposición al análisis confirmatorio basado en hipótesis). (Klein, 2010, p. 194).

En esta línea, según Klein, ya se han logrado avances interesantes.

El énfasis en el análisis exploratorio [...] sugiere la posibilidad de una colaboración interdisciplinaria dinámica entre los investigadores en neuroimágenes y los filósofos. Como señala Harrison, “está claro que tenemos trabajo conceptual que hacer antes de encender el escáner” (2008, p. 321). Ese trabajo conceptual asegura que los filósofos puedan ayudar a la nascente ciencia de NI tanto como pueden aprender de ella. (Klein, 2010, p. 195)

Las inferencias en neuroimagen funcional, entonces, presentan desafíos fundamentales para su interpretación. Mientras que la inferencia directa permite identificar correlaciones entre activación cerebral y procesos cognitivos, la inferencia inversa es mucho más problemática, ya que asume causalidades que no siempre están justificadas. La incorporación de enfoques bayesianos ha intentado mitigar estas dificultades, pero sigue existiendo un debate abierto sobre la validez de estas técnicas y su impacto en la credibilidad de los resultados neurocientíficos.

3.7. Sesgos de publicación

A partir del teorema de Bayes (que abordamos previamente) y una serie de “suposiciones razonables” sobre las investigaciones, John Ioannidis publicó en 2005 su famoso artículo “Why most published research findings are false” (“Por qué la mayoría de los resultados de investigaciones publicados son falsos”).

En este texto afirma que “la probabilidad de que un resultado de investigación sea verdadero depende de la **probabilidad previa** de que sea verdadero (antes de hacer el estudio), del **poder estadístico** del estudio y del nivel de **significancia estadística**”³⁴.

Es decir:

- Antes de la realización de un estudio existe una probabilidad previa, basada en la evidencia existente hasta el momento. Si esta probabilidad es baja, es posible que un resultado estadísticamente significativo no sea suficiente para permitirnos afirmar la hipótesis.
- El poder estadístico del estudio es la capacidad para detectar un efecto real, si existe. Un estudio con bajo poder estadístico aumenta la probabilidad de obtener falsos negativos.

³⁴ El resaltado es propio.

- Finalmente, el nivel de significancia α permite establecer el umbral de un resultado estadísticamente significativo. Reducir el nivel de significancia disminuirá la probabilidad de obtener falsos positivos, pero también puede aumentar la probabilidad de obtener falsos negativos.

Kelly y Hoptman (2022) sintetizan la posición de Ioannidis señalando que, ante el resultado positivo de una prueba, la probabilidad de que este hallazgo sea verdadero (VPP, valor predictivo positivo) depende de tres variables:

1. el nivel α de significancia estadística (donde α es la probabilidad de un resultado positivo de la prueba, si la hipótesis es falsa, o tasa de error tipo I),
2. el poder del estudio ($1 - \beta$, donde β es la probabilidad de una prueba negativa, dado que la hipótesis sea verdadera o tasa de error tipo II), y
3. la probabilidad de que la hipótesis sea verdadera (R, la razón entre la probabilidad de que la hipótesis sea verdadera y la probabilidad de que sea falsa).

Esta relación se expresa con la ecuación

$$VPP = \frac{R (1 - \beta)}{[\alpha + R(1 - \beta)]}$$

De esta ecuación se deduce que cualquier hipótesis probablemente será falsa, incluso después de una prueba positiva, cuando $R < \alpha$. Se infiere también que, cuando la potencia de estudio es igual a α , la probabilidad de que la hipótesis sea verdadera sigue siendo la misma que antes de la prueba. Por lo tanto, los estudios con un poder estadístico inadecuado carecen de la capacidad de aumentar nuestra confianza en las hipótesis probadas.

Pero, además, el VPP también puede verse reducido por los sesgos que elevan el valor real de α por encima de su valor nominal.

Ioannidis (2005, p. 617) define el sesgo como “la combinación de varios factores de diseño, datos, análisis y presentación que tienden a producir hallazgos de investigación cuando no deberían producirse”. El sesgo en el contexto de una investigación puede implicar la manipulación del análisis o el informe de los resultados, la presentación de información selectiva o distorsionada³⁵.

³⁵ También se presenta el sesgo en el *peer review* o revisión por pares. Una investigación de 2020 arrojó como resultado que “el 12% (179) de los conjuntos de comentarios incluyó al menos un comentario no profesional hacia el autor o su trabajo, y el 41% (611) contenía críticas incompletas, inexactas o no fundamentadas”. Gerwing, T. G.,

A partir de esta noción, Ioannidis observa diversos elementos para tener en cuenta a la hora de identificar la presencia o no de sesgos, entre ellos:

1. El tamaño de la muestra de la investigación.
2. La dimensión del efecto que se pretende medir.
3. La flexibilidad en los diseños de investigación, las definiciones, los resultados y las metodologías.
4. Los intereses financieros involucrados.

Kelly y Hoptman (2022) y Head, Holman, Lanfear, Kahn y Jennions (2015) señalan, en el universo de los sesgos, un conjunto de fenómenos frecuentes en la investigación científica y, sin duda alguna, epistémicamente peligrosos.

1. El **“efecto cajón de archivo”** (expresión acuñada por Rosenthal en 1979): los estudios con resultados no significativos tienen tasas de publicación más bajas.

Para cualquier área de investigación dada, no se puede decir cuántos estudios se han realizado pero nunca informado. La visión extrema del "problema del cajón de archivos" es que las revistas se llenan con el 5% de los estudios que muestran errores de tipo I, mientras que los cajones de archivo se llenan con el 95% de los estudios que muestran resultados no significativos. (Rosenthal, 1979)

2. El **sesgo de inflación**, también conocido como *p-hacking*, “dragado de datos” o “informe selectivo”, que consiste en manipular los datos de la investigación para que el resultado supere el umbral de error estadístico (usualmente, 0,05). Es posible, por ejemplo, realizar una gran cantidad de pruebas estadísticas con los datos de la investigación e informar solo aquellas cuyos resultados resulten significativos. Los incentivos académicos para la publicación de hallazgos “interesantes” en revistas de alto impacto pueden inclinar la investigación hacia la producción de falsos positivos³⁶.
3. Después de ver los resultados de un estudio, los investigadores pueden sentirse inclinados a reconsiderar sus hipótesis originales para que coincidan con los datos observados, lo que se conoce como **HARKing** (término derivado del acrónimo HARK, *hypothesizing after the results are known*: realizar las hipótesis después de conocer los resultados).

Allen Gerwing, A. M., Avery-Gomm, S., Choi, C. Y., Clements, J. C. y Rash, J. A. (2020). Quantifying professionalism in peer review. *Research integrity and peer review*, 5(1), 1-8.

³⁶ “Publicar o perecer” (*publish or perish*) es una frase usada con frecuencia para hacer referencia a la presión que experimentan los investigadores para publicar los resultados de sus investigaciones y así ser “relevantes” en la comunidad académica. Cf, entre otros, Arredondo, G. P. (2022). Publicar o perecer. *Salud Quintana Roo*, 9(35), 5-5.

Los sesgos de publicación y el *p-hacking* constituyen un problema serio en la investigación en neuroimagen. La presión por obtener resultados estadísticamente significativos puede llevar a prácticas que inflan artificialmente la evidencia a favor de ciertas hipótesis, afectando la confiabilidad de los estudios. Esto plantea la necesidad de una mayor transparencia en los análisis y de estrategias metodológicas que minimicen la manipulación selectiva de datos.

3.8. Discusión y proyecciones

El problema relativo a la pregnancia de lo visual excede el ámbito de las neuroimágenes. En efecto: en la alegoría de la caverna, Platón muestra que lo visible no siempre es verdadero. Los prisioneros toman por real lo que en realidad son sombras proyectadas sobre la pared, copias de objetos que no pueden ver. Solo cuando uno de ellos logra liberarse y ascender al mundo exterior comienza a percibir, no sin dificultad, la verdadera realidad.

La enseñanza es clara: **no todo lo que se ve es real, ni todo lo real puede verse sin mediación**. Y, como antes sucediera con la sombra en la caverna, hoy la imagen neurocientífica se nos presenta como portadora de verdad, por lo que es necesaria una mirada crítica y advertida.

No es posible establecer un isomorfismo —es decir, una correspondencia “uno a uno”— entre lo que se ve y lo que es. La imagen no equivale al fenómeno, sino que lo reconfigura o lo reinterpreta bajo determinados supuestos. Esta distancia entre visualización y realidad es clave para comprender por qué las neuroimágenes deben leerse críticamente, incluso cuando su apariencia visual sugiera evidencia directa.

Esta crítica se vincula con la distinción propuesta por Barberis, Ginnobili y Roffé (2025), quienes identifican diferentes niveles en los que puede surgir desacuerdo sobre lo que se observa en un experimento. Su **taxonomía de la inconmensurabilidad empírica** distingue tres planos: experimental, perceptual y conceptual. La inconmensurabilidad experimental se presenta cuando las diferencias en el diseño o en el objeto de estudio impiden la comparación directa entre resultados. La inconmensurabilidad perceptual, más profunda, implica que, incluso ante el mismo objeto y con la misma técnica, los científicos pueden literalmente ver cosas distintas, modelados por sus compromisos teóricos. Finalmente, la inconmensurabilidad conceptual se da cuando, aun compartiendo lo percibido, los investigadores lo interpretan de forma incompatible por pertenecer a marcos teóricos diferentes. Esta clasificación permite comprender por qué, en neurociencias, el uso de técnicas estandarizadas como la RMf no garantiza una base empírica común ni tampoco

una interpretación neutra de los datos: lo que se considera evidente depende del lugar que ocupan los modelos y conceptos en el entramado teórico del que forma parte la observación.

Por otra parte, el análisis de la RMf ha revelado el carácter problemático de su validez como fuente de evidencia. Uno de los motivos principales es que la interpretación de los datos no es directa, sino que se encuentra mediada por modelos estadísticos sumamente complejos y por supuestos teóricos que influyen en la interpretación final de las imágenes obtenidas.

Los problemas vinculados con la inferencia directa y la inferencia inversa también plantean desafíos fundamentales a la hora de validar los estudios en neurociencia cognitiva. La inferencia directa puede proporcionar correlaciones relativamente claras, si bien, como vimos, no se encuentra exenta de problemas. La inferencia inversa, en cambio, que parte de un patrón de activación cerebral y a partir de él busca deducir un proceso cognitivo, puede dar lugar a correlaciones espurias y la confiabilidad de las conclusiones puede verse seriamente afectada. En este marco, el empleo de estrategias estadísticas —incluyendo los modelos bayesianos— tampoco debe entenderse como garante de acceso a la verdad, porque el concepto de probabilidad surgió a partir de una transformación cultural que introdujo una nueva forma de autoridad epistémica. Valores como medias y desviaciones estándar pasaron a funcionar como “marcadores de objetividad” y desplazaron otras formas de juicio. Ian Hacking (2010) denomina a este proceso “domesticación del azar”:

La probabilidad no es capaz de dictar valores. Sin embargo, en la práctica contemporánea, se ha vuelto la base de toda elección considerada razonable por los funcionarios. Ninguna decisión pública, análisis de riesgo, evaluación de impacto ambiental ni estrategia militar puede hoy llevarse a cabo sin una teoría de la decisión formulada en términos probabilísticos. **Al cubrir la opinión con un barniz de objetividad, reemplazamos el juicio por el cálculo.** (Hacking, 2010, p. 4, el destacado es propio.)

Considero que, en el contexto de este trabajo, cuestionar la naturalización de este “imperio de la estadística” implica volver —una y otra vez— a preguntarse qué clase de verdad produce, qué condiciones pueden hacerla epistémicamente válida, y permitir también que la ética actúe como una instancia crítica, una suerte de “conciencia de la ciencia”, para que los procedimientos formales propios de la neurociencia no terminen incurriendo en automatismos acríticos o clausurando cualquier forma de deliberación conceptual.

Este señalamiento ya fue advertido, entre otros, por Ioannidis (2005), quien mostró que un alto porcentaje de los hallazgos reportados como estadísticamente significativos en la investigación biomédica son en verdad falsos positivos o resultados no replicables. Esta advertencia pone una vez más en cuestión la relación entre evidencia numérica y validez epistémica.

Toda esta complejidad se profundiza si atendemos a cómo la filosofía de la ciencia ha concebido históricamente la relación entre datos, modelos y teorías. Desde la concepción semántica formulada por Patrick Suppes, una teoría científica no se reduce a un conjunto de enunciados, sino que se define como una clase de modelos que median entre los datos empíricos y los fenómenos que se pretende representar. En esta misma línea, van Fraassen sostiene que los modelos científicos no muestran cómo es el mundo en sí, sino cómo se comporta bajo ciertas condiciones observacionales. Para este autor, la función de una teoría es “salvar los fenómenos”, es decir, construir representaciones que se ajusten a lo observable, sin necesidad de asumir que dichas representaciones reflejan directamente la realidad.

A partir de esta perspectiva, podría ser eficaz a los fines de nuestro trabajo pensar la distinción entre **dato**, **modelo** y **significado**. El **dato** corresponde a la señal física recogida —por ejemplo, la variación de oxigenación registrada por la RMf—; el **modelo** es la estructura matemática y estadística que organiza e interpreta esos datos; el **significado** es la atribución conceptual que se hace del resultado final, en términos de funciones cognitivas o procesos mentales. Esta distinción, frecuente en filosofía de las ciencias, puede ponerse en paralelo con los niveles de inferencia señalados por Adina Roskies:

Dato → parte de la corriente causal (de la actividad neural a los datos en bruto)

Modelo → transición dentro de la corriente causal (del dato a la imagen)

Significado → corriente funcional

Si bien los dos primeros niveles no están exentos de dificultades, la inferencia funcional (del patrón de imagen a una función mental) resulta ser la más problemática.

Esta problematización de lo que se da por “evidente” a través de imágenes será aún más apremiante en el siguiente capítulo, donde se analizará la relación entre neurociencia y psicología cognitiva. Se evaluará en qué medida los constructos psicológicos —frecuentemente inestables o ambiguos— pueden ser operacionalizados en estudios de neuroimagen, y cómo las limitaciones epistemológicas de la RMf influyen en nuestra capacidad para validar modelos psicológicos. En este contexto, las cuestiones inferenciales se amplían, pues no solo se trata de interpretar patrones cerebrales, sino también de determinar si los conceptos psicológicos que utilizamos para hacerlo son lo suficientemente robustos y bien definidos como para sostener conclusiones fiables.

4. Neuroimágenes y psicología cognitiva

“... la pieza más crítica del equipo sigue siendo el propio cerebro del investigador. Todo el equipo del mundo no nos ayudará si no sabemos cómo usarlo correctamente, lo que requiere algo más que saber cómo operarlo. Aristóteles no habría sido necesariamente más profundo si hubiera tenido una computadora portátil y hubiera sabido programar. Lo que se necesita con urgencia ahora, con todos estos escáneres zumbando, es comprender exactamente lo que estamos observando, viendo, midiendo y sobre qué estamos preguntándonos.”

(Endel Tulving³⁷)

En el capítulo anterior me referí a la resonancia magnética funcional como fuente de evidencia, identificando sus limitaciones metodológicas e inferenciales. A partir de esto, en el presente capítulo examino su relación con la psicología cognitiva, explorando cómo los constructos psicológicos han sido utilizados en investigaciones neurocientíficas y hasta qué punto pueden integrarse en una ontología cognitiva coherente. En particular, abordo la oposición entre el enfoque modularista/localizacionista y teorías más dinámicas sobre la cognición, además de los desafíos que plantea la estabilidad de los constructos psicológicos en la investigación con neuroimágenes.

Finalmente, analizo las condiciones de confiabilidad y validez en el diseño experimental.

4.1. Psicología cognitiva

“Originariamente, la filosofía abarcaba la totalidad del conocimiento y los filósofos eran polímatas. Por ejemplo, Aristóteles trabajó en problemas de física, biología, psicología y ciencia política, así como en problemas de lógica y ética; y Descartes se interesó en la matemática, la física, la biología y la psicología tanto como en la filosofía propiamente dicha. Hoy en día, la filosofía es una rama de las

³⁷ Entrevista en *Cognitive Neuroscience* (2002, Gazzaniga, Ivry y Mangun, Eds., NY: Norton, p. 323, en <http://www.fmri4newbies.com/>)

humanidades, y los filósofos limitan su atención a problemas conceptuales de un cierto tipo. [...]

La psicología era una rama de la filosofía, de la cual se dice que se independizó alrededor de 1850, con el nacimiento de la psicofísica. ¿Por qué los psicólogos contemporáneos habrían de preocuparse por la filosofía? Porque, lo sepan o no, les guste o no, los psicólogos se basan en y utilizan una cantidad de ideas filosóficas, sobre todo ideas acerca de la naturaleza de la mente y la ciencia. Todo psicólogo, por tanto, no sólo es un científico o un terapeuta, sino un filósofo aficionado, en general *malgré lui*.”

(Mario Bunge, *Filosofía de la psicología*)

4.1.1. Las ciencias de la cognición

Ahora bien, ¿cómo se inscribe la psicología en el marco de las ciencias de la cognición?

Nicolás Venturelli y Sergio Barberis (2023) realizan un recorrido por la historia de las ciencias cognitivas. Por razones de extensión, me limitaré a sintetizar en un breve cuadro las ideas fundamentales de su artículo.

Tabla 4-1

Esquema de la historia de las ciencias cognitivas, según Venturelli y Barberis (2023)

1	Antecedentes	Cibernética como antecedente de revolución cognitiva (década de 1940)		
2	Nacimiento	Quiebre con el conductismo hasta el momento imperante con relación al abordaje de fenómenos psicológicos en los años cincuenta		
3	Consolidación	Establecimiento y consolidación de las ciencias cognitivas como programa multidisciplinario en los años sesenta y setenta		
4	Redescubrimiento del cerebro	Viraje que hacia fines del siglo XX dieron las ciencias cognitivas hacia el estudio del cerebro	Advenimiento del conexionismo en los años ochenta	El conexionismo consiste en el intento de estudiar la cognición a través del uso de redes neuronales artificiales. Hubo muchos intentos durante la posguerra (primeros años de la cibernética) conocidos como un “primer conexionismo”. Hay heterogeneidad dentro del programa conexionista, pero la segunda oleada de redes neuronales artificiales estuvo centrada en el modelado de habilidades cognitivas particulares. Entre las críticas, se plantea que las redes conexionistas traen inspiración sobre la arquitectura del cerebro “pero son extremadamente poco realistas respecto de la complejidad de las neuronas y redes neuronales biológicas”.

			Desarrollo y de la proliferación de las técnicas de neuroimagen funcional que redundaron en el florecimiento de las neurociencias cognitivas	Las neuroimágenes funcionales han sido el principal motor de un programa centrado inicialmente en la localización cerebral de operaciones y procesos cognitivos delimitados. Posteriormente la tendencia fue virando hacia los patrones de conectividad distribuidos a lo largo de grandes áreas del cerebro. Es aquí donde se puede ver más claramente el tipo de vínculo simbiótico que la psicología cognitiva y la neurociencia han tenido para la conformación de las neurociencias cognitivas.
--	--	--	--	--

Estos autores señalan que las ciencias cognitivas suelen caracterizarse como un proyecto inter o multidisciplinario. Pero, de las disciplinas involucradas, es dable pensar que algunas vayan adquiriendo mayor visibilidad y, por qué no, incluso ambicionar cierta hegemonía sobre las otras. Más aún, en sus versiones de divulgación, pueden pretender convertirse en modelo omniexplicativo de la realidad.

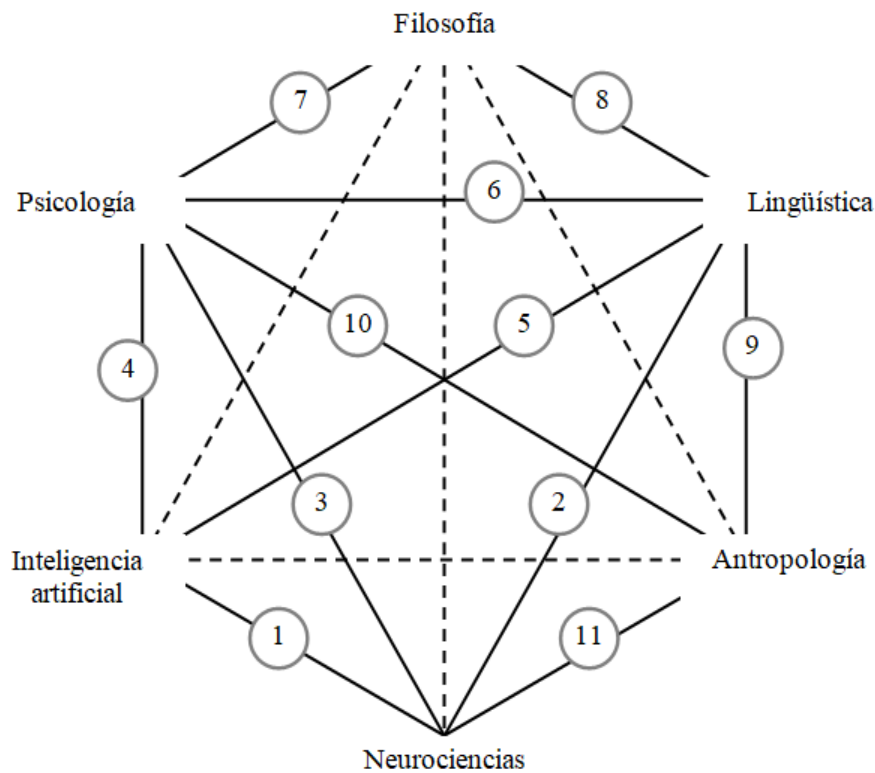
El desarrollo de las ciencias cognitivas marcó una ruptura con el conductismo, permitiendo el estudio de los procesos mentales internos. Sin embargo, a medida que la neurociencia cognitiva ha ido evolucionando, han surgido interrogantes sobre la compatibilidad entre modelos psicológicos y las evidencias obtenidas mediante neuroimagen. La tensión entre psicología y neurociencia es clave para entender los desafíos epistemológicos de la integración de estos enfoques.

4.1.2. El hexágono de Sloan

Dentro del período considerado como “de consolidación” y más específicamente en 1978, George Miller, Samuel Jay Keyser y Edward Walker elaboran un informe para la Alfred P. Sloan Foundation, que financiaba un programa para la Ciencia Cognitiva. En él se valían de una figura hexagonal para analizar los posibles vínculos entre disciplinas y es justamente “el patrón de interconexiones entre estas subdisciplinas lo que fundamenta la idea de una ciencia autónoma de la cognición” (Venturelli, Barberis, 2023, 13).

Figura 4-1.

El “hexágono de Sloan” o hexágono cognitivo.



Nota: tomado de Barberis, Wright, 2023.

Las once líneas continuas de esta figura representan aquellos dominios de investigación interdisciplinaria mejor establecidos hasta el momento de la presentación del informe:

- 1: cibernética
- 2: neurolingüística
- 3: neuropsicología
- 4: simulación de los procesos cognitivos
- 5: lingüística computacional
- 6: psicolingüística
- 7: filosofía de la psicología
- 8: filosofía del lenguaje
- 9: lingüística antropológica
- 10: antropología cognitiva
- 11: evolución del cerebro

Las cuatro líneas quebradas representaban dominios de interés para la ciencia cognitiva que hasta ese momento se encontraban menos desarrollados.

Obsérvese que, entre las líneas que representaban interconexiones débiles o *in status nascendi*, el informe incluye la filosofía de la neurociencia (o la neurociencia de la filosofía), que cuyo desarrollo se produjo sobre todo en las últimas décadas.

El “hexágono de Sloan” ilustró la interdependencia entre diferentes disciplinas dentro de las ciencias cognitivas, pero con el tiempo la neurociencia ha ganado predominancia, relegando en cierta medida a la psicología cognitiva. La cuestión central es si la psicología cognitiva puede mantenerse como un campo autónomo o si su integración con la neurociencia la llevará a una reformulación de sus constructos fundamentales. En las siguientes secciones, examinaré cómo la neuroimagen ha influido en esta transformación.

4.1.3. Caracterización de la psicología cognitiva

La psicología cognitiva ha sido definida como el análisis científico de las estructuras y los procesos mentales con el propósito de comprender la conducta humana (Mayer, 1985, en Meza, 2005).

Ahora bien, al intentar dar cuenta de sus características, encontramos que numerosos autores (Meza, 2005, Riviere, 1987, entre otros) señalan que no puede considerarse un cuerpo integrado y homogéneo de conocimientos:

[...] lo que más bien existe es un conjunto de aproximaciones a lo cognitivo, cada una con ciertos matices peculiares, pero compartiendo todas en menor o mayor medida ciertas características fundamentales (por ejemplo, el ser centralistas o interesados en procesos internos del sujeto, buscar la validez ecológica de sus propuestas). (Meza, 2005, p. 153)

Riviere (1987) llega incluso a considerar que el concepto de psicología cognitiva tiene un carácter difuso (p. 21).

Esta falta de unidad teórica en la psicología es un problema añejo que no ha sido resuelto por ninguna escuela o corriente psicológica (Bunge y Ardila, 1988) y difícilmente la psicología cognitiva conseguiría finiquitarlo, por su excesivo énfasis en el procesamiento de la información (Arias Gallegos, 2021, p. 190).

Ardila y Mario Bunge (2002, 285-286) proponen un ejemplo para dar cuenta del carácter fuertemente interdisciplinario de la psicología científica, que nos permitimos resumir y ampliar. Supongamos que alguien escribe en una pizarra una oración con contenido cognitivo:

- En tanto la oración porta un elemento de conocimiento, tanto su producción como su comprensión corresponden a la esfera de la psicología cognitiva.
- Dado que es una oración, su estudio atañe a la lingüística.
- El acto motor de escribir la oración se inscribe dentro del estudio de la conducta.
- La escritura implica percepciones visuales y táctiles que involucran a la psicología de la percepción.
- En tanto la oración puede servir para comunicar, también entra en el dominio de la psicología social.
- La motivación que llevó a escribirla podría ser interés de la neuroendocrinología.
- Eventuales errores o dificultades de escritura, de ortografía o de lectura involucrarían a la neurología, la psicopedagogía, la neurología.
- En la elaboración de la frase intervienen funciones como la memoria y la creatividad.
- Un especialista en psicodiagnóstico podría realizar inferencias a partir del tamaño, la forma, el estilo de la línea, la presión de los trazos.
- También la producción implica la puesta en marcha de una serie de procesos cerebrales (áreas de Wernicke y de Broca, sensorial, motora, corteza visual) que son de interés de las neurociencias cognitivas.

[...] en la psicología científica no hay ramas autónomas. La división de la psicología en diferentes campos (por ejemplo, de acuerdo con las diversas facultades clásicas) es extremadamente artificial, como también lo es la separación, hoy de moda, de la psicología cognitiva con respecto a las otras ramas de la ciencia de la conducta y de la mente. [...] En realidad, escribir o leer una oración es un proceso unitario con una cantidad de aspectos diferentes, aunque interconectados. La división no está en el proceso mismo, sino en los ojos de su espectador. El proceso de formar y escribir una oración es un proceso biosocial y, aunque es legítimo —y aun indispensable— distinguir sus distintos aspectos y enfatizarlos uno por vez, tales aspectos no debieran separarse unos de otros. En general, la voz de orden es: distinguir sin separar; unir sin confundir. (Ardila y Bunge, 2002, 285-286)

La psicología cognitiva no constituye un campo teóricamente homogéneo, sino que abarca múltiples enfoques con diferentes grados de formalización. Esto plantea un problema cuando se intenta integrar sus constructos con la neurociencia, ya que la falta de uniformidad dificulta su operacionalización en estudios con neuroimagen. En la siguiente sección, exploraremos hasta qué

punto los constructos psicológicos pueden considerarse estables y si la psicología puede generar modelos compatibles con los hallazgos neurocientíficos.

4.2. Sobre los constructos de la psicología cognitiva

El debate sobre la unidad de las ciencias mente-cerebro ha experimentado en los últimos años una fuerte revitalización, en gran medida debido al desarrollo de las técnicas de neuroimagen.

Señala Sullivan (2016, p. 662) que “las versiones anteriores del debate se centraban en si las teorías psicológicas podrían reducirse a teorías neurocientíficas”, mientras que en la actualidad, según esta autora, la polémica gira en torno a la pregunta acerca de si es posible una ciencia de cognición unificada que integre explicaciones psicológicas y neurocientíficas.

Aunque la psicología cognitiva y la neurociencia suelen considerarse empresas científicas distintas, Piccinini y Craver (2011) han argumentado que los dos campos no son explicativamente autónomos.

Los defensores de la explicación mecanicista (Piccinini y Craver, 2011) argumentan que la psicología cognitiva no es autónoma de la neurociencia porque las explicaciones de las capacidades cognitivas mediante el análisis funcional son simplemente explicaciones mecanicistas incompletas. (Sullivan, 2016, p. 662)

La posición de Piccinini y Craver pueden ser vista como situando las dos formas de explicación en diferentes puntos dentro de un *continuum* explicativo. Una vez que la neurociencia completa “los aspectos estructurales que faltan en un análisis funcional”, “se convierte en una explicación mecanicista más completa” (Piccinini y Craver, 2011, p. 308).

Argumenta Sullivan que la práctica experimental en psicología cognitiva y neurociencia no es propicia para el tipo de integración explicativa propuesta por Piccinini y Craver, dado que “la integración de los análisis funcionales y las explicaciones mecanicistas requiere que los componentes de los dos tipos de explicación, a saber, las capacidades cognitivas, sean estables” (Sullivan, 2016, p. 662).

La inestabilidad de los constructos de la psicología cognitiva representa un desafío para su integración con la neurociencia. Si los conceptos psicológicos varían según el paradigma o el enfoque teórico, su correspondencia con la actividad neuronal se vuelve problemática. Esto genera

dudas sobre la posibilidad de construir una ciencia de la cognición verdaderamente unificada, lo que constituirá el objetivo del análisis en las próximas secciones.

4.2.1.1. *¿Es posible una ciencia de la cognición unificada?*

En primer lugar, cabe señalar que los defensores de la unidad de la ciencia generalmente se valen de la **reducción teórica**. (E. Nagel, 1981)

Una reducción, en el sentido en el que empleamos aquí la palabra, es la explicación de una teoría o de un conjunto de leyes experimentales establecidas en un campo de investigación por otra teoría formulada habitualmente, aunque no invariablemente, para otro dominio. (Nagel, 1981, p. 445)

Y establece las siguientes condiciones formales:

... cuando las leyes de la ciencia secundaria contienen algún término “A” que está ausente de las suposiciones teóricas de la ciencia primaria, hay dos condiciones formales necesarias para reducir la primera a la segunda: (1) Deben introducirse suposiciones de algún tipo que postulen relaciones adecuadas entre lo significado por “A” y características indicadas en términos teóricos ya presentes en la ciencia primaria. Queda por examinar la naturaleza de tales suposiciones; pero, sin prejuzgar el resultado del ulterior examen, será conveniente llamar a esta condición “**condición de conectabilidad**”. (2) Con ayuda de estas suposiciones adicionales, todas las leyes de la ciencia secundaria, inclusive las que contienen el término “A”, deben ser lógicamente deducibles de las premisas teóricas y de las definiciones coordinadoras asociadas con ellas en la disciplina primaria. Llamemos a esta condición “**condición de deducibilidad**”. (E. Nagel, 1981, pp. 464-465)

La condición de conectabilidad implicaría que, para que pueda tener lugar una reducción interteórica exitosa, debe establecerse una “ley de puente” que permita que los referentes de los términos en la teoría a reducir puedan ponerse en una relación de equivalencia biunívoca con los referentes de los términos en la primera teoría. (Sullivan, 2016, p. 665).

Pero resulta prácticamente trivial pensar que, para que se cumpla este requisito de la conectabilidad, se requiere la **estabilidad de constructo**, sobre todo entre los términos propios de la psicología cognitiva.

Sullivan señala que hay motivos para pensar que los constructos de la psicología cognitiva no son estables.

1. En primer lugar, porque es poco común que las teorías formales de la psicología cognitiva sean usadas como guía en el diseño de tareas de investigación. Cita a Russell Poldrack cuando señala que “las comparaciones de tareas en muchos estudios se basan en juicios

intuitivos con respecto a los procesos cognitivos involucrados en una tarea en particular” (2010, p. 149). En otras palabras, el análisis de tareas, que es un componente de la explicación y de la evaluación del constructo, no es algo que ocurra actualmente entre laboratorios o investigadores de manera consistente y coordinada.

2. Otro factor que favorece la inestabilidad de los constructos cognitivos es el pluralismo metodológico. Al hacer un análisis comparativo de las tareas en los diferentes laboratorios de neurociencia cognitiva, encontramos “una multiplicidad de protocolos experimentales” (Sullivan, 2009).

[...] en la medida en que los investigadores a menudo no están de acuerdo sobre qué paradigmas experimentales deben usarse para investigar una capacidad cognitiva dada, y tienen libertad para diseñar las tareas que consideren más adecuadas a sus objetivos explicativos. (Sullivan, 2016, p. 670)

Como ya hemos visto, varios neurocientíficos reconocen la inestabilidad de los constructos de la psicología cognitiva. Price y Friston aspiran a una taxonomía más abstracta que pueda englobar las diversas tareas específicas. Poldrack, por su parte, invita a

[...] desarrollar tareas cognitivas más apropiadas para la localización funcional (2006) y participar en un análisis de tareas más riguroso (2010), aboga por el uso de técnicas de metaanálisis y minería de datos³⁸ como base para evaluar la fuerza de las hipótesis acerca de qué funciones están realizando áreas específicas del cerebro (2006). (Sullivan, 2016, p. 670)

4.2.1.2. *Pluralismo coordinado*

La inestabilidad de constructos a la que hice referencia da lugar a un “pluralismo descoordinado”. Para superarlo, Sullivan (2017) propone el “**pluralismo coordinado**”, que implicaría tanto la estandarización de definiciones y métodos como la colaboración interdisciplinaria.

Algunos de los aspectos más importantes de este enfoque son:

- Definiciones y métodos estandarizados: para superar la ambigüedad conceptual y la multiplicidad de métodos, el pluralismo coordinado se apoya en iniciativas como

³⁸ La minería de datos o *data mining* es la aplicación de algoritmos específicos para extraer patrones de los datos. Se puede consultar al respecto: Fayyad, U., Piatetsky-Shapiro, G., y Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37-37.

Cognitive Atlas, CogPo y Experiment Factory³⁹. Estas plataformas ofrecen recursos para crear definiciones más precisas y consensuadas de los constructos psicológicos, así como para estandarizar los protocolos experimentales.

- Fomento de la colaboración interdisciplinaria, a través de iniciativas como CNTRICS y RDoC⁴⁰.
- Énfasis en la validación del constructo, para garantizar que los paradigmas experimentales midan aquello que pretenden medir.
- Equilibrio entre estabilidad e innovación: se reconoce también la necesidad de flexibilidad para adaptarse a nuevos descubrimientos, por lo que se busca un equilibrio entre contar con una base sólida para la investigación y estar abiertos a la revisión de los constructos a medida que el conocimiento avanza.

4.2.2. Modelos explicativos

¿Qué significa **explicar** un proceso cognitivo a partir de la neuroimagen? Desde la filosofía analítica, la explicación científica no se limita a describir regularidades empíricas, sino que implica un proceso de ampliación y enriquecimiento del conocimiento, en el cual los datos deben integrarse en un marco teórico que permita comprender su estructura y su significado.

Díez y Moulines (2017) enfatizan que toda observación científica está mediada por teorías previas, lo que significa que interpretar una neuroimagen no es un acto neutral, sino un proceso en el que intervienen hipótesis sobre la función cognitiva, modelos neurocientíficos y supuestos metodológicos. En este sentido, explicar la cognición a partir de la resonancia magnética funcional (RMf) no es simplemente asociar áreas cerebrales activadas con tareas específicas, sino construir una estructura explicativa que relacione los patrones observados con principios teóricos que den cuenta de su función.

Sosteniendo esta misma perspectiva, Barberis, Roffé y Ginnobili (2022) han señalado que el uso de las neuroimágenes para explicar la cognición debe ir más allá de la mera correlación entre patrones de activación y tareas cognitivas. Así, retoman la propuesta de Díez (2014) al argumentar que una explicación adecuada en neurociencia debe ser un **proceso de enriquecimiento**

³⁹ Ver el Apéndice 1 para una breve descripción de estas iniciativas.

⁴⁰ Ver el Apéndice 1.

conceptual, en el cual los hallazgos de la RMf se articulen con marcos teóricos explícitos que permitan interpretar su significado.

En este marco, la relación entre la psicología cognitiva y la neurociencia ha sido históricamente compleja, oscilando entre enfoques reduccionistas, que buscan subsumir la primera en la segunda, y posturas que defienden su autonomía explicativa. En el fondo de esta disputa subyace el problema mente-cuerpo de la filosofía de la mente, que conlleva preguntas clave sobre las relaciones explicativas entre ambas disciplinas: ¿las neuroimágenes y la psicología cognitiva explican los mismos fenómenos desde enfoques distintos (por ejemplo, mediante explicaciones mecanicistas en contraste con explicaciones funcionales)? ¿O, en cambio, están abordando fenómenos de distinta naturaleza, lo que implicaría que la discusión no es solo epistemológica, sino también ontológica?

Branca y otros (2015) proponen tres modelos de explicación en neurociencias (cf. también Sullivan, 2009). Los presentaremos brevemente a continuación.

4.2.2.1. *Explicación funcional*

En psicología cognitiva, desde hace un largo tiempo —e incluso hasta la fecha— predominan explicaciones de tipo funcionalistas a la hora de dar cuenta de capacidades psicológicas, tales como atención, memoria o percepción. Este tipo de explicaciones, de tipo psicológicas, se proponen de manera independiente respecto a las explicaciones de tipo neurocientíficas.

Un tipo de explicación clásica dentro del marco funcionalista es el análisis funcional propuesto por Cummins (1983, 2000). Dicho análisis puede definirse como el análisis de una capacidad en términos de las propiedades funcionales del sistema y de su organización. Este autor afirma que los fenómenos psicológicos/cognitivos deben entenderse en términos de *propiedades disposicionales complejas* (capacidades) y que dichas propiedades, por lo general, deben explicarse mediante el análisis tanto de la capacidad misma como del sistema que la posee (especificando las partes del mismo). En el caso de que la propiedad analizada sea una capacidad, entonces el análisis se denominará funcional. (Branca, 2015, p. 179)

Las explicaciones y los modelos funcionales representan la consecución lógica, a nivel explicativo científico, de aquellos postulados funcionalistas y autonomistas apoyados en la realizabilidad múltiple.

4.2.2.2. Reduccionismo “implacable”

Bickle (2003, 2006) propuso el denominado reduccionismo *ruthless*, un modelo reduccionista que propone reducir la psicología a la neurociencia de nivel molecular.

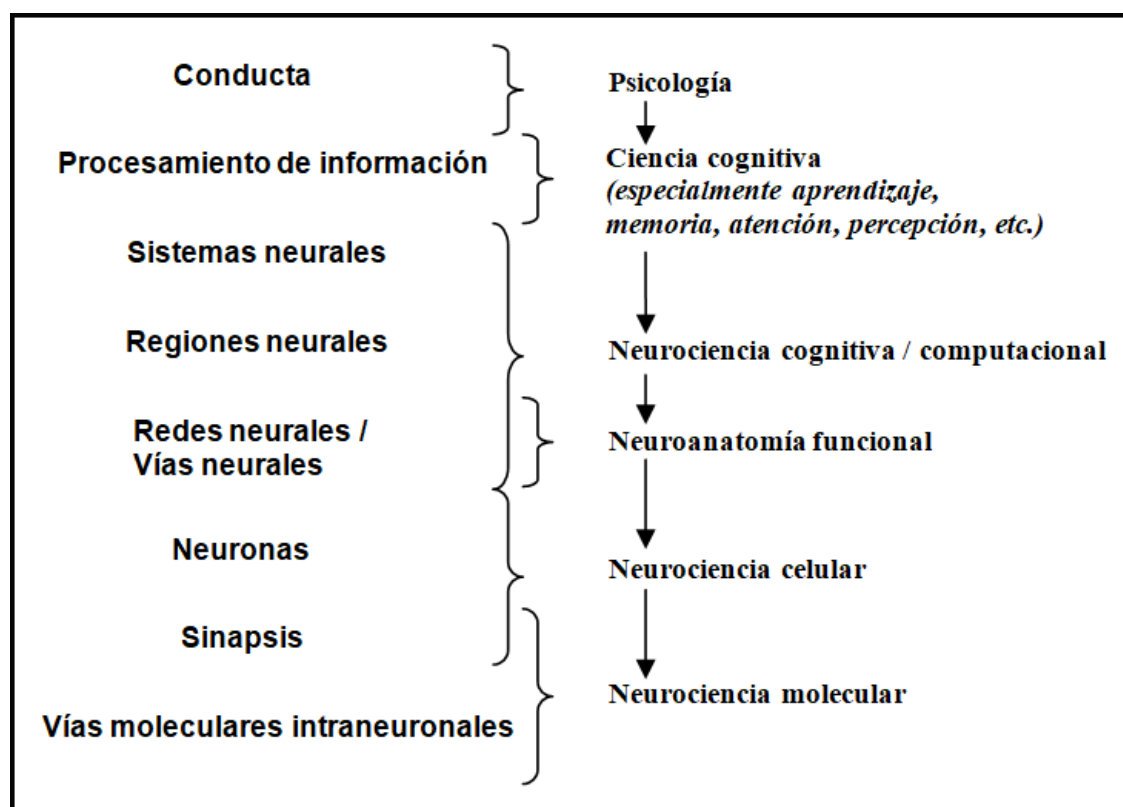
Su presunto éxito se basa en demostrar cómo las intervenciones en el nivel bajo (celular/molecular) producen cambios causales y significativos en el nivel conductual sin necesidad de estudiar otros niveles explicativos.

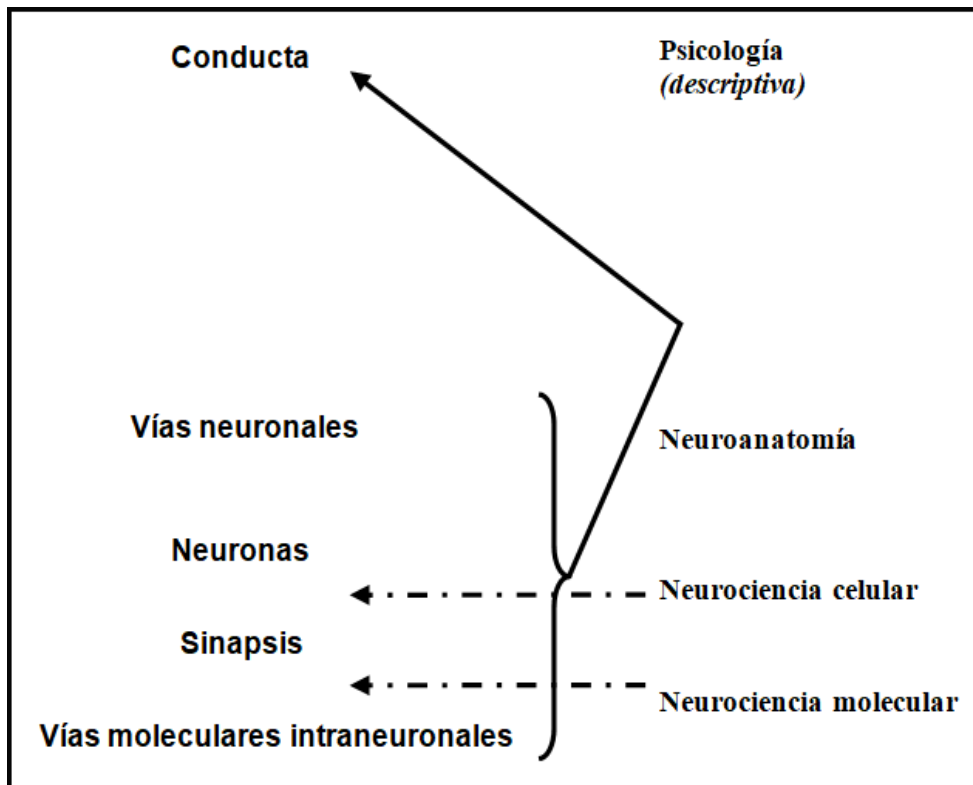
De esta manera, se sostiene que la neurociencia y la psicología cognitivas solo poseerán un valor heurístico, puesto que no serían disciplinas necesarias para poder explicar cambios causales de la conducta.

Figura 4-2.

Figura 4-3.

Traducción de dos esquemas presentados por Bickle para explicar su reduccionismo “ruthless”. La primera figura muestra el estándar de organización neural, con una vista de reducción “paso a paso” (flechas hacia abajo). La segunda imagen representa el reduccionismo implacable. Las flechas discontinuas representan niveles de intervención experimental; la flecha sólida, el nivel en el que se miden estas intervenciones. Según esta formulación, la psicología “es un esfuerzo descriptivo que genera datos conductuales, en lugar de explicaciones”.





Nota: traducidos de Bickle, 2006.

4.2.2.3. *Mecanicismo constitutivo*

En la explicación mecanicista constitutiva, el comportamiento o propiedades de un sistema son explicados por las actividades y propiedades de sus partes (Craver, 2006). Las totalidades constitutivas —la suma de los mecanismos— se encuentran conformadas en cada nivel mecanicista, por mecanismos causales-etiológicos que se relacionan componencialmente con mecanismos de nivel superior. Los niveles bajos no *causan* a los niveles altos; los niveles altos tampoco *causan* a los bajos, sino que están *compuestos* por estos. A la apariencia causal que tienen las relaciones internivel, se les dará el nombre de “*efectos mecanísticamente mediados*” (Craver y Bechtel, 2006), es decir: de aparente causalidad internivel debida a una relación componencial.

El mecanicismo constitutivo sostiene que, en la práctica actual de neurociencias, los científicos apuntan a establecer puentes entre niveles antes que a reducirlos, y que para dar cuenta efectiva de ciertos fenómenos, deben tenerse en cuenta (como efectivamente lo hacen numerosas líneas de investigación) aspectos relativos a la organización sistémica global de los mecanismos involucrados. (Branca, 2015, p. 183)

Por su parte, Sergio Barberis (2012) advierte que, si bien Piccinini y Craver rechazan un determinismo estricto al admitir que los mecanismos pueden ser probabilísticos, esta postura genera tensiones dentro del mecanicismo, ya que si los niveles inferiores no determinan

completamente a los superiores, entonces la relación explicativa entre niveles sigue sin estar bien definida.

4.2.3. Un cerebro pluripotencial

Por su parte, Colin Klein señala:

Diferentes partes del cerebro hacen diferentes contribuciones a la cognición. Sin embargo, ninguna parte del cerebro funciona sola: una tarea cognitiva siempre es llevada a cabo por una red de regiones cerebrales que trabajan en concierto. Ambas afirmaciones no deberían ser destacables, pero cuando se trata de neuroimágenes, estas perogrulladas han recibido énfasis marcadamente diferentes. Los análisis orientados a redes son cada vez más comunes (Pessoa 2008; Ramsey et al. 2009). Sin embargo, durante aproximadamente la primera década y media de la neuroimagen (NI), la atención se centró en la función de las regiones del cerebro consideradas de forma aislada del contexto neuronal más amplio. (2012, p. 952)

Propone entonces una analogía: muchos sistemas operativos incluyen una subrutina para hacer transformadas rápidas de Fourier⁴¹ (S). La función de S dependerá del contexto: en compresión de imágenes, servirá para “identificar componentes de alta frecuencia que se pueden eliminar sin

⁴¹ La **transformada discreta de Fourier** o DFT (del inglés *discrete Fourier transform*) descompone una secuencia de datos discretos en su representación en términos de diferentes frecuencias. En otras palabras, podríamos decir que transforma una función matemática en el dominio del tiempo en otra, que permite una representación en el dominio de la frecuencia. Por su parte, la **transformada rápida de Fourier** o FFT (*fast Fourier transform*) es un conjunto de algoritmos para calcular más rápidamente la DFT.

Los algoritmos FFT se utilizan tan habitualmente para calcular la DFT que coloquialmente la primera sigla se utiliza a menudo en lugar de la segunda, si bien existe una diferencia formal: la DFT es una función matemática, mientras que la FFT es un algoritmo o familia de algoritmos.

Los datos obtenidos a través de la iRMf constituyen series temporales que representan la actividad en diferentes regiones cerebrales. Para analizarlas y extraer información (frecuencias dominantes, patrones de actividad) la DFT es una herramienta clave del proceso, dentro de las técnicas de análisis de frecuencia.

Para profundizar el tema de la DFT y la FFT, puede consultarse: Spilbury, M. J. y Euceda, A. (2016). Transformada Rápida de Fourier. *Revista De La Escuela De Física*, 4(2), 45–52. <https://doi.org/10.5377/ref.v4i2.8276>. También: De la Fraga, L. (2001). La transformada discreta de Fourier y la transformada rápida de Fourier. Recuperado de [https://delta.cs.cinvestav.mx/~fraga/Cursos/PDI/tdf.pdf].

Para comprender el uso de las DFT y FFT en la resonancia magnética funcional, ver: Rowe, D. B., Nencka, A. S. y Hoffmann, R. G. (2007). Signal and noise of Fourier reconstructed fMRI data. *Journal of Neuroscience Methods*, 159(2), 361-369.

pérdida de fidelidad”; en los programas de audio, se usará para separar canales de audio.

La subrutina S será pluripotencial dependiendo del nivel de abstracción.

Señala entonces que, para Price y Friston las ontologías cognitivas son deficientes, en primer lugar, porque la meta sería encontrar, dada una región, la función que explicara “todos los patrones de activación” (Klein, 2012, p. 954). En segundo lugar, porque incluyen muchas etiquetas funcionales específicas que pueden resultar engañosas (como “identificación visual de formas de palabras”). Ahora bien, Klein argumenta que, en varias ocasiones, las descripciones funcionales abstractas pueden resultar vagas y no ser demasiado reveladoras.

Por ello, considera necesario repensar las categorías cognitivas, pero sin descartar las atribuciones más específicas: “Debido a que las atribuciones específicas son sensibles al contexto, debemos usarlas con cuidado: solo estamos justificados para apelar a ellas si estamos seguros de que estamos en el mismo contexto que la atribución original” (956). Y llama a esto “inferencia inversa sensible al contexto”.

[...] es cierto que las regiones del cerebro están funcionalmente especializadas. Pero el enfoque en la función de las regiones del cerebro ha sido, sugiero, prematuro. Primero, debemos observar más de cerca las redes cerebrales y descubrir con qué tareas están asociadas. Eso tiene el potencial de mostrar si nuestras categorizaciones existentes de tareas cognitivas son incorrectas y, de ser así, dónde debemos revisar. Solo una vez que hayamos hecho esto podremos volver con confianza a las regiones individuales del cerebro y descubrir cuáles podrían ser sus contribuciones distintivas a la cognición. (Klein, 2012, p. 959)

Los diferentes modelos explicativos de la relación mente-cerebro han dado lugar a enfoques contrastantes en la neurociencia cognitiva. Mientras que el funcionalismo mantiene cierta autonomía para la psicología cognitiva, el reduccionismo ‘implacable’ propone eliminarla en favor de explicaciones puramente neurocientíficas. En un punto intermedio, el mecanicismo constitutivo sugiere que la psicología cognitiva puede ser integrada en una visión más amplia de los mecanismos cerebrales. En la siguiente sección, analizaremos cómo estos modelos afectan la construcción de una ontología cognitiva coherente.

4.3. Ontología cognitiva

“If neuroimaging is the answer, what is the question?”

(Steven Kosslyn)

Toda teoría científica implica un conjunto de compromisos ontológicos, aun cuando estos no sean explícitamente reconocidos. Como sostiene Daniel Dennett (1991, p. 36), “La ontología de una teoría es el catálogo de cosas y tipos de cosas que la teoría considera que existen”, lo que implica que cualquier empresa científica presupone una determinada concepción acerca de qué entidades forman parte de su marco explicativo y cómo se articulan entre sí.

En este sentido, la ciencia no opera en un vacío filosófico, sino que está atravesada por supuestos ontológicos que, de no ser examinados, se trasladan sin cuestionamiento al desarrollo teórico. En *Darwin's Dangerous Idea* (1995, p. 21), Dennett advierte: “No existe tal cosa como una ciencia libre de filosofía; solo hay ciencia cuyo bagaje filosófico ha sido transferido al otro lado de la frontera sin ser examinado”. Entonces, la tarea del filósofo no es imponer una ontología ajena a la ciencia, sino analizar los presupuestos ontológicos ya presentes en ella.

De esta manera, las ciencias cognitivas y sus modelos teóricos implican compromisos sobre qué entidades existen en la mente y el cerebro y cómo deben conceptualizarse. La investigación no solo describe procesos: los construye dentro de marcos ontológicos específicos, para determinar qué se considera una explicación válida del fenómeno cognitivo.

Ahora bien, Poldrack señaló en una ocasión que “el problema fundamental es nuestra ontología psicológica de la Edad de Piedra” (Bunzl, Hanson y Poldrack, 2010, p. 54). Al respecto, Daniel C. Burnston (2022, p. 259) sugiere que las categorías mentales usadas de manera estándar en las ciencias psicológicas son demasiado simples, generales y rudimentarias para captar cómo el cerebro implementa la conducta, y que tal vez “esas categorías necesiten ser revisadas, refinadas o incluso abandonadas para comprender la función cerebral”.

4.3.1. Una expresión, más de tres significados

La expresión “ontología cognitiva” fue introducida al debate contemporáneo por Cathy Price y Karl Friston (2005). En este artículo señalan: “Lo que necesitamos es una definición sistemática de las relaciones estructura-función, de modo que las estructuras predigan las funciones y las funciones predigan las estructuras”. Y agregan: “Estas descripciones sistemáticas se denominan ontologías” (p. 262).

De acuerdo con esta formulación y siguiendo a Marco Viola (2016), una ontología completa debería constar de:

- a. Una ontología F de las funciones cognitivas que describa todas las operaciones mentales ($f_1, f_2 \dots f_n$);
- b. Una ontología S de las estructuras neurales S que describa todas las partes del cerebro ($s_1, s_2 \dots s_n$);
- c. Una correspondencia biunívoca entre cada elemento de F y cada elemento de S ($f_1 \leftrightarrow s_1, f_2 \leftrightarrow s_2 \dots f_n \leftrightarrow s_n$).

Obsérvese que esta representa la versión “fuerte” de una ontología completa aplicada al campo de la cognición, en tanto implica una correspondencia biunívoca, es decir, uno a uno, entre los elementos estructurales y los funcionales.

Pero la expresión “ontología cognitiva” puede admitir definiciones menos restrictivas y así cobrar diferentes significados.

1. Una ontología cognitiva podría ser una **nomenclatura**, es decir, “un conjunto de términos estandarizados que los investigadores pretenden utilizar de manera sistemática para promover la comprensión mutua” (Janssen et al., 2017, p. 124). Proyectos como el *Cognitive Atlas* (Poldrack et al., 2011) se inscriben en la intención de capturar el estado de la neuropsicología cognitiva en un marco estandarizado como el de esta primera acepción.
2. Una ontología cognitiva podría referirse a un dominio, a un conjunto de entidades a las que se refiere una teoría cognitiva.

Esto es “ontología” en el sentido familiar para los filósofos de la ciencia (aunque ampliado más allá del sentido quineano tradicional para incluir propiedades y agrupaciones taxonómicas). Los debates sobre si la memoria de trabajo está unificada o es específica de la modalidad, o si la esquizofrenia es realmente una categoría taxonómica apropiada, o si la conversión de ortografía a sonido sigue una ruta única o doble, pertenecen a esta categoría. Aunque Price y Friston mencionan el sentido informático de “ontología”, su preocupación es principalmente con el sentido del dominio. Argumentan que postular nuevos términos (y asegurar su uso estandarizado) no es suficiente para dar sentido a los hallazgos de superposición en distintos estudios de imágenes de resonancia magnética funcional. En lugar de una nueva terminología para los procesos existentes, lo que se requiere es **postular nuevos procesos psicológicos, nuevas formas de dividir o estructurar los sistemas cognitivos**. (Janssen et al., 2017, p. 124. El resaltado es nuestro.)

3. En tercer lugar, una ontología cognitiva podría ser un conjunto de categorías metafísicas básicas.

Entonces, por ejemplo, podríamos debatir si una buena teoría cognitiva debe referirse a mecanismos, procesos dinámicos [...] o conceptos clínicos. Además, el dominio cognitivo suele estar estratificado en niveles. Una ontología cognitiva en este tercer sentido debería indicar si la relación entre niveles es de composición, constitución u otra cosa. Estas preguntas, a su vez, afectan la forma en que cree que deberían desarrollarse y verificarse las teorías cognitivas. (Janssen et al., 2017, p. 124)

Estas maneras diferentes de comprender la ontología cognitiva se hacen notar también en el campo de la filosofía de la psiquiatría, por ejemplo, en relación con el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM), que procura “proporcionar una nomenclatura uniforme para los profesionales de la salud mental” (Janssen et al., 2017, p. 125). ¿Deberían eliminarse las categorías psiquiátricas en favor de “categorías neurales tratables”? ¿O podrían coexistir diferentes taxonomías? Aquí se impone rescatar la posición de Sullivan, quien propone que la construcción de conocimiento se puede realizar a través de un “pluralismo coordinado”.

4.3.2. Alcance y niveles de análisis

Janssen, Klein y Slors se preguntan entonces: “¿Cuál es la unidad de análisis adecuada a la hora de reflexionar sobre cuestiones de ontología cognitiva?” Y aquí realizan una interesante observación:

Los debates tradicionales sobre la **mente extendida** obviamente entran en esta categoría. Estos han girado durante mucho tiempo sobre cuestiones de cómo delimitar lo mental (Rupert, 2009; Menary, 2010). Pero mientras que los debates anteriores eran relativamente abstractos, el trabajo de neuroimagen brinda ejemplos concretos. Hutto et al., por ejemplo, argumentan que la verdadera lección que se puede extraer del trabajo de neuroimagen es que los cerebros son “**proteicos**”. Es decir, varias regiones neuronales cambian su función sobre la marcha para adaptarse a las circunstancias (ver Anderson, 2014 para una visión similar). Como tal, los cerebros por sí solos no son la unidad de análisis correcta; también deberíamos incluir el mundo que codetermina la funcionalidad del cerebro. (Janssen, 2017, p. 125).

No es obvio que haya un solo nivel cognitivo. Entonces, ¿podemos aceptar más de una ontología para lo mental? ¿Podemos aceptar algún tipo de realismo mental? ¿O los postulados de nuestras mejores teorías psicológicas deberían ser tratados solo de manera instrumental?

En esta línea, Muhammed Ali Khalidi (2023) plantea que la psicología cognitiva desarrolló categorías taxonómicas basadas en la descripción funcional de procesos como la memoria, la atención o la toma de decisiones, mientras que la neurociencia busca los correlatos neuronales de estos procesos.

A diferencia de la posición de Piccinini y Craver, entre otros autores que sostienen que neurociencia y psicología cognitiva no son disciplinas completamente autónomas, Khalidi señala que muchas categorías cognitivas no pueden correlacionarse de manera directa con estructuras neuronales discretas, porque suelen estar definidas por su papel funcional y su relación con el entorno. Esto generaría un mapeo “muchos a muchos” entre las categorías neuronales y las cognitivas, lo que impediría una reducción simple de la psicología a la neurociencia.

La ontología cognitiva, entonces, tendría el papel de establecer un marco conceptual en el cual ambas disciplinas puedan dialogar sin forzar correspondencias arbitrarias⁴². Mientras la neurociencia aporta información sobre las estructuras y los procesos cerebrales, la psicología cognitiva proporciona las categorías y conceptos iniciales relativos a la cognición. La ontología cognitiva revisa estas categorías en busca de las “clases reales” (*real kinds*) de la cognición; la reducción solo sería válida cuando se demuestra que los elementos de la taxonomía carecen de base empírica sólida o son vestigios de lo que Khalidi llama “psicología popular”.

Si las categorías ontológicas que estructuran el estudio de la cognición son formuladas únicamente desde la neurociencia, el hexágono cognitivo de Sloan corre el riesgo de transformarse en un pentágono, eliminando a la psicología cognitiva o subordinándola a modelos exclusivamente neurobiológicos. Este proceso no solo empobrecería la interacción interdisciplinaria dentro de las ciencias cognitivas, sino que también podría generar un sesgo en la comprensión de fenómenos

⁴² En su propuesta, Khalidi enumera cuatro categorías ontológicas de lo mental: individuos (u objetos), estados, procesos y capacidades. Los **individuos** serían entidades cognitivas discretas, por ejemplo, un recuerdo; los **estados** constituirían configuraciones temporales del sistema cognitivo, como estar ansioso; los **procesos** serían transformaciones dinámicas (reconocer a una persona o realizar una inferencia); las **capacidades** estarían dadas por disposiciones generales del sistema cognitivo, tales como la memoria episódica o el lenguaje. A manera de ejemplo: el recuerdo individual de un encuentro con amigos implicaría poner en marcha un proceso de recuperación de información que, a su vez, dependería de la memoria episódica en tanto capacidad; finalmente, la evocación podría provocar un estado emocional de alegría o de nostalgia.

como la conciencia, la toma de decisiones y las emociones, reduciéndolos a sus correlatos neurales y dejando de lado niveles de análisis más amplios. Este escenario es consistente con una **perspectiva reduccionista fuerte**, en la que los niveles explicativos superiores son considerados meras descripciones parciales de fenómenos que, en última instancia, deben ser explicados en términos de sus mecanismos neurobiológicos subyacentes. Leemos, por ejemplo, en Bickle (2003, p. 130):

¿Qué ocurre con la reducción de la ciencia cognitiva a la neurociencia celular frente a la supuesta "autonomía" de las metodologías, explicaciones y teorías de nivel superior?

[...] Para nosotros, las teorías y métodos de nivel superior cumplen un papel *heurístico* útil y aparentemente ineliminable en la búsqueda de mecanismos y reducciones de nivel inferior. Aumentan considerablemente la base descriptiva para la experimentación y la teorización en niveles inferiores. Nos indican en qué parte del cerebro buscar y nos orientan en la construcción de tareas conductuales que aislen las variables dependientes e independientes cruciales. *Pero eso es todo lo que hacen, y todo lo que pueden hacer.*

Una vez que agotan esa función descriptiva y metodológica, desaparecen, como la escalera de Wittgenstein.

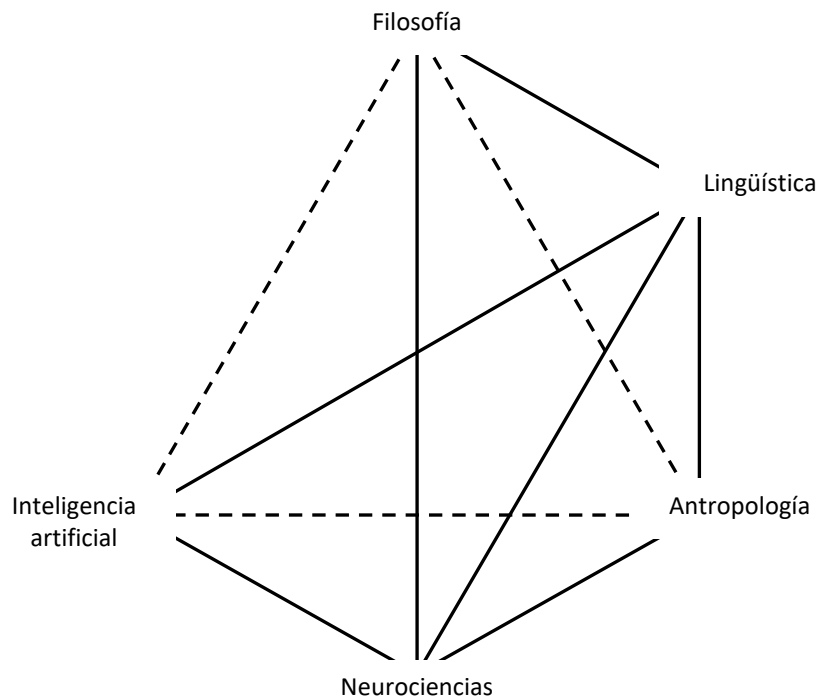
En una nota al final del capítulo, Bickle (2003, p. 160) explica a los no filósofos esta última referencia:

Ludwig Wittgenstein concluyó su primera obra importante, el *Tractatus Logico-Philosophicus*, con la siguiente observación: “Mis proposiciones sirven como elucidaciones de la siguiente manera: quien me entienda acabará por reconocerlas como carentes de sentido, después de que las ha usado —como peldaños— para subir más allá de ellas. (Debe, por así decirlo, tirar la escalera después de haber subido por ella)” ([1919] 1961).

Y agrega: “Cabe señalar que mi referencia a Wittgenstein es metafórica. Como debería quedar claro a partir de la discusión en el texto, no estoy acusando literalmente a los neurocientíficos cognitivos de decir disparates.” *And yet, and yet...* Como solían decir los antiguos juristas romanos: *Excusatio non petita, accusatio manifesta*.

Figura 4-4.

El hexágono de Sloan “empobrecido” por la pérdida de la psicología cognitiva.



Nota: elaboración propia.

Si la ontología cognitiva, entonces, busca redefinir las categorías de la cognición en términos de mecanismos neurobiológicos, podría pretender subsumir a la psicología cognitiva. Los procesos mentales que tradicionalmente han sido descritos en términos de representaciones, computaciones u operaciones funcionales serían reconceptualizados en términos de actividades de redes neuronales específicas, desplazando así la necesidad de explicaciones psicológicas autónomas.

Dicho de otro modo: si la ontología cognitiva se formula bajo el supuesto de que toda función cognitiva tiene una contrapartida neuronal estricta, podría considerarse que la psicología cognitiva es solo una etapa transitoria en la historia de las ciencias cognitivas, destinada a ser absorbida completamente por la neurociencia.

Para que esto no suceda:

- La ontología cognitiva debe evitar caer en el reduccionismo eliminativo, es decir, la idea de que las explicaciones psicológicas pueden ser completamente reemplazadas por explicaciones neurocientíficas. Neurociencia y psicología cognitiva deben coexistir en un esquema de niveles explicativos diferenciados.
- Si la ontología cognitiva se basa en un modelo mecanicista como el propuesto por Craver y Piccinini, no puede limitarse a la suposición de que solo los mecanismos neuronales son explicativos.

- La psicología cognitiva se ha valido históricamente de distintos marcos explicativos. Una ontología cognitiva no reduccionista debe reconocer la legitimidad de explicaciones no mecanicistas (Barberis, 2012).

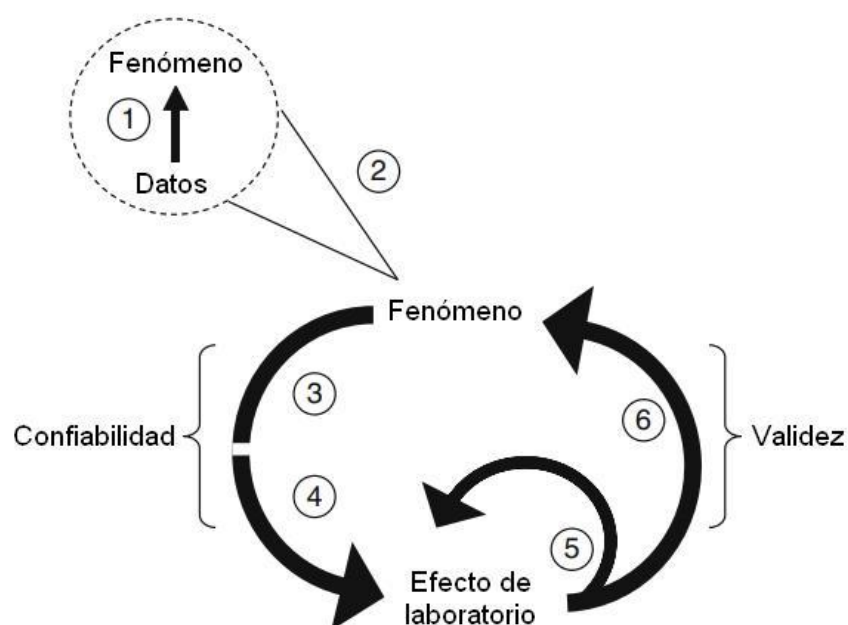
El debate sobre la ontología cognitiva pone en evidencia que no existe un consenso claro sobre cómo definir los procesos mentales en términos de actividad neuronal. Mientras algunos investigadores sostienen que la neurociencia debería reformular los constructos psicológicos, otros argumentan que estos constructos tienen valor explicativo propio y no pueden reducirse sin perder información relevante. En este contexto, surge la pregunta de si la neurociencia cognitiva puede prescindir de la psicología o si ambas disciplinas deben mantener un diálogo constante.

4.4. El diseño experimental

En capítulos anteriores ya hemos abordado las características técnicas de una investigación que involucre las neuroimágenes funcionales. Jaqueline Sullivan (2009, 2015) representa, a través del siguiente gráfico, las características del proceso experimental, para luego referirse a las condiciones de confiabilidad y validez.

Figura 4-5.

El proceso de investigación según J. Sullivan.



Nota: traducido de Sullivan, 2015.

El proceso experimental según Sullivan comienza, de acuerdo con este gráfico, cuando un investigador se formula una pregunta empírica sobre un fenómeno de interés (1): por ejemplo, “¿La región X del cerebro se encuentra involucrada en la conducta Y”? Este interrogante se redefinirá para procurar producir un determinado efecto en el laboratorio experimental (2). El fenómeno deberá ser definido operacionalmente. De esta manera, comienza el proceso de diseño (3) y de implementación de la producción de datos (4). Los paradigmas experimentales generalmente se basan en estímulos presentados a través de una computadora, dado que los resonadores magnéticos imponen, por su estructura, ciertas limitaciones físicas. Los sujetos usualmente responden presionando un botón y, durante esta tarea, el investigador escanea el cerebro centrándose en una o varias ROI.

Los paradigmas experimentales incluyen, generalmente, procedimientos de producción, de medición y de detección, empleados para especificar los estímulos, las variables de respuesta y las medidas comparativas necesarias para atribuir una función cognitiva, ubicar la función en una región cerebral específica o identificar cambios plásticos en una sinapsis o conjunto de sinapsis. (Sullivan, 2015, p. 37).

Los datos son luego procesados para eliminar artefactos⁴³ experimentales y reducir el ruido; a los resultados se les aplican luego técnicas de análisis estadístico. Así, de esta manera se busca “discriminar entre hipótesis funcionales en competencia sobre las áreas del cerebro bajo investigación” (Sullivan, 2015, p. 32).

La hipótesis discriminada se trata entonces como una afirmación y se tomará como verdadera o falsa tanto para el efecto producido en el laboratorio (5) como para el fenómeno original de interés en el mundo (6).

Ahora bien, dentro de un diseño experimental se deben tener siempre presentes las condiciones de confiabilidad y validez.

4.4.1. Confiabilidad y validez

La confiabilidad es la piedra angular de cualquier empresa científica. Los problemas de validez y significatividad de la investigación carecen relativamente de sentido si los resultados de nuestros experimentos no son dignos de confianza. Es cierto que la confiabilidad puede variar mucho según las herramientas que se utilicen y lo que se mida.

⁴³ Se conoce como “artefacto experimental” a cualquier efecto no deseado o error en los datos de un experimento que no está relacionado con la variable que se está estudiando.

Por lo tanto, es imperativo que cualquier esfuerzo científico sea consciente de la fiabilidad de sus mediciones.

Sorprendentemente, la mayoría de los investigadores de resonancia magnética funcional (iRMf) solo tienen una vaga idea de cuán confiables son sus resultados. (Bennett y Miller, 2010, p. 133)

Como ya se señaló, los investigadores en neurociencias buscan, en su diseño experimental, producir los datos necesarios para discriminar entre afirmaciones contrapuestas referidas a un único fenómeno de interés. Para ello, la evidencia tiene que ser adecuada a este propósito.

Esta adecuación está asociada a dos conceptos clave: la confiabilidad y la validez. Tradicionalmente, la confiabilidad de un experimento está asociada a su reproducibilidad; la validez, al nivel de certeza de que dicho experimento mide lo que dice medir. Pero, en la práctica, resultan conceptos mucho más complejos.

De acuerdo con Sullivan, la confiabilidad puede asociarse al “criterio de severidad” de Deborah Mayo⁴⁴:

Para que la prueba de una hipótesis sea confiable, debe pasar una prueba severa: debe ser altamente probable que los datos que surjan de la prueba de una hipótesis no proporcionen evidencia en apoyo de esa hipótesis si esa hipótesis fuera de hecho falsa. (Sullivan, 2015, p.40).

En otros términos, el "criterio de severidad" establece que una prueba de hipótesis es confiable si es capaz de distinguir claramente entre la hipótesis que se está probando y otras hipótesis alternativas. Esto implicaría que la prueba estadística debe ser capaz de identificar de manera

⁴⁴ El “criterio de severidad” de Deborah Mayo es una herramienta estadística que va más allá de la significancia estadística. Se utiliza principalmente para evaluar la solidez de la evidencia cuando una prueba estadística no logra rechazar una hipótesis nula. Se centra en la magnitud de la diferencia observada en un estudio, en lugar de simplemente determinar si un resultado es estadísticamente significativo o no.

En lugar de solo decir "no se encontró ningún efecto" cuando una prueba no rechaza la hipótesis nula, el criterio de severidad pregunta: "¿Cuán probable sería obtener una diferencia tan grande (o mayor) como la observada, si la hipótesis alternativa (que sí hay un efecto o riesgo) fuera cierta?". Esta probabilidad calculada es la severidad de la prueba contra la hipótesis alternativa.

Para ampliar:

Mayo, D. G., y Spanos, A. (Eds.). (2010). *Error and inference: Recent exchanges on experimental reasoning, reliability, and the objectivity and rationality of science*. Cambridge University Press.

Mayo, D. G. (2018). *Statistical inference as severe testing: How to get beyond the statistics wars* (Reprint ed.). Cambridge University Press.

efectiva las situaciones en las que la hipótesis nula es falsa. Si la hipótesis nula es incorrecta, la prueba debe ser lo suficientemente rigurosa como para encontrar evidencia que sugiera que la hipótesis nula es falsa.

De manera similar, el proceso de producción de datos puede considerarse confiable “si y solo si da como resultado datos analizados estadísticamente que pueden usarse para discriminar una hipótesis de un conjunto de hipótesis en competencia sobre un efecto producido en el laboratorio” (Sullivan, 2015, p. 40).

Ahora bien, una segunda característica deseable en el proceso experimental es la **validez**, de ella, científicos y filósofos distinguen dos dimensiones: la **validez externa** y la **validez interna**.

Los investigadores no solo desean que las conclusiones de sus resultados se apliquen a los efectos que se estudian en el laboratorio (validez interna), sino que también esperan que se apliquen a los fenómenos de interés a los que se dirigieron originalmente sus preguntas empíricas (validez externa).

Tomemos la explicación de Francesco Guala (2003):

La validez interna se logra cuando el experimentador comprende correctamente la estructura y el comportamiento de un sistema de laboratorio (sus principales factores causales, las formas en que interactúan y los fenómenos que provocan). Por ejemplo: el resultado de un experimento E es internamente válido si el experimentador atribuye la producción de un efecto B a un factor (o conjunto de factores) A, y A es realmente la (o una) causa de B en E. Además, es externamente válida si A causa B no sólo en E, sino también en un conjunto de otras circunstancias de interés, F, G, H, etc.

Los problemas de validez interna son antecedentes cronológicos y epistémicos de los problemas de validez externa: no tiene mucho sentido preguntarse si un resultado es válido fuera de las circunstancias experimentales a menos que estemos seguros de que lo es en ellas. (Guala, 2003, p. 1198)

La confiabilidad y la validez son, ambas, restricciones normativas en el proceso experimental. Pero mientras la confiabilidad busca “simplificar” el proceso en el contexto del laboratorio, restringe el proceso de producción de datos y “la extensión de las pretensiones interpretativas”, la validez se orienta en la dirección opuesta, dado que prescribe que en un diseño experimental se incorporen las dimensiones de complejidad que acompañan al fenómeno de interés en el mundo que se investiga.

Adherirse a las prescripciones normativas de validez conducirá inevitablemente a una disminución de la simplicidad del efecto producido en el laboratorio y a una expansión del

conjunto de hipótesis en competencia que pertenecen a ese efecto. En otras palabras, conducirá a una disminución de la confiabilidad. Sin embargo, sin confiabilidad, nada se gana, ya que, si se pierde el control en el laboratorio, no se puede decir nada verdadero sobre el efecto producido allí: se perderá también la validez interna. (Sullivan, 2015, p. 41).

En este complejo escenario de restricciones contradictorias, también se pueden mencionar otros tipos de validez que operan en el mismo sentido que la validez interna/externa.

4.4.2. Validez ecológica

La primera restricción es la **validez ecológica**. Señala Sullivan que, a diferencia de la validez externa, que se ocupa de si una afirmación interpretativa a la que se llega en un estudio determinado puede extenderse al mundo real, *la validez ecológica se ocupa de si el contexto, los estímulos empleados y las respuestas provocadas en el contexto experimental son similares a los que se encontrarían en el mundo*. Por ejemplo, realizar una tarea cognitiva en un escáner iRMf es diferente a participar en una actividad cognitiva en un entorno menos restringido, por lo que podríamos decir que **los experimentos que utilizan iRMf no son ecológicamente válidos**.

Según Schmuckler (2001), la definición clásica de validez ecológica propuesta por Bronfenbrenner (1977, 1979) establece que esta:

“se refiere a la medida en que el ambiente experimentado por los sujetos en una investigación científica tiene las propiedades que el experimentador supone o asume que tiene”. (Bronfenbrenner, 1977, p. 516). En este sentido, la validez ecológica implica mantener la integridad de la situación de la vida real en el contexto experimental mientras permanece fiel al contexto social y cultural más amplio. Por supuesto, una cuestión importante y no resuelta consiste en determinar a quién debe representar el entorno. (Schmuckler, 2001, p. 421).

A manera de ejemplo, Ardila y Solís (2008) explican que los lóbulos frontales intervienen en dos tipos de funciones ejecutivas interrelacionadas pero diversas:

(1) solución de problemas, planeación, formación de conceptos, desarrollo e implementación de estrategias, memoria de trabajo, etc. (funciones ejecutivas ‘metacognitivas’); es decir, las funciones ejecutivas tal y como se conciben en las neurociencias contemporáneas; y (2) coordinación de la cognición y emoción/ motivación

(funciones ejecutivas ‘emocionales’): es decir, satisfacer las necesidades biológicas de acuerdo a las condiciones existentes. (2008, p. 1)

Las funciones ejecutivas metacognitivas están asociadas con las áreas prefrontales dorsolaterales, mientras que las emocionales se vinculan con el área orbitofrontal y medial frontal. Pero, como las pruebas que evalúan funciones ejecutivas se centran en las correspondientes al primer tipo y la habilidad para resolver problemas cotidianos (aplicación funcional de las funciones ejecutivas) requiere más del segundo tipo de habilidades, es posible afirmar que las investigaciones usuales de las funciones ejecutivas carecen de validez ecológica.

Un ejemplo muy claro de lo que atañe a la validez ecológica puede observarse en un artículo en *Neuroscience News* de comienzos de 2024, llamado “La escritura a mano aumenta la conectividad cerebral y el aprendizaje”. En el resumen, es posible leer:

A medida que la escritura a mano tradicional es reemplazada progresivamente por dispositivos digitales, es esencial investigar las implicaciones para el cerebro humano. Se registró la actividad eléctrica cerebral en 36 estudiantes universitarios mientras escribían a mano palabras presentadas visualmente usando un bolígrafo digital y tecleaban las palabras en un teclado. [...] Nuestros resultados sugieren que el patrón espaciotemporal de información visual y propioceptiva obtenida a través de los movimientos precisos y controlados de las manos al usar un bolígrafo contribuyen ampliamente a los patrones de conectividad cerebral que promueven el aprendizaje. Instamos a que los niños, desde una edad temprana, estén expuestos a actividades de escritura a mano en la escuela para establecer los patrones de conectividad neuronal que proporcionan al cerebro condiciones óptimas para el aprendizaje. Aunque es vital mantener la práctica de la escritura a mano en la escuela, también es importante mantenerse al día con los avances tecnológicos en continuo desarrollo. Por lo tanto, tanto los maestros como los estudiantes deben ser conscientes de qué práctica tiene el mejor efecto de aprendizaje en qué contexto, por ejemplo, cuando toman notas de conferencias o cuando escriben un ensayo. (Van der Weel y Van der Meer, 2024, p. 1)

Se trata, como vemos, de una muestra sumamente exigua y bastante uniforme, ya que está solo integrada por estudiantes universitarios. Pero observemos el diseño experimental:

Los participantes recibieron instrucciones para hacer una de dos cosas: (a) escribir en cursiva con su mano derecha la palabra presentada con un bolígrafo digital directamente en la pantalla, o (b) escribir la palabra presentada en un teclado utilizando el dedo índice derecho. Antes de cada ensayo, aparecía la instrucción “escribir” o “escribir a máquina”

antes de que apareciera una de las palabras objetivo y se les daba a los participantes 25 segundos para escribir la palabra a mano o teclear la palabra varias veces, separadas por un espacio. (Van der Weel y Van der Meer, 2024, p. 2)

Se nota aquí con claridad cómo las tareas realizadas por los participantes del estudio no reflejan las situaciones de la vida real en las que las personas generalmente escriben a mano o en un teclado. En particular, la tarea de escribir en un teclado es simplificada al utilizar solo el dedo índice de la mano derecha, lo cual se aleja totalmente de la experiencia real de escribir en un teclado.

4.4.3. Validez de constructo

El segundo tipo de validez que puede limitar el proceso experimental es la **validez de constructo**. La idea básica aquí es que los investigadores en neurociencia y neurobiología cognitivas están interesados en desarrollar paradigmas experimentales que permitan individualizar capacidades cognitivas específicas, porque quieren poder hacer afirmaciones de estructura-función o mecanicistas sobre esas capacidades. Esto significa que debería darse el caso de que **el efecto bajo estudio en el laboratorio sea una instancia real de los fenómenos seleccionados por una construcción dada** (por ejemplo, "atención").

Se debe tener en cuenta de que las fallas en la validez de constructo pueden plantear un problema de confiabilidad, ya que para poder elegir entre afirmaciones contrapuestas vinculadas a una capacidad producida en el laboratorio es necesario que el efecto producido por el paradigma elemental sea una instancia real de dicha capacidad. (Sullivan, 2015, p. 41).

4.4.3.1. Algo más sobre la validez de constructo

Para comprender la importancia de la validez de constructo, es necesario recordar lo que señalan Streiner y otros (2015):

Atributos como la altura o el peso son fácilmente observables o pueden ser ‘definidos operativamente’; es decir, definidos por la forma en que se miden. [...] Sin embargo, una vez que nos alejamos del ámbito de los atributos físicos hacia otros más “psicológicos” como la ansiedad, la inteligencia o el dolor, comenzamos a tratar con variables más abstractas, que no se pueden observar directamente. No podemos “ver” la ansiedad; todo lo que podemos observar son comportamientos que, según nuestra teoría de la ansiedad, son el resultado de ella. (p. 235)

Existen numerosos factores en psicología que no se ven directamente, no son observables, sino solo a través de sus manifestaciones hipotéticas. “Se puede pensar en un constructo como una ‘miniteoría’ para explicar las relaciones entre varios comportamientos o actitudes” (Streiner, 2015, p. 236).

No hay un solo experimento que pueda “probar” inequívocamente un constructo (Streiner, 2015, p. 237). La validación de constructo implica un proceso continuo, de aprender más sobre él, hacer nuevas predicciones y luego probarlas. Señalan los autores que Albert Einstein dijo una vez que ha habido cientos de experimentos que respaldan su teoría de la relatividad, pero solo se necesitaría un estudio no confirmatorio para refutarla y que lo mismo ocurre con la validación de constructos; cada estudio solo sirve para fortalecer lo que Cronbach y Meehl (1955) llamaban la "red nomológica" de las predicciones entrelazadas de una teoría, pero un solo experimento bien diseñado con hallazgos negativos bastaría para poner en cuestión la totalidad del constructo.

4.5. Discusión y proyecciones

Como vimos, Jacqueline Sullivan (2015) propone la existencia de un *trade-off* metodológico entre confiabilidad y validez en el diseño experimental en neurociencias. Esta tensión puede entenderse como una revisión crítica del clásico conflicto entre validez interna y externa⁴⁵, si bien, en su formulación más general, Sullivan la presenta como una oposición estructural entre dos grandes principios normativos que guían la investigación científica.

Si bien confiabilidad y validez interna no son equivalentes, la primera suele considerarse una condición de posibilidad de la segunda: sin un procedimiento experimental confiable —esto es, replicable, controlado y consistente—, no es posible establecer inferencias válidas, ni siquiera dentro del entorno controlado del laboratorio.

A partir del texto de Sullivan, podemos ilustrar la imbricación entre validez interna y confiabilidad mediante el siguiente cuadro, que resume las condiciones que determinan la primera:

⁴⁵ La distinción entre validez interna y validez externa fue formulada en términos clásicos por autores como Campbell y Stanley (1963) y posteriormente ampliada por Cook y Campbell (1979) y Shadish, Cook y Campbell (2002).

Tabla 4-2*Condiciones de la validez interna*

Condición	Descripción breve	Referencia en Sullivan (2015)
Confiabilidad del proceso de producción de datos	El experimento debe ser replicable y consistente, y permitir discriminar entre hipótesis.	p. 40
Control de variables extrañas o confusoras	Es necesario eliminar causas alternativas del efecto observado.	p. 40
Adecuada formulación de las hipótesis	Las hipótesis deben ser claras, específicas y contrastables.	p. 40
Calidad del diseño experimental	El diseño debe tener protocolos estandarizados, con tareas e instrumentos bien definidos.	pp. 38–39
Procesamiento estadístico adecuado	Se deben usar técnicas estadísticas apropiadas para analizar e interpretar los datos obtenidos.	p. 39
Ausencia de sesgos sistemáticos	Es clave evitar errores de recolección, interpretación o análisis que distorsionen los resultados.	p. 35
Vínculo claro entre el efecto producido y el fenómeno de interés	El efecto observado debe ser una instancia válida del fenómeno que se busca estudiar.	pp. 40–41

Nota: elaboración propia a partir de Sullivan (2015).

Un ejemplo puede servirnos para ilustrar el *trade-off* mencionado. Supongamos que se diseña un estudio de resonancia magnética funcional para investigar la actividad cerebral involucrada en la toma de decisiones morales. El experimento incluye una serie de escenarios breves, presentados en pantalla mientras los participantes se encuentran dentro del escáner. Entre ellos se incluyen variaciones del dilema del tranvía, un clásico experimento mental en ética normativa⁴⁶.

Cada ensayo experimental presenta una versión distinta del dilema y el participante debe seleccionar entre dos opciones mediante un botón. Se registran tanto los tiempos de respuesta como los patrones de activación cerebral.

Este diseño cumple con varios criterios metodológicos:

- Es altamente **confiable**: el protocolo está estandarizado, es replicable y genera datos consistentes.

⁴⁶ En la versión más conocida de este dilema, un tranvía sin frenos se dirige hacia cinco personas atadas a las vías. Un observador, situado junto a una palanca, puede desviar el tranvía hacia otra vía, donde hay una sola persona. La decisión moral consiste en elegir entre no intervenir y permitir que mueran cinco personas o intervenir activamente y provocar la muerte de una sola persona para salvar a las otras cinco.

- Presenta **validez interna aceptable**, esto es, permite el control de variables, la comparación de condiciones y el establecimiento de correlaciones entre la versión del dilema y la actividad cerebral.

Sin embargo, el estudio enfrenta limitaciones significativas en términos de validez ecológica: los dilemas son abstractos, textuales y sin consecuencias reales, alejados de las situaciones morales del día a día; además, el contexto experimental minimiza la implicación emocional. Por otro lado, también se encuentra afectada la validez de constructo: la tarea se reduce a una respuesta binaria, simplificando decisiones que en la vida real son mucho más ambiguas y dinámicas.

Así, aunque el experimento sea confiable y válido internamente, su capacidad para explicar cómo las personas toman decisiones morales en contextos reales es limitada. Este caso ejemplifica con claridad el argumento de Sullivan sobre el *trade-off* entre confiabilidad y validez. Para asegurar datos confiables, el diseño experimental simplifica el fenómeno al punto de dejarlo vaciado de sus dimensiones éticas, contextuales y experienciales.

Formular el *trade-off* en términos de confiabilidad y validez resulta, además, metodológicamente más fructífero que mantener el binomio clásico de validez interna y externa. Esta última oposición se limita al problema de la generalización de los resultados, mientras que la formulación de Sullivan permite analizar el proceso experimental en su totalidad, desde la producción del dato hasta su interpretación. Además, incluye formas de validez que el enfoque clásico deja fuera, como la validez de constructo, que resulta clave en investigaciones sobre procesos cognitivos. Por otra parte, la distinción entre confiabilidad y validez permite establecer jerarquías metodológicas: mientras que las distintas formas de validez pueden entrar en tensión entre sí, la confiabilidad es una condición necesaria para cualquiera de ellas.

Esta perspectiva ofrece, en definitiva, una herramienta más potente para pensar los límites epistémicos del conocimiento neurocientífico, ya que articula los requisitos técnicos del diseño experimental con la fidelidad teórica de las afirmaciones que se desprenden de los datos.

Por otra parte, a lo largo del capítulo se ha puesto de manifiesto que la relación entre neurociencia y psicología cognitiva sigue siendo objeto de debate, principalmente debido a la inestabilidad de los constructos psicológicos y a la dificultad de integrarlos sistemáticamente en modelos experimentales de la neurociencia. Aunque la neurociencia cognitiva ha aportado información valiosa sobre la actividad cerebral, la falta de estabilidad conceptual en la psicología cognitiva limita la capacidad de las neuroimágenes para proporcionar explicaciones satisfactorias de los procesos mentales. Desde una perspectiva filosófica, esto plantea interrogantes sobre qué significa “explicar” en el contexto de la neurociencia cognitiva: si la explicación debe dar cuenta de

relaciones causales, permitir la formulación de predicciones o simplemente proporcionar descripciones funcionales de la actividad cerebral. En este sentido, la dificultad para estabilizar constructos psicológicos se traduce en una falta de criterios consensuados sobre qué tipo de explicación puede ofrecer la neuroimagen y qué alcance tiene en la comprensión de la cognición.

Esta dificultad no se limita al plano técnico, dado que la validez de constructo depende directamente de la estabilidad conceptual. Muchos de los términos con los que operan los diseños experimentales en neurociencia cognitiva provienen de la psicología, pero no cuentan con definiciones operativas consistentes ni con un marco teórico unificado. Sullivan ha señalado que esta situación se ve agravada por la multiplicidad de protocolos experimentales utilizados en los distintos laboratorios, lo cual dificulta la comparación sistemática de resultados y debilita la posibilidad de acumular conocimiento replicable. En la medida en que cada grupo de investigación opera con definiciones propias, tareas particulares y supuestos implícitos, la psicología cognitiva pierde su estatuto de disciplina empírica coherente y se transforma en un reservorio de intuiciones conceptuales sin capacidad de orientar la investigación de manera unificada.

Esta dificultad se agrava si se considera lo que en filosofía de la ciencia se conoce como el **problema de Duhem–Quine**⁴⁷. Según este planteo, **ningún experimento permite contrastar una hipótesis de manera aislada**, ya que siempre se la prueba junto con un conjunto de supuestos auxiliares, teorías previas y condiciones metodológicas. En el contexto de la neurociencia cognitiva, esto significa que un resultado inesperado —como la ausencia de activación en una región cerebral supuestamente asociada a una tarea— no permite saber con certeza si lo que falló fue la hipótesis, el diseño experimental, el análisis estadístico o el constructo psicológico en el que se basó el estudio.

Este problema tiene consecuencias directas sobre la **inferencia inversa**, ya que si la función atribuida a una región cerebral se apoya en un concepto inestable o mal definido, entonces la operación de deducir una actividad mental a partir de un patrón de activación cerebral se vuelve especialmente vulnerable. En última instancia, el riesgo no está solo en cómo se interpreta la imagen, sino en la fragilidad del marco conceptual que le da sentido.

⁴⁷ Este problema fue formulado originalmente por Pierre Duhem (1906) y ampliado por W. V. O. Quine (1951), y ha sido ampliamente discutido por autores contemporáneos como Chalmers (1999), quien lo presenta como una dificultad estructural para determinar qué hipótesis debe rechazarse cuando un experimento no arroja los resultados esperados. En particular, muestra que las hipótesis no se contrastan de forma aislada, sino junto a un entramado de supuestos auxiliares que dificultan una evaluación empírica directa.

Esta deriva debilita a la psicología como interlocutora epistémica y abre la posibilidad de que sea absorbida por una ontología formulada exclusivamente desde parámetros neurobiológicos. Bickle lo expresa con crudeza, al señalar que la psicología cumple apenas un rol heurístico en la fase de diseño experimental y que, una vez cumplida esa función, puede ser descartada como una escalera que ya no se necesita. Desde esta perspectiva, los modelos psicológicos no tendrían valor explicativo autónomo, sino que serían transitorios, reemplazables por una caracterización puramente neuronal de la conducta.

Frente a este escenario, Sullivan propone una alternativa que denomina “pluralismo coordinado”, que busca compatibilizar la diversidad conceptual y metodológica propia de la psicología cognitiva con ciertas condiciones de estandarización. Este enfoque apunta a generar definiciones más precisas, promover el uso compartido de plataformas experimentales y fomentar la colaboración interdisciplinaria en torno a marcos comunes. Se trata de una estrategia para mantener la riqueza teórica de la psicología sin resignar su capacidad de participar activamente en la construcción de explicaciones científicas compartidas con la neurociencia.

Este intento de articulación propuesto por Sullivan se vuelve particularmente relevante si se considera, tal como advierte Steven Kosslyn (1999), que las técnicas de neuroimagen pueden ofrecer respuestas válidas solo bajo ciertas condiciones teóricas y metodológicas. En términos de esta “aplicabilidad”, me parece esclarecedor por su sencillez el artículo cuyo título figura como epígrafe al inicio del apartado 4.3., sobre ontología cognitiva. Sintetizo en un cuadro mínimo las preguntas que, según este autor, pueden abordar las neuroimágenes y cuáles exceden su alcance interpretativo.

Tabla 4-3

Preguntas a las que la RMf puede responder.

Pregunta	¿Pueden responderla las neuroimágenes?	Justificación
¿Dónde sucede algo en el cerebro?	Sí	Localización funcional
¿Difieren dos tareas cognitivamente?	Sí	Comparación de patrones de activación
¿Una región respalda un modelo funcional?	Sí (con marco teórico)	Confirmación teórica basada en hipótesis
¿Qué siente o piensa alguien exactamente?	No	No accede al contenido subjetivo
¿Una activación equivale a un proceso mental?	No necesariamente	Las regiones son multifuncionales
¿Una región causa una conducta?	No	Solo hay correlación, no causalidad

Nota: elaboración propia basada en Kosslyn, 1999.

Si la ontología cognitiva se consolida como el marco dominante para el estudio de la mente, la psicología cognitiva podría desaparecer o, al menos, ser reconfigurada dentro de este nuevo sistema conceptual. Sin embargo, la idea de inconmensurabilidad parcial sugiere que este proceso no puede ser una simple sustitución acumulativa, sino una transformación que implicará pérdidas y reestructuraciones en la manera en que se conceptualizan los procesos cognitivos.

La psicología y la neurociencia cognitiva no operan con un mismo conjunto de categorías ontológicas. Mientras la primera ha desarrollado modelos funcionales basados en la observación del comportamiento y la experiencia subjetiva, la segunda busca redefinir estos fenómenos en términos de dinámicas cerebrales. Sin embargo, la traducción entre ambos niveles de análisis no es directa ni simétrica: a medida que los conceptos psicológicos son reformulados en el lenguaje neurocientífico, no siempre conservan su significado original, sino que son redefinidos en función de nuevas técnicas de medición y marcos teóricos. Esto implica que la neurociencia cognitiva no simplemente absorbe a la psicología, sino que la transforma, dejando fuera ciertas dimensiones que esta última abordaba desde su propio campo disciplinar.

Más allá del hecho que la psicología cognitiva podría perder su autonomía teórica frente a la ontología cognitiva, considero que es probable que, más que una eliminación absoluta, lo que emerja sea una reconfiguración de la psicología dentro del discurso neurocientífico, donde su rol quedará supeditado a las categorías y modelos de la ontología cognitiva. En este sentido, la neurociencia no solo avanza sobre el campo de la psicología, sino que también impone un horizonte de inteligibilidad que determina qué aspectos de la mente son considerados legítimos objetos de estudio y cuáles quedan fuera del ámbito explicativo predominante.

Finalmente, las dificultades para definir lo que se mide y cómo se interpreta también afectan la forma en que los hallazgos de la neurociencia cognitiva son aplicados en contextos institucionales, como la clínica, la justicia o la producción de identidades. Por ello, en el próximo capítulo examino estos usos desde una perspectiva crítica, para examinar qué supuestos epistemológicos y normativos operan cuando las imágenes cerebrales son presentadas como evidencia en decisiones que afectan directamente a las personas.

5. El desafío de la neuroética

En los capítulos anteriores se abordaron cuestiones vinculadas al diseño experimental en estudios con neuroimágenes, su estatus epistemológico y su articulación con la psicología cognitiva. Estos desarrollos habilitan una nueva dimensión de análisis: el impacto que los resultados de las neuroimágenes pueden tener en el plano ético. En este capítulo, me propongo explorar cómo el uso de estas tecnologías incide en aspectos como la privacidad mental, la autonomía y la identidad personal y analizo sus implicancias en contextos clínicos, jurídicos y sociales.

5.1. Contexto y delimitaciones del campo

La palabra “neuroética” entró en el vocabulario de las neurociencias a principios del siglo XXI, acuñada por William Safire (1929-2009), escritor, periodista y por aquel entonces presidente de la *Dana Foundation*⁴⁸ (Farah, 2012, p. 572)⁴⁹.

En el año 2002, la Fundación Dana realizó una reunión que congregaba neurocientíficos, especialistas en ética y otros pensadores en San Francisco. Dicho encuentro recibió el nombre de *Neuroethics, mapping the field*⁵⁰ y podemos afirmar que a partir de allí la neuroética como disciplina empezó a recibir una atención creciente.

En el mismo año Adina Roskies escribió “Neuroética para el nuevo milenio” (Roskies, 2002), artículo en el que proponía una división bipartita de la neuroética en la “ética de la neurociencia”, y “la neurociencia de la ética”, lo que sugiere una extensión del alcance de la neuroética para abarcar la comprensión de la base biológica del pensamiento y el comportamiento éticos y las formas en que esto podría en sí mismo influir e informar nuestro pensamiento ético.

En las páginas que siguen, examino el primero de los elementos de esta clasificación, “la ética de la neurociencia” y, de modo particular, dado el tema de este trabajo, aquellas dimensiones de la ética especialmente asociadas a las técnicas de neuroimagen.

⁴⁸ La **Fundación Dana** se define en su página web www.dana.org como una organización filantrópica privada con sede en Nueva York dedicada al avance de la neurociencia y la sociedad a través del sostén de iniciativas interdisciplinarias entre neurociencia, ética, derecho, política, humanidades y artes.

⁴⁹ Otros autores señalan que el término “neuroética” se usaba anteriormente, pero sin el estatus disciplinario que surgió de la reunión en la Fundación Dana.

⁵⁰ El PDF con las conferencias presentadas allí pueden ser descargadas de la página de la Fundación: www.dana.org.

Martha Farah señala: “Están surgiendo nuevos problemas éticos a medida que la neurociencia nos brinda formas sin precedentes de comprender la mente humana y de predecirla, influir en ella e incluso controlarla.” (Farah, 2010, p. 2)

De acuerdo con Farah, es posible distinguir, por un lado, ciertas **cuestiones bioéticas clásicas** — que aparecen *también* en el campo de la neuroética pero cuyo problema de fondo no es diferente del que podría haber surgido en otra área— y, por otro, aquellas cuestiones éticas que implican nuevos desafíos surgidos a partir de las neurociencias contemporáneas.

Algunos de los ejemplos más relevantes del primer caso podrían ser los siguientes:

1. Desarrollo de pruebas predictivas para enfermedades neurodegenerativas:

Para las compañías de seguros, los departamentos de personal e incluso para el propio individuo, la predicción de la enfermedad de Alzheimer permitiría una planificación más racional para el futuro. La pregunta ética, por supuesto, es qué precio tiene esta capacidad adicional de planificación.

El conocimiento de que uno está destinado a desarrollar la enfermedad de Alzheimer es una carga terrible, sobre todo porque no hay cura. Las principales preocupaciones éticas son los derechos de privacidad (¿debería su compañía de seguros o jefe saber los resultados de la prueba?) y calidad de vida (¿cuáles son los efectos sobre el bienestar del paciente de saber versus no saber?). (Farah, 2010, p. 4)

2. En el caso de la RMf, la investigación actual implica colocar al sujeto humano en un campo magnético de fuerza de 1,5 o 3 teslas⁵¹ y todo parece indicar que es seguro. Hasta hace poco, las limitaciones técnicas impedían el uso de campos más potentes, sin embargo, hoy es posible realizar escaneos a 7 teslas y más.

Los fuertes campos magnéticos estáticos pueden afectar la presión arterial, la función cardíaca y la actividad neuronal. [...] Los sujetos en escáneres de alto campo a veces informan haber visto luces como resultado de corrientes inducidas en sus retinas o nervios ópticos. Aunque los estudios de seguridad han sugerido que son benignos, se sabe poco sobre los efectos a largo plazo de estos protocolos de exploración más nuevos y potentes. (Farah, 2010, p. 6)

⁵¹ El Tesla es el valor del total del flujo magnético dividido por el área, de lo que puede deducirse que si reducimos el área afectada se incrementa la densidad del flujo magnético.

3. Otra cuestión bioética relacionada a la iRMf es el hallazgo de anomalías cerebrales encontradas en el curso de una investigación. Esto plantea la cuestión de qué deben hacer los investigadores con estos hallazgos incidentales. Actualmente no existe un consenso general sobre el procedimiento para tratar estos hallazgos incidentales en la investigación (Farah, 2010, p. 6).

En cuanto a los **nuevos desafíos éticos propios de la neurociencia**, hacen referencia a las tecnologías para explorar, monitorear y manipular el cerebro y cómo “estos desarrollos se cruzarán con nuestras creencias fuertemente arraigadas sobre el valor de la privacidad, la libertad, la equidad y la responsabilidad” (Farah, 2010, p. 7).

Podríamos pensar, entonces, junto con Roskies (2021) a la neuroética como “centrada en los problemas e inquietudes éticas surgidas a partir de nuestra comprensión cada vez mayor del cerebro, con la consecuente posibilidad de influir e incluso de controlar la función cerebral y, de allí, potencialmente también la conducta humana.”

Las técnicas de neuroimagen funcional, hemos visto en los capítulos anteriores, son complejas tanto en diseño como en implementación e interpretación de los datos. Pero, curiosamente y a diferencia de otros tipos de investigaciones científicas, esta sofisticación tecnológica va de la mano con una enorme facilidad de acceso para académicos (Roskies, 2015b, p. 660) e incluso para legos. Son tantos los desafíos que nos presenta la neuroética aplicada a las neuroimágenes funcionales que haré solo algunas referencias muy breves para poder encuadrar el posible impacto de los desarrollos que se actualizan día a día. Parafraseando a Bickle (2019): “La ficción especulativa del año pasado es la realidad científica de hoy”.

Por otra parte, Haselager y Mecacci (2021) plantean lo que suele denominarse el “dilema de Collingridge”⁵²: en etapas tempranas la tecnología es sencilla de regular, pero sus consecuencias son impredecibles. En cambio, en momentos de desarrollo avanzado, las consecuencias son claras, pero regular la tecnología es difícil y costoso, porque ya está arraigada en la sociedad.

Este dilema subraya la necesidad de anticipar y adaptarse a los impactos de las tecnologías disruptivas, como la inteligencia artificial o la biotecnología, mediante enfoques proactivos y flexibles.

⁵² Llamado así por David Collingridge, quien lo formuló por primera vez en su libro *The Social Control of Technology* (1980) con el nombre de “dilema del control”: “Cuando el cambio es fácil, no se puede prever la necesidad de este; cuando la necesidad de cambio es evidente, el cambio se ha vuelto costoso, difícil y lento”.

5.2. Neuroimágenes y conciencia

Los **trastornos cuantitativos de la conciencia** afectan a numerosas personas y hasta hace muy poco tiempo, solo podían diagnosticarse sobre una base conductual y con un enorme porcentaje de error. El desarrollo de técnicas de neuroimagen puede tener “implicaciones profundas para las estrategias médicas, el manejo del dolor y las decisiones sobre el final de la vida” (Roskies, 2015, p. 660).

Quizás lo más emocionante es la vía potencial que brindan las tecnologías de imágenes cerebrales para comunicarse con pacientes conscientes y cognitivamente intactos, pero incapaces de responder verbal o conductualmente a las instrucciones. La neuroimagen abre la posibilidad de que se pueda restablecer la comunicación con estos pacientes mediante la asociación de respuestas con diferentes tareas imaginativas que tienen distintas “firmas” neuronales. El poder potencial de esta técnica de lectura de la mente se combina con un conjunto de desafíos éticos únicos con respecto a cómo interpretar la capacidad de toma de decisiones autónoma y potencialmente trascendental en pacientes cuyas habilidades y lucidez se ven afectadas en grados inciertos y fluctúan con el tiempo. (Roskies, 2015, p. 660)

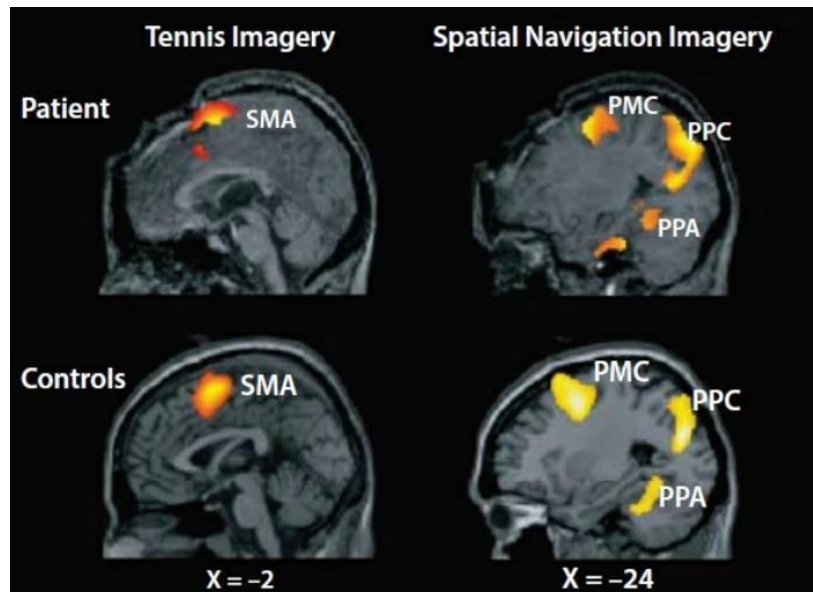
Russell Poldrack (2018, pp. 82-84) describe el “síndrome de enclaustramiento”, causado por un daño en el tronco encefálico, que deja a los pacientes completamente conscientes pero incapaces de realizar ningún movimiento fuera de parpadear y mover los ojos. A partir de este síndrome, surge la pregunta: ¿qué pasaría si algunas de las personas que supuestamente están en estado vegetativo fueran realmente conscientes pero incapaces de expresarse? El neurocientífico Adrian Owen encontró una solución simple: colocó a las personas en el escáner de resonancia magnética y luego les pidió que se imaginaran caminando por su casa o jugando al tenis.

Probaron el método en una mujer de 23 años que había sufrido una lesión cerebral grave en un accidente automovilístico cinco meses antes y que permanecía en estado vegetativo, completamente insensible a la estimulación. Si bien su falta de respuesta sugirió que no tenía una función cognitiva intacta, la resonancia magnética funcional mostró una historia diferente. Cuando se le dijo que se imaginara jugando al tenis, hubo actividad en su corteza premotora, mientras que hubo actividad en toda la red que se activa durante la navegación espacial en personas sanas cuando se le pidió que imaginara que se paseaba por su casa [...] El hallazgo histórico inspiró un análisis mucho más amplio, que demostró que la proporción de personas en estado vegetativo que pasan la prueba de conciencia es

relativamente baja: en un estudio posterior dirigido por Martin Monti, 5 de 54 personas en estado vegetativo mostraron evidencia de conciencia. (Poldrack, 2018, p. 83)

Figura 5-1.

Los resultados del artículo de Adrian Owen de 2006, que muestran en rojo/amarillo las áreas del cerebro activas mientras se imagina jugando al tenis (izquierda) e imaginando navegar por la casa (derecha) en el paciente vegetativo (arriba) y los controles sanos (abajo).



Nota: de “Detecting Awareness in the Vegetative State” por Adrian M. Owen, Martin R. Coleman, Melanie Boly, Matthew H. Davis, Steven Laurys y John D. Pickard. Obtenido de Poldrack, 2018.

Estos estudios, más allá del enorme avance que representan, plantean también algunas cuestiones éticas. Si las personas en estado vegetativo pero con evidencia de conciencia pueden responder preguntas significa que, en principio, podría decirnos si desean seguir viviendo. ¿Cómo decidiríamos si tienen suficiente capacidad de razonamiento para hacer este juicio? ¿Cómo responderíamos?

El uso de neuroimágenes en el estudio de la conciencia ha permitido avanzar en la comprensión de los estados mentales y las funciones cognitivas en pacientes con trastornos de conciencia. Sin embargo, estas aplicaciones presentan dilemas éticos importantes, ya que la interpretación de los datos sigue siendo incierta y depende de modelos teóricos en construcción. En los siguientes apartados, analizaremos cómo estas cuestiones afectan la toma de decisiones en contextos clínicos y legales.

5.3. Neuroimágenes y lectura de la mente

En el estado actual de la cuestión, si bien las limitaciones de las técnicas de neuroimagen nos permiten ser relativamente escépticos ante la posibilidad de “leer la mente” en el sentido de **“conocer los contenidos de pensamiento”** —más todavía si pensamos en hacerlo en forma encubierta o sin el consentimiento del individuo—, de acuerdo con Roskies, hay inferencias menos específicas que podrían plantear cuestionamientos éticos acerca del derecho a la privacidad y sus límites (2015, p. 661).

Al respecto, señalan Haselager y Mecacci (2021, p. 205):

La combinación de inteligencia artificial y neuroimagen (AINI), específicamente la combinación de aprendizaje automático (ML) e imágenes de resonancia magnética funcional (iRMf), ha aumentado las posibilidades para la lectura del cerebro, es decir, la decodificación del contenido mental o la identificación de disposiciones conductuales a partir de la actividad cerebral. Aunque la lectura del cerebro se encuentra actualmente en una etapa inicial, compuesta principalmente por pruebas de principio⁵³, un mayor progreso podría llevar a aplicaciones socialmente relevantes.

De hecho, ha llegado a acuñarse el acrónimo ELSI (*Ethical, Legal and Social Implications*) para referirse a las implicaciones éticas, legales y sociales de la investigación científica; en particular, se hace referencia a la *ELSI gap* o “brecha ELSI” en campos donde el desarrollo tecnológico avanza rápidamente (como la inteligencia artificial o las neurociencias), para aludir a la brecha entre estos avances y la consideración de sus aspectos éticos, legales y sociales.

Las técnicas de IA, especialmente el aprendizaje automático, en combinación con datos de neuroimagen, pueden predecir fenómenos psiquiátricos y/o respuestas al tratamiento, de maneras cada vez mejores y más novedosas. En una revisión reciente de aplicaciones de AINI para la psiquiatría, Nielsen et al. indican que el aprendizaje supervisado puede aplicarse para predecir, por ejemplo, el trastorno por déficit de atención e hiperactividad (TDAH), el autismo, la depresión, la esquizofrenia y el síndrome de Tourette. Pueden identificarse desviaciones relacionadas con la edad del desarrollo estándar, y esto puede usarse para predecir comportamientos de riesgo en el mundo real. (Haselager y Mecacci, 2021, p. 209)

⁵³ Una prueba de principio es una demostración inicial destinada a validar si una idea, concepto o teoría es técnicamente factible o viable en un nivel básico; es decir, si puede funcionar bajo ciertas condiciones controladas.

Pero es claro que estos avances, lejos de cerrar el debate, abren nuevas zonas de incertidumbre y exigen una discusión más profunda sobre sus consecuencias sociales y epistémicas.

Paulus et al. y Jollans et al. discuten las complejidades de AINI aplicadas a la predicción del comportamiento individual, en particular para encontrar un punto óptimo en el compromiso entre sesgo y varianza, es decir, evitando tanto modelos excesivamente complejos que se ajustan (demasiado bien) a los datos pero no generalizan, como modelos excesivamente simples que no se ajustan a los datos. (Haselager y Mecacci, 2021, p. 210)

Como vemos, aparece una vez más el *trade-off* entre confiabilidad y validez que analizamos precedentemente.

El aumento de la capacidad para la lectura cerebral debido a la mejora rápida de las AINI, así como el potencial riesgo de malinterpretar las posibilidades actuales, traen aparejados cuestionamientos diversos, entre ellos, los referidos a las prácticas actuales de consentimiento informado.

En síntesis, la posibilidad de utilizar la neuroimagen para acceder a los estados mentales de una persona plantea serios problemas de privacidad y autonomía. Si bien estas técnicas aún están en desarrollo, su uso en la detección de intenciones o recuerdos podría dar lugar a vulneraciones de derechos fundamentales. En los próximos apartados, analizo cómo estas tecnologías impactan en ámbitos como la detección de mentiras y la predicción del comportamiento.

5.4. Neuroimágenes y detección de mentiras

Otro ámbito de la neuroética aplicada a la neuroimagen es el de la **detección de mentiras**. Los experimentos realizados hasta el presente en relación con este ámbito son de cuestionable validez (sobre todo en lo que hace a la validez ecológica). No obstante, cabe mencionar aquí que un estudio realizado en 2011 (McCabe et al., 2011) muestra que la presentación de pruebas de IRMf en relación con la detección de mentiras producía en el jurado, a la hora de dictar el veredicto, una influencia mayor que cualquier otra condición probatoria. Esto se asocia, sin duda, a la pregnancia de la imagen visual, ya analizada por Roskies al comparar las neuroimágenes y las fotografías.

Figura 5-2.

Imagen de la técnica de "Potenciales relacionados con eventos".

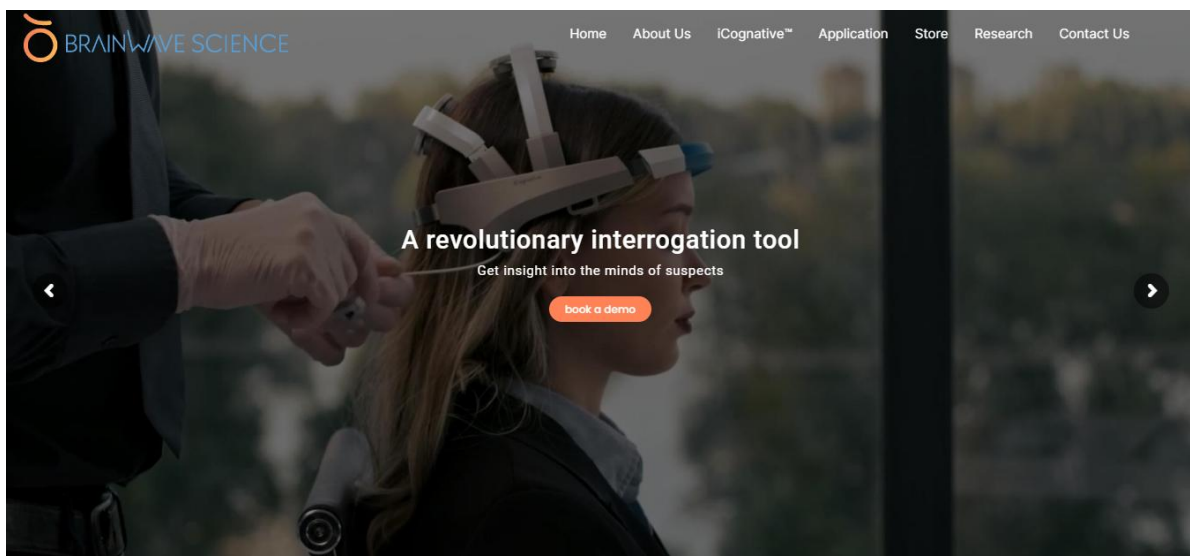
Si bien la técnica ilustrada se basa en electroencefalografía (EEG) y no en resonancia magnética, la incluimos aquí porque ilustra paradigmáticamente la lógica de la 'lectura de mente' y la pregnancia visual que también —y con mayor fuerza anatómica— caracteriza a los litigios que involucran neuroimágenes por RMf.

El texto dice: “Una herramienta de interrogación revolucionaria. Entre en la mente de los sospechosos”. El método, llamado Brain Fingerprinting por su desarrollador (más tarde rebautizada “iCognitive”, ha sido admitido como evidencia en un juicio y está siendo promovido como un medio de detección de terroristas. (cf. Farah, 2005, p. 35)



En las FAQ de la página, se puede leer: “La diferencia fundamental entre el autor de un acto o delito terrorista y una persona inocente es que el autor, después de haber cometido el delito, tiene un conocimiento específico del mismo almacenado en su cerebro. Un sospechoso inocente no lo tiene. De manera similar, un terrorista intelectual o un terrorista entrenado tiene un conocimiento específico relacionado con el terrorismo almacenado en su cerebro. Esto es lo que las pruebas iCognitive detectan, de manera científica y precisa. Las palabras, frases o imágenes relevantes para un acto terrorista, un delito, un entrenamiento terrorista o un conocimiento o experiencia específicos se presentan en una pantalla de computadora, en una

serie con otras palabras, frases o imágenes irrelevantes. Las respuestas de las ondas cerebrales de un sospechoso se miden de manera no invasiva utilizando un auricular equipado con sensores de EEG. Luego, un programa de computadora patentado analiza los datos para determinar si la información relevante para el delito está almacenada en el cerebro. Una respuesta cerebral específica "P300" es emitida por el cerebro de un autor que tiene los detalles de un delito almacenados en su cerebro, pero no por un sospechoso inocente que carece de este registro en su cerebro. Esto está demostrado por una extensa investigación científica publicada y revisada por pares”.



Nota: tomado de la página <https://www.brainwavescience.com/>

El uso de neuroimágenes para detectar mentiras ha despertado expectativas en el ámbito judicial, pero presenta importantes limitaciones técnicas y éticas. La falta de precisión y la posibilidad de falsos positivos dificultan su validez como prueba legal. Además, su aplicación sin el consentimiento informado de la persona evaluada podría atentar contra derechos fundamentales. En la siguiente sección, se examina cómo la neuroimagen se está utilizando para predecir comportamientos y qué riesgos implica este uso.

5.5. Neuroimágenes y predicción de comportamientos

Uno de los puntos más espinosos en el ámbito de la ética de la neuroimagen es el “carácter prometedor” de esta a la hora de **predecir tanto enfermedades como comportamientos**. En el primero de los casos, podría tener consecuencias discriminatorias, por ejemplo, en el ámbito laboral. Pero la segunda posibilidad, la de predecir comportamientos, comporta el riesgo de la llamada **peligrosidad predelictual**. A la manera de la distopía presentada por la película *Minority Report*, se abre un gran abanico de cuestionamientos éticos que incluyen la presunción de inocencia y, sobre todo, el libre albedrío.

Señala González Lagier (2012, p. 500):

Uno de los problemas que muchos neurocientíficos creen haber desentrañado es el del libre albedrío. De acuerdo con estos autores, el análisis del cerebro ha mostrado que la libertad humana es meramente una ilusión, una ficción generada por nuestro propio cerebro. Determinados experimentos —los famosos experimentos de la década de los ochenta de B. Libet y otros en la misma línea— parecen permitir concluir que el cerebro “se pone en marcha” para realizar una determinada acción antes de que seamos conscientes de haber tomado la decisión correspondiente, de lo que suele inferirse (1) que nuestra decisión consciente no tiene efecto causal en la acción realizada y (2) que la sensación o impresión de actuar libremente y a consecuencia de nuestra decisión consciente es un efecto causal de la propia actividad cerebral.

Señala el autor, además, que el problema que las neurociencias plantean para la noción de libre albedrío no es más que “un nuevo ropaje para un viejo problema persistente de la filosofía” (p. 501). En su artículo, expresa los argumentos de G. H. Von Wright “en contra de la posibilidad de demostrar empíricamente el determinismo universal entendido como una ley irrestricta (válida

también para las acciones humanas) y contra la reducción de lo mental a lo neuronal” (González Lagier, 2012, 500).

El intento de predecir comportamientos mediante neuroimágenes introduce serias preocupaciones sobre el determinismo biológico y la discriminación. Si bien estas herramientas pueden contribuir a la identificación temprana de ciertos trastornos, su uso en la justicia penal o en la evaluación de riesgos individuales podría derivar en medidas preventivas injustificadas. A continuación, analizaremos cómo la neuroimagen también ha sido utilizada para justificar diferencias de género y sus implicaciones éticas.

5.6. Neuroimágenes y género

Podríamos afirmar que, durante el siglo XX, la interpretación de las diferencias biológicas entre los géneros se vio afectada por tres grandes factores. Los primeros dos son el descubrimiento del gen SRY por el endocrinólogo francés Alfred Jost en 1940 y la consolidación de la neuroendocrinología como disciplina científica en 1959. De ellos se desprendieron los llamados “dogma clásico” y “teoría organizacional/activacional” (O/A).

El dogma clásico hace referencia a las ideas tradicionales sobre la determinación del sexo en aves y mamíferos, que sostienen que “el sexo genético del animal determina el tipo de gónadas, y luego el fenotipo del tejido no gonadal es controlado exclusivamente por la secreción de dichas gónadas” (Ciccia, 2017, p. 163).

Por su parte, la teoría organizacional/activacional propone que las hormonas organizan y activan patrones de comportamiento y características físicas en el desarrollo y que estos procesos son influenciados por factores genéticos y ambientales.

La teoría O/A extendió a los cerebros la interpretación dimórfica de la diferenciación sexual y planteó que a partir de un cerebro monomórfico (inicialmente “hembra”) ocurre una diferenciación (“masculinización”) causada por la testosterona. Dicha diferenciación, que organizaría permanentemente el cerebro de manera sexo-específica, se activaría en la vida posnatal y explicaría así las diferencias esenciales entre machos y hembras. Tanto el dogma clásico como la mencionada teoría sugieren que los procesos de diferenciación activos serían solo característicos del macho. (Ciccia, 2021)

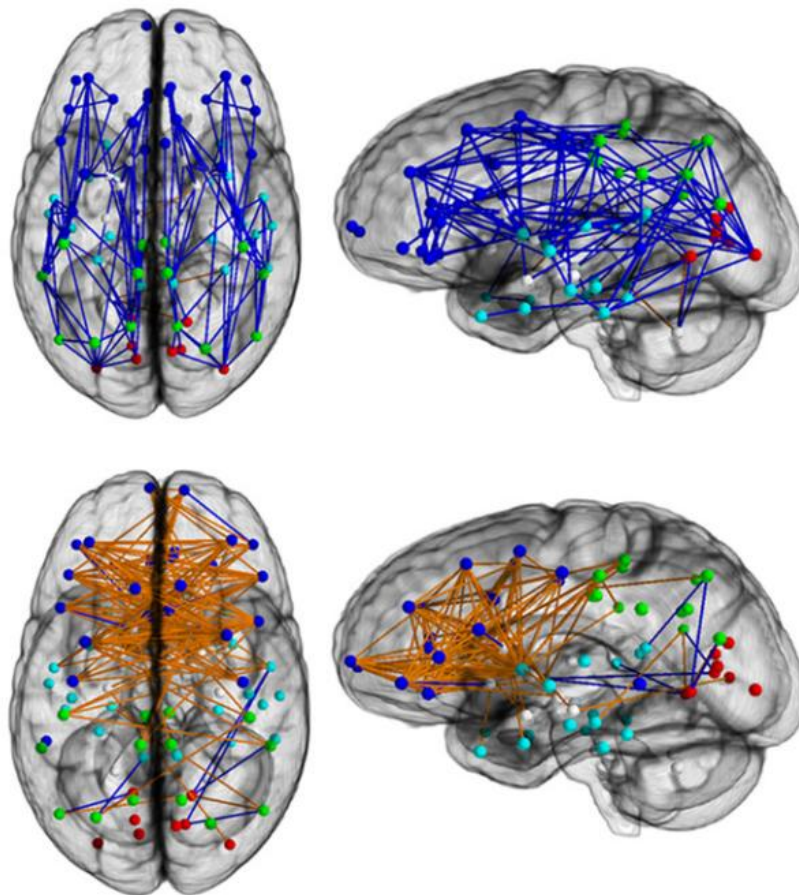
¿En qué consistirían estas diferencias sexo-específicas? Entre otras, estereotipos de género como los que sostienen que los hombres superan a las mujeres en matemáticas y pruebas espaciales, mientras que las mujeres superan a los hombres en pruebas verbales. (Hyde, 2016)

El tercero de los factores —y el que reviste importancia para nosotros aquí— es el advenimiento de las neuroimágenes y, en particular, de las técnicas funcionales.

De las investigaciones realizadas, podemos mencionar el de Madhura Ingalhalikar, publicado en 2014. Según los resultados obtenidos, las mujeres exhibían una mayor conectividad interhemisférica, mientras que los hombres mostraban una mayor conectividad intrahemisférica. De acuerdo con Ciccía (2017), la interpretación de Ingalhalikar sugiere que el cerebro de los hombres está estructurado para favorecer la conectividad entre percepción y la acción coordinada, mientras que el de las mujeres facilitaría la comunicación entre modos de procesamiento analítico e intuitivo (Ingalhalikar, 2014, p. 823).

Figura 5-3.

Análisis de conexiones. Las conexiones intrahemisféricas se muestran en azul y las interhemisféricas se muestran en naranja.



Nota: tomado de Ingalhalikar, 2014.

Ciccia (2017) analiza la investigación de Ingallhalikar en relación con el discurso científico que interpreta las diferencias cerebrales en hombres y mujeres como fenómenos binarios y jerárquicos, lo que, según esta autora, ayudaría a perpetuar la discriminación y la marginación de ciertas identidades de género.

Dicho paradigma discursivo parte de asumir como válidas las diferencias y procura encontrar los sustratos neuroanatómicos y neurofisiológicos que las expliquen. En efecto, el estudio comienza señalando:

Las diferencias en el comportamiento humano entre sexos muestran complementariedad adaptativa: los hombres tienen mejores habilidades motoras y espaciales, mientras que las mujeres tienen habilidades superiores de memoria y cognición social. Varios estudios también muestran diferencias sexuales en los cerebros humanos, pero no explican esta complementariedad. (Ingallhalikar, 2014, p. 1)

Los aspectos culturales y sociales condicionan las diferencias de género en el cerebro y el significado de estas posibles diferencias, cuando aparecen, debe ser interpretado críticamente a fin de evitar caer en el sesgo de confirmación que no hace sino reproducir los valores patriarcales y heterosexistas. Se debe tener en cuenta que las identidades binarias de género no reflejan de manera precisa la diversidad de la experiencia humana y no deben ser utilizadas como categorías limitantes en la investigación o la práctica clínica.

La elevada plasticidad cerebral que nos caracteriza como especie vuelve altamente permeable el sistema nervioso central a nuestra práctica cultural. En este sentido, las estructuras cerebrales asociadas a la química y mecánica de la reproducción no se encuentran exentas de dichas prácticas. En efecto, ninguna constitución genética-hormonal y/o genital, es capaz de predecir un tipo de cerebro particular. (Ciccia, 2017, p. 171)

En resumidas cuentas, la investigación sugiere que las diferencias en la activación cerebral entre géneros podrían estar relacionadas con factores biológicos y socioculturales, sustentadas por experiencias y expectativas de género y no exclusivamente por factores neurobiológicos.

La tendencia a esencializar las diferencias de género puede reforzar estereotipos y justificar desigualdades sin una base científica sólida. En la siguiente sección, examinaré cómo la divulgación de resultados neurocientíficos en los medios de comunicación influye en la percepción pública de estos y otros estudios

5.7. Neuroimágenes y opinión pública

La amplia difusión de la investigación en neurociencias ha trascendido las comunidades científicas para llegar a los medios de comunicación y el público en general.

Sin embargo, ¿los límites de lo que esta tecnología puede y no puede lograr se comunican de manera efectiva al público? ¿Se comprenden sus limitaciones? ¿Se consideran útiles y significativas las aplicaciones de la tecnología? ¿Son algunos estudios más propicios a la mala interpretación que otros? ¿Cuáles son los riesgos asociados para la sociedad? (Racine, 2005, p. 1)

Ya hemos visto cómo la investigación en iRMf está guiada por una serie de supuestos y de limitaciones de orden tanto técnico como epistemológico. No obstante, ¿qué sucede cuando los resultados de una investigación se publicitan?

Enmarcamos nuestros argumentos de acuerdo con tres tendencias que hemos observado en la cobertura de prensa de iRMf: 'neurorrealismo', 'neuroesencialismo' y 'neuropolítica', y exploramos cómo la neuroética puede atender a los aspectos éticos legales y sociales relacionados con la promoción de la comunicación multidireccional en neurociencia. (Racine, 2005, p. 2)

En las investigaciones realizadas por Racine y otros sobre la cobertura de prensa de los estudios de iRMf, los autores observaron:

- muy limitada explicación de las limitaciones de iRMf (67 %).
- tono mayoritariamente optimista (79 %).
- Escaso abordaje de las cuestiones éticas (7 %).

A partir de allí, desarrollaron las tres tendencias mencionadas que explicaremos a continuación:

Neurorrealismo:

Nuestro concepto de “neurorrealismo” describe cómo la cobertura de las investigaciones de iRMf puede hacer que un fenómeno sea real, objetivo o efectivo acriticamente a los ojos del público. [...] [Por ejemplo] Debido a que la investigación de iRMf muestra la activación en los centros de recompensa cuando los sujetos ingieren alimentos ricos en grasas, se lee: “La grasa realmente brinda placer”. [...]

El neurorrealismo refleja la forma acrítica en que una investigación de iRMf puede tomarse como validación o invalidación de nuestra visión ordinaria del mundo. Por lo tanto, el neurorrealismo se basa en la creencia de que la resonancia magnética funcional nos permite

capturar una “prueba visual” de la actividad cerebral, a pesar de las enormes complejidades de la adquisición de datos y el procesamiento de imágenes. (Racine, 2005, p. 3)

Nótese cómo esta explicación del “neurorealismo” guarda estrecha afinidad con lo desarrollado por Roskies sobre la pregnancia visual de las neuroimágenes.

Neuroesencialismo:

El concepto de "neuroesencialismo" refleja cómo la investigación de iRMf puede representarse equiparando la subjetividad y la identidad personal con el cerebro. En este sentido, el cerebro se utiliza implícitamente como atajo para conceptos más globales como la persona, el individuo o el yo. Este es el caso de muchas expresiones en las que el cerebro se utiliza como sujeto gramatical. Los ejemplos principales de este fenómeno son: “El cerebro puede desterrar los recuerdos no deseados”, “Cómo almacena el cerebro los idiomas” [...]

Queda claro que esta mirada neuroesencialista representa, intencionalmente o no, una forma de reduccionismo: identidad, yo, mente quedan reducidas a la noción de cerebro. Bennett y Hacker (2022), en su libro *Philosophical foundations of neuroscience*, describen esto en términos de “falacia mereológica”. Plantean que las primeras generaciones de neurocientíficos (que incluían a figuras como Charles Sherrington, Edgar Adrian, John Eccles y Wilder Penfield) se caracterizaron por sostener una postura cartesiana. Las nuevas generaciones, en cambio, tienden a rechazar el dualismo cartesiano de sus predecesores y trasladar entonces la atribución de la mente al cerebro. Así, se comienza a afirmar que el cerebro piensa, siente, decide... cayendo así en la falacia mereológica que consiste en atribuir a una parte (el cerebro) lo que solo puede ser predicado del todo.

La cuestión que estamos enfrentando es una cuestión filosófica, no científica. Exige una clarificación conceptual, no una investigación experimental. No se puede investigar experimentalmente si los cerebros piensan, creen, adivinan, razonan, forman hipótesis, etc., hasta que se sepa qué implicaría que un cerebro hiciera eso; es decir, hasta que tengamos claro el significado de estas frases y sepamos qué (si es que algo) puede contar como el hecho de que un cerebro haga estas cosas y qué tipo de evidencia respaldaría la atribución de tales atributos al cerebro. (No se pueden buscar los polos de la Tierra hasta que se sepa qué es un polo; es decir, qué significa la expresión 'polo' y qué cuenta como encontrar un polo de la Tierra. De lo contrario, como Winnie Pooh, uno podría embarcarse en una expedición hacia el Polo Este). La cuestión debatida es: ¿tiene sentido atribuir tales

atributos al cerebro? ¿Existe algo así como el pensamiento, la creencia, etc., de un cerebro? (¿Hay algo así como el Polo Este?)

En las *Investigaciones Filosóficas*, Wittgenstein hizo una observación profunda que se relaciona directamente con nuestras preocupaciones. “Solo de un ser humano y de lo que se asemeja (se comporta como) un ser humano vivo se puede decir: tiene sensaciones; ve, es ciego; oye, es sordo; es consciente o inconsciente.” (Bennett y Hacker, 2022, p. 82)

Los autores argumentan que no tiene sentido atribuir predicados como estos (o sus negaciones) al cerebro, porque este no es un sujeto lógicamente apropiado para predicados psicológicos, salvo que se lo haga de una manera metafórica o metonímica⁵⁴.

Neuropolítica:

La 'neuropolítica' describe los intentos de utilizar los resultados de iRMf para promover agendas políticas y personales, como en el caso de los grupos de interés que defienden la investigación de problemas sociales utilizando iRMf. Por ejemplo, la Lighted Candle Society, una organización sin fines de lucro con sede en Utah que se dedica a la mejora de los valores morales, aboga por el uso de resonancia magnética funcional para demostrar que la pornografía es adictiva. (Racine, 2005, p. 3)

Este caso —como otros que podrían identificarse también en el contexto local— ilustra cómo los hallazgos neurocientíficos pueden extenderse más allá del ámbito estrictamente académico y ser movilizados estratégicamente en el debate público para legitimar o justificar determinadas políticas.

Figura 5-4.

Algunos ejemplos de la difusión de la neurociencia en la prensa.

⁵⁴ De la misma manera, las teorías localizacionistas que obvian la complejidad estructura y funcional del cerebro y asumen que una función puede estar totalmente localizada en una región determinada pueden ser consideradas otras formas de falacia mereológica en neurociencias.



Nota: imágenes tomadas de https://elpais.com/economia/2016/07/17/actualidad/1468776267_359871.html
<https://www.larazon.es/gente/poder/20221211/gqsyod762bebbno3vjhwtq273a.html>

Racine (2005) promueve, entonces, un enfoque proactivo de comunicación multidireccional, con intercambios entre neurocientíficos y el público, dado que el modelo de comunicación tradicional a través de la prensa tiene deficiencias significativas.

En resumen, la manera en que los medios de comunicación presentan los estudios de neuroimagen influye en la percepción pública de la neurociencia y sus aplicaciones. La simplificación excesiva o la interpretación errónea de los datos pueden terminar generando expectativas poco realistas o reforzando narrativas reduccionistas sobre la conducta humana. En el siguiente apartado, analizo cómo la neurociencia y la inteligencia artificial están convergiendo y qué implicaciones éticas surgen de esta interacción.

5.8. Neuroimágenes e inteligencia artificial

En tiempos recientes, el avance vertiginoso de la inteligencia artificial —especialmente en su vertiente generativa— ha generado no solo entusiasmo tecnológico, sino también interrogantes filosóficos que vale la pena revisar.

Más allá de su impacto tecnológico, resulta pertinente examinar cómo la terminología de la inteligencia artificial generativa ha adoptado términos de la neuropsicología para usarlos de manera metafórica, con el peligro de la cristalización o lexicalización de la metáfora, que difumina

los límites de la acepción original. Conceptos como inteligencia, red neuronal, aprendizaje y alucinación han sido incorporados al lenguaje de la IA, resignificando sus significados y desafiando su conceptualización original.

Por otra parte, de acuerdo con Yang et al. (2024), la inteligencia artificial está transformando el diseño experimental con iRMf, proporcionando herramientas avanzadas que permiten optimizar cada fase del proceso. A manera de ejemplo: a través del aprendizaje automático (ML o *machine learning*), es posible analizar grandes volúmenes de neuroimágenes, identificando patrones que anteriormente eran difíciles de detectar. No obstante, si bien estas herramientas prometen mejorar aspectos técnicos —desde la selección precisa de la ROI hasta la corrección de artefactos experimentales—, conviene preguntarse si esta eficiencia técnica se traduce en una mayor claridad interpretativa.

La aplicación de la IA en neurociencias plantea una serie de desafíos éticos y técnicos que requieren un abordaje riguroso. Una de las principales preocupaciones es la **privacidad y seguridad de los datos**. La enorme cantidad de información sensible derivada de la neuroimagen genera inquietudes respecto a su almacenamiento y posibles usos indebidos. Existe el riesgo de que estos datos sean vulnerables a ataques cibernéticos o a accesos no autorizados, lo que podría comprometer la privacidad de los individuos. (Haselager y Mecacci, 2021)

Otro desafío significativo es el **sesgo algorítmico** (Luxton y Hudlicka, 2024). Los modelos de aprendizaje automático se entrenan con conjuntos de datos que, en muchas ocasiones, reflejan sesgos preexistentes en la sociedad. “La IA solo puede ser tan buena como los datos con los que fue entrenada” (Krauss, 2024, p. 172). Esto puede derivar en diagnósticos inexactos o en decisiones que afecten de manera desigual a distintos grupos poblacionales.

Adicionalmente, se presenta el problema de las "**alucinaciones**" en los modelos de lenguaje, que implica que los sistemas de IA pueden generar información falsa o imprecisa. Este fenómeno pone de manifiesto la ausencia de un mecanismo interno de verificación en estos modelos, lo que podría comprometer la validez de los análisis en neurociencia cognitiva. (Krauss, 2024, p. 158)

A su vez, la opacidad de los algoritmos continúa representando una barrera importante. Muchos modelos de redes neuronales profundas operan como "cajas negras", lo que significa que sus procesos de toma de decisiones no son completamente comprensibles para los investigadores. Esta falta de transparencia dificulta la detección de errores y plantea interrogantes sobre la fiabilidad de los resultados obtenidos. Si no es posible explicar cómo un sistema de IA ha llegado a una conclusión determinada, su utilidad en la investigación científica y en la práctica clínica se ve seriamente limitada. (Krauss, 2024, p. 175)

Cabe señalar, con todo, que la opacidad de los modelos no implica necesariamente una limitación epistémica insalvable. Como plantea Duede (2023), la falta de interpretabilidad de los modelos de aprendizaje profundo solo se vuelve problemática cuando sus resultados se utilizan como afirmaciones que requieren justificación, como sucede en contextos clínicos o decisionales. En cambio, cuando estos modelos se emplean como herramientas exploratorias —por ejemplo, para identificar patrones, generar hipótesis o guiar la atención de los investigadores—, su opacidad puede no ser un obstáculo, siempre que se delimite con claridad su función dentro del proceso de investigación. En este sentido, la clave no reside únicamente en la transparencia del modelo, sino en **el lugar que ocupan sus resultados dentro del circuito epistémico y decisional**.

En este escenario de creciente sofisticación técnica, en el que los modelos se vuelven más complejos y difíciles de interpretar, incluso cuando se los utiliza con fines exploratorios, se plantea una paradoja: cuanto más avanzadas se vuelven las técnicas, más urgente parece la necesidad de límites normativos claros. Pero, ¿es posible regular lo que aún no comprendemos del todo?

La aplicación de inteligencia artificial en el campo de las neuroimágenes abrió, sin duda alguna, nuevas posibilidades en el estudio del cerebro y la mente. Sin embargo, también plantea interrogantes éticos sobre la autonomía, la privacidad y el uso de tecnologías predictivas en la toma de decisiones. En la última sección de este capítulo, abordaremos cómo la neuroimagen influye en la construcción de la identidad personal y qué dilemas éticos emergen de su uso en este ámbito.

5.9. Neuroimágenes e identidad

El uso de las neuroimágenes funcionales y, en particular, de la resonancia magnética funcional, transformó la investigación en neurociencias pero también incidió en nuestra concepción de la identidad personal. Propongo analizar esta idea a partir del concepto de desacoplamiento (*decoupling*), entendido como la separación o independencia entre diferentes niveles de descripción, sistemas o fenómenos. En este caso, hablaré de **desacoplamiento identitario**, es decir, la creciente brecha entre la experiencia subjetiva y las representaciones tecnológicas de la mente.

Daniel Dennett (1991) con su concepto de “yo como centro de gravedad narrativa”, postula que la identidad, lejos de ser una entidad fija, es un proceso dinámico que se construye narrativamente⁵⁵. Sin embargo, esta construcción se ve reconfigurada con el uso de las neuroimágenes, que introducen una representación externa y objetivada del yo.

Por su parte, Nikolas Rose, en *Políticas de la vida. Biomedicina, poder y subjetividad en el siglo XXI* (2012), subraya que, en las últimas décadas, las concepciones sobre la relación mente-cerebro “se han visto moldeadas por desarrollos relativos a estudios por imágenes del cerebro, las neurociencias, la psicofarmacología y la genética de la conducta” (p. 370). Estos avances han contribuido a la consolidación de lo que podría llamarse un *régimen neurocientífico de la identidad*, en el cual la subjetividad es reinterpretada a partir de modelos biomédicos. Sin embargo, Rose advierte que estas transformaciones no son meramente teóricas o metodológicas, sino que afectan a la constitución misma del sujeto: “¿qué ocurre cuando es el yo lo que se encuentra sujeto a transformaciones por parte de la tecnología biomédica, cuando la cognición, la emoción, la volición, el estado de ánimo y el deseo se abren a la intervención?” (p. 370).

Esta pregunta resuena en el contexto del desacoplamiento identitario, ya que sugiere que la identidad personal no solo es interpretada desde una perspectiva neurocientífica, sino que también es reformulada a partir de esta última. La neuroimagen introduce una **representación externa y objetivada del yo**, que modifica la forma en que los individuos se comprenden a sí mismos. Al traducir la actividad cerebral en patrones visuales y asignar correlatos neurales a emociones, pensamientos y acciones, se produce un desajuste entre la experiencia subjetiva del individuo y su interpretación neurocientífica. Pero, al mismo tiempo, esta mediación tecnológica transforma la percepción de la agencia personal, desplazando el foco desde la narrativa subjetiva hacia un modelo basado en mecanismos cerebrales.

En esta perspectiva, Rose señala que la antigua distinción entre lo orgánico y lo funcional—tradicional en la psiquiatría y la psicología—ha ido disolviéndose. Según él, lo funcional (psicológico) queda subsumido en lo orgánico (biológico), hasta el punto de que toda alteración

⁵⁵ Esta perspectiva dialoga con la propuesta de Anthony Giddens en *Modernity and Self-Identity: Self and Society in the Late Modern Age* (1991), donde desarrolla el concepto de “sistemas de reflexividad de la identidad del yo”. Para Giddens, en la modernidad tardía la identidad ya no es determinada por estructuras tradicionales como la familia, la religión o la clase social, sino que se configura de manera reflexiva, a través de un proceso continuo en el que los individuos reinterpretan sus experiencias y ajustan su auto-narrativa. Este proceso no ocurre en el vacío, sino que está atravesado por múltiples mediaciones, es decir, por la influencia de diversos sistemas simbólicos, tecnológicos y sociales que moldean la forma en que las personas se perciben a sí mismas y a los demás. Introduce así la noción de desanclaje (*disembedding*), que describe cómo las relaciones y la identidad han dejado de estar ligadas a un contexto local o tradicional, para volverse globales y mediadas por tecnologías.

psicológica termina siendo concebida como una perturbación fisiológica. En este sentido, cita a Michael R. Trimble en *Biological Psychiatry*: “con nuestro conocimiento actual, la distinción entre ‘orgánico’ y ‘funcional’ se disuelve, despojada de su dualismo cartesiano” (p. 373). Sin embargo, esta supuesta resolución del dualismo cartesiano—que Rose describe más como una *disolución* que como una síntesis dialéctica—no conduce a una superación integradora de la dicotomía, sino a una posición afín al reduccionismo fisicalista.

Ian Hacking (1983) refuerza esta idea al señalar que las herramientas científicas no solo representan la realidad, sino que también la construyen. Esto es especialmente relevante en el caso de la RMf, donde la medición de la actividad cerebral no es inmediata ni directa, sino que pasa por múltiples niveles de procesamiento estadístico antes de generar una imagen interpretable. Este proceso de mediación da lugar a lo que podríamos llamar *desacoplamiento metodológico*: la brecha entre los datos cerebrales y los procesos cognitivos inferidos, mediada por modelos computacionales y decisiones teóricas previas. Como ya se expresó en capítulos anteriores, la construcción de una imagen neurocientífica del pensamiento humano no es un reflejo neutral de la actividad cerebral, sino el resultado de una historia de ideas y tecnologías que han modelado su desarrollo.

Desde una perspectiva neuroética, el uso de las neuroimágenes como prueba científica no solo refuerza la idea de que la identidad personal puede reducirse a estados cerebrales específicos, sino que también puede derivar en lo que podríamos llamar un *desacoplamiento ontológico*. En este caso, el *yo narrativo* se subordina a una interpretación materialista y externalizada del sujeto, donde la actividad cerebral se convierte en el criterio último para definir quiénes somos.

Además, la problematización de la identidad se amplifica en el plano sociocultural. La apropiación de las neuroimágenes por parte del discurso mediático y la opinión pública introduce un *desacoplamiento interpretativo*, en el que la identidad personal es concebida en términos puramente neurobiológicos, sin considerar la complejidad de los factores psicológicos y sociales que la configuran. Con el avance de la inteligencia artificial y los sistemas automatizados de diagnóstico, emergen nuevas capas de mediación en la interpretación de la mente y el comportamiento humano, lo que refuerza aún más la externalización y objetivación del yo.

En definitiva, el impacto de las neuroimágenes en la identidad personal no puede entenderse solo como un avance en la investigación neurocientífica, sino como un proceso que opera en múltiples niveles de desacoplamiento: identitario, metodológico, ontológico e interpretativo. Estas separaciones, lejos de ser meras cuestiones técnicas, tienen implicancias profundas en nuestra manera de comprender la subjetividad, el libre albedrío y la agencia personal en un mundo cada vez más mediado por las neurociencias y la tecnología.

5.10. Discusión y proyecciones

El uso de neuroimágenes ha transformado profundamente la forma en que concebimos la identidad humana, desde la definición de patologías hasta la noción de autenticidad de nuestro pensamiento. Sin embargo, la tendencia a equiparar la actividad cerebral con la identidad personal introduce problemas filosóficos y éticos de gran envergadura. Estas tecnologías, que tienen el potencial de ampliar nuestra comprensión del cerebro y la mente, también pueden ser utilizadas de manera indiscriminada, lo que incrementa el riesgo de vulnerar derechos fundamentales y de reforzar interpretaciones reduccionistas de la conducta humana.

Por otro lado, las neuroimágenes no solo son herramientas científicas, sino que también actúan como dispositivos culturales que moldean la manera en que nos comprendemos a nosotros mismos. La influencia de la neurociencia en la opinión pública y en la toma de decisiones políticas sugiere que estas tecnologías pueden reforzar perspectivas que disminuyen la autonomía y la responsabilidad moral de los individuos. En particular, interpretar las neuroimágenes como una “fotografía de la mente” no solo es inexacto, sino que puede conducir a narrativas simplificadoras que afectan nuestra percepción del ser humano.

En las páginas que siguen retomo las discusiones y reflexiones realizadas hasta aquí, para sintetizar los hallazgos principales y también para explorar sus posibles implicaciones epistemológicas y sociales.

6. Conclusiones

Este trabajo tuvo como objetivo analizar los fundamentos epistemológicos, las implicaciones experimentales y los desafíos neuroéticos de las técnicas de neuroimagen funcional, con el fin de evaluar su impacto en la construcción del conocimiento acerca de la cognición y de la mente. A lo largo de sus páginas, se exploró cómo la evolución histórica de estas técnicas ha estado condicionada por modelos teóricos previos, cómo su interpretación se encuentra mediada por supuestos metodológicos y cómo su uso en diferentes ámbitos plantea desafíos éticos fundamentales.

A lo largo del recorrido realizado, los objetivos planteados al inicio fueron abordados de la siguiente manera:

- Se clasificaron y discutieron críticamente los principales problemas epistemológicos que surgen del uso de técnicas de neuroimagen funcional, especialmente en torno al localizacionismo, la carga teórica del instrumental y la caracterización y validez de las inferencias. Se analizaron también los supuestos filosóficos y conceptuales que subyacen a la interpretación de la RMf, prestando particular atención a las diferencias entre los enfoques modularistas y dinamicistas y a los presupuestos que informan el diseño experimental. Se realizó la crítica de los principios de inserción pura y de sustracción cognitiva y se presentaron modelos alternativos.
- Se esclarecieron las nociones de validez y confiabilidad experimental, mostrando cómo estas operan como restricciones normativas en tensión. Se mostró además cómo el modelo experimental clásico enfrenta limitaciones al intentar capturar fenómenos complejos como la cognición.
- Se exploraron los efectos de estas tecnologías en el campo de la neuroética, destacando su impacto en la configuración de discursos sobre la conciencia, el género, la opinión pública y la predicción de comportamientos, entre otros. Se evaluó además cómo repercuten estos discursos en la opinión pública.

En las secciones que siguen, desarrollo una mirada más detallada a partir de tres ejes de análisis —epistemológico, experimental y neuroético— que se corresponden con los objetivos específicos.

Se distingue, dentro de cada eje de análisis, distintos planos o niveles que permiten organizar de manera más precisa los principales desafíos identificados a lo largo del trabajo.

Cabe aclarar que, si bien esta clasificación facilita la organización y sistematización del material, en muchos casos las fronteras entre un eje y otro, o entre distintos planos de análisis, no son del todo definidas o tienen cierto grado de solapamiento. Las dimensiones epistemológica, experimental y ética se entrelazan frecuentemente en la práctica científica y en la interpretación filosófica y la segmentación adoptada aquí responde a fines analíticos y no ontológicos.

Fundamentos epistemológicos

Los fundamentos epistemológicos de las técnicas de neuroimagen funcional fueron abordados desde cuatro planos analíticos: ontológico, conceptual, lingüístico y filosófico-antropológico. Estas dimensiones fueron principalmente desarrolladas en los capítulos 3 y 4.

En el plano **ontológico** (capítulo →[4.](#)), se planteó la fragilidad de los constructos psicológicos como “memoria”, “atención” o “emoción”, lo que llevó a discutir el proyecto de una ontología cognitiva (→[4.3](#)). Se advirtió que, si dicha ontología se construye exclusivamente a partir de los resultados de la investigación mediante neuroimágenes funcionales y luego, a su vez, se usa para interpretar nuevos datos, se corre el riesgo de caer en un cierre epistémico autorreferencial y en un razonamiento circular (→[4.2-4.3](#)), en el que la neurociencia cognitiva define sus propios marcos explicativos sin contrastarlos con otras perspectivas.

Esto plantea el riesgo de que la ontología cognitiva se consolide como un marco cerrado, en el que la neurociencia no solo redefine, sino que sustituye los conceptos psicológicos sin permitir un diálogo interdisciplinario. Esto comprometería la posibilidad de una psicología cognitiva autónoma, con capacidad explicativa propia, y podrían conducir a un cierre epistémico, en el que la complejidad de lo mental queda subsumida en un único nivel de análisis, clausurando el diálogo interdisciplinario que exige la comprensión de la cognición (→[4.3.2](#)). La inconmensurabilidad parcial entre ambas disciplinas que sostengo en este trabajo (→[2.1.7](#); [3.8](#)) impediría una traducción directa de los constructos psicológicos a términos neurobiológicos sin alteraciones en su significado o pérdida de sentidos a la hora de describir la cognición.

Este peligro se vincula, además, con formas de reduccionismo fisicalista o eliminativista que tienden a identificar los procesos mentales con mecanismos neuronales, disolviendo así su especificidad conceptual.

En el aspecto **conceptual** (capítulos →[3.](#) y [4.](#)), la revisión histórica permitió reconocer cómo la interpretación de las neuroimágenes funcionales está condicionada por modelos previos sobre la localización cerebral, partiendo desde enfoques tempranos como la frenología hasta las perspectivas funcionalistas más recientes. Esta evolución teórica continúa influyendo en los debates actuales entre enfoques modularistas y dinámicos y en cómo se entienden los constructos psicológicos investigados (→[2.1-2.2](#)).

En este marco, se examinó el difícil equilibrio entre la necesidad de simplificar para poder modelar fenómenos mentales y la necesidad de preservar su complejidad inherente. Esta tensión se manifiesta tanto en la elección de modelos teóricos como en la interpretación de los resultados obtenidos mediante neuroimágenes funcionales. Un ejemplo paradigmático de este problema es la multifuncionalidad de las regiones cerebrales (→[3.2](#)), que impide establecer correspondencias unívocas entre estructuras cerebrales y funciones cognitivas específicas. Esta dificultad se acentúa por la inestabilidad de los constructos psicológicos (→[4.2](#)), que carecen muchas veces de una definición operativa clara y estable a lo largo de los estudios. De allí surgen los problemas de correspondencia uno a uno entre categorías cognitivas y patrones de activación cerebral, lo cual pone en cuestión algunos de los supuestos implícitos en el uso de técnicas como la RMf. En este contexto, cobran relevancia las tensiones entre enfoques modularistas, que presuponen cierta localización funcional estática, y modelos dinámicos de redes cerebrales, que postulan una organización distribuida, plástica y contextual de la actividad cognitiva (→[3.2](#), [4.2](#)).

Es fundamental recordar, aquí, como uno de los principios guía, que correlación no es causalidad: la coincidencia entre la activación de una región cerebral y un proceso cognitivo no implica necesariamente que dicha región sea la causa de dicho proceso. Sin una diferenciación clara entre ambos conceptos, se corre el riesgo de asumir relaciones causales que pueden no existir, lo que compromete la validez de las inferencias neurocientíficas.

En el plano **del lenguaje** (→[3.4](#)), se señalaron las dificultades asociadas al uso de ciertas metáforas ampliamente difundidas en el campo de la neuroimagen funcional, tales como “mapa cerebral” o “fotografía de la mente”. Estas expresiones, si bien resultan efectivas desde el punto de vista comunicativo, pueden inducir a interpretaciones erróneas al sugerir una correspondencia directa, precisa y estable entre la imagen producida y los procesos mentales subyacentes. La pregnancia visual de estas metáforas contribuye a reforzar la ilusión de transparencia empírica y puede oscurecer la naturaleza estadística, inferencial y construida de las imágenes obtenidas por RMf.

Además, se problematizó en particular la falacia mereológica (→[5.7](#)), entendida como la reificación impropia de una relación metonímica. Atribuir a una región cerebral propiedades que

solo tendrían sentido en el sujeto en su conjunto puede derivar en un error ontológico, al validar de manera implícita posiciones reduccionistas que disuelven la agencia, la intencionalidad o la experiencia transformándolas discursivamente en meras funciones localizadas.

En cuanto al nivel **filosófico-antropológico** (→[5.2-5.9](#)), se analizó el modo en que la investigación en neuroimágenes puede promover una reducción de la subjetividad a patrones de activación cerebral. Esta tendencia contribuye a una transformación profunda en la autocomprensión del sujeto, al desplazar la noción de experiencia hacia la idea de representación cerebral. En este marco, fenómenos complejos como la intencionalidad, la conciencia o el sentido de agencia tienden a ser reformulados en clave neurobiológica, lo que empobrece la comprensión fenomenológica y contextual de la mente.

Este giro reduccionista plantea interrogantes filosóficos fundamentales sobre la libertad, la responsabilidad moral y la identidad personal. La posibilidad de interpretar las decisiones humanas como simples productos de activaciones cerebrales específicas pone en cuestión nociones como el libre albedrío y la agencia, así como la continuidad narrativa del yo. De este modo, la neuroimagen funcional no solo ofrece nuevos modos de investigar la mente, sino que incide directamente sobre los marcos ontológicos y normativos con los que pensamos lo humano (→[5.9](#)).

Las neuroimágenes no son solo herramientas científicas, sino también, en términos foucaultianos, constituyen dispositivos conceptuales que moldean nuestra manera de pensar sobre la mente y el cerebro. Estos dispositivos no solo producen conocimiento, sino que también regulan las formas de interpretación legítimas y las que no lo son. Así, la neurociencia contemporánea puede llevar a reforzar una visión reduccionista y fisicalista de la mente, invisibilizando otras formas de comprender la relación mente-cerebro.

El análisis epistemológico permitió clasificar y discutir críticamente los principales problemas asociados a la interpretación de las neuroimágenes, atendiendo a la carga teórica de la observación y a la inconmensurabilidad parcial entre psicología y neurociencia. Se mostró que la neurociencia no solo intenta validar los constructos psicológicos, sino que los transforma, lo que genera discontinuidades en la integración de ambas disciplinas. Con ello, se verificó el objetivo de analizar los problemas epistemológicos vinculados a la identificación de correlatos neuronales de procesos cognitivos-conductuales, mostrando que esta relación no es directa ni libre de supuestos teóricos.

Principios experimentales

Los principios experimentales fueron abordados desde dos planos: técnico y metodológico. Ambos fueron principalmente desarrollados en los capítulos 2 y 4.

En el **plano técnico** (→ [2.2-2.2.6](#)), se expusieron las limitaciones de la RMf en cuanto a resolución espacial y temporal, así como la variabilidad intersujeto. Estas condiciones afectan la precisión de los datos obtenidos y dificultan la atribución unívoca de funciones a regiones específicas del cerebro. Sin embargo, los desafíos técnicos no se reducen al momento de la adquisición de la señal, sino que atraviesan todo el proceso de transformación desde la señal fisiológica hasta la generación de la imagen estadística.

Este proceso incluye etapas complejas de posprocesado, como la corrección de movimiento, el suavizado espacial y la normalización a un espacio anatómico de referencia mediante plantillas estándar del cerebro, como MNI o Talairach. Cada una de estas operaciones implica decisiones metodológicas que condicionan el resultado final y pueden introducir distorsiones si no son controladas críticamente.

En el **nivel metodológico** (→ [4.4](#)), se señalaron los problemas vinculados al diseño experimental, como el uso de muestras pequeñas y poco representativas, el *p-hacking* y la selección sesgada de regiones cerebrales. Estos factores limitan la potencia estadística y la validez generalizable de muchos estudios y contribuyen a la reproducción de sesgos en la publicación. Además, se discutieron críticamente los principios de inserción pura y sustracción cognitiva, que han guiado históricamente el diseño de tareas en neuroimagen funcional, aunque enfrentan cada vez mayores cuestionamientos teóricos y empíricos.

Se destacó también la distinción entre inferencia directa e inferencia inversa en la interpretación de datos de RMf (→ [3.6](#)), subrayando los desafíos metodológicos y epistemológicos relacionados con la segunda. No obstante, como señala Klein, la utilidad de la RMf no se agota en estos dos tipos de inferencias, sino que también radica en su capacidad para generar hipótesis exploratorias, identificar patrones de conectividad funcional y refinar modelos teóricos en neurociencia cognitiva (→ [3.6.3](#)).

Asimismo, se discutieron los desafíos epistemológicos asociados tanto a los modelos bayesianos como a los modelos frecuentistas o tradicionales ([3.6.4](#)), observando cómo ambos enfoques inferenciales dependen de supuestos previos que guían la interpretación estadística de los datos.

Se destacó también la distinción entre inferencia directa e inversa, así como los desafíos epistemológicos asociados tanto a los modelos bayesianos como a los modelos frecuentistas o

tradicionales. Ambos enfoques inferenciales dependen de supuestos previos que, si no son adecuadamente explicitados o justificados, pueden conducir a conclusiones sesgadas o sobregeneralizadas.

En este marco, se destacó la tensión estructural entre confiabilidad y validez, especialmente la validez ecológica, dado que los procedimientos experimentales tienden a simplificar artificialmente los fenómenos estudiados (→ [4.4.2](#)), limitando así su relevancia para contextos naturales.

El análisis realizado permitió esclarecer cómo los supuestos técnicos y metodológicos condicionan la obtención e interpretación de los datos experimentales en neuroimágenes funcionales. Se problematizaron procesos técnicos aparentemente neutrales, como la corrección de movimiento o la normalización mediante plantillas anatómicas estándar (MNI o Talairach), en tanto implican elecciones metodológicas que no son triviales ni automáticas. Cada una de estas decisiones puede introducir distorsiones o sesgos en el resultado final si no son controladas críticamente. Algo similar sucede con aspectos propiamente metodológicos, como el tamaño reducido o la baja representatividad de las muestras, la selección sesgada de regiones cerebrales o la aplicación acrítica de los principios de inserción pura y sustracción cognitiva. El análisis permite inferir que todas estas decisiones inciden sobre la validez y la confiabilidad de los resultados. De este modo, se cumplió con el objetivo planteado de clarificar las implicaciones experimentales que atraviesan la utilización de RMf en la investigación cognitiva, destacando sus limitaciones y enfatizando la necesidad de abordajes críticos e informados en su aplicación.

Desafíos neuroéticos

Los desafíos neuroéticos se abordaron desde cuatro planos: normativo, epistémico, discursivo y subjetivo. Estas dimensiones fueron desarrolladas principalmente en el capítulo 5.

En el **plano normativo** (→ [5.2-5.6](#)), se discutieron los dilemas que surgen del uso de neuroimágenes en contextos clínicos, educativos, jurídicos y de seguridad, donde los resultados obtenidos pueden adquirir un estatus probatorio que excede sus posibilidades reales de interpretación. Se abordaron particularmente sus aplicaciones en entornos diagnósticos y de predicción de conductas, lo que plantea riesgos importantes si se soslaya la complejidad del fenómeno cognitivo o se simplifican las inferencias sobre la intención o la verdad.

Asimismo, se discutieron los debates sobre la privacidad mental —es decir, sobre los límites éticos de acceder, registrar o interpretar la actividad cerebral— y la necesidad de garantizar consentimiento informado, especialmente en poblaciones vulnerables. También se alertó sobre el uso potencial de estas tecnologías para justificar decisiones judiciales o políticas públicas, sin una adecuada contextualización científica ni filosófica. En todos estos casos, se subrayó la importancia de establecer marcos regulatorios sensibles y específicos, que eviten aplicaciones abusivas, simplificadoras o discriminatorias del conocimiento neurocientífico (→[5.1](#)).

En el **nivel epistémico** (→[5.7](#)), se puso de relieve la percepción infundada de que las neuroimágenes constituyen pruebas objetivas e irrefutables, cuando en realidad están atravesadas por modelos teóricos, decisiones metodológicas y tratamientos estadísticos complejos. Esta apariencia de objetividad se ve reforzada por la fuerza visual de las imágenes y por su presentación pública en formatos que sugieren evidencia directa y transparente.

Se discutieron los problemas derivados de la carga teórica de la observación y del instrumental y de los métodos de procesamiento de datos, subrayando que la neuroimagen no es una 'fotografía del pensamiento', sino una representación construida a partir de modelos estadísticos y conceptuales.

Se destacó, además, la dificultad de comunicar adecuadamente el grado de incertidumbre asociado a los datos obtenidos por resonancia magnética funcional. Las inferencias realizadas dependen de múltiples supuestos, umbrales estadísticos y procedimientos de procesamiento que rara vez son explicitados fuera del ámbito técnico. En este contexto, se problematizó también la idea de “ver la mente” en el cerebro como una metáfora totalizante, que clausura otras formas de comprender lo mental y refuerza una lectura reductora de los procesos subjetivos y sociales implicados en la cognición.

En cuanto al **plano discursivo** (→[5.7](#)), se analizó cómo las neuroimágenes —al circular en medios masivos y espacios de divulgación— contribuyen a consolidar narrativas deterministas o biologicistas sobre el comportamiento humano. Estas narrativas tienden a presentar explicaciones cerebrales como más legítimas o concluyentes que otras formas de comprensión psicológica, social o cultural, lo cual refuerza una jerarquía epistémica en favor de lo neurocientífico.

Se discutió, en este sentido, el papel de la divulgación científica en la construcción de estereotipos acerca del género, la inteligencia o la conducta, así como el uso retórico del prefijo “neuro” para conferir autoridad a afirmaciones que carecen de suficiente respaldo empírico o conceptual. Se advirtió también que ciertos discursos mediáticos y de autoayuda se apoyan acríticamente en hallazgos neurocientíficos, promoviendo una imagen del cerebro como fundamento último de toda

explicación humana. En este contexto, la neurociencia funciona como un soporte epistémico que legitima, a veces de forma implícita, discursos normativos o políticos, desplazando otras voces y complejidades del debate público.

Por otra parte, la neurociencia ha adquirido un estatus privilegiado en la validación del conocimiento, imponiendo jerarquías que favorecen explicaciones cerebrales sobre otras formas de comprensión de la mente (→[5.7](#)). En ámbitos como la divulgación científica y la toma de decisiones institucionales, la presencia de imágenes cerebrales puede conferir legitimidad a discursos que, sin este respaldo visual, serían sometidos a mayor escrutinio. Frente a esta tendencia, es clave desarrollar una actitud crítica que evite que el discurso neurocientífico se convierta en un marco epistémico excluyente.

En el **nivel subjetivo** (→[5.9](#)), se advirtió que estas tecnologías pueden incidir en la autocomprensión de las personas, al reforzar visiones de sí mismas ancladas en mecanismos cerebrales, desplazando así dimensiones fenomenológicas, históricas o sociales de la experiencia. Esta transformación implica una progresiva reducción de la identidad a patrones neuronales, donde el sujeto es concebido menos como agente intencional que como producto de configuraciones cerebrales observables.

El uso de neuroimágenes en distintos ámbitos sociales y jurídicos ha planteado riesgos éticos significativos. En este capítulo, se analizó cómo la aplicación de estas técnicas en la detección de mentiras, la predicción del comportamiento y la construcción de la identidad, entre otros ámbitos, puede terminar reforzando un determinismo neurocientífico que minimiza la agencia y la autonomía individual. Además, hemos visto que la neurociencia no solo describe el cerebro, sino que también transforma la manera en que los individuos se comprenden a sí mismos y toman decisiones sobre sus vidas.

Asimismo, se señaló un fenómeno de externalización de la agencia, por el cual el cerebro —más que la persona— pasa a ser el verdadero sujeto de la acción. Frases como “mi cerebro decidió por mí” ejemplifican esta inversión, que puede tener efectos significativos sobre el sentido de responsabilidad moral y la narrativa del yo. En este escenario, lo que está en juego no es solo una teoría sobre la mente, sino el modo en que los individuos se comprenden, se explican y se posicionan como actores en el mundo social (→[5.9](#)).

El estudio de los desafíos neuroéticos permitió examinar cómo las neuroimágenes no solo tienen implicancias metodológicas, sino que también afectan la manera en que se conceptualizan la identidad, la agencia y la subjetividad. Se abordó la influencia del discurso neurocientífico en la

opinión pública, la consolidación de la neuroautoridad y el riesgo de reduccionismo neurobiológico en la autocomprensión del sujeto. Con ello, se verificó el objetivo de analizar las implicancias sociales, jurídicas y culturales de las neuroimágenes, mostrando cómo estas tecnologías no solo describen la mente, sino que también configuran nuevas formas de interpretarla.

Los tres ejes de análisis que acabamos de abordar permiten comprender la resonancia magnética funcional no como una herramienta meramente técnica, sino como un dispositivo complejo cuya interpretación exige también un enfoque filosófico atento a sus fundamentos, sus prácticas y sus proyecciones sociales. Considerar, además, los distintos planos implicados en cada dimensión permite responder de manera más clara y ordenada a los objetivos formulados y también abrir nuevas preguntas en el cruce entre filosofía, psicología cognitiva y neuroimagen funcional.

Hacia una hermenéutica de la neuroimagen funcional

A modo de cierre, considero necesario señalar que el aporte de la filosofía de las neuroimágenes no se reduce a indicar límites metodológicos, falacias inferenciales o riesgos epistemológicos asociados al uso de la resonancia magnética funcional. Por el contrario, a partir de los análisis aquí presentados, su contribución más relevante reside en la elaboración de una perspectiva constructiva, enfocada a fortalecer la legitimidad epistemológica de los resultados y a promover prácticas científicas más rigurosas y responsables.

La propuesta es, entonces, avanzar hacia una **hermenéutica rigurosa de la neuroimagen funcional**. Esta propuesta no supone un desplazamiento hacia enfoques hermenéuticos de tradición continental ni una teoría del significado de corte subjetivista, sino una elucidación metateórica de segundo orden, orientada a clarificar las condiciones bajo las cuales las prácticas de neuroimagen producen significado y adquieren estatus epistémico. Ello exige reconocer que la neuroimagen es un texto complejo, en el sentido de un artefacto epistémico, cuyo significado se construye a través de una larga cadena de interpretaciones, supuestos teóricos y decisiones estadísticas. La noción de “texto” remite aquí a la explicitación de la mediación inferencial y conceptual que interviene a lo largo de todo el proceso de producción de la neuroimagen, desde el diseño experimental y el registro de los datos hasta las conclusiones cognitivas que se derivan de ellos.

Esta perspectiva hermenéutica constituye un **programa de investigación abierto** que permite trazar nuevas líneas de indagación en tres dimensiones fundamentales, ya que, como todo sistema de signos, su interpretación requiere un análisis en tres niveles interdependientes: el de los significados que se le atribuyen (semántica), el de las reglas técnicas que gobiernan su producción (sintáctica) y el de los usos y efectos que genera en el mundo social (pragmática). Estas distinciones no describen niveles ontológicos de la realidad cognitiva ni pretenden ofrecer una teoría general del significado, sino que funcionan como herramientas analíticas útiles para organizar los distintos tipos de problemas que emergen en la práctica científica. No resulta casual que los tres niveles problemáticos identificados en este trabajo —el conceptual, el metodológico y el ético— se correspondan estructuralmente con estas dimensiones hermenéuticas. Esta convergencia permite reinterpretar el recorrido de esta tesis como el análisis sistemático de las condiciones de producción de sentido de la neuroimagen funcional.

A continuación, delinearé cómo este programa hermenéutico abre líneas de investigación futura en estas tres dimensiones:

1. Hermenéutica del sentido (dimensión semántica)

El primer desafío es semántico: ¿de qué hablamos cuando vemos una activación cerebral? Como hemos visto, antes de asociar patrones de activación a estados mentales o a procesos cognitivos, resulta necesario interrogarse acerca de los constructos psicológicos involucrados. Si los conceptos de base son inestables o se encuentran insuficientemente delimitados, el mapa cerebral carece de una referencia clara y la atribución de significado a la imagen se vuelve problemática.

Este problema se hace particularmente visible en el pasaje desde los modelos que utilizaban la localización anatómica como criterio semántico hacia los enfoques basados en redes funcionales. Ambos modelos coexisten actualmente en la práctica neurocientífica y cumplen funciones heurísticas distintas, lo que refuerza la necesidad de explicitar los criterios semánticos que orientan la atribución de significado a la activación en cada caso, precisamente para resistir visiones reduccionistas que pretenden abandonar la autonomía explicativa de la psicología cognitiva en favor de mecanismos moleculares exhaustivos.

En este contexto, el desplazamiento desde ROI estructurales hacia patrones distribuidos de coactivación no simplifica la interpretación de la activación, sino que la complejiza, en tanto desplaza la referencia desde regiones relativamente estables hacia entidades funcionales cuya identificación depende de supuestos teóricos y decisiones analíticas previas. Esta complejización vuelve aún más problemática la relación entre psicología cognitiva y ontología cognitiva. Asimismo, puede dar lugar a correlaciones espurias y a esquemas de circularidad causal, cuando

los compromisos ontológicos guían la interpretación de los datos y los resultados son luego utilizados para reafirmar esos mismos compromisos.

La investigación futura deberá, por tanto, evitar la acumulación de datos descontextualizados y enfocarse en una articulación crítica entre psicología cognitiva y ontología cognitiva. Lejos de asumir la eliminación de la psicología, se trata de defender una coevolución de ambas disciplinas, en la que los hallazgos neurobiológicos puedan refinar los conceptos psicológicos sin disolver la riqueza fenomenológica de la mente humana.

2. Hermenéutica de la producción (dimensión sintáctica)

El segundo nivel es sintáctico: ¿bajo qué reglas se construye la imagen? Ante la crisis de replicabilidad y la heterogeneidad de los protocolos, se impone desarrollar estándares metodológicos que transparenten la “gramática” de la producción de datos. Aquí, el desafío más interesante reside en las innovaciones técnicas que buscan romper la sintaxis rígida del laboratorio, como la espectroscopía funcional de infrarrojo cercano (fNIRS).

Técnicas como la fNIRS portátil (*wearable*), que permiten mediciones en entornos naturales con mayor validez ecológica —desde estudios en niños en movimiento hasta monitoreo de fatiga cognitiva en contextos cotidianos—, obligan a redefinir la “gramática” de la evidencia: ¿qué precisión inferencial sacrificamos por realismo contextual y cómo la filosofía de la técnica puede equilibrar ambos sin comprometer el rigor? La filosofía de la técnica deberá vigilar que este cambio de sintaxis experimental no sacrifique aún más el rigor inferencial en pos de una supuesta “naturalidad”.

3. Hermenéutica de la recepción (dimensión pragmática)

Finalmente, el nivel pragmático se pregunta por los efectos: ¿qué hace la imagen con nosotros? La interpretación de neuroimágenes en contextos clínicos, jurídicos y mediáticos encierra riesgos de sobreinterpretación que requieren un enfoque de “ciencia responsable”. Esto implica diseñar marcos normativos que protejan la privacidad mental y eviten la vigilancia cognitiva indiscriminada.

Un caso particularmente revelador en esta dimensión es el de la así llamada “resonancia magnética funcional en tiempo real” (rt-fMRI, por su sigla en inglés, también denominada *closed-loop fMRI* o RMf de bucle cerrado), a menudo usada en técnicas de *neurofeedback*. A diferencia de la RMf tradicional, permite al sujeto modular activamente su propia actividad cerebral —por ejemplo, reducir la activación asociada al *craving* en adicciones (como tabaquismo o trastorno por uso de videojuegos), aumentar la actividad en regiones de control emocional en depresión resistente, o

incrementar la atención selectiva en adultos mayores— mediante retroalimentación inmediata y adaptativa.

No obstante, pese a su denominación, sabemos que la rt-fMRI no constituye un acceso inmediato o simultáneo a la actividad neuronal, sino una representación mediada por retardos temporales, filtrados estadísticos y decisiones algorítmicas. Su interés filosófico reside en los efectos pragmáticos y performativos que introduce al acoplar registro, procesamiento e intervención en un mismo dispositivo técnico.

Por ello, sin desconocer la importancia de los marcos regulatorios, la investigación futura debe centrarse también en la performatividad de estas imágenes. Se trata de investigar no solo cómo la técnica lee al sujeto, sino cómo su circulación social y aplicación clínica reconfiguran nuestra autocomprensión, agencia y responsabilidad moral —especialmente en una era donde las intervenciones *closed-loop* prometen personalización terapéutica pero arriesgan narrativas reduccionistas que narran al humano primordialmente como biología modulable.

Desde esta perspectiva, el aporte de la filosofía no se sitúa en el plano de la explicación empírica, sino en un nivel metateórico compatible con la filosofía analítica de la ciencia, orientado a distinguir entre datos, inferencias y supuestos conceptuales.

En definitiva, el objetivo de esta investigación no ha sido invalidar el inmenso valor heurístico y clínico de las neuroimágenes, sino mostrar la necesidad de una prudencia interpretativa y de una reflexividad crítica sostenidas. Este análisis nos recuerda que, entre la actividad cerebral registrada y la comprensión de la mente humana, siempre existirá un espacio ineludible para la interpretación. Lejos de constituir un déficit, ese espacio es una condición constitutiva del conocimiento en este campo. Y es precisamente allí donde la filosofía se vuelve no solo relevante, sino indispensable.

7. Apéndice

Sitios web (colaborativos) sobre psicología cognitiva y neuroimágenes

https://www.cognitiveatlas.org/	Dirigido por Russell Poldrack.
http://www.humanconnectomeproject.org/	Dependiente del Instituto de Informática y Neuroimagen Mark y Mary Stevens de la University of Southern California.
https://www.brainmap.org/	Base de datos de experimentos de RMf. Proyecto dependiente del Instituto de Investigación de Imágenes del Centro de Ciencias de la Salud de la Universidad de Texas en San Antonio.
https://portal.brain-map.org/	Del Allen Institute de Seattle.
https://thevirtualbrain.org	Software gratuito de simulación cerebral.
https://www.nitrc.org/	Plataforma de recursos colaborativos para la neuroimagen y la neurociencia computacional.
https://openneuro.org/	Plataforma abierta para el intercambio de datos de neuroimágenes.
https://neurovault.org/	Plataforma para compartir mapas de activación cerebral.
https://www.humanconnectome.org/	Ofrece acceso a conjuntos de datos de neuroimágenes funcionales, estructurales y de conectividad.
https://www.incf.org/	Promueve la colaboración internacional en neuroinformática y proporciona una plataforma para compartir datos y herramientas.
http://cogpo.org/	Proyecto de Ontología del Paradigma Cognitivo.
https://www.expfactory.org/	Experimentos de comportamiento de código abierto.
https://cntrics.ucdavis.edu/	Iniciativa "Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia" (Investigación en neurociencia cognitiva para mejorar la cognición en la esquizofrenia), que busca desarrollar enfoques de medición basados en la neurociencia cognitiva, social y afectiva, para implementar tratamientos que mejoren las funciones cognitivas deterioradas en individuos con esquizofrenia.
https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc	Página oficial del Research Domain Criteria (RDoC) del Instituto Nacional de Salud Mental de los Estados Unidos (NIMH, por sus siglas en inglés). El RDoC es un marco de investigación diseñado para estudiar los trastornos mentales de una manera que integra múltiples niveles de información.
https://neurosynth.org/ https://compose.neurosynth.org/	Neurosynth es una plataforma automatizada que sintetiza datos fMRI de múltiples estudios publicados. Neurosynth Compose permite realizar metaanálisis de neuroimagen sin conocimientos de programación avanzada.

8. Referencias bibliográficas

- Abelson, R. P. (1995). *Statistics as principled argument*. Lawrence Erlbaum Associates.
- Abelson, R. P. (1997). On the surprising longevity of flogged horses: Why there is a case for the significance test. *Psychological Science*, 8(1), 12–15. <http://www.jstor.org/stable/40062838>
- Abrahamsen, A. y Bechtel, W. (2012). From Reactive to Endogenously Active Dynamical Conceptions of the Brain. In *Philosophy of Behavioral Biology*. K. S. Plaisance & T. A. C. Reydon (Eds.) (pp. 329–366).
- Alsina, J. (1970). Hipócrates. Sobre la enfermedad sagrada. *Boletín del Instituto de Estudios Helénicos*, 4(1), 87-96.
- Ardila, A. A. y Solís, F. O. (2008). Desarrollo histórico de las funciones ejecutivas. *Revista Neuropsicología, Neuropsiquiatría y Neurociencias*, 8(1), 1-21.
- Ardila, R. y Bunge, M. (2002). *Filosofía de la psicología*. Siglo XXI Editores.
- Arias Gallegos, W. L. (2021). Antecedentes, desarrollo y consolidación de la psicología cognitiva: un análisis histórico, *Revista Tesis Psicológica*, vol. 16, núm. 2, 2021, enero-junio, pp. 172-198. Fundación Universitaria Los Libertadores DOI: <https://doi.org/10.37511/tesis.v16n2a9>
- Aristóteles. (1994). *Metafísica* (Libro A, 980a21–27) [Ed. bilingüe]. Madrid: Gredos.
- Aristóteles. (2010). *Obra biológica: De Partibus Animalium, De Motu Animalium, De Incessu Animalium* (R. Bartolomé, Trad.; A. Marcos, Introd. y notas). Madrid: Luarna Ediciones.
- Armony, J. L., Trejo-Martínez, D. y Hernández, D. (2012). Resonancia Magnética Funcional (RMf): principios y aplicaciones en Neuropsicología y Neurociencias Cognitivas. *Neuropsicología Latinoamericana*, 4(2), 36-50. <https://dx.doi.org/10.5579/rnl.2012.010>
- Barberis, S. D. (2012). Un análisis crítico de la concepción mecanicista de la explicación. *Revista Latinoamericana de Filosofía*, Vol. XXXVIII N° 2, 233-265.
- Barberis, S. D. (2024). La revolución de Geschwind en afasiología y sus implicaciones para la evolución del lenguaje. En F. Giri, A. Gonzalo, G. Parera y D. Blanco (Eds.), *Darwin y la evolución del lenguaje humano* (pp. 205-222). Ediciones UNL.
- Barberis, S. D., Ginnobili, S., y Roffé, A. J. (2025). *Cuestionando el dogma experimentalista: Incommensurabilidad empírica en la neurociencia temprana*. Estudios de Filosofía, 72. <https://doi.org/10.17533/udea.ef.359301>
- Barberis, S. D., Roffé, A. y Ginnobili, S. (2022). Explicación funcional y análisis sistémico: el caso de la afasiología moderna. *Cuadernos Filosóficos. Segunda época*, (18), 69–92.
- Barberis, S. D. y Wright, C. (2022). *History of Behavioral Neurology*. En S. Della Sala (Ed.), *Encyclopedia of behavioral neuroscience* (Vol. 1, pp. 1–13). Elsevier.

- Bareither, I. (2014). Historia de las imágenes cerebrales. *Mente y cerebro*, 69, 41-48.
- Beaney, M. (2021). Analysis. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition). Stanford University.
<https://plato.stanford.edu/archives/sum2021/entries/analysis/>
- Beaumont, J. G. (2008). *Introduction to neuropsychology*. Guilford Press.
- Bennett, M. R. y Hacker, P. M. S. (2022). *Philosophical foundations of neuroscience* (primera edición publicada en 2003). John Wiley & Sons.
- Bennett, C. M. y Miller, M. B. (2010). How reliable are the results from functional magnetic resonance imaging? *Annals of the New York Academy of Sciences*, 1191, 133–155.
<https://doi.org/10.1111/j.1749-6632.2010.05446.x>
- Bickle, J. (2003). *Philosophy and neuroscience: A ruthlessly reductive account*. Springer Science & Business Media.
- Bickle, J. (2006b). Reducing mind to molecular pathways: explicating the reductionism implicit in current cellular and molecular neuroscience. *Synthese*, 151(3), 411-434.
[doi:10.1007/s11229-006-9015-2](https://doi.org/10.1007/s11229-006-9015-2)
- Bickle, J., Mandik, P. y Landreth, A. (2019). The philosophy of neuroscience. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition). Stanford University.
<https://plato.stanford.edu/archives/fall2019/entries/neuroscience/>
- Bowers, J. S. y Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological bulletin*, 138(3), 389.
- Bradford Hill, A. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*, 58(5), 295-300.
- Branca, M. I., Ramírez, A. O. y Vilatta, M. E. (2015). Modelos de explicación en psicología cognitiva y neurociencias. *Anuario de Investigaciones de la Facultad de Psicología*. UNC. Año 2015, Vol. 2, N°1, 176-191
- Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth. Fuente:
<https://wellcomecollection.org/works/vrnkkxtj>
- Brodmann, K. y Garey, L. J. (2005). *Brodmann's Localisation in the cerebral cortex*. Springer.
- Bunzl, M., Hanson, S. y Poldrack, R. (2010). An Exchange about Localism. En Hanson, Stephen José y Bunzl, Martin (eds.) (2010) *Foundational Issues in Human Brain Mapping*, The MIT Press Cambridge.

- Burnston, D. C. (2022). Cognitive Ontologies, Task Ontologies, and Explanation in Cognitive Neuroscience. En J. Bickle, C. F. Craver y A.-S. Barwich (Eds.), *The Tools of Neuroscience: Experimental and philosophical perspectives* (pp. 259-283). Routledge.
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J. y Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. En *Nat. Rev. Neurosci.* 14, 365–376.
- Carver, R. P. (1978). The case against statistical significance testing. *Harvard Educational Review*, 48, 378-399.
- Chalmers, A. F. (1999). *¿Qué es esa cosa llamada ciencia?* (3.^a ed. corregida y aumentada). Siglo XXI de España Editores.
- Ciccia, L. (2017). La ficción de los sexos: Hacia un pensamiento Neuroqueer desde la Epistemología Feminista (Tesis de doctorado). Universidad de Buenos Aires.
- Ciccia, L. (2021). Dimorfismo sexual, ¿natural? Una reinterpretación crítica de las diferencias biológicas. *Revista Bioética*, 29, 66-75.
- Clark, L. A. y Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7, 309–19.
- Cronbach, L. J. y Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- De Quervain, D. J. et al. (2004) The neural basis of altruistic punishment. *Science* 305, 1254–1258.
- Demertzi, A. y Laureys, S. (2015) Detecting Levels of Consciousness. En Clausen, Jens; Levy, Neil (eds.) *Handbook of Neuroethics*. Springer Dordrecht Heidelberg, New York-London.
- Dennett, D. (1991). *Consciousness Explained*. Little, Brown & Co.
- Díez, J. A., y Moulines, C. U. (2017). *Fundamentos de filosofía de la ciencia*. Barcelona: Ariel.
- Domènech, E. (1977). *La frenología: análisis histórico de una doctrina psicológica organicista*. Universidad de Barcelona.
- Duede, E. (2023). Deep learning opacity in scientific discovery. *Philosophy of Science*, 90(5), 1089–1099. <https://doi.org/10.1017/psa.2023.8>
- Farah, M. J. (2005). Neuroethics: the practical and the philosophical. En *Trends in cognitive sciences*, 9(1), 34–40. <https://doi.org/10.1016/j.tics.2004.12.001>
- Farah, M. J. (2010). Neuroethics: An Overview. In M. Farah (Ed.), *Neuroethics, An Introduction with Readings*. Cambridge, MA: MIT Press, 2010.
- Farah, M. J. (2012). Neuroethics: The Ethical, Legal, and Societal Impact of Neuroscience. *Annu. Rev. Psychol.* 2012. 63:571–91

- Farah, M. J., Gillihan, S. (2013). Neuroimaging in clinical psychiatry. En Chatterjee, A., Farah, M. J. (eds.) *Neuroethics in Practice*. Oxford University Press, New York.
- Faro, S. H., Mohamed, F. B. (Eds.). (2010). *BOLD fMRI: A guide to functional imaging for neuroscientists*. Springer Science & Business Media.
- Field, A. P. y Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, 63(3), 665-694.
- Filippi, M. (2016). *fMRI techniques and protocols*. M. Filippi (Ed.). Milan, Italy: Humana Press.
- Fodor, J. (1997). Special Sciences: Still Autonomous After All These Years. *Philosophical perspectives: Mind, Causation, and World* 11 (1997): 149–163.
- Fontana, H. (2010). Resumen del libro de "Teoría de la localización comparativa de la corteza cerebral" expuesta en sus principios en base a la estructura celular. *Revista argentina de neurocirugía*, 24 (Supl. 1), S1-S22. Recuperado en 24 de enero de 2023, de http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1850-15322010000400004&lng=es&tlng=es.
- Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S. y Dolan, R. J. (1996). The Trouble with Cognitive Subtraction. *Neuroimage* 4.2 (1996): 97–104.
- Friston, K. J., Rotshtein, P., Geng, J. J., Sterzer, P. y Henson, R. N. (2006). A critique of functional localisers. *Neuroimage*, 30(4), 1077-1087.
- García de Lucas, Á. (09/09/2020). La breve, breve historia del estudio del cerebro a través de las imágenes. *Psicología y neurociencia*. <https://www.investigacionyciencia.es/blogs/psicologia-y-neurociencia/107/posts/la-breve-breve-historia-del-estudio-del-cerebro-a-travs-de-las-imgenes-18858>
- Gazzaniga, M. S., Ivry, R. B. y Mangun, G. R. (2019). *Cognitive neuroscience: The biology of the mind* (5th ed.). W. W. Norton & Company.
- Gold, I. y Roskies, A. (2008) "Philosophy of Neuroscience" en Ruse, Michael (ed.) *The Oxford Handbook of Philosophy of Biology*. DOI: 10.1093/oxfordhb/9780195182057.003.0016
- González Álvarez, J. (2012) *Breve historia del cerebro*. Crítica, Barcelona.
- González Lagier, D. (2012). ¿La tercera humillación? (Sobre neurociencia, filosofía y libre albedrío). *DOXA, Cuadernos de Filosofía del Derecho*, 35 (2012) ISSN: 0214-8676 pp. 499-510.
- Guala, F. (2003). Experimental localism and external validity. *Philosophy of Science Supplement*, 70, 1195–1205.
- Hacking, I. (1983). *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge University Press.

- Hacking, I. (2010). *The taming of chance* (Ideas in Context, No. 17). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511819766>
- Hanson, N. R. (1961). *Patterns of discovery: An inquiry into the conceptual foundations of science*. Cambridge University Press.
- Haselager, P. y Mecacci, G. (2021). The importance of expiry dates: Evaluating the societal impact of AI-based neuroimaging. En F. Jotterand y M. Ienca (Eds.), *Artificial intelligence in brain and mental health: Philosophical, ethical & policy issues* (pp. 205–216). Springer.
- Haueis, P y Slaby, J. (2017). Connectomes as Constitutively Epistemic. En N. Rose, T. Mahfoud y S. McLean (Eds.), *Modeling Brains: The Making and Use of Animal Models in Neuroscience and Psychiatry* (pp. 3-22). Elsevier.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L. y Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539), 2425–2430. <https://doi.org/10.1126/science.1063736>
- Head, M. L., Holman, L., Lanfear, R., Kahn, A. T. y Jennions, M. D, (2015) The Extent and Consequences of P-Hacking in Science. *PLoS Biol* 13(3): e1002106.
- Henson, R. (2005). What can functional neuroimaging tell the experimental psychologist? En *Quarterly Journal of Experimental Psychology* 58A:193–233
- Hillary, F. G. y DeLuca, J. (Eds.). (2007). *Functional neuroimaging in clinical populations*. The Guilford Press.
- Hipócrates, 400 a. C. *On the Sacred Disease*. Disponible en <http://classics.mit.edu/>
- Huettel, S. A., Song, A. W. y McCarthy, G. (2014). *Functional magnetic resonance imaging* (3a ed.). Sinauer Associates.
- Hutzler, F. (2014). Reverse inference is not a fallacy per se: Cognitive processes can be inferred from functional imaging data. *NeuroImage*, 84, 1061–1069.
- Hyde, J. S. (2016). Sex and cognition: gender and cognitive functions. *Current Opinion in Neurobiology*, 38, 53-56. <https://doi.org/10.1016/j.conb.2016.03.016>
- Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T. D., Elliott, M. A., Ruparel, K., ... y Verma, R. (2014). Sex differences in the structural connectome of the human brain. *Proceedings of the National Academy of Sciences*, 111(2), 823-828.
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS medicine*, 2(8), e124.
- Janssen, A., Klein, C. y Slors, M. (2017). What is a cognitive ontology, anyway? *Philosophical Explorations*, 20(2), 123-128.

- Joyce, J. (2021). Bayes' theorem. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition). Stanford University.
<https://plato.stanford.edu/archives/fall2021/entries/bayes-theorem/>
- Karenberg, A. (2015). Blood, phlegm and spirits: Galen on stroke. *Hist Med*, 2(2), 160-8.
- Kelly, R.E., Jr. y Hoptman, M.J. (2022) Replicability in Brain Imaging. *Brain Sci.* 2022, 12, 397.
<https://doi.org/10.3390/brainsci12030397>
- Khalidi, M. A. (2023). *Cognitive ontology*. Cambridge University Press.
- Klein, C. (2010). Philosophical Issues in Neuroimaging. En *Philosophy Compass* 5/2 (2010): 186–198, 10.1111/j.1747-9991.2009.00275.x
- Klein, C. (2010b). Images Are Not the Evidence in Neuroimaging. En *Brit. J. Phil. Sci.* 61 (2010), 265–278.
- Klein, C. (2012). Cognitive Ontology and Region- versus Network-Oriented Analyses. *Philosophy of Science*, 79(5), 952–960. <https://doi.org/10.1086/667843>
- Kosslyn, S. M. (1999). If neuroimaging is the answer, what is the question? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 354(1387), 1283–1294.
<https://doi.org/10.1098/rstb.1999.0475>
- Krauss, P. (2024). *Artificial intelligence and brain research: Neural networks, deep learning and the future of cognition*. Springer.
- Luxton, D. D. y Hudlicka, E. (2024). Intelligent virtual agents in behavioral and mental healthcare: Ethics and application considerations. En F. Jotterand y M. Ienca (Eds.), *Artificial intelligence in brain and mental health: Philosophical, ethical & policy issues* (pp. 41–56). Springer.
- Maestú-Unturbe, F., Ríos Lago, M. y Cabestrero Alonso, R. (2007). *Neuroimagen. Técnicas y procesos cognitivos*. Elsevier España.
- Martín-Araguz, A., Bustamante-Martínez C., Emam-Mansour M. T. y Moreno-Martínez, J. M. (2002). Neurociencia en el Egipto faraónico y en la escuela de Alejandría. *Rev. Neurol* 2002;34 (12):1183-1194 doi: [10.33588/rn.3412.2001461](https://doi.org/10.33588/rn.3412.2001461)
- McCabe, D. P., Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition*, 107(1), 343–352.
- McCabe, D. P., Castel, A. D. y Rhodes, M. G. (2011). The influence of fMRI lie detection evidence on juror decision-making. *Behavioral Sciences & the Law*, 29(4), 566-577.
- Meza, A. (2005). El doble estatus de la psicología cognitiva: como enfoque y como área de investigación. *Revista de investigación en psicología*, 8(1), 145-163.
- Nagel, E. (1981). *La estructura de la ciencia*. Barcelona: Paidós.

- Nickerson, R. S. (2000). Null hypothesis significance testing: a review of an old and continuing controversy. *Psychological methods*, 5(2), 241-301.
- Ongaro, G. y Kaptchuk, T. J. (2019). Symptom perception, placebo effects, and the Bayesian brain. *Pain*, 160(1), 1–4. <https://doi.org/10.1097/j.pain.0000000000001367>. Traducción: intramed.net
- Outes, D. L. y Orlando, J. C. (2008). Alcmeón de Crotona. El cerebro y las funciones psíquicas. *Neuropsiquiatría y Salud Mental. Alcmeón: Revista Argentina de Clínica Neuropsiquiátrica*, 15(1), 34–49.
- Pessoa, L. (2014). *The Cognitive-Emotional Brain: From Interactions to Integration*. MIT Press.
- Piccinini, G. (2015). Foundational Issues in Cognitive Neuroscience. En Clausen, Jens; Levy, Neil (eds.) *Handbook of Neuroethics*. Springer Dordrecht Heidelberg, New York-London.
- Piccinini, G., Craver, C. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, 183(3), 283-311.
- Platón. (1986). República (Libro VII, 514a–520a) [Edición bilingüe]. Madrid: Gredos.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, Vol.10 No.2 February 2006.
- Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social cognitive and affective neuroscience*, 2(1), 67-70.
- Poldrack, R. A. (2010). Subtraction and Beyond: The Logic of Experimental Designs for Neuroimaging. En *Foundational Issues in Human Brain Mapping*, ed. Stephen Hanson and Martin Bunzl, 147–59. MIT Press.
- Poldrack, R. A. (2010b). Mapping mental function to brain structure: How can cognitive neuroimaging succeed? *Perspectives on Psychological Science*, 5(5), 753–761. <https://doi.org/10.1177/1745691610388777>
- Poldrack, R. A. (2018) *The New Mind Readers. What Neuroimaging Can and Cannot Reveal about Our Thoughts*. Princeton University Press, New Jersey.
- Poldrack, R. A. (2019). The costs of reproducibility. En *Neuron*, 101(1), 11-14.
- Poldrack, R. A., Mumford, J. y Nichols, T. (2011). *Handbook of Functional MRI Data Analysis*. Cambridge University Press.
- Poldrack, R. A. y Mumford, J. A. (2010). On the proper role of nonindependent ROI analysis: A commentary on Vul and Kanwisher. En S. J. Hanson y M. Bunzl (Eds.), *Foundational issues in human brain mapping* (pp. 93-96). The MIT Press.
- Price, C. J. y Friston, K. J. (2005). Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology*, 22(3-4), 262-275.

- Rábano, A. (2019). El “error” de Aristóteles: estructura y función del cerebro en los tratados biológicos. *Neurosciences and History*, 6(4), 138-143.
- Racine, E., Bar-Ilan, O. e Illes, J. (2005). fMRI in the public eye. *Nature Reviews Neuroscience*, 6(2), 159-164.
- Rangel, E. N. (2015). ¿Qué es la Estadística Bayesiana? *Journal of Basic Sciences*, 1(1).
- Riviere, A. (1987). *El sujeto de la psicología cognitiva*. Alianza editorial.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Roskies, A. L. (2002). Neuroethics for the New Millenium. In *Neuron*, Vol. 35, 21–23, July 3, 2002.
- Roskies, A. L. (2008). Neuroimaging and Inferential Distance. *Neuroethics* 1, 19–30 (2008). <https://doi.org/10.1007/s12152-007-9003-3>
- Roskies, A. L. (2010). Neuroimaging and inferential distance: The perils of pictures. En S. J. Hanson y M. Bunzl (Eds.), *Foundational Issues in Human Brain Mapping* (pp. 195–215). MIT Press. <https://doi.org/10.7551/mitpress/9780262014021.003.0017>
- Roskies, A. L. (2015) Mind Reading, Lie Detection, and Privacy. En Clausen, Jens; Levy, Neil (eds.) *Handbook of Neuroethics*. Springer Dordrecht Heidelberg, New York-London. 679-695
- Roskies, A. L. (2015b) Neuroimaging Neuroethics: Introduction. En Clausen, Jens; Levy, Neil (eds.) *Handbook of Neuroethics*. Springer Dordrecht Heidelberg, New York-London.
- Roskies, A. L. (2021), Neuroethics. *The Stanford Encyclopedia of Philosophy* (edición de primavera de 2021), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2021/entries/neuroethics/>>.
- Salazar, J. L. R. (2016). La controversia sobre la sede del alma en la antigüedad. *Boletín Científico Sapiens Research*, 6(2), 20-28.
- Saxe, R., Brett, M. y Kanwisher, N. (2006). Divide and conquer: a defense of functional localizers. *Neuroimage*, 30(4), 1088-1096.
- Schmuckler, M. (2001). What is ecological validity? A dimensional analysis. *Infancy*, 2(4), 419–436
- Sternberg, S. (2011). Modular processes in mind and brain. *Cognitive Neuropsychology*, 28(3-4), 156–208. <http://doi.org/10.1080/02643294.2011.557231>
- Streiner, D. L., Norman, G. R. y Cairney, J. (2015). Health measurement scales: a practical guide to their development and use. Oxford University Press, USA.

- Sullivan, J. (2015) Experimentation in Cognitive Neuroscience and Cognitive Neurobiology, en Clausen, Jens; Levy, Neil (eds.) *Handbook of Neuroethics*. Springer Dordrecht Heidelberg, New York-London. pp-32-46.
- Sullivan, J. A. (2009). The multiplicity of experimental protocols: a challenge to reductionist and non-reductionist models of the unity of neuroscience. *Synthese* (2009) 167:511–539 DOI 10.1007/s11229-008-9389-4
- Sullivan, J. A. (2016). Construct Stabilization and the Unity of the Mind-Brain Sciences. *Philosophy of Science*, 83 (December 2016) pp. 662–673.
- Sullivan, J. A. (2017). Coordinated pluralism as a means to facilitate integrative taxonomies of cognition. *Philosophical Explorations*, 20(2), 129-145.
- Tappin, B. M. y Gadsby, S. (2019). Biased belief in the Bayesian brain: A deeper look at the evidence. *Consciousness and cognition*, 68, 107-114.
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L. y Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, 331, 1279–1285.
- Uttal, W. R. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain*. The MIT Press.
- Van der Weel, F. R. y Van der Meer, A. L. (2024). Handwriting but not typewriting leads to widespread brain connectivity: A high-density EEG study with implications for the classroom. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1219945>
- Vargas Sanabria, M. (2014). Evolución histórica de las autopsias y situación actual en Costa Rica. *Medicina Legal de Costa Rica*, 31(2), 42-54. Recuperado el 3 de enero de 2023, de http://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S1409-00152014000200005&lng=en&tlng=es.
- Venturelli, A. N. (2021). Manuscrito – *Rev. Int. Fil. Campinas*, v. 44, n. 1, pp. 01-36, Jan.-Mar. 2021.
- Venturelli, A. N. y Branca, I. (2016) Evidencia y neurociencias cognitivas: El caso de la resonancia magnética funcional. *Tópicos, Revista de Filosofía* 50 (2016), 177-207.
- Venturelli, N. (2015). Consideraciones epistemológicas alrededor del enfoque dinamicista en las neurociencias cognitivas. *Ciências & Cognição*, 20(1): 18-28.
- Venturelli, N., Barberis, S. (2023) Historia y filosofía de las ciencias cognitivas. En Liza Skidelsky, *Introducción a la filosofía de las ciencias cognitivas* / (compil.) Bogotá: Ediciones Uniandes. <https://doi.org/10.51573/Andes.9789587985191.9789587985207.1>

- Viola, M. (2016). L'agenda ontologica della neuroscienza cognitiva: le neuroscienze come "arbitro" delle categorie psicologiche (e viceversa). *Rivista internazionale di Filosofia e Psicologia*, 7(2), 144-165.
- Westbrook, C. y Kaut, C. (1998). *MRI in Practice*. Blackwell Science Ltd.
- Vul, E., Harris, C., Winkielman, P. y Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition ["Voodoo correlations" in social neuroscience]. *Perspectives on Psychological Science*, 4(3), 274-290. <https://doi.org/10.1111/j.1745-6924.2009.01125.x>
- Yang, H., Wang, G., Li, Z., Li, H., Zheng, J., Hu, Y., Cao, X., Liao, C., Ye, H. y Tian, Q. (2024). *Artificial intelligence for neuro MRI acquisition: A review. Magnetic Resonance Materials in Physics, Biology and Medicine*, 37(4), 567–589.