

MOISÉS PÉREZ-MARCOS
Universidad de Sevilla , España
mperez43@us.es

Recibido: 28/10/2025
Aprobado: 24/11/2025
DOI: 10.24310/stheg.12.2026.22518

La influencia kantiana en las teorías del procesamiento predictivo¹

The kantian influence on predictive processing theories

RESUMEN: El objetivo es presentar brevemente la discusión sobre las raíces kantianas de la teoría del procesamiento predictivo (propuesta entre otros por Karl J. Friston y Andy Clark), para examinar en qué medida hay en ella elementos de la filosofía de Kant. Tras evaluar las posturas de los principales autores que han intervenido en la polémica, concluyo que, en efecto, pueden encontrarse en las teorías del procesamiento predictivo algunas ideas aparentemente kantianas, y han servido por eso de apoyo a los defensores de la “conexión kantiana”. Pero si tenemos en cuenta el sentido y el carácter general de la filosofía de Kant, tomada en su conjunto, además de otros aspectos centrales de la misma, que no aparecen de ninguna manera en las teorías del procesamiento predictivo, podemos ver también que las diferencias entre las teorías del procesamiento predictivo y el pensamiento del filósofo de Königsberg son en realidad demasiado grandes como para establecer una conexión filosófica muy relevante.

PALABRAS CLAVE: KANT, TEORÍAS DEL PROCESAMIENTO PREDICTIVO, CIENCIA COGNITIVA, FILOSOFÍA DE LAS NEUROSCIENCIAS.

ABSTRACT: The paper's aim is to briefly present the discussion on the Kantian roots of predictive processing theory (proposed among others by Karl J. Friston and Andy Clark), in order to examine to what extent there are elements of Kant's philosophy in it. After evaluating the positions of the main authors involved in the controversy, I conclude that some apparently Kantian ideas can indeed be found in predictive processing theories and have thus served as support for the advocates of the 'Kantian connection'. But if we take into account the meaning and general character of Kant's philosophy, taken as a whole, as well as other central aspects of it, which do not appear in any way in the predictive processing theories, we can also see that the differences between these theories and the thought of the Königsberg's philosopher are in fact too great to establish a strong philosophical connection.

KEYWORDS: KANT, PREDICTIVE CODING THEORIES, COGNITIVE SCIENCE, PHILOSOPHY OF NEUROSCIENCES.

[1] Este texto ha sido posible gracias a un permiso de la Universidad de Sevilla para una estancia de investigación en la Universidad de Heidelberg, y ha contado con el apoyo económico de la Fundación Tatiana dentro del programa de fomento de la difusión y comunicación de resultados científicos para becarios de investigación asociados al Centro Internacional de Neurociencia y Ética (CINET).

I. LAS TEORÍAS DEL PROCESAMIENTO PREDICTIVO

Immanuel Kant es sin duda uno de los filósofos que más impacto ha tenido en el pensamiento posterior, pero ¿hasta qué punto su influencia se deja notar en las ciencias cognitivas?¹. Quizá el mayor defensor de la existencia de esta influencia sea el filósofo canadiense Andrew Brook,² que llega a calificar a Kant de “padrino intelectual de la ciencia cognitiva”.³ Brook cree que el modelo de la mente que ha heredado la ciencia cognitiva es esencialmente el kantiano. En un interesantísimo estudio sobre la recepción de Kant en la filosofía de la mente contemporánea, Pedro Jesús Teruel ha puesto en cuestión que esta interpretación sea adecuada.⁴ Teruel distingue tres etapas en el pensamiento del filósofo de Königsberg: la fase dogmático-teórica, la del cierre escéptico y finalmente la apertura crítica. En la filosofía de la mente contemporánea se puede rastrear la presencia de ideas de cada una de estas etapas del pensamiento kantiano, y cada una de ellas debe ser evaluada individualmente. Pero en general podemos decir, siguiendo a Teruel, que los filósofos de la mente no han sido respetuosos con la existencia de estas etapas en el pensamiento de Kant, además de que no siempre han tenido en cuenta el conjunto de dicho pensamiento. En definitiva, lo que se puede objetar a este uso de Kant por parte de la actual filosofía de la mente es que está utilizando a un pensador claramente contrario al materialismo para justificar opciones materialistas (o emergentistas, que quizás no sean muy diferentes del materialismo). La trayectoria intelectual del propio Kant puede verse precisamente como un intento de elaborar un pensamiento que no termine en un materialismo reduccionista (otro asunto sería el de hasta qué punto Kant consigue lo que pretende, pero ahora no podemos abordarlo).

[1] SCHLICHT, T. y NEWEN, A., “Kant and Cognitive Science Revisited” en *Logical Analysis and History of Philosophy / Philosophiegeschichte und logische Analyse*, Vol. 18 (2015) pp. 87-113. Estudios recientes han explorado su relevancia para otras ciencias, como la mecánica cuántica: KAUARK-LEITE, P., “¿Podría Kant arrojar luz sobre la indeterminación onto-lógica en la teoría cuántica?” en *Naturaleza y Libertad. Revista de estudios interdisciplinares*, n.º 19, 2025, 107-127.

[2] BROOK, A., *Kant and the Mind*. Cambridge: Cambridge University Press, 1994; y BROOK, A., “Kant and Cognitive Science” en *Teleskop. Revista de pensamiento y cultura* I, 4 (2004) pp. 13-21. Traducido al castellano por Ester Hernández en TERUEL, P. J., *Kant y las ciencias*. Madrid: Biblioteca Nueva, 2011, pp. 227-236, por donde se cita.

[3] BROOK, “Kant and Cognitive Science”, op. cit. p. 227.

[4] TERUEL, P. J., “La recepción de Kant en la *Philosophy of Mind*. Una revisión crítica de las fuentes kantianas”, en TERUEL, P. J. (ed.), *Kant y las ciencias*. Madrid: Biblioteca Nueva, 2011, pp. 244-302. Más sobre la relevancia del Kant del Opus postumum para la antropología filosófica en TERUEL, P. J., “La síntesis final. implicaciones de la filosofía natural contenida en el Opus postumum para la cuestión «¿Qué es el ser humano?»”, en *Naturaleza y Libertad. Revista de estudios interdisciplinares*, nº 19, 2025, 83-105.

Dentro del variado grupo de ideas que en filosofía de la mente o en ciencia cognitiva se han caracterizado como kantianas, está la teoría del procesamiento predictivo. Más que de una teoría estaríamos ante un grupo o familia de ellas, que comparten algunos supuestos básicos y difieren en detalles quizá secundarios. No todo el mundo entiende lo mismo por procesamiento predictivo,⁵ pero para satisfacer nuestro propósito bastará con una caracterización general de este marco teórico o, quizás también se pueda llamar así, este programa de investigación (en el sentido que Lakatos da a esta expresión).⁶ El procesamiento predictivo ha sido defendido desde el punto de vista neurocientífico por Karl Friston,⁷ y efectivamente su contribución es insoslayable. Desde un punto de vista más filosófico, han abordado su exposición y defensa Andy Clark⁸ y Jakob Hohwy.⁹

Las teorías del procesamiento predictivo proponen un marco conceptual que, según sus defensores, unificaría el estudio de la percepción, la acción, la cognición y las relaciones entre ellas. Este marco no es una propuesta a nivel neurofisiológico, sino a nivel conceptual, formulado en términos computacionales. Es cierto que algunos defensores del procesamiento predictivo han elaborado hipótesis sobre cuál sería el fundamento neurofisiológico o el correlato cerebral del procesamiento predictivo, pero conviene insistir en que se trata sobre todo de un modelo teórico.

La idea básica del modelo teórico es que el cerebro es una máquina de hacer predicciones. La concepción, llamémosle clásica, entendía que el cerebro recibía los datos provenientes de los sentidos de un modo pasivo y, si acaso en un momento posterior, reelaboraba o procesaba esos datos, esa información. Lo que afirman los defensores del procesamiento predictivo es que el cerebro realiza anticipaciones o predicciones sobre qué información le va a llegar. Además, intenta minimizar constantemente el error, es decir, la diferencia entre sus predicciones y lo que realmente le llega. Es tanto como decir que el cerebro funciona intentando hacer cada vez predicciones mejores. La clave para comprender el funcionamiento del cerebro es ese procesamiento descendente (*top-down processing*). Lo relevante, por ejemplo, en la percepción, no es tanto

[5] SPREVAK, M., “Predictive coding I: Introduction” en *Philosophy Compass*, e12950, 2024.

[6] LAKATOS, I., *Philosophical papers*. Cambridge: Cambridge University Press, 1980.

[7] FRISTON, K., “Learning and inference in the brain” en *Neural Networks*, 16 (9) 2003, pp. 1325-1352; FRISTON, K., “The free-energy principle: A rough guide to the brain?” en *Trends in Cognitive Sciences*, 13 (7) 2009, pp. 293-301; FRISTON, K., “The free-energy principle: A unified brain theory?” en *Nature Reviews Neuroscience*, 11 (2) 2010, pp. 127-138.

[8] CLARK, A., “Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognition Science” en *Behavioral and Brain Sciences* 36.3 (2013), pp. 181-204; CLARK, A., *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. New York: Oxford University Press, 2016.

[9] HOHwy, J., *The Predictive Mind*. Oxford: Oxford University Press, 2013.

lo que llega al cerebro desde los sentidos, sino lo que el cerebro elabora previamente a eso que recibe. Esta característica del cerebro afecta a todo lo que hace (no solo a la percepción) y todo el tiempo (no solo en las ocasiones cuando lo que viene de los sentidos no está claro). Las predicciones que el cerebro realiza son de naturaleza estadística, es decir, asigna probabilidades a las diferentes posibilidades, que funcionan como representaciones de lo que ocurre de hecho en el mundo, y esas predicciones están jerárquicamente organizadas, es decir, implican un modelo jerárquico en el que “las variables de los niveles adyacentes son condicionalmente independientes”.¹⁰

Se pueden decir muchas más cosas sobre el modelo del procesamiento predictivo, como que el tipo de inferencia que realiza el cerebro es bayesiana (lo que da una forma y expresión fundamentalmente matemática a la teoría), o que las predicciones del cerebro se aplican no solamente a lo recibido por los sentidos sobre el mundo exterior (exterocepción) sino también a la información que llega del propio organismo (interocepción), lo que ayuda al mantenimiento de la homeostasis y, en último término, a la supervivencia. Se puede decir, también, que el cerebro funciona básicamente como una máquina de producir y testar hipótesis, o que con esta conducta lo que el cerebro busca es minimizar la energía libre (esta es la forma concreta que adopta la teoría en manos de Friston). Con lo dicho, sin embargo, basta para hacerse una idea general sobre los elementos conceptuales básicos de la teoría.

II. IDEAS «KANTIANAS» EN EL PROCESAMIENTO PREDICTIVO

La mayor parte de los autores que se refieren al procesamiento predictivo citan al filósofo Immanuel Kant como uno de los antecedentes más notables de esta perspectiva teórica. Swanson llega a afirmar que “el PP [procesamiento predictivo] puede incluso ser visto como un paso importante en la evolución de la psicología trascendental de Kant”.¹¹ El filósofo alemán habría sido el primero en afirmar el carácter activo de la mente humana, gracias a su famoso “giro copernicano”. Según Kant:

[10] WIESE, W. y METZINGER, Th., “Vanilla PP for Philosophers: A Primer on Predictive Processing”, en T. METZINGER y W. WIESE (Eds.), *Philosophy and Predictive Processing: 1*. Frankfurt am Main: MIND Group, 2017, p. 3.

[11] SWANSON, L., “The Predictive Processing Paradigm Has Roots in Kant” en *Front Syst Neurosci.* 2016 Oct 10;10:79, p. 10. Una respuesta directa a Swanson puede verse en BENI, M. D., “Commentary: The Predictive Processing Paradigm Has Roots in Kant” en *Frontiers in Systems Neuroscience* 11: 98 (2018). Beni considera que la interpretación de Swanson es unilateral y negligente, pues ignora las profundas diferencias que existen entre Kant y los teóricos del procesamiento predictivo. Es cierto que ve cierta similitud entre las ideas de Kant y las de Jakob Hohwy, pero no entre las del pensador alemán y Andy Clark.

Si la intuición tuviera que regirse por la naturaleza de los objetos, no veo cómo podría conocerse algo *a priori* sobre esa naturaleza. Si, en cambio, es el objeto (en cuanto objeto de los sentidos) el que se rige por la naturaleza de nuestra facultad de intuición, puedo representarme fácilmente tal posibilidad.¹²

La percepción, desde esta perspectiva, no se puede explicar de modo pasivo (como aquello que es recibido en los sentidos), sino que supone una forma de lo percibido independiente de la experiencia (*a priori*) y que moldea esta. Kant, por lo tanto, sería el responsable de haber anticipado la idea básica del procesamiento predictivo, que enfatiza la naturaleza constructiva de la percepción (el procesamiento de los estímulos está fuertemente influenciado, si no moldeado en su mayoría, por las influencias descendentes, que crean constantemente predicciones sobre lo que se espera percibir).

Salvo esta indicación de carácter muy general, son pocos los autores que han intentado examinar sistemáticamente o con más profundidad la relación entre el pensamiento de Kant y las teorías del procesamiento predictivo. Link R. Swanson es uno de ellos.¹³ No solamente identifica como kantiana la idea del papel activo o descendente de la percepción, sino que señala, además: el papel de los *hyperpriors*, la función general de los “modelos generativos”, el papel crucial de la imaginación en la percepción, el proceso de “análisis vía síntesis”, y el uso del tipo de razonamiento por “ingeniería inversa” o análisis descendente. Comentaré brevemente cada una de ellas.

II.1. LOS HYPERAPRIORS Y LAS CONDICIONES DE TODA EXPERIENCIA

Ya hemos dicho que una de las ideas básicas de las teorías del procesamiento predictivo es que la percepción algo parecido a las predicciones (por lo general inconscientes) que realiza el cerebro. Los sistemas para realizar predicciones requieren una serie de *restricciones* sobre un conjunto de probabilidades previas que ayuden a elegir las opciones más probables, pues sería prácticamente imposible contemplar todas las posibilidades realmente existentes. Estas restricciones probabilísticas se conocen en la literatura técnica como “*priors*”, expresión que quizás podríamos traducir como “*aprioris*”. Estos antecedentes formarían un conjunto jerarquizado, dentro del cual a los más abstractos se les conoce como “*hyperpriors*” o “hiperaprioris”¹⁴. Los hiperaprioris más elementales serían el

[12] KANT, KrV, BXVII, p. 20. El número de página hace referencia a la edición española (Kant, I., *Crítica de la razón pura*, traducción de Pedro Ribas. Madrid: Alfaguara, 1998).

[13] SWANSON, “The Predictive Processing”, op. cit.

[14] Otra posibilidad de traducción sería decir “antecedentes” en vez de “*priors*” y “superantecedentes” en vez de “*hyperpriors*”. La traducción que he elegido muestra mejor su relación con los conceptos kantianos, pero no me gustaría que, precisamente por eso, prejuzgase la cuestión.

espacio y el tiempo. De entre todas las posibilidades lógicas existentes, el cerebro solo hace predicciones que tengan en cuenta que solo puede haber un objeto en el mismo sitio a la vez, que algo solo se puede mover a izquierda o a derecha pero no a ambos lados a la vez, etc. Este tipo de restricciones, que limitan el ámbito de lo que es razonable esperar, pueden interpretarse como condiciones necesarias a priori de toda experiencia, es decir, como las condiciones formales (en terminología kantiana) de toda percepción. Si este planteamiento llevó a algunos a pensar que Kant estaba defendiendo un idealismo con respecto al espacio y el tiempo, otro tanto ha ocurrido en la perspectiva neurocientífica. No son pocos los que creen que espacio y tiempo no serían, entonces, algo existente de modo objetivo en la realidad, sino más bien creaciones de nuestro cerebro. Volvemos, en definitiva, a la idea de que la realidad es una gran alucinación.

II.2. LOS MODELOS GENERATIVOS Y EL ESQUEMATISMO

Uno de los problemas más difíciles de resolver es el de explicar cómo es posible el reconocimiento de los objetos. Para nosotros es una realidad cotidiana que, cuando vemos un animal corriendo a nuestro alrededor, inmediatamente podemos identificarlo como nuestro perro. Pero aún no sabemos bien cómo es posible semejante cosa. Hemos sido conscientes de la dificultad de resolver este tipo de problemas cuando hemos intentado imitar artificialmente estas capacidades. La identificación de objetos es uno de los campos más desafiantes en inteligencia artificial y neurociencia computacional. Una de las propuestas es la basada en “modelos generativos”. Estos modelos son sistemas capaces de, por un lado, identificar ciertos patrones en la información de entrada y, por otro lado, elaboran predicciones o combinaciones de situaciones hipotéticas que podrían dar lugar a combinaciones semejantes a las dadas en la entrada. Los modelos generativos tienen la capacidad de producir “internamente” una gran variedad de estados que podrían corresponderse con la información recibida del exterior, y de elegir cuáles de esos estados se corresponden de modo más probable con eso recibido. Todo esto se consigue gracias a sistemas de inteligencia artificial basados en redes neuronales de diferentes capas (que dicho sea de paso, nadie sabe muy bien por qué consiguen lo que consiguen), pero no necesitamos ahora entrar en los detalles.

La capacidad de identificar imágenes de los sistemas artificiales está aún muy lejos de la de cualquier ser humano, pero los defensores del procesamiento predictivo habitualmente comparan estos modelos generativos con los esquematismos kantianos. Cuando Kant intentó explicar cómo era posible que la mente humana aplicase los conceptos puros del entendimiento (categorías) a la experiencia, elaboró para ello la doctrina del esquema, que sería un elemento mediador entre la categoría, por un lado, y la apariencia o fenómeno (la expe-

riencia), por otro lado. Este elemento mediador o representación mediadora es en cierto sentido intelectual y en cierto sentido sensible. El esquema es un producto de la imaginación, pero no es una imagen, sino la idea de un procedimiento universal de la imaginación. El concepto de perro, por ejemplo, es aplicado a la realidad percibida, no porque la mente sea capaz de comparar lo percibido con una imagen interna de un perro, sino porque la propia imaginación es capaz de dibujar la figura de un animal cuadrúpedo en general. Naturalmente, para identificar algo como “perro” hace falta más que la imagen generada por la imaginación de “un concepto de animal cuadrúpedo en general”, pero el ejemplo, del propio Kant, sirve para saber de qué estamos hablando. Esta parte de la *Critica* de Kant ha sido una de las más discutidas, y el mismo Kant reconoce que “El verdadero funcionamiento de este arte difícilmente dejará la naturaleza que lo conoczamos y difícilmente lo pondremos al descubierto”.¹⁵

Podríamos pensar que los modelos generativos de la inteligencia artificial, utilizados por las teorías del procesamiento predictivo, son un modo de conocer lo que para Kant sería difícilmente conocido. En ese caso, Kant, con su doctrina de los esquemas, habría anticipado la existencia de esos modelos generativos en dos sentidos. Primero, porque, como ocurre en la doctrina del esquematismo kantiano, el reconocimiento de objetos requiere la confluencia de dos movimientos, uno ascendente (lo que viene de los sentidos) y otro descendente (lo que genera el propio, la anticipación). Segundo, porque como ocurre con los esquemas kantianos, no reconocemos los objetos porque comparemos lo que recibimos sensorialmente de estos con un conjunto de imágenes que previamente poseemos en algún lugar, sino porque se identifican un conjunto de reglas abstractas endógenas que se usarían para generar esos mismos patrones de lo percibido en la imaginación. O sea, lo que permite que apliquemos a un cierto patrón de perceptos un concepto general es la mediación de un modelo o regla (esquema) de cómo se podría haber producido ese patrón, modelo o regla que es producido por el propio sistema y no recibido por él desde fuera.

II.3. LA IMAGINACIÓN

Con lo que acabamos de decir es fácil suponer que la “imaginación”, en el sentido de la actividad de la que son capaces los modelos generativos, desempeña un papel fundamental en los procesos de percepción. Podríamos pensar, si estamos buscando la similitud con las teorías del procesamiento predictivo, que esta “imaginación” desempeña más o menos el mismo papel que la imaginación en el pensamiento de Kant: “La *imaginación* es la facultad de representar un objeto en la intuición incluso cuando este no se halla presente”.¹⁶ Dado que

[15] KANT, KrV, B181, p. 185.

[16] KANT, KrV, B151, p. 166.

representa un objeto en la intuición, es una facultad que pertenece a la sensibilidad, pero no determinable por ella sino determinante de ella, es decir, “la imaginación es una facultad que determina *a priori* la sensibilidad; la síntesis de las intuiciones efectuada por esta facultad tiene que ser una síntesis trascendental de la *imaginación de acuerdo con las categorías*”.¹⁷ Podemos decir, por traducirlo al modelo del procesamiento predictivo, que un modelo generativo crea espontáneamente, aunque de acuerdo con ciertos *priors* e *hyperpriors*, unas expectativas que, posteriormente, podrían aplicarse a los datos de entrada (percepción). Esas expectativas, pues, son generadas por la “imaginación” y condicionan o moldean la experiencia de acuerdo con ciertos elementos *a priori*.

II.4. ANÁLISIS VÍA SÍNTESIS

En las teorías del procesamiento predictivo, una sensación (una entrada sensorial) es analizada mediante una comparación con los procesos internos que podrían sintetizar o dar lugar a un patrón semejante al de la entrada sensorial. Es decir, los modelos del procesamiento predictivo son capaces de analizar las entradas porque son capaces de generar internamente procesos que darían lugar a un patrón similar en los datos de esas entradas. Por eso podemos hablar del “análisis vía síntesis”, es decir: el análisis es posible porque previamente se puede hacer una síntesis (la generada en la propia mente de modo independiente de la experiencia). Es de sobra conocida la importancia que Kant concede a la síntesis en su epistemología. El conocimiento es posible o acontece como síntesis, que puede serlo de la diversidad de lo intuido (en la sensopercepción), de la diversidad de los conceptos, o de lo intuido con los conceptos.¹⁸ La síntesis es el aspecto central en cualquier investigación sobre la cognición, pues como dice Kant:

La síntesis de algo diverso (sea empírico o dado *a priori*) produce ante todo un conocimiento que, inicialmente puede ser todavía toscos y confuso y que, por ello mismo, necesita un análisis. Pero es propiamente la síntesis la que recoge los elementos en orden al conocimiento y los reúne con vistas a cierto contenido. Ella constituye, pues, lo primero que debemos atender si queremos juzgar sobre el origen primero de nuestro conocimiento”.¹⁹

[17] KANT, KrV, B152, p. 166.

[18] En la primera edición de la *Critica de la razón pura*, Kant habla de tres tipos de síntesis que “necesariamente tienen lugar en el conocimiento”, que son: “aprehensión de las representaciones, como modificaciones del psiquismo en la intuición; reproducción de dichas representaciones en la imaginación y reconocimiento de las misma en el concepto” (KANT, KrV, A 97, p. 131).

[19] KANT, KrV, B103, p. 112.

Kant, en coherencia con su giro copernicano, cree que no podemos ligar nada en los objetos que no haya sido previamente ligado por nosotros mismos. Somos capaces de analizar algo, sí, pero solo porque previamente lo hemos sintetizado. Dice Kant sobre el acto de la síntesis: “Se advierte fácilmente que este acto ha de ser originariamente uno, indistintamente válido para toda combinación y que la disolución, el *análisis*, que parece ser su opuesto, siempre lo presupone”.²⁰ Parece, pues, que Kant estaría haciendo dos afirmaciones que encontramos después en las teorías del procesamiento predictivo: la primera, que la síntesis es previa al análisis y viene exigida por este; y segunda, que la síntesis es lo primero que hemos de tener en cuenta en cualquier investigación sobre cómo funciona la cognición.

II.5. INGENIERÍA INVERSA Y MÉTODO TRASCENDENTAL

Las teorías del procesamiento predictivo comparten con la filosofía kantiana otro rasgo, esta vez más general, que en cierto modo ataña a la metodología. En el campo de la ingeniería se conoce como “ingeniería inversa” al método de estudio de los sistemas informáticos que trata de averiguar cómo es posible que la máquina realice su función utilizando para ello un análisis descendente (de arriba hacia abajo), que clarifica la función de cada una de las partes y cómo estas contribuyen a la función final o global. Si quiero saber cómo funciona una pieza de hardware que no ha sido diseñada por mí, puedo partir de la función que realiza e intentar saber cómo cada una de las partes de las que está formada contribuye a esa función. La otra metodología, la ascendente o de abajo hacia arriba, parte de los elementos simples, e intenta poco a poco reconstruir las funciones que se intentan explicar, tratando de dar cuenta de cómo surgen de esos elementos simples las funciones superiores. Se parte, en el caso de la ciencia cognitiva, de los mecanismos neurales, y se intenta dar cuenta de cómo pueden surgir las características funcionales del conjunto. Por ejemplo, puedo construir una red neuronal inspirada biológicamente e intentar averiguar de qué funciones cognitivas es capaz y cómo consigue realizarlas.

En la ciencia cognitiva se pueden utilizar ambas estrategias, pero en las teorías del procesamiento predictivo la más común es la primera, la estrategia descendente²¹. Podemos pensar, puestos a buscar analogías, que Kant habría

[20] KANT, KrV, B130, p. 153.

[21] Hay que mencionar, no obstante, que algunos han visto que el enfoque del procesamiento predictivo es un buen candidato a “metodología intermedia” o de enfoque que puede integrar ambas perspectivas. Así, por ejemplo, Hohwy dice que: “El programa de la predicción del error parece razonablemente bien situado entre dos corrientes opuestas en ciencia cognitiva. Por un lado, hay un enfoque descendente [*top-down*], que parte del análisis conceptual y funcional de los procesos cognitivos y busca después elaborar un modelo del cerebro por ingeniería inversa.

adoptado también esta perspectiva: no otra cosa sería su famoso método trascendental. Como es sabido, Kant adoptó una perspectiva trascendental en buena parte de su pensamiento, entendiendo por trascendental el estudio de las condiciones de posibilidad. El método trascendental aplicado al problema del conocimiento buscará esclarecer cómo es posible este, es decir, busca dilucidar qué condiciones tendrían que darse para que exista semejante cosa como el conocimiento. En sus propias palabras: “llamo trascendental todo conocimiento que se ocupa, no tanto de los objetos, cuanto de nuestro modo de conocerlos, en cuanto que tal modo ha de ser posible *a priori*”.²² Así, la pregunta “¿qué es el conocimiento?” se transforma en esta otra: “¿cómo es posible el conocimiento?”. Kant habría sido, pues, un pionero del enfoque descendente.

III. LA INFLUENCIA INDIRECTA DE HELMHOLTZ

La similitud aparente entre algunos planteamientos de las ciencias cognitivas y, de manera particular, entre las teorías del procesamiento predictivo y algunas ideas kantianas es, ciertamente, asombrosa. Pero podría tener su razón de ser en la genealogía de las propias ciencias cognitivas. La inmensa mayoría de quienes abordan estas cuestiones citan al científico alemán Hermann von Helmholtz (1821-1894) como el primer eslabón de una cadena que explicaría la conexión entre la filosofía kantiana y las ciencias cognitivas actuales. En efecto, muchos defensores del paradigma del procesamiento predictivo o de algunas de las ideas que lo conforman, han invocado a Helmholtz como uno de sus antecedentes.

En el pensamiento de Helmholtz hay algunas ideas que han dejado un gran impacto en la ciencia cognitiva actual. Quizá la más importante es la concepción de la percepción como inferencia. Para el fisiólogo alemán, la mente realiza una serie de ajustes o “inferencias inconscientes” que construyen una imagen coherente de las experiencias. Percibir algo es en cierto modo haberlo previsto. Esto explica que frecuentemente nuestra mente vea cosas que de hecho no hay (como cuando completa información que falta en lo que recibimos de los sentidos, o ve como continuos fenómenos que en realidad son discretos) y que en algunas ocasiones concretas solo vemos lo que esperamos ver. La idea de que la percepción es una “inferencia inconsciente” le llevó a afirmar, por ejemplo, que la posición espacial, que en ocasiones se ha utilizado para individualizar un objeto, es en realidad algo que pone el sujeto que percibe, es, por decirlo así, una interpretación que nuestro sistema perceptivo realiza de las sensaciones.

Por el otro lado, hay un enfoque ascendente [*bottom-up*], que construye redes neuronales biológicamente inspiradas y busca aprender sobre qué funciones cognitivas son capaces de realizar” (HOHRY, *The Predictive Mind*, op. cit., p. 8).

[22] KANT, KrV, B25, p. 58.

Igualmente, las cualidades de lo percibido pertenecen solo a nuestro sistema nervioso, y no a la realidad. Se deja entrever en lo ya dicho, pero conviene explicitar que otra de las ideas de Helmholtz tiene que ver con la relevancia de los procesos descendientes (o top-down) en la percepción. Si de algún modo lo que esperamos ver influye en lo que de hecho vemos, parece claro que las expectativas (ya sean conscientes o inconscientes) moldean o desempeñan un papel activo en la percepción.

El propio Helmholtz reconoció la influencia del pensamiento de Kant en el suyo propio. Su padre, Ferdinand Helmholtz, era amigo del filósofo Fichte, uno de los discípulos de Kant (y uno de los responsables de la interpretación idealista de la filosofía de su maestro), por lo que no es extraño pensar que debió familiarizarse con las ideas del filósofo de Königsberg desde muy pronto. En algunos asuntos, de hecho, Helmholtz creyó que, gracias a la ciencia, había corregido a Kant, pero no para desacreditarle, sino para ser más coherentemente kantiano. Según Lenoir: “En muchos sentidos, Helmholtz consideraba que su trabajo en epistemología no contradecía a Kant, sino que actualizaba los puntos de vista de Kant a la luz de los nuevos avances de la fisiología experimental [...]. Aunque discutido en su época, Helmholtz se consideraba más kantiano que Kant”.²³

El padre de Helmholtz fue amigo de Fichte, pero conviene aclarar que la interpretación idealista de este no era compartida por Helmholtz. Si (siendo muy reduccionista) existen dos corrientes interpretativas de Kant, la idealista y la positivista, hay quienes han atribuido a Helmholtz ser precisamente el pionero de la segunda. El fisiólogo alemán puede ser visto como uno de los primeros en poner en marcha un proceso de “naturalización del sujeto trascendental kantiano”,²⁴ a los que seguirían Wundt y Baldwin, considerados junto con Helmholtz padres de la psicología empírica.

En definitiva, la relación entre las ideas de Kant y la moderna ciencia cognitiva, y de modo más concreto su relación con las teorías del procesamiento predictivo, parecen más verosímiles si tenemos en cuenta la conexión histórica que existe entre ellas a través de Helmholtz, que habría infectado o sembrado, según cómo se mire, las ciencias cognitivas contemporáneas con el espíritu kantiano (aunque quizás aquí la palabra “espíritu” no sea precisamente acertada).

[23] LENOIR, T., “Operationalizing Kant: manifolds, models and mathematics in Helmholtz’s theories of perception”, en FRIEDMAN, M. y NORDMANN, A. (Eds.), *Kantian Legacy in Nineteenth-Century Science*. Cambridge, MA: MIT Press, 2006, p. 141.

[24] SÁNCHEZ, J.C.; FERNÁNDEZ, T. R. y LOY, I., “La génesis de la intuición. Helmholtz y la naturalización del sujeto trascendental kantiano” en *Revista de Historia de la Psicología*, Vol. 16, Nº 3-4 (1995) pp. 375-382.

IV. ALGUNOS COMENTARIOS CRÍTICOS

Escapa al propósito de este estudio realizar un discernimiento pormenorizado de cada una de las ideas presuntamente kantianas que se encuentran en las teorías del procesamiento predictivo, pero continuaré con algunos comentarios críticos sobre el asunto. Un primer grupo de ellos se dirigen a las propias teorías del procesamiento predictivo. Un segundo grupo a la comparación entre estas y Kant.

En primer lugar, hay que tener en cuenta que, como es normal, las teorías del procesamiento predictivo son muy cuestionadas por algunos neurocientíficos y filósofos. Algunos han llegado a decir que el procesamiento predictivo es una teoría no suficientemente madura como para ser puesta a prueba por la experiencia, e incluso hay cierta polémica (quizá exagerada) sobre si la teoría es o no científica. Es verdad que se trata de un modelo (fundamentalmente matemático), y que algunos neurocientíficos defienden su carácter meramente instrumental, desentendiéndose de cualquier interpretación ontológica. Pero también es cierto que puede ser visto como una descripción realista de lo que el cerebro hace, de qué papel desempeña en el conjunto del organismo y de cómo funciona. Y aquí ya empiezan los problemas, porque son cada vez más los que cuestionan que el cerebro haga algo así como procesar algo o realizar predicciones. A pesar del sentido técnico que se da a estas palabras en las teorías del procesamiento predictivo, parece bastante claro que se está atribuyendo características propias del nivel personal al nivel sub-personal. Esto hace que las teorías del procesamiento predictivo tengan una gran relevancia filosófica, pero también aparecen muchas dudas sobre la legitimidad del uso de esos términos para describir lo que en realidad hace el cerebro. Decir que el cerebro es capaz de procesar y realizar predicciones, pero no explicar cómo es posible que haga esas cosas es casi tanto como resolver un misterio apelando para ello a otro (o a otros dos). Por dar algún ejemplo, un autor crítico con la idea de procesamiento tal y como la utilizan las neurociencias en la actualidad es el neurocientífico teórico Romain Brette.²⁵ Y una crítica a la idea de que el cerebro hace predicciones puede verse en el psiquiatra y filósofo Thomas Fuchs.²⁶

Las teorías del procesamiento predictivo son, en ocasiones, altamente abstractas. Y, aunque existen algunos intentos en esta dirección, aún no se ha dilucidado cuál podría ser el fundamento fisiológico de los conceptos que utilizan. En ocasiones la ciencia comienza en un estadio muy abstracto y es-

[25] BRETTÉ, R., “Is coding a relevant metaphor for the brain?” en *Behavioral and Brain Sciences* 42, e215, 2019, pp. 1–58; BRETTÉ, R., “Brains as Computers: Metaphor, Analogy, Theory or Fact?” en *Front. Ecol. Evol.* 10:878729, 2022.

[26] FUCHS, Th., *Ecology of the Brain. The Phenomenology and Biology of the Embodied Mind*. Oxford: Oxford university Press, 2018.

peculatorio, es cierto, pero sería deseable que, en la medida de lo posible, los modelos sobre el funcionamiento del cerebro tuviesen precisamente en cuenta eso: el funcionamiento del cerebro, es decir, su fisiología, y no solo los modelos abstractos matemáticos. No niego que el paradigma del procesamiento predictivo sea útil heuristicamente hablando, pero sería deseable que sus resultados se pudiesen ver también en este terreno.

No cabe duda de que el procesamiento predictivo y, en general, buena parte de las ciencias cognitivas contemporáneas, hace bien en reconocer un papel activo que caería del lado del sujeto en la percepción (así como en otras actividades mentales). Pero el asunto es hasta dónde llega ese papel activo, es decir, en qué medida nos acerca o aleja de una postura idealista. Algunos han llegado a defender que una consecuencia lógica del procesamiento predictivo es que el mundo es una alucinación producida por el cerebro.²⁷ Es lo que Wiese y Metzinger han denominado “aislamiento medioambiental”.²⁸ Sin embargo, reconocer el papel constructivo del cerebro o de la mente en ciertas circunstancias o para ciertos aspectos de la realidad no es algo que deba necesariamente impedir el realismo. Al fin y al cabo, dado que la realidad es experimentada por un sujeto, este tendrá necesariamente que elaborar su perspectiva subjetiva. Creo que las neurociencias no conducen necesariamente a la afirmación de un idealismo del tipo mencionado anteriormente, es decir, a la idea de que el mundo es solo una alucinación de nuestro cerebro.

En segundo lugar, podemos decir que la comparación entre las teorías del procesamiento predictivo y la filosofía kantiana no está exenta de cierta verdad. Parece innegable que en ocasiones algunas ideas del procesamiento predictivo evocan una “sensación kantiana”.²⁹ Como hemos visto, esta conexión entre Kant y algunas partes de las ciencias cognitivas podría explicarse apelando para ello a la figura de Hermann von Helmholtz: el fisiólogo habría intentando naturalizar el sujeto trascendental kantiano para ser más fiel que el propio Kant a sus ideas. Pero resulta también claro que hay muchos aspectos del procesamiento predictivo que no se encuentran en Kant, como las perspectivas probabilista, evolucionista, computacional y neurocientífica. Hay, también, muchos elementos, centrales para la filosofía kantiana, que no aparecen en las teorías del procesamiento predictivo, como la concepción antropológica de fondo (que no es materialista, como sí lo es la de la mayor parte de los defensores del procesamiento predictivo), o la idea de un sujeto propiamente tal, irreductible a lo fisiológico. La cuestión, entonces, es si las teorías del procesamiento predictivo

[27] METZINGER, Th., *Being no one: The self-model theory of subjectivity*. Cambridge, MA: MIT Press, 2004.

[28] WIESE y METZINGER, “Vanilla PP for Philosophers”, op. cit.

[29] CLARCK, “Whatever Next?”, op. cit., p. 196.

pueden considerarse legítimamente, tal y como sugieren algunos, una especie de desarrollo lógico y científico de la filosofía trascendental kantiana. Si lo que se quiere decir con esto es que la neurociencia contemporánea estaría haciendo algo así como desarrollar o confirmar experimentalmente lo que Kant quiso hacer en filosofía, creo que es afirmar demasiado. Decir que el procesamiento predictivo tiene “sabor kantiano” puede llevar a engaño, porque de hecho las concepciones filosóficas e incluso el mismo propósito de las empresas kantiana y la de los defensores del procesamiento predictivo son radicalmente diferentes, en el sentido de que parten de supuestos diferentes y conducen a afirmaciones diferentes. Las posibles semejanzas no deben hacernos olvidar las profundas diferencias que existen entre ambos planteamientos.

Ocurre algo similar a lo que encontramos en la discusión sobre quién realizó una mejor interpretación de Kant, si los idealistas o los positivistas. Parece bastante claro que, si nos fijamos en los ataques de Kant contra la legitimidad de la metafísica dogmática, o si tenemos en cuenta el elevado ideal que tuvo de la ciencia, podemos encontrar en Kant elementos del posterior pensamiento positivista. Y también parece claro que, si tenemos en cuenta las reflexiones de Kant sobre la apercepción trascendental o sobre lo inalcanzable de la realidad nouménica, podemos ver también ahí elementos que utilizará posteriormente el idealismo. Pero no podemos olvidar que Kant no es ni un idealista ni un positivista. Es cierto que sus ideas se pueden utilizar en uno u otro sentido, pero si hemos de ser rigurosos, no podemos decir que el idealismo o el positivismo sean kantianos, a lo sumo que son posibilidades teóricas que pueden surgir de determinadas interpretaciones de su pensamiento. Otro asunto es hasta qué punto Kant consiguió evitar el idealismo con el que no estaba de acuerdo (pues nunca llegó a negar la existencia real de lo nouménico, cosa que le parecía de sentido común), o evitar el positivismo científico que tampoco hubiese aceptado (dado que el uso no teórico de la razón ofrece un conocimiento que, claramente, va más allá del científico).

V. CONCLUSIÓN

A la pregunta sobre si Kant puede ser considerado el padrino de las ciencias cognitivas se puede responder diciendo que sí y que no. Sí, si tenemos en cuenta algunas de sus afirmaciones de un modo más bien aislado del resto del contexto de su filosofía. No, si entendemos que la filosofía de Kant contiene o lleva por sí misma a las ciencias cognitivas actuales (o al planteamiento sobre la mente que vemos en ellas). Y otro tanto puede decirse sobre el uso que se hace de Kant en las teorías del procesamiento predictivo: es cierto que fuera del contexto de su filosofía, algunas afirmaciones de Kant suenan semejantes a las que hacen los defensores de este modo de entender el funcionamiento

del cerebro, y también es verdad que algunos de sus defensores se han basado en Kant. Pero decir que las teorías del procesamiento predictivo son algo así como el desarrollo científico de lo que Kant había hecho en filosofía, es afirmar algo que puede llevar a confusión, dadas las enormes diferencias que hay de fondo entre lo que afirman las teorías del procesamiento predictivo y lo que dice la filosofía kantiana. Creo, en definitiva, que debemos acercarnos a este tipo de filiaciones conceptuales con cautela. Las mismas cautelas aplicadas por el profesor Pedro Jesús Teruel al caso de la recepción de Kant en la filosofía de la mente contemporánea³⁰ pueden ser aplicadas al caso más concreto de las teorías del procesamiento predictivo.

VI. BIBLIOGRAFÍA

- BENI, M. D., “Commentary: The Predictive Processing Paradigm Has Roots in Kant” en *Frontiers in Systems Neuroscience* 11: 98, 2018. DOI: [10.3389/fnsys.2017.00098](https://doi.org/10.3389/fnsys.2017.00098).
- BRETTE, R., “Is coding a relevant metaphor for the brain?” en *Behavioral and Brain Sciences* 42, e215, 2019, pp. 1-58. DOI: <https://doi.org/10.1017/S0140525X19000049>.
- BRETTE, R., “Brains as Computers: Metaphor, Analogy, Theory or Fact?” en *Front. Ecol. Evol.* 10:878729, 2022. DOI: <https://doi.org/10.3389/fevo.2022.878729>.
- BROOK, A., *Kant and the Mind*. Cambridge: Cambridge University Press, 1994.
- BROOK, A., “Kant and Cognitive Science” en *Teleskop. Revista de pensamiento y cultura* I, 4, 2004, pp. 13-21. Traducido al castellano por Ester Hernández en Teruel, P. J. (Ed.), *Kant y las ciencias*. Madrid: Biblioteca Nueva, 2011, pp. 227-236.
- CLARK, A., “Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognition Science” en *Behavioral and Brain Sciences* 36.3, 2013, pp. 181-204. DOI: <https://doi.org/10.1017/S0140525X12000477>.
- CLARK, A., *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. New York: Oxford University Press, 2016. DOI: <https://doi.org/10.1093/acprof:oso/9780190217013.001.0001>.
- FRISTON, K., “Learning and inference in the brain”, *Neural Networks*, 16 (9) 2023, pp. 1325-1352. DOI: <https://doi.org/10.1016/j.neunet.2003.06.005>.
- FRISTON, K., “The free-energy principle: A rough guide to the brain?” en *Trends in Cognitive Sciences*, 13 (7), 2009, pp. 293-301. DOI: <https://doi.org/10.1016/j.tics.2009.04.005>.
- FRISTON, K., “The free-energy principle: A unified brain theory?” en *Nature Reviews Neuroscience*, 11 (2), 2020, pp. 127-138. DOI: <https://doi.org/10.1038/nrn2787>.
- FUCHS, Th. *Ecology of the Brain. The Phenomenology and Biology of the Embodied Mind*, Oxford: Oxford university Press, 2018. DOI: <https://doi.org/10.1093/med/9780199646883.001.0001>.
- HOHWHY, J., *The Predictive Mind*. Oxford: Oxford University Press, 2013. DOI: <https://doi.org/10.1093/acprof:oso/9780199682737.001.0001>.
- KANT, I., *Critica de la razón pura*, traducción de Pedro Ribas. Madrid: Alfaguara, 1998.

[30] TERUEL, “La recepción de Kant”, op. cit.

- KAUARK-LEITE, P., “¿Podría Kant arrojar luz sobre la indeterminación onto-lógica en la teoría cuántica?”, en *Naturaleza y Libertad. Revista de estudios interdisciplinares*, nº 19, 2025 , 107-127. <https://doi.org/10.24310/nyl.19.2025.21210>.
- LAKATOS, I., *Philosophical papers*. Cambridge: Cambridge University Press, 1980.
- LENOIR, T., “Operationalizing Kant: manifolds, models and mathematics in Helmholtz’s theories of perception”, en FRIEDMAN, M. y NORDMANN, A. (Eds.), *Kantian Legacy in Nineteenth-Century Science*. Cambridge, MA: MIT Press, 2006, pp. 141-210. DOI: <https://doi.org/10.7551/mitpress/4048.003.0008>.
- METZINGER, Th., *Being no one: The self-model theory of subjectivity*. Cambridge, MA: MIT Press, 2004.
- SÁNCHEZ, J.C.; FERNÁNDEZ, T.R. y LOY, I., “La génesis de la intuición. Helmholtz y la naturalización del sujeto trascendental kantiano” en *Revista de Historia de la Psicología*, Vol. 16, Nº 3-4, 1995, pp. 375-382.
- SCHLICHT, T. y NEWEN, A., “Kant and Cognitive Science Revisited” en *Logical Analysis and History of Philosophy / Philosophiegeschichte und logische Analyse*, Vol. 18, 2015, pp. 87-113. DOI: <https://doi.org/10.30965/26664275-01801008>.
- SPREVAK, M., “Predictive coding I: Introduction” en *Philosophy Compass*, e12950, 2024. DOI: <https://doi.org/10.1111/phc3.12950>.
- SWANSON, L.R., “The Predictive Processing Paradigm Has Roots in Kant” en *Front Syst Neurosci.* 10 (79), 2016. DOI: <https://doi.org/10.30965/26664275-01801008>.
- TERUEL, P.J., *Kant y las ciencias*. Madrid: Biblioteca Nueva, 2011.
- TERUEL, P.J., “La recepción de Kant en la *Philosophy of Mind*. Una revisión crítica de las fuentes kantianas”, en TERUEL, P. J. (ed.), *Kant y las ciencias*. Madrid: Biblioteca Nueva, 2011, pp. 244-302.
- TERUEL, P. J., “La síntesis final. implicaciones de la filosofía natural contenida en el Opus postumum para la cuestión «¿Qué es el ser humano?»”, en *Naturaleza y Libertad. Revista de estudios interdisciplinares*, nº 19, 2025, pp. 83-105. <https://doi.org/10.24310/nyl.19.2025.20334>.
- WIESE, W. y METZINGER Th., “Vanilla PP for Philosophers: A Primer on Predictive Processing”, en Metzinger, T y Wiese, W. (Eds.), *Philosophy and Predictive Processing*: 1, 2017. Frankfurt am Main: MIND Group. DOI: <http://doi.org/10.25358/openscience-624>.