

Rodrigo Díaz Olguín
Psicoterapeuta

Cipra
Psicoterapia Constructivista

De la entropía a la psicología: introducción al marco de inferencia activa.

Los organismos buscan minimizar la energía libre mediante la construcción de modelos predictivos del mundo. Estos modelos permiten a los organismos percibir, aprender y actuar de manera que reducen la incertidumbre y logran sus objetivos. Se explora la conexión entre la entropía termodinámica y la entropía de la información, y cómo los organismos resisten la desorganización. Se introducen conceptos como la energía libre variacional, la energía libre esperada, las mantas de Markov y las redes bayesianas, que sustentan el marco de Inferencia Activa. Finalmente, junto con las principales críticas, se discuten las implicaciones de este marco en otras áreas, como la teoría de las emociones y el estrés, revisando los posibles aportes para la psicología.

I. Introducción

Este artículo ofrece una síntesis del marco de Inferencia Activa, un modelo teórico desarrollado por Karl Friston que conecta principios de física estadística, teoría de la información y neurociencia cognitiva para explicar cómo los organismos vivos perciben, aprenden y actúan en su entorno. Aunque el alcance explicativo de este marco es promisorio, aún se requieren varias direcciones de investigación para seguir refinando su carácter empírico y el poder explicativo de sus algoritmos, además de compararlo con otros modelos no bayesianos (Hodson et al., 2023). El objetivo de este artículo es ofrecer una introducción que pueda ser útil para iniciar el estudio del marco.

El artículo se divide en cuatro partes. Primero, se revisarán las principales nociones que sirven como antecedente al marco. Se presentará la relación entre entropía termodinámica y de la información, que sustentan las premisas de Ashby. Además se presentará una visión general del concepto de manta de Markov como herramienta para definir límites estadísticos entre sistemas. Finalmente se presenta una una noción general de la forma que adopta el teorema de Bayes y la inferencia variacional en el enfoque de Friston. En la segunda parte se introducirán las nociones centrales de *energía libre variacional* y de *energía libre esperada*, como elementos que subyacen a la minimización de la incertidumbre tanto en la percepción como en la acción presente y futura. En la tercera parte se comentarán algunas de las áreas en que este marco se

ha desarrollado, con especial atención los posibles aportes en psicología, particularmente estudio de la emoción y el estrés. En la cuarta parte se discutirán algunos puntos de crítica y debate actual sobre el marco.

Se seguirá de alguna forma el mapa sugerido por Friston, quien propone distintas vías o caminos para la comprensión del marco de inferencia activa. Se condensarán las principales nociones de estas rutas, que conectan la entropía con teoría de probabilidades, intentando restringir las formulaciones matemáticas a las básicas para comprender el marco.

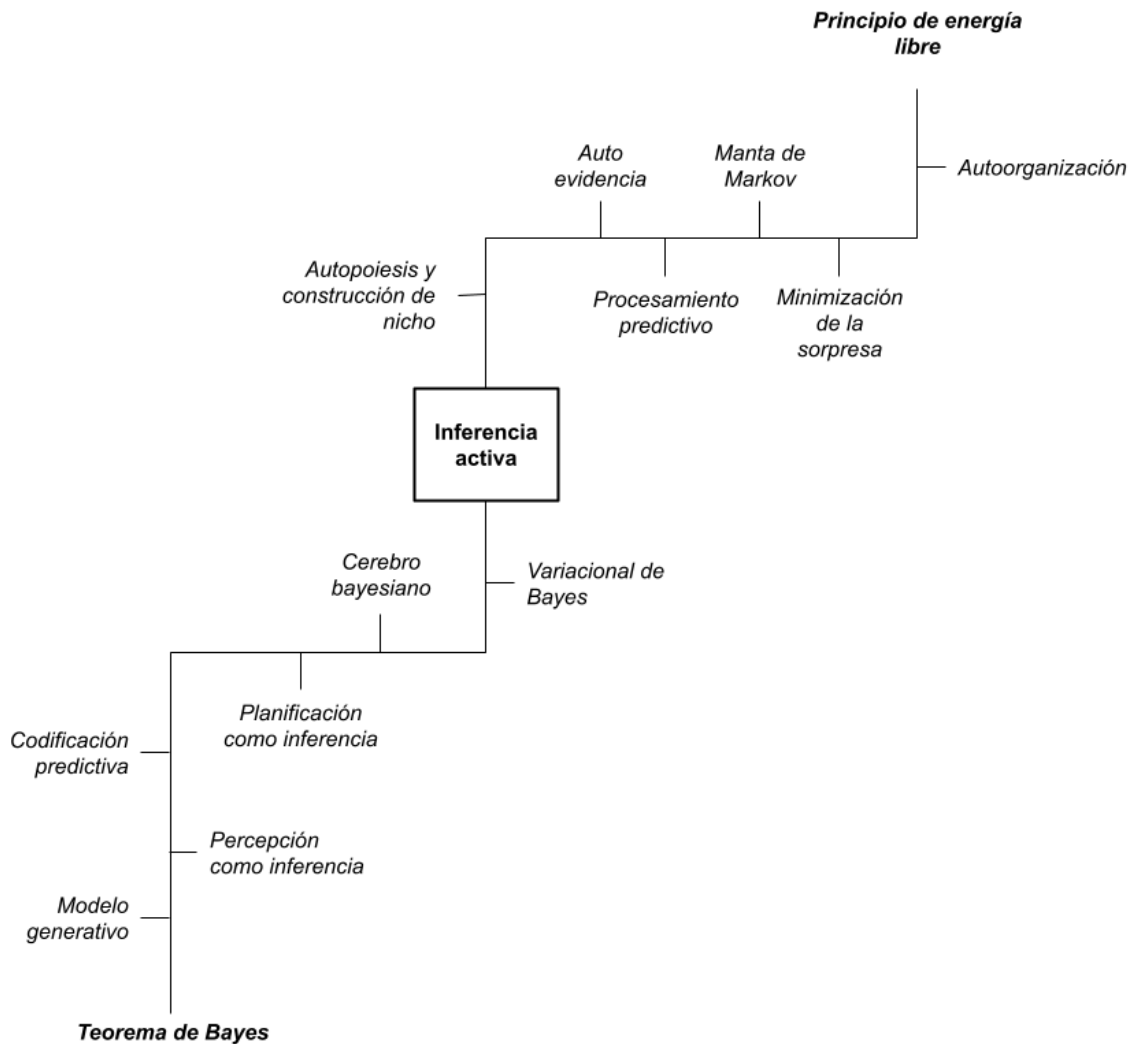


Figura 1. “Dos caminos a la Inferencia Activa”, esquema traducido de Parr, Pezzulo & Friston (2022)

II. Antecedentes

El *principio de energía libre* propone un principio unificador para entender la cognición y el comportamiento adaptativo en los sistemas vivos. Postula que los organismos operan minimizando una cantidad denominada «energía libre» mediante la inferencia activa y el procesamiento predictivo. La energía libre cuantifica la sorpresa o incertidumbre, la discrepancia entre el modelo predictivo y la evidencia del modelo. Según el *principio de energía libre*, los sistemas vivos buscan reducir esta energía libre generando y actualizando los modelos que anticipan las causas de la evidencia. Este proceso implica procesos de inferencia y aprendizaje constantes, ajustando las predicciones internas para alinearlas con las observaciones.

Al minimizar la energía libre, los organismos mantienen un estado de orden que armoniza con sus percepciones y acciones, permitiéndoles formular predicciones más precisas sobre su entorno y disminuir, en promedio, la incertidumbre. Este principio ofrece un marco teórico para comprender la percepción, el aprendizaje y la cognición, destacando la relevancia de la inferencia activa y de los modelos predictivos en la configuración del comportamiento y la adaptación.

Entropía y neguentropía.

La palabra «entropía» procede del griego ἐντροπία [«εν-» + «τροπή», en giro, cambio interno] y significa evolución o transformación. Este término fue acuñado en 1865 por Rudolf Clausius quien desarrolló las nociones planteadas por un joven Sadi Carnot en 1824. En la perspectiva de Clausius, la entropía se define como la parte de la energía de un sistema que no puede utilizarse para realizar trabajo, expresándose como el cociente entre el calor cedido por un cuerpo y su temperatura absoluta. En términos más simples, cuando un sistema transfiere calor a una temperatura constante, no toda esa energía puede convertirse en trabajo mecánico. La entropía cuantifica esta porción de energía no aprovechable, indicando la tendencia natural de los sistemas a evolucionar hacia estados de mayor desorden y menor disponibilidad de energía útil.

La entropía termodinámica —simbolizada como S — es una magnitud física para un sistema en equilibrio que mide el número de microestados compatibles con el macroestado de equilibrio¹. Esto puede entenderse como el grado de organización del sistema: a menor entropía encontraremos una mayor cantidad de estados improbables —a veces entendido como “orden”, aunque esto puede ser impreciso—, a mayor entropía la equivalencia de los microestados respecto al macroestado es mayor, es decir, hay mayor cantidad de estados probables. El curso entrópico, de estados

¹ Esta concepción estadística, de Boltzmann, se revisará más adelante,

improbables a los más probables, la tendencia natural hacia la homogeneidad, es una cualidad central que restringe la reversibilidad de los sistemas termodinámicos y los orienta, en términos de probabilidades, hacia la irreversibilidad² (Schrödinger, 1944; Feynman et al., 1970).

Los sistemas biológicos, al ser termodinámicamente abiertos, intercambian energía y entropía constantemente con su entorno. Operan lejos del equilibrio y son sistemas disipativos —disipan la energía— con comportamientos autoorganizados (Ashby, 1957, 1962; Nicolis & Prigogine, 1977; Haken, 1983; Kauffman, 1992). Sin embargo, van más allá de la autoorganización disipativa al interactuar bidireccionalmente con su entorno, resistiendo la deriva entrópica durante largos periodos.

Para compensar la entropía, estos sistemas emplean diversos mecanismos: incorporar componentes organizados de baja entropía, como proteínas; modular fases de desarrollo, como la metamorfosis en insectos; anticipar fluctuaciones energéticas, por ejemplo, mediante la estivación en animales; o modificar el entorno para generar condiciones favorables, entre otros. El grado de compensación entrópica diferencia a los sistemas biológicos de otros sistemas autoorganizados, permitiéndoles mantener sus estructuras y procesos en medio del cambio constante. Esta noción primero fue introducida como «entropía negativa» por Erwin Schrödinger en su libro *What is life?* (1943). Más tarde, fue el francés Léon Brillouin quien cambió la palabra por «neguentropía», para expresar la noción en una forma mucho más positiva, planteando que un sistema vivo importa neguentropía y la almacena (Brillouin, 1953). Otras formas aceptadas del término son «negentropía», «negantropía» y «sintropía».

En este marco, la optimización de la energía se convierte en una meta central para minimizar la energía libre termodinámica disponible. La energía libre termodinámica es la cantidad de energía en un sistema disponible para realizar trabajo a temperatura y presión constantes. Existen dos tipos principales: la energía libre de Gibbs —para procesos químicos— y la energía libre de Helmholtz —relacionada con sistemas a volumen constante—.

Más adelante se abordará cómo esta noción termodinámica culmina en el concepto de energía libre variacional dentro de la Inferencia Activa de Karl Friston.

² Este es el principio que subyace a la segunda ley de la termodinámica «La cantidad de entropía del universo tiende a incrementarse en el tiempo», que establece que en un sistema aislado, la entropía tiende a aumentar con el tiempo, reflejando una tendencia natural hacia la homogeneidad energética. La segunda ley se ha expresado de muchas maneras. Su primera formulación, antes de una adecuada definición de entropía y basada en la teoría calórica, es el *teorema de Carnot*, formulada por el científico francés Sadi Carnot, en 1824. Basándose en el trabajo de Carnot, la primera definición de la segunda ley basada en el concepto de entropía es la del científico alemán Rudolf Clausius en la década de 1850.

Entropía termodinámica e información.

La neguentropía permite conectar científicamente dos nociones que se mostraban intuitivamente relacionadas (Schrödinger, 1943): la lucha contra la entropía termodinámica y la gestión de la información. Los sistemas biológicos no solo combaten la tendencia a la desorganización, sino que también representan un aumento de densidad de información. En otras palabras, un sistema biológico es un sistema neguentrópico que constituye conocimiento. ¿Cómo se fue perfilando científicamente esta íntima relación entre entropía termodinámica y entropía de la información?

En sus inicios, la palabra «entropía» fue acuñada por el físico alemán Rudolf Clausius (1850) al postular la tesis de que en cualquier proceso mecánico, una pequeña cantidad de energía térmica se disipa gradualmente a través de la frontera del sistema.

Tiempo después, surge una comprensión probabilística de la entropía termodinámica, gracias a Ludwig Boltzmann, quien entre los años 1872 y 1875 desarrolló una perspectiva estadística de la entropía (Schrödinger, 1944; Feynman et al., 1970).

Según su enfoque, las partículas de un sistema aislado, pudiendo ocupar varios microestados posibles, tienden a aumentar su dispersión y eventualmente alcanzarán una configuración de máxima entropía, un estado homogéneo y desorganizado. Esta entropía (S) se define como:

$$S = k * \ln(\Omega)$$

Donde:

- k o k_b es la Constante de Boltzmann, que relaciona temperatura absoluta y energía.
- \ln es un logaritmo natural.
- Ω es el número de microestados accesibles al sistema.

Años después, Max Plank precisaría algunas variables de esa relación —la constante de Boltzmann— a través de un valor exacto.

En 1867 el matemático escocés James Clerk Maxwell propuso un experimento mental donde un ser imaginario, el después conocido como «demonio de Maxwell», podría disminuir la entropía de un sistema sin realizar trabajo adicional, desafiando la segunda ley de la termodinámica. En 1929, Leó Szilárd resolvió esta paradoja en su tesis doctoral, considerando al demonio como parte del sistema (Lanouette & Silard, 1992). Al hacer esto, Szilárd demostró que medir la velocidad molecular requiere un gasto energético, lo cual aumenta la entropía total del sistema —entropía del gas + la entropía del demonio y sus operaciones—. De este modo, vinculó la termodinámica con la teoría

de la información al identificar dos tipos de información (Boyd, Mandal & Crutchfield, 2017):

1. La «sorpresa»³ derivada de la medición.
2. La información antecedente, organizada en algún tipo de memoria o contingencia histórica.

No obstante Szilárd fue reconocido por muchos logros en su carrera y por su participación en varios eventos históricos, la expresión de esta vinculación y su importancia, tardaría años.

Sorpresa, incertidumbre y teoría de la información de Shannon.

Décadas después de la tesis de Szilárd, en lo que se considera un hito, el matemático e ingeniero Claude Shannon desarrolló la noción de lo que hoy se denomina *entropía de la información*⁴ en su famoso artículo *Una teoría matemática de la comunicación* (Shannon, 1948), en el que cuantificó el monto de información nueva que puede esperarse de una fuente, considerando que el máximo de incertidumbre ocurre cuando todos los eventos son igualmente probables de ocurrir. De acuerdo a esto, la probabilidad y la incertidumbre tienen una relación inversa: se define la incertidumbre asociada a una probabilidad particular como su logaritmo negativo, $-\log(p)$, es decir que un aumento en la probabilidad implica una reducción en la incertidumbre y viceversa.

En teoría de la información, el término «sorpresa» describe el grado en que un resultado sería inesperado si ocurriera, siendo mayor para eventos de baja probabilidad. En esta perspectiva, cuando un resultado ocurre, elimina la incertidumbre asociada a su probabilidad, produciendo información. Por lo tanto, cuantificar la incertidumbre permite medir la cantidad de información contenida en el resultado. Shannon (1948) desarrolló un marco general basado en este principio para cuantificar la información. La Entropía de Shannon (H) se expresa como:

$$H(X) = - \sum p(x) * \log_2 p(x)$$

³ En todo el documento, «sorpresa» no guarda relación con el sentimiento subjetivo de asombro, sino alude a obtener un resultado poco probable dada una distribución. En este contexto, algunos vocablos usados comúnmente adquiere un sentido técnico, más sencillo: por ejemplo, “sorpresa” se refiere a la autoinformación, “incertidumbre” a la entropía de la información y “creencias” bayesianas a distribuciones de probabilidad posterior.

⁴ En la nomenclatura inicial de Shannon, esto es *incertidumbre*. La incertidumbre promedio ponderada, también conocida como «sorpresa esperada», fue originalmente denominada "incertidumbre" por Shannon (1948). Más tarde, al notar su equivalencia matemática con la fórmula de la física para describir el desorden en sistemas cerrados, se adoptó el término "entropía" a sugerencia de John von Neumann.

Donde:

- $p(x)$ es la probabilidad de que ocurra el evento x .
- Σ representa la suma sobre todos los posibles valores de x .
- \log_2 es el logaritmo en base 2 (asumiendo un sistema binario de información).

Tiempo después William Ross Ashby, quien además de psiquiatra y neurólogo fue un destacado cibernético, reinterpretó la entropía de Shannon a través de su *ley de la variedad requerida*⁵ (Ashby, 1957; Boyd et al., 2017). Esta ley afirma que los sistemas deben poseer un número de respuestas al menos igual al número de desafíos que enfrentan (Ashby, 1957). Este enfoque permitió integrar la teoría de la información en la comprensión de sistemas biológicos y cognitivos.

Desde la perspectiva de Ashby, los sistemas biológicos restringen sus estados fisiológicos a un repertorio limitado de baja entropía (Ashby, 1957, 1962), cumpliendo con el objetivo de minimizar la variabilidad e incertidumbre. Esto supone un *límite* desde el cuál se pueden hacer inferencias, que separa básicamente qué es atribuible al organismo y qué no.

Este límite no solo es importante para fundamentar cálculos probabilísticos, sino que permite definir qué es un sistema, distinto de otro. Es útil notar que estrictamente, los límites de un sistema están dados por los criterios de observación utilizados para determinar cuáles son sus operaciones y cuáles operaciones corresponden a otros sistemas. Es decir, en una red de operaciones que están anidadas las unas con las otras, la definición de parámetros en las operaciones estudiadas son las que definen los límites del sistema⁶. En redes bayesianas, una Manta de Markov puede ser útil para formalizar esta noción de límite. Antes de revisar la noción de Manta de Markov, se desarrollará brevemente el rol que juega el teorema de Bayes en el marco de Inferencia Activa

⁵ «La variedad absorbe la variedad», se refiere a que los sistemas biológicos deben tener un número de mecanismos de control o respuestas que sea al menos igual o mayor que el número de perturbaciones/desafíos potenciales que deben enfrentar. Cualquier sistema complejo debe tener un cierto nivel de variedad interna –o complejidad– para ser capaz de funcionar y adaptarse dentro de su entorno. Para afrontar adecuadamente la diversidad de problemas que el mundo plantea, es necesario contar con un repertorio de respuestas que sean al menos tan matizadas como los problemas que enfrentas. Aunque esta ley se ideó originalmente en relación con la forma en que los organismos se adaptan, rápidamente se acopló al teorema de la información de Claude Shannon y a los sistemas en general.

⁶ Por ejemplo, algunas teorías de «nicho ecológico» consideran esta noción (Letten et al., 2017).

El Teorema de Bayes.

El Teorema de Bayes es una herramienta fundamental en estadística y probabilidad que nos permite actualizar nuestras creencias sobre un evento en función de nueva evidencia. Es muy utilizado en áreas como la inteligencia artificial, la medicina, la psicología y el aprendizaje automático.

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

En donde:

- $P(A|B)$: Probabilidad posterior. Es la probabilidad de que ocurra el evento A dado que se sabe que ha ocurrido B.
- $P(B|A)$: Probabilidad condicional. Es la probabilidad de que ocurra B si sabemos que ha ocurrido A.
- $P(A)$: Probabilidad a priori de A, es decir, la probabilidad de A antes de considerar la información de B.
- $P(B)$: Probabilidad marginal de B, que es la probabilidad total de que ocurra B, considerando todos los eventos posibles. En estadística bayesiana, representa la probabilidad de generar la muestra observada para todos los valores posibles de los parámetros; puede entenderse como la probabilidad del modelo en sí y, por lo tanto, a menudo se la denomina *evidencia del modelo* o simplemente *evidencia*. Más adelante en este artículo se revisará cómo el problema de marginalización es un tópico importante en este marco.

Por ejemplo, consideremos un médico psiquiatra que ajusta la probabilidad de que un paciente tenga un diagnóstico particular, como la depresión, basado en nueva información. Entonces el médico evalúa al paciente y registra la siguiente información inicial:

1. La prevalencia de la depresión, que notaremos “D”, en la población general es del 10%. Por lo tanto $P(D)=0,1$.
2. Si un paciente tiene depresión, hay un 80% de probabilidad de que informe insomnio, que notaremos “i”, es decir $P(i|D)=0,8$.
3. Si un paciente no tiene depresión, que notaremos con la negación $\neg D$, aún hay un 20% de probabilidad de que informe insomnio. Es decir, $P(i|\neg D)=0,2$.
4. El paciente informa que tiene insomnio. El terapeuta quiere saber cuál es la probabilidad de que el paciente tenga depresión dado que ha reportado insomnio, es decir, quiere saber cuánto es $P(D|i)$.

Ahora aplicando el teorema de Bayes:

$$P(D|i) = \frac{P(i|D) * P(D)}{P(i)}$$

Si los valores son:

- P(D) = 0,1
- $\neg D$ = 0,9 (probabilidad de no tener depresión)
- P(i|D) = 0,8
- P(i| $\neg D$) = 0,2

Primero se necesita saber cuál es P(i), es decir cuál es la probabilidad total de que el paciente tenga insomnio considerando todas las maneras en que ese evento puede ocurrir. En el ejemplo, el insomnio puede darse en dos escenarios posibles:

- El paciente tiene depresión (“D”), con una probabilidad de P(D)
- El paciente no tiene depresión (“ $\neg D$ ”), con una probabilidad de P($\neg D$)

La probabilidad total de que el paciente tenga insomnio entonces, se calcula como la suma ponderada de las probabilidades de estos dos escenarios. La notación simple para esto es:

$$P(i) = \sum_{i \in I} P(i|D) * P(D)$$

Donde “I” (“i” mayúscula) son todos los posibles estados de “i”.

Esto se puede calcular de la siguiente forma. La probabilidad del primer escenario —la probabilidad de que el paciente tenga insomnio dado que tiene depresión— se expresa P(i|D) * P(D) y la del segundo escenario —la probabilidad de que el paciente tenga insomnio dado que no tiene depresión— se expresa P(i| $\neg D$) * P($\neg D$). Entonces la suma de las probabilidades, es:

$$P(i) = P(i|D) * P(D) + P(i|\neg D) * P(\neg D)$$

Sustituyendo los valores del ejemplo:

- P(D) = 0,1 (10% de probabilidad de tener depresión)
- P($\neg D$) = 1 - P(D) = 0,9 (90% de probabilidad de no tener depresión)
- P(i|D) = 0,8 (80% de probabilidad de insomnio si tiene depresión)
- P(i| $\neg D$) = 0,2 (20% de probabilidad de insomnio sin depresión)

Entonces, el cálculo de la probabilidad total de que el paciente tenga insomnio en todos los escenarios:

$$P(i) = P(i|D) * P(D) + P(i|\neg D) * P(\neg D)$$

$$P(i) = (0,8 * 0,1) + (0,2 * 0,9)$$

$$P(i) = 0,08 + 0,18 = 0,26$$

La probabilidad total de que el paciente informe insomnio es del 26%. Esto considera tanto a los pacientes con depresión como a los que no la tienen, ponderando cada caso según su probabilidad.

Tenemos:

$$P(D|i) = \frac{P(i|D) * P(D)}{P(i)}$$

$$P(D|i) = \frac{0,8 * 0,1}{0,26} = \frac{0,08}{0,26} \approx 0,3077$$

Es decir, la probabilidad de que el paciente tenga depresión dado que reporta insomnio es aproximadamente del 30.77%. Aunque el insomnio es un síntoma común de la depresión, también puede ocurrir por otras razones —estrés, problemas físicos, etc.—, lo que reduce la probabilidad.

Por supuesto, ningún médico psiquiatra basará el diagnóstico solo en el conteo de síntomas, pero un resultado así podría ayudarle a considerar más profundamente el historial del paciente, realizar más evaluaciones y explorar otros síntomas relacionados con la depresión.

El teorema de Bayes es un método fundamental en el marco de energía libre, que permite calcular la probabilidad de un estado dado una observación. La relación con el enfoque de Friston radica en que el marco de energía libre utiliza el teorema de Bayes como base para la inferencia perceptual. Friston propone que el cerebro funciona como una máquina de inferencia bayesiana, constantemente actualizando sus creencias sobre el mundo a partir de las observaciones sensoriales. Sin embargo, calcular la probabilidad de una observación puede ser computacionalmente costoso, especialmente cuando hay muchas posibles observaciones. El enfoque de Friston utiliza una aproximación al teorema de Bayes, minimizando la "energía libre", que es una medida de la discrepancia entre las creencias del agente y las observaciones sensoriales

Inferencia variacional y teorema de Bayes.

En términos simples, el teorema de Bayes en el marco de Inferencia Activa combina la probabilidad previa de un estado —lo que se sabe antes de la observación— con la probabilidad de la observación dado ese estado —la «verosimilitud»—, para obtener finalmente la probabilidad posterior del estado —la probabilidad actualizada después de la observación—.

Para probabilidades de observación y estados, consideremos el teorema de Bayes:

$$P(s|o) = \frac{P(o|s) * P(s)}{P(o)}$$

En donde:

- $P(s|o)$, *probabilidad posterior*, la probabilidad de un estado “s” dada la observación “o”
- $P(o|s)$, es la *probabilidad condicional* de observar “o” dado que el estado es “s”. Esto mide qué tan probable es que ocurra la observación “o” si asumimos que el estado actual es “s”. Es una forma de evaluar qué tan consistente es la observación con el estado.
- $P(s)$, es la *probabilidad a priori* del estado s, es decir, la probabilidad de que el estado s ocurra antes de observar cualquier dato o evidencia. Representa el conocimiento previo sobre los estados.
- $P(o|s) * P(s)$, es el cálculo de una medida conjunta que combina qué tan bien el estado “s” explica la observación “o” — $P(o|s)$ — y la creencia inicial en el estado “s” — $P(s)$ —. En términos prácticos, es el numerador del teorema de Bayes, que se utiliza para actualizar nuestra creencia en el estado “s” después de observar “o”. En el teorema de Bayes, esta actualización se formaliza luego al dividirla por $P(o)$, la probabilidad total de la observación, para finalmente obtener la probabilidad a posteriori $P(s|o)$.

En el marco de Inferencia Activa, un problema que a menudo se considera de interés involucra la probabilidad de observación. Suponiendo que no esté simplemente dada de antemano, tiene que ser derivada. Pero como en situaciones prácticas hay infinitas observaciones potenciales, este cálculo es muy problemático sino imposible, dado que implica iterar sobre todas las situaciones que potencialmente se aplican⁷.

Frente a este problema de calcular directamente la distribución posterior exacta, se han propuesto métodos sencillos, como por ejemplo:

⁷ Esto se denomina “marginalización”

- *Aproximación de Monte Carlo*: utiliza muestras aleatorias a través de diferentes procedimientos para aproximar la distribución posterior. Entre las técnicas más comunes se encuentran el muestreo de Gibbs, el muestreo por importancia y el método de aceptación y rechazo.
- *Aproximación determinista*: aborda la estimación de la distribución posterior mediante métodos deterministas, como la cuadratura de Gauss o la factorización basada en momentos.
- *Inferencia variacional*: trata la inferencia como un problema de optimización, donde se busca identificar una distribución que se asemeje lo más posible a la distribución posterior verdadera.

Los métodos variacionales transforman problemas complejos en problemas más simples, los tratan como problemas de optimización. La inferencia variacional consiste en maximizar una medida conocida como la "evidencia límite inferior" (ELBO, por sus siglas en inglés). En términos simples, este enfoque transforma los problemas de inferencia probabilística —es decir, determinar el valor de una variable dada la información sobre otra— en problemas de optimización. En este contexto, optimizar significa encontrar los valores que maximicen o minimicen una función. En el caso de la inferencia variacional, el objetivo es maximizar la ELBO. De manera equivalente, esto también puede interpretarse como minimizar otra cantidad llamada «divergencia de Kullback-Leibler» (KL), que mide qué tan diferente es una distribución aproximada de la verdadera distribución posterior. A lo largo del tiempo, se han propuesto diversos métodos para lograr esta maximización de la ELBO, lo que permite aproximar soluciones a problemas de inferencia probabilística de manera más eficiente y manejable.

Divergencia Kullback-Leibler.

La distribución posterior no solo debe alcanzar su punto máximo, sino hacerlo en el lugar correcto. Es decir, los estados inferidos deben ser mínimamente inciertos y al mismo tiempo correctos. La teoría de la información ofrece una solución mediante la divergencia de Kullback-Leibler (KL), que mide la divergencia informativa entre dos distribuciones. La divergencia KL es una medida asimétrica que cuantifica la diferencia o similitud entre dos distribuciones de probabilidad, P y Q . Esta mide el número esperado de bits adicionales requeridos para codificar muestras de P utilizando un modelo basado en Q , en lugar de uno basado en P . En este contexto, P suele representar la distribución «verdadera» de los datos, observaciones o una distribución teórica, mientras que Q suele corresponder a una teoría, modelo, descripción o

aproximación de P . Aunque frecuentemente se la denomina métrica o distancia, la divergencia KL no cumple con los criterios formales para ser considerada como tal. Por ejemplo, no es simétrica: la divergencia KL de P a Q no equivale necesariamente a la divergencia de Q a P . La divergencia KL pertenece a una categoría más amplia de medidas conocidas como divergencias f . Fue introducida originalmente en 1951 por Solomon Kullback y Richard Leibler como una medida direccionada para comparar distribuciones.

Este enfoque permite cuantificar el costo informativo tanto de la incertidumbre como de la incorrección en una distribución. Geoffrey Hinton y sus colaboradores demostraron que este enfoque puede aplicarse a la optimización de la inferencia bayesiana aproximada (Hinton & Zemel, 1993; Hinton, 2002). Este método implica la minimización de una cantidad combinada, utilizando una analogía física: inferir el estado que produce una observación se asemeja a determinar la configuración de un sistema físico. Según Dayan et al. (1995), *"las explicaciones alternativas de un ejemplo pueden verse como configuraciones alternativas de un sistema físico"*. El método busca reducir la incertidumbre, descrita mediante probabilidades logarítmicas negativas. Sin embargo, los conceptos se expresan con la terminología de la física estadística. Por ejemplo, el logaritmo negativo de la probabilidad de predicción se describe como «energía de la observación», mientras que la cantidad crítica a minimizar es el exceso de la «energía esperada sobre la sorpresa esperada». Este exceso, conocido en física estadística como *energía libre de Helmholtz*, cuantifica la incertidumbre introducida en las predicciones debido a la ambigüedad.

En términos informativos, la energía libre refleja el grado en que las predicciones son innecesariamente inciertas, considerándola una forma de ruido estadístico. El método también elimina la necesidad de observaciones previas, ya que se basa únicamente en probabilidades de estado y predicción. Esto asegura que las distribuciones posteriores aproximadas converjan hacia sus contrapartes verdaderas. Además, la energía libre calculada de esta manera siempre supera la derivada de la distribución posterior verdadera, asegurando una aproximación progresivamente más precisa y menos incierta.

Mantas de Markov y redes bayesianas.

Los sistemas biológicos se organizan de forma autónoma al minimizar su energía libre. La manta de Markov es una herramienta estadística fundamental para entender cómo estos sistemas logran dicha minimización mediante procesos de inferencia (Ashby, 1957, p. 195).

Una manta de Markov actúa como un límite estadístico que separa los estados internos y externos de un sistema. Este límite se define como un conjunto de variables que protege a una variable aleatoria de la influencia directa de otros nodos en una red bayesiana⁸. En otras palabras, la manta de Markov de una variable corresponde a un subconjunto de elementos que interactúan directamente con ella. Conocer todo sobre este subconjunto hace innecesario cualquier conocimiento adicional sobre elementos fuera de él para entender la variable de interés. Esto es relevante porque permite realizar inferencias sobre una variable en un modelo gráfico utilizando únicamente la información local proporcionada por su manta de Markov.

Las mantas de Markov pueden autoensamblarse en sistemas globales, creando capas de límites anidados. Esto sugiere que cualquier sistema vivo está compuesto por mantas de Markov, y sus límites no necesariamente coinciden con los límites biofísicos de un organismo.

La capacidad de mantener esta separación permite a los sistemas realizar inferencias activas, es decir, actuar sobre el entorno para reducir la incertidumbre acerca de las causas de sus sensaciones. Este proceso de minimización de la *sorpres*a⁹ a través de la acción es crucial para preservar la integridad de la manta de Markov y, por ende, para garantizar la existencia del sistema mismo.

Desde un enfoque probabilístico, este principio puede formalizarse con la función $F(s,a,r)$ que opera como una función objetivo en un modelo estadístico:

$$F(s, a, r) = - \ln p(s, a|m) + D_{KL} [q(\varphi|r) || p(\varphi|s, a)]$$

Donde:

- s representa un estado y a representa una acción

⁸ La inferencia bayesiana es un tipo de inferencia estadística en la que las evidencias u observaciones se emplean para actualizar o inferir la probabilidad de que una hipótesis pueda ser cierta. El nombre «bayesiana» proviene del uso frecuente que se hace del teorema de Bayes durante el proceso de inferencia. La inferencia bayesiana usa un estimador numérico del grado de creencia en una hipótesis aún antes de observar la evidencia y calcula un estimador numérico del grado de creencia en la hipótesis después de haber observado la evidencia. La inferencia bayesiana generalmente se basa en grados de creencia, o probabilidades subjetivas, en el proceso de inducción y no necesariamente declara proveer un método objetivo de inducción.

⁹ El término «sorpresa» se refiere a la medida de cuánto difieren las observaciones actuales respecto a sus observaciones preferidas, aquellas que preservan su integridad. Esto no se puede lograr pasivamente, sino que se debe actuar, controlar de manera adaptativa, los bucles de acción-percepción para obtener las observaciones sensoriales deseadas. Entonces, aunque no se puede minimizar directamente la sorpresa, se puede minimizar un proxy llamado *energía libre variacional* (Friston et al., 2006).

- r representa una recompensa o resultado
- m modelo o contexto en el que ocurren s , a
- φ representa una variable latente (desconocida) que está siendo inferida
- $p()$ es una distribución de probabilidad a priori (lo que se asume antes de ver los datos)
- $q()$ es una distribución de probabilidad aproximada (posiblemente una inferencia posterior).
- D_{KL} es la *divergencia de Kullback-Leibler*, que mide la diferencia entre dos distribuciones de probabilidad.

La función $F(s, a, r)$ toma una forma de *función objetivo* (o de pérdida) en un modelo probabilístico. Es decir, se busca minimizar F para lograr dos cosas:

- Minimizar la sorpresa: $-\ln p(s,a|m)$, donde s representa el estado, a la acción y m el modelo. Este término mide la improbabilidad de observar una combinación estado-acción dado un modelo. Si $p(s,a|m)$ es alto —es decir, el estado y la acción son probables bajo el modelo—, la sorpresa será baja. En términos prácticos, este enfoque busca evitar eventos inesperados en el modelo, haciendo que el modelo sea consistente con los datos.
- Regularizar las inferencias: $D_{KL} [q(\varphi | r) || p(\varphi|s,a)]$, en donde la divergencia Kullback-Leibler compara la distribución aproximada $q(\varphi|r)$ —es decir, cómo inferimos la variable latente φ después de observar r — con la distribución original o esperada $p(\varphi|s,a)$. Busca que q sea lo más similar posible a p , penalizando grandes discrepancias, para garantizar coherencia entre las inferencias y el modelo original —asegura que la distribución aproximada de las variables latentes sea lo más parecida posible a la distribución esperada, evitando discrepancias significativas—.

III. El marco de Inferencia Activa

En el apartado anterior se revisaron algunos elementos básicos que pueden servir de introducción para el marco de Inferencia Activa. Este marco, inicialmente concebido para explicar datos provenientes de investigaciones en neurociencia, ha comenzado a expandir su alcance hacia otras áreas. Estas incluyen teoría de la emoción, estudio del estrés, la contextualización de procesos cognitivos (Butz et al., 2024), la neurofenomenología de la meditación (Lutz et al., 2024), el estudio del uso de sustancias (Hakimi et al., 2024), la comprensión de sesgos optimistas y la actualización de creencias (Karnick & Brick, 2024), el modelado del desarrollo cognitivo infantil (Theuer et al., 2024), y la biomimética del aprendizaje estructurado (Friston et al., 2024). También abarca consideraciones sobre la inferencia activa en el contexto de actores sociales y teorías neuro-bio-sociales (Cheadle et al., 2024), así como el estudio de narrativas y sus funciones adaptativas cognitivas y sociales (Bouizegarene et al., 2024).

En este apartado se sintetizarán los aspectos centrales del marco propiamente tal, y se considerarán dos de sus elementos: la *energía libre variacional* y la *energía libre esperada*, para luego comentar algunas de las implicaciones de estas premisas.

Energía libre variacional

El término energía libre se utiliza en física estadística para caracterizar sistemas termodinámicos. Sin embargo, en el marco de la Inferencia Activa, aunque se emplean las mismas ecuaciones, estas describen el estado de creencias de un agente en relación con un modelo generativo. Por lo tanto, al hablar de minimización de energía libre, nos referimos a procesos que ajustan las creencias del agente, no a cambios en las partículas que constituyen un sistema biológico. Para evitar malentendidos, se adopta el término *energía libre variacional*, una denominación común en el ámbito del aprendizaje automático. A diferencia de la energía libre en la termodinámica estadística de equilibrio, la Inferencia Activa se aplica a organismos vivos o sistemas fuera de equilibrio, los cuales interactúan continuamente con su entorno (Friston, 2019).

Durante la década de 1990, algunos investigadores como Geoffrey Hinton y Karl Friston comenzaron a estudiar el concepto de *energía libre* como una medida calculable y manejable de la discrepancia entre distribuciones conjuntas y evidencias de un modelo (Hinton & Zemel, 1993; Hinton, 2002; Friston, 2005). Karl Friston (2005), propuso un marco en el que el cerebro bayesiano emerge de un principio general de minimización de la energía libre (Friston et al., 2006).

De acuerdo a Friston, realizar una inferencia bayesiana exacta resulta computacionalmente inviable en la mayoría de los casos debido a que en modelos complejos puede existir una gran cantidad de estados ocultos que deben marginalizarse y porque la operación de marginalización puede requerir resolver integrales analíticamente intratables.

Para abordar estas limitaciones, la inferencia activa introduce una *aproximación variacional* de inferencia bayesiana¹⁰, que simplifica los cálculos reemplazando las cantidades intratables —la probabilidad posterior $P(x|y)$ y la evidencia del modelo $P(y)$ — por aproximaciones computables: una posterior aproximada $Q(x)$, y la *energía libre variacional* F , transformando el problema original de inferencia bayesiana en un problema de optimización¹¹. Esto no solo facilita los cálculos, sino que también proporciona una herramienta poderosa para integrar percepción y acción en sistemas complejos.

Para la *energía libre variacional*, Karl Friston sintetiza la siguiente ecuación:

$$\begin{aligned}
 F[Q, y] &= \underbrace{-\mathbb{E}_{Q(x)}[\ln P(y, x)]}_{\text{energía}} - \underbrace{H[Q(x)]}_{\text{entropía}} \\
 &= \underbrace{D_{KL}[Q(x)||P(x)]}_{\text{complejidad}} - \underbrace{\mathbb{E}_{Q(x)}[\ln P(y|x)]}_{\text{precisión}} \\
 &= \underbrace{D_{KL}[Q(x)||P(x|y)]}_{\text{divergencia}} - \underbrace{\ln P(y)}_{\text{evidencia}}
 \end{aligned}$$

¹⁰ La inferencia variacional es una técnica matemática utilizada en estadística bayesiana y aprendizaje automático para aproximar distribuciones posteriores complejas que son difíciles de calcular directamente. Esta función objetivo aparece en inferencia variacional para ajustar $Q(x)$ una distribución propuesta, de modo que sea una buena aproximación a $P(x|y)$, la verdadera posterior. En otras palabras, es un método para resolver problemas de inferencia (determinar las probabilidades de eventos o parámetros desconocidos) utilizando técnicas de optimización.

¹¹ El concepto de energía libre variacional se introdujo en la física estadística para convertir problemas difíciles de integración de densidad de probabilidad en problemas de optimización más fáciles. Es una cantidad de información teórica (como la sorpresa), a diferencia de una cantidad termodinámica. La energía libre variacional se ha explotado en el aprendizaje automático y las estadísticas para resolver muchos problemas de inferencia y aprendizaje. En este contexto, la sorpresa se llama la evidencia -negativa- del modelo. Esto significa que minimizar la sorpresa es lo mismo que maximizar la evidencia sensorial de la existencia de un agente, si consideramos al agente como un modelo de su mundo.

Esta ecuación representa tres distintas maneras de expresar la función objetivo $F[Q, y]$ en inferencia variacional. Cada forma resalta un aspecto diferente del proceso de optimización y provee de intuiciones para la ciencia cognitiva.

La primera línea de la ecuación indica que minimizar con respecto a Q implica garantizar la coherencia con el modelo generativo —energía— y, al mismo tiempo, mantener una alta entropía posterior: en ausencia de datos o creencias¹² previas claras, se adoptan creencias de máxima incertidumbre sobre los estados ocultos del mundo. En otras palabras, cuando no contamos con información suficiente, es preferible permanecer en un estado de incertidumbre y adoptar creencias de alta entropía.

La segunda línea resalta que la minimización de la energía libre puede interpretarse como el proceso de encontrar la mejor explicación para los datos sensoriales, privilegiando aquella que sea lo suficientemente simple —mínima complejidad— y al mismo tiempo capaz de explicar los datos con precisión. Desde una perspectiva cognitiva, esta preferencia por explicaciones simples es crucial, ya que actualizar las creencias previas para acomodar nuevos datos implica un costo cognitivo (Zénon et al., 2019). Por ello, es deseable adoptar explicaciones que difieran lo menos posible de las creencias anteriores.

La tercera línea de la ecuación describe la energía libre como un límite superior de la evidencia logarítmica negativa. Este límite está determinado por la divergencia entre Q —la distribución posterior aproximada— y la verdadera probabilidad posterior $P(x|y)$, que solo sería accesible mediante cálculos exactos en lugar de variacionales. A medida que esta divergencia disminuye, la energía libre se aproxima a la evidencia logarítmica negativa —sorpresa— y se iguala a ella si Q coincide exactamente con $P(x|y)$. Esto proporciona una base formal para la inferencia perceptiva, que puede interpretarse como el proceso de minimizar la energía libre optimizando Q para acercarlo al posterior exacto.

Esta tercera expresión $D_{KL}[Q(x)||P(x|y)] - \ln P(y)$ es frecuente en la inferencia probabilística y el aprendizaje basado en modelos bayesianos. Por ejemplo en inferencia bayesiana variacional, $Q(x)$ es la aproximación a la posterior $P(x|y)$ en donde el objetivo es minimizar esta expresión para encontrar la mejor aproximación posible $Q(x)$. En modelos generativos probabilísticos, se utiliza como parte del cálculo de una función objetivo, como la evidencia marginal negativa o el ELBO (Evidence Lower

¹² En este contexto, las «creencias» se entienden como distribuciones de probabilidad previas. Existen diferentes tipos de distribuciones, y en un modelo generativo basado en la inferencia bayesiana, estas pueden interpretarse como sobre-distribuciones (de naturaleza más aleatoria) que explican evidencias particulares.

Bound). La expresión $D_{KL}[Q(x)||P(x|y)] - \ln P(y)$ combina conceptos de la teoría de la información —la divergencia de Kullback-Leibler (DKL)— y la probabilidad. La divergencia de Kullback-Leibler es una medida de qué tanta información se pierde cuando se usa una distribución de probabilidad Q para aproximar una distribución de probabilidad P. En otras palabras, cuantifica la diferencia entre dos distribuciones de probabilidad. No es una métrica en el sentido matemático estricto, ya que no es simétrica ($DKL(P|Q) \neq DKL(Q|P)$).

La expresión $D_{KL}[Q(x)||P(x|y)] - \ln P(y)$ calcula la divergencia entre $Q(x)$ —una distribución de probabilidad sobre la variable x— y $P(x|y)$ —la distribución de probabilidad condicional de x dado y, es decir, la probabilidad de x sabiendo que y ha ocurrido—. Por lo tanto, mide la diferencia entre la distribución $Q(x)$ y la distribución de x dado y. En segundo lugar, la expresión incluye un negativo logarítmico $-\ln P(y)$, en donde $P(y)$ es la probabilidad que ocurra el evento “y”, por lo que $-\ln P(y)$ es el negativo del logaritmo natural de $P(y)$. Esto se conoce como la «sorpresa» o la «auto-información» de “y”, que significa básicamente que cuanto menor sea la probabilidad de y, mayor será su sorpresa.

Sin embargo, la inferencia perceptiva no es la única vía para minimizar la energía libre. También es posible modificar el término de evidencia logarítmica actuando sobre los datos sensoriales, es decir, cambiando el entorno que genera las observaciones. Desde una perspectiva cognitiva, esta descomposición resulta interesante, ya que establece una relación directa entre los dos objetivos complementarios de la inferencia activa: percepción —minimizar la divergencia— y acción —maximizar la evidencia—.

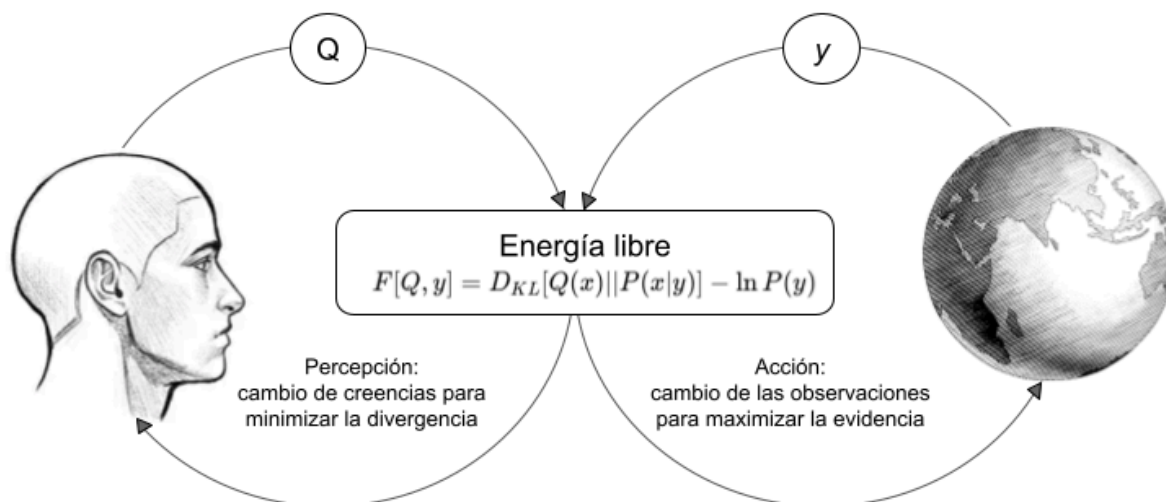


Fig. 02. Adaptado de Parr et al., 2022, roles complementarios de la percepción y acción en la minimización de la *energía libre variacional*.

En otras palabras, los sistemas biológicos, en su reducción de la energía libre, pueden cambiar la densidad de reconocimiento para modificar las expectativas condicionales sobre lo que se muestrea o pueden cambiar las muestras sensoriales —es decir, la entrada sensorial— para que se ajusten a las expectativas. Esta actividad bidireccional es denominada *inferencia activa*.

Energía libre esperada

La inferencia activa es un marco teórico que explica cómo los organismos vivos aseguran su existencia minimizando la sorpresa, representada operativamente mediante la *energía libre variacional*. Este enfoque integra la percepción y la acción en un único proceso inferencial, permitiendo que los organismos adapten su modelo generativo interno a las condiciones del entorno.

La *energía libre esperada* extiende este marco al ámbito prospectivo y contrafactual, proporcionando un criterio para evaluar políticas futuras en términos de su capacidad para reducir la sorpresa. Al considerar tanto el riesgo como la ambigüedad asociados a diferentes políticas, los organismos pueden anticipar las consecuencias de sus decisiones y seleccionar cursos de acción que optimicen su supervivencia y adaptación.

La *energía libre esperada* representa una forma avanzada de cognición prospectiva, siendo la planificación su expresión más prominente. Planificar una secuencia de acciones implica anticipar las observaciones futuras que se espera obtener. En este contexto, cada secuencia posible de acciones recibe el nombre de política. Es importante diferenciar entre acción y política. Mientras que la acción tiene un impacto directo sobre el entorno, la política constituye una hipótesis acerca de un patrón o forma de comportamiento.

Los resultados asociados a una política no son directamente accesibles, ya que pertenecen al ámbito del futuro. Sin embargo, pueden ser previstos al integrar dos componentes del modelo generativo:

1. *Creencias sobre cómo cambian en el tiempo los estados ocultos en función de una política específica.* Se utiliza la notación \tilde{x} para denotar una secuencia o trayectoria de estos estados, condicionados por la política π . Este componente entonces se describe mediante $P(\tilde{x}|\pi)$, que define la dinámica del modelo.
2. *Distribución de probabilidad usual.* Describe las observaciones que se esperan en cada posible estado. Este componente complementa al anterior al vincular estados con observaciones a través de $P(y|x)$.

Combinando estos elementos, un organismo puede activar su modelo generativo de forma indirecta para formular preguntas del tipo «¿qué pasaría si?». Esto permite realizar simulaciones contrafactuales, evaluando las posibles consecuencias de diversas políticas.

Mediante la marginalización de los estados, se obtiene una probabilidad marginal o evidencia de una política, $P(\tilde{y}|\pi)$, que puede aproximarse utilizando la energía libre esperada asociada (G). En otras palabras, al conocer cómo las políticas afectan las transiciones de estado, es posible calcular la probabilidad de una secuencia de observaciones bajo una política dada. La inferencia activa descompone el proceso de planificación en dos operaciones sucesivas:

1. *Cálculo de una puntuación para cada política.* Se establece una creencia previa acerca de las políticas, en donde las políticas óptimas se asocian con altas probabilidades, mientras que las menos favorables tienen probabilidades bajas. En este contexto, la calidad de una política se evalúa a través de su energía libre esperada negativa. Este criterio es análogo al utilizado en modelos generativos, donde la calidad del ajuste se mide por la energía libre negativa. La *energía libre esperada* (G) de una política difiere de la *energía libre variacional* (F), ya que el cálculo de G implica considerar observaciones futuras dependientes de la política. En contraste, F se limita a las observaciones presentes y pasadas. Por

tanto, el cálculo de G requiere del modelo generativo para predecir observaciones futuras asociadas a cada política en un horizonte de planificación definido. Además, dado que una política se extiende a lo largo de múltiples periodos de tiempo, su energía libre esperada debe integrar todas las etapas temporales futuras.

2. *Formación de creencias posteriores sobre las políticas.* La energía libre esperada de cada política puede transformarse en una puntuación de calidad al tomar su valor negativo. Esta puntuación se convierte en una probabilidad previa para los agentes que realizan inferencia activa. Las políticas asociadas con menores valores de energía libre esperada obtienen mayores probabilidades y se convierten en aquellas que el organismo espera seguir.

La energía libre esperada puede descomponerse en múltiples formas matemáticamente equivalentes, cada una proporcionando una perspectiva diferente sobre esta cantidad. Estas descomposiciones son útiles no solo para interpretar la energía libre esperada, sino también para entender cómo los sistemas de inferencia activa pueden integrar diferentes estrategias de decisión, desde maximizar la utilidad pragmática hasta reducir la incertidumbre epistémica, o viceversa. Tres interpretaciones son descritas a continuación:

$$\begin{aligned}
 G(\pi) &= \underbrace{-\mathbb{E}_{Q(\tilde{x}, \tilde{y}|\pi)}[D_{KL}[Q(\tilde{x}|\tilde{y}, \pi)||Q(\tilde{x}|\pi)]]}_{\text{ganancia informativa}} - \underbrace{\mathbb{E}_{Q(\tilde{y}|\pi)}[\ln P(\tilde{y}|C)]}_{\text{valor pragmático}} \\
 &= \underbrace{\mathbb{E}_{Q(\tilde{x}|\pi)}[H[P(\tilde{y}|\tilde{x})]]}_{\text{ambigüedad esperada}} + \underbrace{D_{KL}[Q(\tilde{y}|\pi)||P(\tilde{y}|C)]}_{\text{riesgo (resultados)}} \\
 &\leq \underbrace{\mathbb{E}_{Q(\tilde{x}|\pi)}[H[P(\tilde{y}|\tilde{x})]]}_{\text{ambigüedad esperada}} + \underbrace{D_{KL}[Q(\tilde{x}|\pi)||P(\tilde{x}|C)]}_{\text{riesgo (estados)}} \\
 &= \underbrace{-\mathbb{E}_{Q(\tilde{x}, \tilde{y}|\pi)}[\ln P(\tilde{y}, \tilde{x}|C)]}_{\text{energía esperada}} - \underbrace{H[Q(\tilde{x}|\pi)]}_{\text{entropía}} \\
 Q(\tilde{x}, \tilde{y}|\pi) &\triangleq Q(\tilde{x}|\pi)P(\tilde{y}|\tilde{x})
 \end{aligned}$$

1. *Ganancia informativa y valor pragmático.* Esta descomposición resalta el valor de la exploración —ganancia informativa— y la explotación —valor pragmático— utilizando las mismas unidades. Al minimizar la energía libre esperada, el balance entre estos términos determina si el comportamiento de un organismo será predominantemente *exploratorio* o *explotador*. El valor pragmático se deriva de las creencias previas sobre observaciones deseadas, formalizadas en el parámetro C , que refleja las preferencias del organismo. Por otro lado, la ganancia informativa se asocia con la divergencia esperada, fomentando políticas que maximicen la adquisición de información útil. Este enfoque contrasta con la minimización de divergencias observadas en la energía libre variacional, destacando la importancia de las observaciones futuras aún no realizadas.
2. *Ambigüedad esperada y riesgo (resultados).* Otra descomposición de la ecuación de *energía libre esperada* considera la interacción entre la ambigüedad y el riesgo —riesgo en términos de resultados—. La ambigüedad esperada corresponde a la inexactitud anticipada, la falta de claridad, en la relación ambigua entre estados y resultados —la dispersión o entropía en la distribución de resultados anticipados—, incluso cuando los estados se conocen con precisión. El riesgo refleja la complejidad esperada al seleccionar políticas con múltiples posibles resultados —obtener varios resultados diferentes, por casualidad, aleatoriamente, bajo la misma política—. En situaciones reales, estos dos factores suelen coexistir, exigiendo políticas que minimicen simultáneamente la ambigüedad y el riesgo para optimizar la calidad de las decisiones.
3. *Ambigüedad esperada y riesgo (estados).* Una tercera formulación alternativa considera el riesgo como una divergencia entre las creencias sobre los estados y las preferencias del organismo, mientras que la ambigüedad se relaciona con la dispersión de las probabilidades esperadas de los estados. Esta perspectiva permite reorganizar la energía libre esperada en términos de energía y entropía esperadas, ofreciendo una analogía directa con la energía libre variacional. Aunque esta formulación supone que las preferencias pueden definirse directamente en términos de estados, su utilidad práctica radica en la capacidad de aprender tanto el espacio de estados como las preferencias simultáneamente.

La energía libre variacional y la energía libre esperada representan componentes esenciales de la inferencia activa, interactuando en dos niveles principales:

1. *Energía libre variacional.* Este componente mide el ajuste entre el modelo generativo interno y las observaciones presentes y pasadas. Minimizar la energía libre variacional permite a los organismos optimizar sus percepciones y creencias, ajustando su modelo interno al entorno y reduciendo la sorpresa asociada con los estados observados.
2. *Energía libre esperada.* En contraste, la energía libre esperada es fundamentalmente prospectiva y contrafactual. Evalúa las políticas posibles en función de su capacidad para minimizar la sorpresa futura, considerando las observaciones esperadas bajo cada política. Este proceso permite a los organismos anticipar consecuencias y seleccionar políticas asociadas con un mínimo de riesgo y ambigüedad.

La interacción entre estos dos niveles se puede conceptualizar como un sistema de bucles inferenciales, donde la minimización de la energía libre variacional optimiza las creencias actuales —bucle externo— y la evaluación de la energía libre esperada permite planificar acciones futuras —bucle interno—.

Síntesis de las derivaciones centrales

De acuerdo a este marco, se pueden establecer las siguientes derivaciones:

- Los sistemas biológicos buscan mantener un equilibrio dinámico caracterizado por orden, control y previsibilidad, incluso frente a fluctuaciones constantes y fuerzas entrópicas. Este esfuerzo neguentrópico se manifiesta en diversos fenómenos, desde la homeostasis fisiológica básica hasta comportamientos complejos. Mientras que la homeostasis asegura que parámetros fisiológicos críticos se mantengan dentro de rangos viables, los organismos más avanzados han desarrollado repertorios de comportamiento abiertos e innovadores para cumplir este objetivo (Sterling, 2012). Además, los sistemas biológicos no solo responden al entorno, sino que también lo moldean al crear nichos ecológicos que minimizan la incertidumbre, como refugios, ciudades, sociedades y leyes.

La minimización de la energía libre puede requerir aceptar aumentos temporales en la entropía con el fin de alcanzar una disminución más sustancial en el mediano o largo plazo. Este enfoque promueve la exploración como estrategia clave para reducir la incertidumbre global. Los imperativos epistémicos se manifiestan en dos dinámicas esenciales (Parr et al., 2022; Pezzulo et al., 2024):

- Minimizar la energía libre variacional, lo que implica maximizar la entropía de las creencias actuales (posterior aproximada).
- Minimizar la energía libre esperada, un proceso que equilibra dos factores:

La entropía predictiva posterior, que representa la incertidumbre sobre los resultados esperados al tomar decisiones y debe ser maximizada. La entropía condicional de los resultados dados los estados, que refleja la ambigüedad asociada a una política y debe ser minimizada. En términos prácticos, mientras que al minimizar la energía libre variacional el enfoque está en maximizar la entropía de las creencias presentes, al minimizar la energía libre esperada el objetivo es seleccionar acciones que reduzcan la ambigüedad de las creencias futuras. Estas dinámicas promueven conductas epistémicas como la exploración, la búsqueda de novedad y la recopilación de información. Dichas conductas contribuyen a mejorar el modelo generativo del organismo, reduciendo la sorpresa y la incertidumbre a largo plazo (Seth, 2013; Friston et al., 2015; Schwartenbeck et al., 2019).

La inferencia activa abarca tanto los sistemas biológicos más simples, que se centran en minimizar la energía libre variacional, como aquellos más complejos, que también minimizan la energía libre esperada. Estos últimos poseen la capacidad de planificar y considerar explícitamente el futuro, lo que les permite elegir entre escenarios alternativos en lugar de reaccionar únicamente al presente. Esta habilidad introduce una dimensión de profundidad temporal y psicológica, abriendo interrogantes empíricas sobre cómo los seres vivos manifiestan distintas formas de inferencia activa a lo largo de un continuo de complejidad.

En este marco, percepción y acción se conciben como procesos complementarios. La percepción minimiza la energía libre mediante la actualización de creencias bayesianas, mientras que la acción lo hace al modificar el entorno para alinearlo con dichas creencias. Esta integración de funciones cognitivas distingue la inferencia activa de otros enfoques. Además, el aprendizaje, desde esta perspectiva, es simplemente otra forma de minimizar la energía libre, operando en una escala temporal más lenta, pero esencialmente análoga a la percepción.

En el marco de la Inferencia Activa, tanto la percepción como el aprendizaje son procesos intrínsecamente dinámicos y proactivos. El cerebro funciona como una máquina predictiva que genera constantemente anticipaciones sobre los estímulos entrantes, en lugar de limitarse a recibirlos pasivamente. Esto es crucial, ya que los procesos perceptivos y de aprendizaje siempre se encuentran modulados por predicciones previas: los estímulos esperados e inesperados influyen de manera diferente en cómo percibimos y aprendemos. Además, los agentes buscan activamente

observaciones sensoriales que les permitan resolver incertidumbres. Este enfoque activo contrasta con muchas teorías contemporáneas que tienden a considerar la percepción y el aprendizaje como procesos esencialmente pasivos.

Por su parte, la acción en la Inferencia Activa es intencionada y orientada a metas. Parte de un resultado deseado, que se codifica como una predicción previa (similar al concepto de punto de ajuste en la cibernética). La planificación, entonces, implica inferir una secuencia de acciones que cumpla con esta predicción, o dicho de otro modo, que reduzca el error entre la predicción previa y el estado actual. Este enfoque de acción dirigida a metas se alinea con formulaciones tempranas de la cibernética, pero difiere de muchas teorías actuales, que explican el comportamiento en términos de asociaciones estímulo-respuesta o políticas estado-acción. En la Inferencia Activa, estas respuestas o hábitos se consideran casos particulares dentro de una familia más amplia de políticas (Friston et al., 2015; Parr et al., 2022).

IV. Prospectivas del marco de Inferencia Activa

La Inferencia Activa es una teoría en neurociencia cognitiva que explica cómo los sistemas biológicos perciben, aprenden y actúan para minimizar activamente la incertidumbre (Friston et al., 2006). Este proceso se basa en predicciones constantes y en la selección de acciones diseñadas para confirmar esas predicciones (Clark, 2013; De Lange et al., 2018). Así, los organismos interactúan con el mundo vivido alineando sus experiencias sensoriales y acciones con sus modelos, minimizando la *energía libre variacional* (Pezzulo et al., 2024).

En este marco, el cerebro es conceptualizado como una «máquina de predicción» que construye un modelo generativo del mundo que permite predecir cómo se desarrollará la situación actual —aprendizaje y percepción— y tomar decisiones efectivas mediante la selección de acciones, la planificación y el equilibrio entre exploración y explotación (Rosen, 2012; Clark, 2013; De Lange et al., 2018; Yon et al., 2019; Pezzulo et al., 2024).

Esto acerca el enfoque de Inferencia Activa a la comprensión de la percepción como una inferencia estadística bayesiana —similar en algunos sentidos a los estudios de Helmholtz, 1866, aunque con consideraciones importantes¹³— y su encarnación moderna en la *hipótesis del cerebro bayesiano* (Doya, 2014). Con algunas consideraciones, el modelo de Friston puede considerarse cercano con modelos ampliamente aceptados en neurociencia, como la *codificación predictiva* (Mumford, 1992; Rao & Ballard, 1999; de Gardelle et al., 2013; De Lange et al., 2018) y la «máquina de Helmholtz» propuesta por Peter Dayan, Geoffrey Hinton, Radford Neal y Richard Zemel (Dayan et al., 1995).

Además, la Inferencia Activa extiende el enfoque inferencial a otros dominios de la cognición, como la corporalidad enactiva y la construcción de contextos o nichos ecológicos, integrándolos:

¹³ A menudo se citan los estudios epistemológicos de Hermann Helmholtz (1821-1894) como los que sentaron las bases de considerar la percepción como una inferencia bayesiana. Helmholtz, en sus escritos maduros, intentó reconciliar las nociones idealistas alemanas de la realidad como hipótesis con la búsqueda de las leyes de la naturaleza por parte de los científicos, y abrazó la opinión de que «*alcanzamos el conocimiento del orden legal en el reino de lo real, pero sólo en la medida en que está representado en las fichas dentro del sistema de impresiones sensoriales*». Sus propuestas de inferir objetos a partir de señales sensoriales internas mediante lo que llamó «inferencias inconscientes» han hecho que Helmholtz sea considerado un proto-bayesiano. Pero la yuxtaposición de los escritos originales de Bayes, la formulación moderna de la inferencia bayesiana y las opiniones de Helmholtz sobre la percepción revela sólo una relación tenue.

1. *Enfoques enactivos*: Estas teorías enfatizan la autoorganización del comportamiento y las interacciones autopoiéticas con el entorno, que garantizan que los organismos vivos se mantengan dentro de límites aceptables (Maturana & Varela, 1980). La inferencia activa proporciona un marco formal que explica cómo los organismos logran resistir la dispersión de sus estados a través de la autoorganización de una estructura estadística, la manta de Markov. Esta estructura facilita los intercambios entre el organismo y el entorno, separando y protegiendo la integridad de los estados internos del organismo frente a la dinámica ambiental externa.
2. *Teorías cibernéticas*: Estas describen el comportamiento como intencional y teleológico, regulado internamente mediante un mecanismo que evalúa continuamente si se alcanza un objetivo, dirigiendo acciones correctivas en caso contrario (Wiener, 1961). Los agentes de inferencia activa minimizan la discrepancia entre los estados preferidos y los percibidos a través de la percepción y la acción. Este proceso minimiza una cantidad estadística medible, la energía libre variacional, que bajo ciertas condiciones corresponde a un error de predicción. Así, la inferencia activa describe el control cibernético como un proceso prospectivo.
3. *Control prospectivo basado en modelos*: Según este enfoque, los agentes utilizan un modelo generativo para construir predicciones que guían tanto la percepción como la acción, evaluando sus posibles acciones futuras (Craik, 1967). Este principio es consistente con el *teorema del buen regulador* (Conant & Ashby, 1970), que establece que cualquier controlador debe ser un buen modelo del entorno. Además, la inferencia activa es compatible con la *teoría ideomotora*, que sostiene que la acción es desencadenada por representaciones predictivas de sus consecuencias, no por estímulos externos. Esto implica que las acciones derivan de creencias sobre el futuro, lo que requiere atenuar temporalmente la evidencia sensorial que podría contradecir esas creencias (Brown et al., 2013).

Física estadística	Inferencia bayesiana y teoría de la información	Interpretación cognitiva
Minimizar la energía libre variacional	Maximizar la evidencia del modelo (o probabilidad marginal), minimizar sorpresa (o auto información)	Percepción y acción
Minimizar la energía libre esperada. Principio de mínima acción (Hamilton)	Inferir el más probable curso de acción (o el menos sorprendente)	Planificación como inferencia
Alcanzar un estado estable de no equilibrio	Realizar una inferencia bayesiana aproximada	Autoevidencia
Flujos de gradiente en funciones de energía; gradiente de descenso en energía libre	Ascenso de gradiente en la evidencia del modelo; gradiente de descenso por sorpresa	Dinámica neuronal

Fig. 03. Adaptado de Parr et al.(2022), elementos de física estadística, inferencia bayesiana y teoría de la información e interpretación cognitiva.

Inferencia activa y emoción.

Actualmente, existen tres teorías predominantes sobre la emoción en el marco de la inferencia activa, cada una enfocada en diferentes niveles de análisis y con una coherencia interna sólida.

La primera es una propuesta fenomenológica que relaciona las emociones con trayectorias de energía libre a lo largo del tiempo (Joffily & Coricelli, 2013, #). Según esta perspectiva, la valencia afectiva positiva está asociada con una disminución en la energía libre o la resolución de la sorpresa, mientras que la valencia negativa corresponde a un aumento en la energía libre. Esta aproximación ofrece una taxonomía formal de las emociones, basada en la energía libre y sus cambios temporales de orden superior.

La segunda perspectiva se centra en la inferencia activa interoceptiva (Seth, 2013; (Garfinkel et al., 2014; Seth & Friston, 2016), que corresponde a la inferencia activa sobre señales corporales internas. Las expectativas de alto nivel sobre el cuerpo integran información interoceptiva, exteroceptiva y propioceptiva, y generan predicciones descendentes que contextualizan estas modalidades entre sí. Este enfoque subraya el control predictivo al servicio de la homeostasis y la alostasis, vinculando las emociones con el mantenimiento del equilibrio fisiológico, el comportamiento afiliativo y la teoría de la mente.

Por último, la Teoría de la Emoción Construida (Barrett, 2017a, 2017b, 2022; Barrett & Russell, 2015; Jungilligens et al., 2022) incorpora la inferencia activa como un pilar central para explicar las emociones. Según esta teoría, el cerebro otorga significado a las sensaciones del cuerpo y del entorno mediante el uso de conceptos basados en experiencias pasadas. Las redes cerebrales multimodales, como la red interoceptiva, generan predicciones que contextualizan y dotan de sentido a las entradas sensoriales, creando categorías emocionales situadas. En esencia, cada experiencia emocional es la mejor hipótesis del cerebro sobre cómo el cuerpo debería responder a una situación específica, lo que permite la selección de acciones informadas por el conocimiento previo. Sin conceptos, las sensaciones carecen de significado, lo que resalta la importancia de las expectativas de alto nivel en la experiencia emocional.

Inferencia activa y estrés.

El marco de la Inferencia Activa, desarrollado por Karl Friston (Friston et al., 2006), destaca la minimización de la energía libre como un principio fundamental para los seres vivos. Según este enfoque, los organismos deben mantener de manera activa una serie de procesos orientados tanto al presente como al futuro para alcanzar metas relacionadas con su viabilidad fisiológica, psicológica o social (Peters et al., 2017).

En este contexto, la Inferencia Activa pone énfasis en cómo los organismos generan predicciones y las ajustan continuamente: los sistemas biológicos son comprendidos como «máquinas de inferencia» que construyen y actualizan modelos internos para reducir la discrepancia entre lo que predicen y lo que perciben. Este mecanismo adaptativo subraya la capacidad de los organismos para enfrentar la incertidumbre y optimizar su interacción con el entorno.

Desde esta perspectiva, un modelo óptimo es aquel que anticipa y selecciona acciones que minimizan la energía libre al conducir a resultados predecibles (Seth & Friston, 2016; Friston et al., 2016; Pezzulo et al., 2024). Subjetivamente, la discrepancia entre las anticipaciones y los resultados puede manifestarse como experiencias de incomodidad emocional, como molestia, enojo, irritabilidad, ansiedad o tristeza. Estas emociones a menudo se acompañan de cambios corporales —como alteraciones en la presión arterial, ritmo cardíaco, sudoración, mareos, molestias musculares y gástricas— y de tensiones interpersonales, tales como dificultades en la comunicación, distanciamientos o discusiones (McCarty, 2023; Tompkins, 2024). Estas experiencias suelen estar vinculadas a un estado de incertidumbre respecto a las acciones necesarias para preservar el bienestar físico, mental o social, definido como un *estado de alta entropía*, comúnmente referido como estrés (Peters et al., 2017). El estrés emerge cuando no existen políticas efectivas para alcanzar los estados preferidos o cuando el individuo no logra ajustar la precisión de sus modelos internos, entre otros factores. En estos casos, el comportamiento orientado a metas fracasa en producir los resultados deseados, delegándose el control en una serie de respuestas autónomas. En el marco de un modelo heterárquico de Inferencia Activa, la incontrolabilidad se interpreta como la incapacidad de los niveles superiores del sistema para contextualizar y coordinar los niveles inferiores. Esto sugiere que las políticas disponibles tienen pocas probabilidades de minimizar la energía libre esperada (Mathys et al., 2011; Arnaldo et al., 2022; Hartwig et al., 2022).

En términos generales, el estrés puede implicar diversas consecuencias, como se describe en Peters et al. (2017):

1. Falta de políticas disponibles para alcanzar los estados preferidos: esto implica la incapacidad de encontrar acciones o estrategias efectivas para lograr los objetivos deseados. Se trata de una falta de opciones viables para minimizar la incertidumbre y alcanzar un estado de mayor predictibilidad y control sobre el entorno. Por ejemplo, una persona se encuentra en una situación laboral precaria, con un alto riesgo de despido; si no tiene a su disposición estrategias para mejorar su situación, como buscar un nuevo trabajo o adquirir nuevas habilidades, la incertidumbre sobre su futuro laboral generará un estado de estrés.
2. Incapacidad de ajustar la precisión de las metas: el problema radica en la dificultad para modular las expectativas y objetivos en función de los cambios experimentados. Si una persona se aferra rígidamente a metas inalcanzables o poco probables, la discrepancia entre sus expectativas y la consecución de sus metas generará frustración y estrés. Por ejemplo, un estudiante que se fija la meta de obtener las mejores calificaciones en todas sus asignaturas, sin considerar sus limitaciones de tiempo o capacidad. Si el estudiante no puede ajustar esta meta a una expectativa más probable, la constante discrepancia entre su objetivo ideal y sus resultados implicará un curso de estrés. El caso opuesto es que la dificultad sea de aumentar la precisión de las metas en alguien que se ha habituado a vivir en circunstancias adversas.
3. Dificultad para contextualizar las señales de niveles inferiores: este punto se refiere a la estructura jerárquica del modelo de la Inferencia Activa, donde los niveles superiores del sistema integran la información proveniente de niveles inferiores. La incapacidad de los niveles superiores para interpretar y contextualizar la información de niveles inferiores genera una energía libre elevada, que se traduce en un estado de estrés. Un ejemplo de esto se puede observar en la respuesta al dolor. El dolor físico es una señal de nivel inferior que alerta al sistema sobre un daño o amenaza potencial. Si los niveles superiores del sistema no pueden interpretar correctamente esta señal, por ejemplo, al no identificar la causa del dolor o subestimar su gravedad, la incertidumbre resultante puede amplificarse y generar un estado de estrés crónico.

Si a pesar del esfuerzo de corrección la incertidumbre persiste, se mantiene en el tiempo, se puede generar una crisis energética. Esta crisis se traduce en «carga alostática», que contribuye a disfunciones sistémicas y cerebrales, incluyendo problemas de memoria, aterogénesis, diabetes y eventos cardiovasculares.

Además, dadas las diversas interpretaciones que pueden derivarse del cálculo de energía libre esperada, en un enfoque de Inferencia de la experiencia de estrés, se pueden configurar:

- Posibles desequilibrios entre la ganancia informativa —el valor de explorar diferentes opciones antes de seleccionar una política— y el valor pragmático o explotación. Bajo estrés podría ocurrir alguna forma de *sobreexploración* —por ejemplo, una persona se obsesiona con investigar y analizar diferentes opciones, sin llegar a una conclusión clara sobre la mejor—. Esto produce una baja ganancia informativa y podría interferir con el valor pragmático. También podría ocurrir *sobreexplotación* —por ejemplo, la persona se aferra a la primera opción valorada, sin considerar otras alternativas—. Esto puede resultar en un valor pragmático bajo e interferir con la ganancia informativa.
- Posibles problemas de optimización entre la ambigüedad esperada y el riesgo en términos de resultados. En una experiencia de estrés la ambigüedad —incertidumbre sobre la probabilidad de diferentes resultados— puede aumentarse al mismo tiempo que el riesgo —variabilidad de la calidad de los resultados—, en detrimento de la calidad de la decisión. Esto podría llevar a un resultado muy bueno si la improvisación funciona, pero también a un resultado muy malo si fracasa. Por ejemplo, un inversor estresado por la volatilidad del mercado podría tomar decisiones impulsivas, invirtiendo en activos de alto riesgo con una alta probabilidad de pérdidas.
- Posibles problemas de optimización entre la ambigüedad esperada y el riesgo en términos de estados. El riesgo sobre los estados se refiere a la divergencia entre las creencias sobre los estados futuros y las preferencias del organismo, mientras que la ambigüedad se refiere a la incertidumbre sobre la probabilidad de esos estados futuros. Asociado a la experiencia de estrés, las personas pueden inclinarse por estrategias que aumentan la ambigüedad y el riesgo sobre los estados. Por ejemplo, un empleado que se siente estresado por la posibilidad de un despido podría evitar cualquier iniciativa que pueda poner en riesgo su trabajo, limitando sus oportunidades de crecimiento profesional. Por ejemplo, una persona podría intentar enfrentar una situación de incertidumbre decidiendo consumir alguna sustancia psicoactiva, aumentando la incertidumbre sobre su funcionamiento, al tiempo que aumenta el riesgo de estados impredecibles.

Para resolver la incertidumbre, los organismos pueden recurrir a cambios de atención, al aprendizaje o a la habituación (Feldman & Friston, 2010; Peters et al., 2017):

- **Atención:** Permite obtener información sensorial más precisa para actualizar las creencias bayesianas. La respuesta inmediata al estrés es la activación, que aumenta la atención y la vigilancia. El cerebro hipervigilante busca información precisa para reducir la incertidumbre sobre la selección de estrategias, lo que conlleva un costo energético.
- **Aprendizaje:** Permite actualizar las creencias sobre el mundo para generar mejores predicciones futuras. La plasticidad sináptica, crucial para el aprendizaje, depende de los niveles de glucocorticoides. Existe una concentración óptima de glucocorticoides que favorece la potenciación a largo plazo, permitiendo consolidar el aprendizaje cuando los modelos generativos del cerebro pueden resolver la incertidumbre.
- **Habitación:** La habitación implica una actualización de los estados objetivo y una reducción de la incertidumbre sobre la mejor estrategia a seguir, permite discernir entre situaciones que deben evitarse y situaciones que pueden tolerarse. La habitación al estrés reduce la respuesta neuroendocrina y cardiovascular, disminuyendo la incertidumbre. Sin embargo, la habitación implica compensaciones metabólicas que pueden tener consecuencias a largo plazo (Peters et al., 2017). La habitación al estrés ocurre cuando las personas están expuestas repetidamente al mismo estímulo aversivo y, como resultado, muestran una atenuación de las respuestas al estrés. Esto puede gradualmente cambiar la toma de decisiones. Por ejemplo, una persona tiene que elegir entre un trabajo exigente y uno repetitivo. Inicialmente, prefiere el trabajo exigente, pero después de habituarse al estrés, cambia su preferencia al trabajo repetitivo (Hartwig et al., 2022).

Otras áreas de investigación.

Algunas áreas de investigación en las que el principio de energía libre y el marco de inferencia activa se ha aplicado, son:

- *Predicción y toma de decisiones.*
Parr et al. (2022) y Rens et al. (2023) han investigado cómo los individuos manejan la incertidumbre y el esfuerzo cognitivo durante la toma de decisiones. Friston et al. (2017) modelaron procesos de toma de decisiones y planificación basados en el concepto de energía libre esperada.
- *Psicopatología.*
Van den Bergh et al. (2017) han utilizado el marco de inferencia activa para

entender el comportamiento aberrante en psicopatología. Por su parte Barrett et al. (2016) desarrollaron una teoría de la alostasis e interocepción en la depresión basada en la inferencia activa. Y Corlett y Fletcher (2015) relacionaron errores de predicción con el desarrollo de delirios y psicosis.

- *Cognición y funciones ejecutivas.*

Friston et al. (2017) examinaron la memoria de trabajo y su integración con modelos generativos jerárquicos. Por su parte, Hasson et al. (2015) propusieron un modelo jerárquico de memoria como componente integral del procesamiento de información.

- *Percepción y control motor.*

Rao y Ballard (1999) introdujeron el modelo de codificación predictiva como implementación neuronal de la percepción-inferencia. Adams et al. (2013) propusieron que el sistema motor genera predicciones propioceptivas en lugar de comandos motores tradicionales.

- *Interocepción y regulación emocional*

Distintos investigadores, como Seth et al. (2012), Pezzulo (2014) y Barrett & Simmons (2015) han explorado cómo la inferencia activa explica el control autónomo y la regulación emocional mediante la generación de predicciones interoceptivas.

- *Marcos integradores*

Por último, se han intentado marcos integradores como el propuesto por Clark (2016), que ha ampliado la noción de "cerebro predictivo" como base para entender la cognición humana.

V. Discusión

Hemos esbozado una introducción al marco de Inferencia Activa que pueda servir de guía de estudio. Hemos revisado un marco que conecta la teoría de la información, la entropía y la neurociencia para explicar cómo los organismos minimizan la energía libre a través de procesos de percepción y acción. La figura 04 sintetiza los principales elementos que este artículo revisó. El marco enfatiza el papel de los modelos predictivos en la percepción y el aprendizaje, destacando su aplicación potencial en la investigación y comprensión de las emociones, la toma de decisiones, psicopatología y procesos de estrés.

Con el propósito de nutrir la visión del enfoque, a continuación se comentarán algunas de las principales críticas que ha recibido.

Principales críticas.

Siguiendo las posibles críticas sugeridas por Thornton (2022), éstas se pueden agrupar en la pérdida de distinciones tradicionales en el contexto cognitivo y la dificultad para evaluar empíricamente las predicciones del marco:

1. Pérdida de distinciones tradicionales:

- *El problema del agente desintegrador.* Al equiparar la incertidumbre de las observaciones sensoriales con la entropía de los estados físicos, se crea la predicción de que un agente que experimenta un aumento de la incertidumbre sufrirá un desorden físico creciente. Esta idea implica que un agente cada vez más incierto se desintegrará lentamente, lo cual es contraintuitivo. Este problema surge al eliminar la distinción entre la constitución física y el procesamiento perceptivo (Thornton, 2022). Frente a esta crítica, la noción de gradualidad y formatos de adecuación por estrés —cambios alostáticos, habituación, aprendizaje, etcétera— pueden proveer soluciones a este problema.
- *La falacia de la sorpresa literal.* El término «sorpresa» se utiliza a menudo para describir el logaritmo negativo de una probabilidad, lo que puede llevar a la suposición errónea de que la medición de la sorpresa implica la evitación de la sorpresa en un sentido literal. Si bien los teóricos de la energía libre reconocen la diferencia entre la sorpresa como medida matemática y la experiencia

subjetiva de la sorpresa, este término puede conducir a errores de interpretación (Thornton, 2022). Frente a esta crítica los autores han intentado enfatizar que los términos «sorpresa» y «creencia bayesiana» deben enmarcarse en el lenguaje matemático que se utiliza en la formalización del marco.

- *El problema del mundo perdido*: La fusión del agente y el entorno en una sola estructura de datos implica que no puede haber diferencia entre cómo es el mundo y cómo el agente espera que sea. Esto elimina la posibilidad de sorpresas genuinas y plantea interrogantes sobre la validez del concepto de evitación de sorpresas en el marco (Thornton, 2022). Esta es una crítica importante y que a menudo es recibida por otros marcos similares, como los enfoques enactivos. Es sugerente preguntarse ¿a dónde se ha ido el mundo?, cuando se considera que los agentes pueden participar activamente en la construcción de sensaciones —percepción— y acciones que generan el mundo vivido. Sin embargo, reconocer la complejidad en la construcción de percepciones no indica asumir una postura solipsista, solo indica considerar que el mundo y sus «estímulos» implican una naturaleza más compleja dado lo imbricado de la relación sujeto-objeto, que incluye tantos aspectos individuales como interpersonales, sociales y culturales.

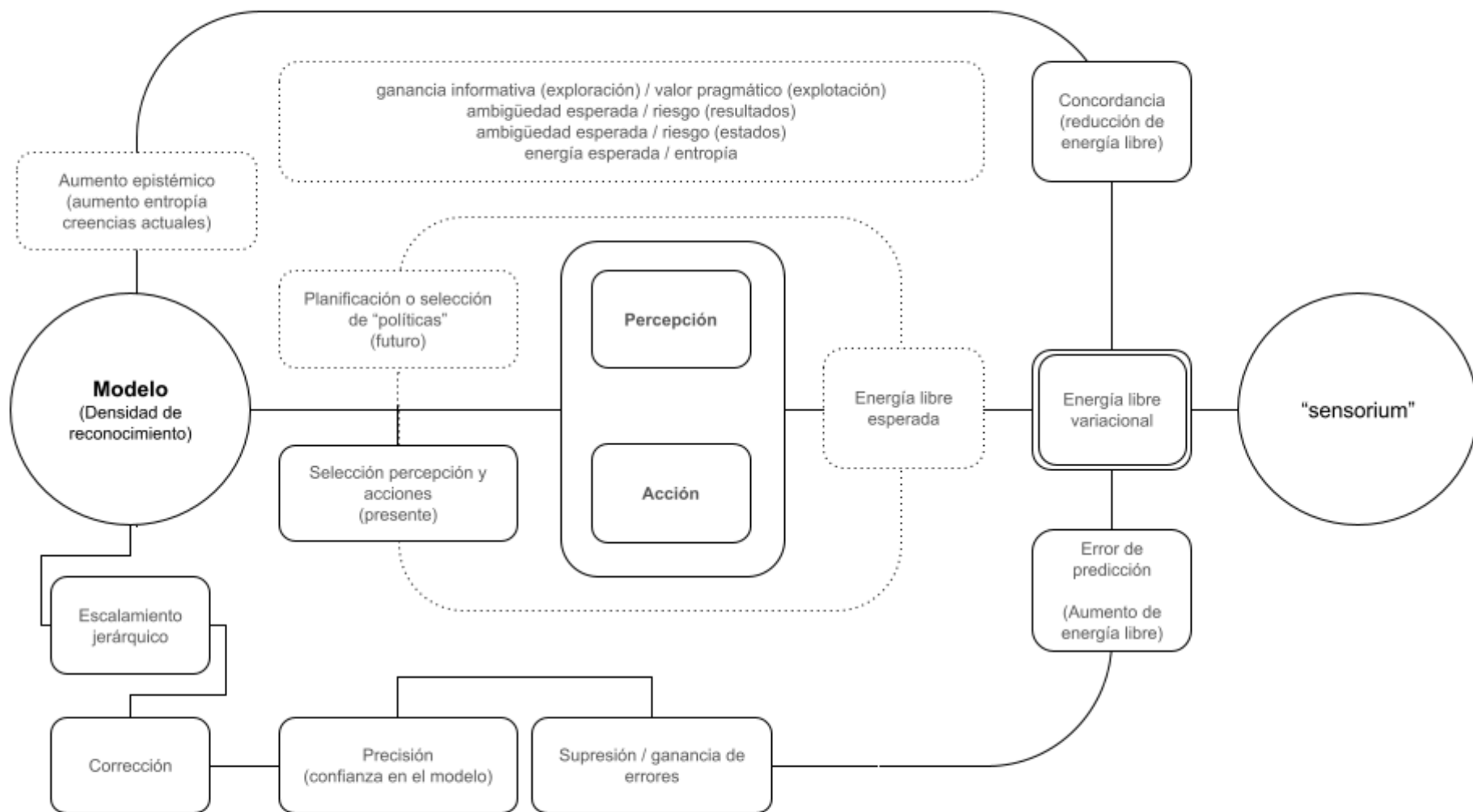


Fig. 04. Esquema de algunos elementos importantes en el marco de Inferencia Activa. Elaboración propia.

2. Dificultades para su evaluación empírica:

- *El problema de la habitación oscura:* Este problema surge al afirmar que la evitación de sorpresas garantiza la supervivencia. Los críticos argumentan que la supervivencia a menudo implica estar alerta y afrontar las sorpresas, en lugar de evitarlas. Si bien los teóricos de la energía libre proponen mecanismos adicionales para explicar la búsqueda de sorpresas en ciertas situaciones, estos argumentos pueden parecer circulares o contradictorios con la idea de que la evitación de sorpresas es el principio rector (Thornton, 2022). Frente a esta crítica los autores del marco argumentan que la minimización de la energía libre puede operar en diferentes escalas de tiempo, lo que puede dar lugar a creencias previas que anulan la necesidad de evitar sorpresas en determinadas situaciones. Es decir, debe incluirse un marco temporal para comprender el *promedio* de la minimización de la energía libre, lo que implica incorporar ciertos grados de aumento de energía libre como formato necesario para las derivas evolutivas —por ejemplo, exploración, juego, etcétera— (Friston et al., 2012).
- *El problema de la caracterización del impacto:* Si la minimización de la energía libre no se puede equiparar con la evitación de sorpresas, surge la pregunta de cómo se debe caracterizar su efecto (Thornton, 2022). Frente a esta crítica, y como no es posible para los agentes evitar directamente la sorpresa, se sugiere dentro del marco la consideración la discrepancia, falta de concordancia, o divergencia entre predicciones y resultados / estados esperados.

Aunque Thornton (2022) reconoce la universalidad del principio de minimización de la energía libre y considera que es un marco en desarrollo, sugiere —en función de afinar su capacidad explicativa y la validez de sus principales predicciones— la necesidad de una mayor validación empírica y una aclaración de sus conceptos clave. De acuerdo a este autor, estos dos aspectos son cruciales para el avance del marco.

Comentarios finales.

El marco de Inferencia Activa, aún en desarrollo, se presenta como un enfoque empíricamente verificable y, de manera incipiente, sustentado en evidencia. No obstante el marco de Inferencia Activa destaca por su capacidad explicativa, al ofrecer no solo un modelo formal y comprobable del procesamiento neurocognitivo, sino también al integrar estos procesos con la acción y la percepción, su utilización por parte de los psicólogos en general ha sido limitada (Badcock & Davey, 2024).

El marco de Inferencia Activa es promisorio. Este enfoque plantea que todo pensamiento y comportamiento están orientados hacia la optimización de predicciones, dirigiendo las acciones hacia su confirmación. También permite una integración única entre los avances en neurociencia cognitiva y las perspectivas evolutivas y de desarrollo. Postula que los sistemas biológicos complejos minimizan la energía libre variacional a través de ciclos dinámicos de acción y percepción, mediados por mecanismos neurocognitivos jerárquicos. Esto se combina con una noción evolutiva dinámica, destacando el papel causal de antecedentes adaptativos específicos de cada especie. Estos van desde predisposiciones innatas moldeadas por la selección natural hasta expectativas abiertas al desarrollo, ensambladas mediante interacciones recurrentes a lo largo del ciclo de vida (Badcock & Davey, 2024).

La Inferencia Activa se posiciona como un modelo unificador con el potencial de sintetizar teoría e investigación en múltiples subdisciplinas de la psicología. No obstante, con la excepción de teorías como la Teoría de la Emoción Construida de Lisa Feldman Barrett (Barrett & Russell, 2015) y ciertos temas clínicos específicos (Putica & Agathos, 2024), el progreso en su integración ha sido lento. Las aplicaciones empíricas han sido limitadas, y la aceptación de este marco por parte de la comunidad psicológica ha enfrentado escepticismo, reflejando una resistencia hacia teorías unificadoras a pesar de su respaldo empírico. A medida que más investigadores adopten este enfoque y lo traduzcan en diseños empíricos innovadores, es plausible que ocurra un cambio de paradigma, aunque este aún parece distante.

VI. Referencias

1. Adams, R., Perrinet, L., & Friston, K. (2012). Smooth Pursuit and Visual Occlusion: Active Inference and Oculomotor Control in Schizophrenia. *PLOS ONE*, 7(10), e47502. doi:10.1371/journal.pone.0047502
2. Adams, R., Shipp, S., & Friston, K. (2013). Predictions not commands: Active inference in the motor system. *Brain Structure & Function*, 218(3), 611-643. <https://doi.org/10.1007/s00429-012-0475-5>
3. Arnaldo, I., Corcoran, A., Friston, K., & Ramstead, M. (2022). Stress and its sequelae: An active inference account of the etiological pathway from allostatic overload to depression. *Neuroscience and biobehavioral reviews*, (135), 104590. doi:10.1016/j.neubiorev.2022.104590
4. Ashby, W. R. (1957). *An Introduction to Cybernetics*. Chapman & Hall.
5. Ashby, W. R. (1962). Principles of the self-organizing system. In *Principles of Self-Organization* (U.S. Office of Naval Research ed., pp. 255-278). Heinz von Foerster & George W. Zopf, Jr.
6. Badcock, P., & Davey, C. (2024). Active Inference in Psychology and Psychiatry: Progress to Date? *Entropy (Basel)*, 26(10), 833. doi:10.3390/e26100833
7. Barrett, L. F. (2017). Categories and Their Role in the Science of Emotion. *Psychol Inq*, 28(1), 20-26. doi:10.1080/1047840X.2017.1261581
8. Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Soc Cogn Affect Neurosci*, 12(1), 1-23. doi:10.1093/scan/nsw15
9. Barrett, L. F. (2022). Context reconsidered: Complex signal ensembles, relational meaning, and population thinking in psychological science. *Am Psychol*, 77(8), 894-920. doi:10.1037/amp0001054
10. Barrett, L. F., Quigley, K., & Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708), 20160011. <https://doi.org/10.1098/rstb.2016.0011>
11. Barrett, L. F., & Russell, J. A. (Eds.). (2015). *The Psychological Construction of Emotion*. Guilford Publications.

12. Barrett, L. F., & Simmons, W. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, 16(7), 419-429. <https://doi.org/10.1038/nrn3950>
13. Bohannon, J. (2016, noviembre 11). *A computer program just ranked the most influential brain scientists of the modern era*. doi:10.1126/science.aal0371. Retrieved noviembre 28, 2024, from <https://www.science.org/content/article/computer-program-just-ranked-most-influential-brain-scientists-modern-era>
14. Bouizegarene, N., Ramstead, M., Constant, A., Friston, K., & Kirmayer, L. (2024). Narrative as active inference: an integrative account of cognitive and social functions in adaptation. *Front Psychol*, 6(15), :1345480. doi:10.3389/fpsyg.2024.1345480
15. Boyd, A.B., Mandal, D., & Crutchfield, J.P. (2017). Leveraging Environmental Correlations: The Thermodynamics of Requisite Variety. *J Stat Phys*, (167), 1555-1585.
16. Brillouin, L. (1953). The Negentropy Principle of Information. *Journal of Applied Physics*, 24(9), 1152-1163. doi:10.1063/1.1721463
17. Brown, H., Adams, R., Parees, I., Edwards, M., & K. (2013). Active inference, sensory attenuation and illusions. *Cognitive processing*, 14(4), 411-427. 10.1007/s10339-013-0571-3
18. Brown, H., & Friston, K. (2012). Free-Energy and Illusions: The Cornsweet Effect. *Frontiers in Psychology*, 3(43). doi:10.3389/fpsyg.2012.00043
19. Butz, M., Mittenbühler, M., Schwöbel, S., Achimova, A., Gumbsch, C., Otte, S., & Kiebel, S. (2024). Contextualizing predictive minds. *Neurosci Biobehav Rev*, 168, 105948. doi:10.1016/j.neubiorev.2024.105948
20. Cheadle, J., Davidson-Turner, K., & Goosby, B. (2024). Active Inference and Social Actors: Towards a Neuro-Bio-Social Theory of Brains and Bodies in Their Worlds. *Kolner Z Soz Sozpsychol*, 73(3), 317-350. doi:10.1007/s11577-024-00936-4
21. Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181-204. doi:10.1017/S0140525X12000477
22. Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.

23. Clausius, R. (1850). Über die bewegende Kraft der Wärme. *Annalen der Physik*, 79(368-397), 500-524.
24. Conant, R., & Ashby, W. R. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, 1(2), 89-97. doi:10.1080/00207727008920220
25. Corlett, P., & Fletcher, P. (2015). Delusions and prediction error: Clarifying the roles of behavioural and brain responses. *Cognitive Neuropsychiatry*, 20(2), 95-105. <https://doi.org/10.1080/13546805.2014.990625>
26. Craik, K. J. W. (1967). *The Nature of Explanation*. Cambridge University Press.
27. Dayan, P., Hinton, G., Neal, R., & Zemel, R. (1995). The Helmholtz machine. *Neural computation*, 7(5), 889-904. doi:10.1162/neco.1995.7.5.889
28. De Gardelle, V., Waszczuk, M., Egner, T., & Summerfield, C. (2013). Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. *Cereb Cortex*, 23(9), 2235-2244. doi: 10.1093/cercor/bhs211
29. De Lange, F., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, 22(9), 764-779. doi:10.1016/j.tics.2018.06.002
30. Doya, K. (2014). *Bayesian Brain: Probabilistic Approaches to Neural Coding* (K. Doya, Ed.). MIT Press.
31. Edwards, M., Adams, R., Brown, H., Parees, I., & Friston, K. (2012). A bayesian account of 'hysteria'. *Brain*, 135(11), 3495–3512. 10.1093/brain/aws129
32. Feldman, H., & Friston, K. (2010). Attention, uncertainty, and free-energy. *Frontiers in human neuroscience*, (4), 215. doi:10.3389/fnhum.2010.00215
33. Feynman, R. P., Leighton, R. B., & Sands, M. L. (1970). *Feynman Lectures On Physics (3 Volume Set)*. Addison-Wesley.
34. Fox, A. S. (Ed.). (2018). *The Nature of Emotion: Fundamental Questions*. Oxford University Press.
35. Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 360, 815-836. 10.1098/rstb.2005.1622
36. Friston, K. (2019). A free energy principle for a particular physics. *arXiv preprint*. doi:10.48550/arXiv.1906.10184

37. Friston, K., Adams, R., Perrinet, L., & Breakspear, M. (2012). Perceptions as Hypotheses: Saccades as Experiments. *Frontiers in Psychology*, 3(151). doi:10.3389/fpsyg.2012.00151
38. Friston, K., Da Costa, L., Tschantz, A., Kiefer, A., Salvatori, T., Neacsu, V., Koudahl, M., C, H., Sajid, N., Markovic, D., Parr, T., Verbelen, T., & Buckley, C. (2024). Supervised structure learning. *Biol Psychol.*, (193), :108891. doi: 10.1016/j.biopsycho.2024.108891
39. Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2016). Active inference and learning. *Neuroscience & Biobehavioral Reviews*, (68), 862-879. doi: 10.1016/j.neubiorev.2016.06.022
40. Friston, K., FitzGerald, T., Rigoli, T., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: A process theory. *Neural Computation*, 29(1), 1-49. doi: https://doi.org/10.1162/NECO_a_00912
41. Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *A free energy principle for the brain*, 100(1-3), 70-87. doi:10.1016/j.jphysparis.2006.10.001
42. Friston, K., Levin, M., Sengupta, B., & Pezzulo, G. (2015). Knowing one's place: a free-energy approach to pattern regulation. *Journal of the Royal Society Interface*, 12(105), 20141383. doi:10.1098/rsif.2014.1383
43. Friston, K., Lin, M., Frith, C., Pezzulo, G., Hobson, J., & Ondobaka, S. (2017). Active inference, curiosity, and insight. *Neural Computation*, 29(10), 2633-2683. doi: https://doi.org/10.1162/neco_a_00999
44. Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, 140(1-2), 137-160. <https://doi.org/10.1007/s00422-011-0424-z>
45. Friston, K., Parr, T., & De Vries, B. (2017). The graphical brain: Belief propagation and active inference. *Network Neuroscience*, 1(2), 381-414. doi: https://doi.org/10.1162/NETN_a_00018
46. Friston, K., Thornton, C., & Clark, A. (2012). Free-energy minimization and the dark-room problem. *Front. Psychology*, 3(130). doi: 10.3389/fpsyg.2012.00130
47. García Andrade, A. (2021). Percepción emocional: sociología neurociencia afectiva. *Revista mexicana de sociología*, 82(4), 835-863. doi:10.22201/iis.01882503p.2020.4.59209

48. Garfinkel, S., Minati, L., Gray, M., Seth, A., Dolan, R., & Critchley, H. (2014). Fear from the heart: sensitivity to fear stimuli depends on individual heartbeats. *J Neurosci*, 7(34), 6573-6582. doi: 10.1523/JNEUROSCI.3507-13.2014
49. Haken, H. (1983). *Synergetics: An Introduction*. Springer Berlin Heidelberg.
50. Hakimi, N., Chou, K., Stewart, J., Paulus, M., & Smith, R. (2024). Computational Mechanisms of Learning and Forgetting Differentiate. *Affective and Substance Use Disorders*, (31), :rs.3.rs-4682224. doi:10.21203/rs.3.rs-4682224/v1
51. Hartwig, M., Bhat, A., & Peters, A. (2022). How Stress Can Change Our Deepest Preferences: Stress Habituation Explained Using the Free Energy Principle. *Frontiers in psychology*, 13, 865203. doi:10.3389/fpsyg.2022.865203
52. Hasson, U., & Honey, C. (2015). Hierarchical process memory: Memory as an integral component of information processing. *Trends in Cognitive Sciences*, 19(6), 304-313. doi: <https://doi.org/10.1016/j.tics.2015.04.006>
53. Heider, F. (1958). *The Psychology of Interpersonal Relations*. John Wiley & Sons, Inc.
54. Hinton, G. (2002). Training products of experts by minimizing contrastive divergence. *Neural Comput*, 14(8), 1771-1800. doi:10.1162/089976602760128018
55. Hinton, G., & Zemel, R. (1993). Autoencoders, minimum description length and Helmholtz free energy. In *Proceedings of the 6th International Conference on Neural Information Processing Systems (NIPS'93)* (pp. 3-10). Morgan Kaufmann Publishers Inc.
56. Hobson, J., & Friston, K. (2012). Waking and dreaming consciousness: Neurobiological and functional considerations. *Progress in Neurobiology*, 98(1), 82-98. doi:10.1016/j.pneurobio.2012.05.003
57. Hodson, R., Mehta, M., & Smith, R. (2023). The empirical status of predictive coding and active inference. *Neurosci Biobehav Rev.*, 157, :105473. doi: 10.1016/j.neubiorev.2023.105473
58. Janson-Schmitt, J., & Rohleder, N. (2024). Interventions to modify the habituation of biological responses to repeated stress in healthy adults: a randomized controlled trial. *Trials*, 25(1), 783. doi:10.1186/s13063-024-08620-w
59. Jaynes, E. (1957a). Information Theory and Statistical Mechanics. *Physical Review*, 106(4), 620-630. doi:10.1103/PhysRev.106.620

60. Jaynes, E. (1957b). Information Theory and Statistical Mechanics II. *Physical Review*, 108(2), 171-190. 10.1103/PhysRev.108.171
61. Joffily, M., & Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLoS Comput Biol.*, 9(6), :e1003094. doi:10.1371/journal.pcbi.1003094
62. Jungilligens, J., Paredes-Echeverri, S., Popkirov, S., Barrett, L. F., & Perez, D. (2022). A new science of emotion: implications for functional neurological disorder. *Brain*, 145(8), 2648-2663. doi: 10.1093/brain/awac204
63. Karnick, A., & Brick, L. (2024). From perception to projection: Exploring neuroaffective advances in understanding optimism bias and belief updating. *Curr Opin Psychol*, (60), :101937. doi:10.1016/j.copsyc.2024.101937
64. Kauffman, S.A. (1992). The Origins of Order: Self-Organization and Selection in Evolution. *Spin Glasses and Biology*, 61-100.
65. Kilner, J., Friston, K., & Frith, C. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, 8(3), 159-166. doi:10.1007/s10339-007-0170-2
66. Lamontagne, S., Zabala, P., & Ballard, E. (2023). Toward objective characterizations of suicide risk: A narrative review of laboratory-based cognitive and behavioral tasks. *Neurosci Biobehav Rev*, (153), 105361. doi:10.1016/j.neubiorev.2023.105361
67. Lanouette, W., & Silard, B. (1992). *Genius in the Shadows: A Biography of Leo Szilard: The Man Behind The Bomb*. Skyhorse Publishing.
68. Lazarus, R. S., & Folkman, S. (1984). *Stress, Appraisal, and Coping*. Springer Publishing Company.
69. Letten, A., Ke, P., & Fukami, T. (2017). Linking modern coexistence theory and contemporary niche theory. *Ecological Monographs*, 87(2), 161-177. doi:10.1002/ecm.1242
70. Lutz, Abdoun, O., Abdoun, Y., Trautwein, F., & Berkovich-Ohana, A. (2024). An Overview of Neurophenomenological Approaches to Meditation and their Relevance to Clinical Research. *Biol Psychiatry Cogn Neurosci Neuroimaging.*, 21(24), 00345-8. doi:10.1016/j.bpsc.2024.11.008
71. Mathys, C., Daunizeau, J., Friston, K., & Stephan, K. (2011). A bayesian foundation for individual learning under uncertainty. *Frontiers in human neuroscience*, 5(39). doi:10.3389/fnhum.2011.00039

72. Maturana, H. & Varela, F. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Springer Netherlands.
73. McCarty, R. (2023). *Stress, Health, and Behavior*. Guilford Publications.
74. Mirza, B., Adams, R., Mathys, C., & Friston, K. (2018). Human visual exploration reduces uncertainty about the sensed world. *PLOS ONE*, 13(1), e0190429. doi:10.1371/journal.pone.0190429
75. Mumford, D. (1992). On the computational architecture of the neocortex. *Biological Cybernetics*, 66(3), 241-251. doi:10.1007/bf00198477
76. Nicolis, G., & Prigogine, I. (1977). *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations*. Wiley.
77. Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. MIT Press.
78. Perrinet, L., Adams, R., & Friston, K. J. (2014). Active inference, eye movements and oculomotor delays. *Biological Cybernetics*, 108(6), 777-801. doi:10.1007/s00422-014-0620-8
79. Peters, A., McEwen, B., & Friston, K. (2017). Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Progress in neurobiology*, (156), 164-188. doi:10.1016/j.pneurobio.2017.05.004
80. Pezzulo, G. (2014). Why do you fear the bogeyman? An embodied predictive coding model of perceptual inference. *Cognitive, affective & behavioral neuroscience*, 14(3), 902-911. doi: <https://doi.org/10.3758/s13415-013-0227-x>
81. Pezzulo, G., Parr, T., & Friston, K. (2024). Active inference as a theory of sentient behavior. *Biological Psychology*, 186(108741). doi:10.1016/j.biopsycho.2023.108741
82. Putica, A., & Agathos, J. (2024). Reconceptualizing complex posttraumatic stress disorder: A predictive processing framework for mechanisms and intervention. *Neurosci Biobehav*, (164), :105836. doi: 10.1016/j.neubiorev.2024.105836.
83. Rao, R., & Ballard, D. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79-87. doi:10.1038/4580
84. Rens, N., Lancia, G., Eluchans, M., Schwartenbeck, P., Cunnington, R., & Pezzulo, G. (2023). Evidence for entropy maximisation in human free choice behaviour. *Cognition*, 232. doi: 05328, 10.1016/j.cognition.2022.105328

85. Rosen, R. (2012). *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations*. Springer.
86. Rudrauf, D., Bennequin, D., Granic, I., Landini, G., Friston, K., & Williford, K. (2017). A mathematical model of embodied consciousness. *Journal of Theoretical Biology*, 428, 106-131. doi:10.1016/j.jtbi.2017.05.032
87. Schmitz, M., Back, s., Seitz, K., Harbrecht, N., Streckert, L., Schulz, A., Herpertz, S., & Bertsch, K. (2023). The impact of traumatic childhood experiences on interoception: disregarding one's own body. *Borderline Personal Disord Emot Dysregul*, 10(1), :5. doi:10.1186/s40479-023-00212-5.
88. Schrödinger, E. (1944). *What is Life? With Mind and Matter and Autobiographical Sketches*. Cambridge University Press.
89. Schwartenbeck, P., Passecker, J., Hauser, T., FitzGerald, T., Kronbichler, M., & Friston, K. (2019). Computational mechanisms of curiosity and goal-directed exploration. *eLife*, 8(e41703). doi:10.7554/eLife.41703
90. Seth, A. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in cognitive sciences*, 17(11), 565-573. doi:10.1016/j.tics.2013.09.007
91. Seth, A. (2023). *La creación del yo: Una nueva ciencia de la conciencia*. Editorial Sexto Piso.
92. Seth, A., & Friston, K. (2016). Active interoceptive inference and the emotional brain. *Philos Trans R Soc Lond B Biol Sci*, 371(1708), :20160007. doi: 10.1098/rstb.2016.0007
93. Seth, A., Suzuki, K., & Critchley, H. (2012). An interoceptive predictive coding model of conscious presence. *Frontiers in psychology*, (2), 395. doi: <https://doi.org/10.3389/fpsyg.2011.00395>
94. Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379-423. doi:10.1002/j.1538-7305.1948.tb01338.x
95. Sinha, R. (2024). Stress and substance use disorders: risk, relapse, and treatment outcomes. *Journal of Clinical Investigation*, 134(16), e172883. <https://doi.org/10.1172/JCI172883>
96. Sterling, P. (2012). Allostasis: a model of predictive regulation. *Physiology & behavior*, 106(1), 5-15. doi.org/10.1016/j.physbeh.2011.06.004

97. Szilárd, L. (1929). Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Zeitschrift Für Physik*, 53(11-12), 840-856. doi:10.1007/bf01341281
98. Theuer, J., Koch, N., Gumbsch, C., Elsner, B., & Butz, M. (2024). Infants infer and predict coherent event interactions: Modeling cognitive development. *PLoS One*, 19(10), :e0312532. doi:10.1371/journal.pone.0312532
99. Tompkins, M. A. (2024). *The Cognitive Behavioral Therapy Workbook: Evidence-Based CBT Skills to Help You Manage Stress, Anxiety, Depression, and More*. New Harbinger Publications.
100. Van den Bergh, O., Witthöft, M., Petersen, S., & Brown, R. (2017). Symptoms and the body: Taking the inferential leap. *Neuroscience & Biobehavioral Reviews*, (74), 185-203. doi: <https://doi.org/10.1016/j.neubiorev.2017.01.015>
101. Wiener, N. (1961). *Cybernetics: or Control and Communication in the Animal and the Machine* (2a ed.). MIT Press.
102. Williford, K., Bennequin, D., Friston, K., & Rudrauf, D. (2018). The Projective Consciousness Model and Phenomenal Selfhood. *Frontiers in Psychology*, 9(2571). doi:10.3389/fpsyg.2018.02571
103. Yon, D., Lange, F., & Press, C. (2019). The Predictive Brain as a Stubborn Scientist. *Trends in Cognitive Sciences*, 23(1), 6-8. doi:10.1016/j.tics.2018.10.003
104. Zénon, A., Solopchuk, O., & Pezzulo, G. (2019). An information-theoretic perspective on the costs of cognition. *Neuropsychologia*, (123), 5-18. doi: 10.1016/j.neuropsychologia.2018.09.013

Nota del autor

Rodrigo Díaz Olguín. Psicólogo, entrenado en psicoterapia postracionalista GIP Barcelona. Diplomado en Tratamiento y Rehabilitación de adicciones para Población General, U de Chile. Psicólogo clínico Unidad de Psiquiatría Hospital Las Higueras, encargado del subprograma Trastornos de la Personalidad. Docente universitario. Autor del libro "Experiencia suicida: comprensión narrativa constructivista", Ril Editores, 2022. Comentarios o consultas sobre este trabajo, escribir a ps.rodrigodiazolguin@gmail.com



Editado por Cipra, Círculo de Psicoterapia Constructivista. Concepción, Chile.
Correo electrónico cipra.psicologia@gmail.com Sitio web <http://www.cipra.cl> Diciembre 2024