

**Donde hay vida, hay mente:  
en apoyo a una tesis fuerte de la continuidad vida-mente<sup>1</sup>**

Michael D. Kirchhoff<sup>1</sup>

Tom Froese<sup>2</sup>

<sup>1</sup> Departamento de Filosofía, Facultad de Derecho, Humanidades y Artes, Universidad de Wollongong, Wollongong 2500, Australia, [kirchhof@uow.edu.au](mailto:kirchhof@uow.edu.au)

<sup>2</sup> Departamento de Ciencias de la Computación, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Ciudad de México 04510, México, [t.froese@gmail.com](mailto:t.froese@gmail.com)

## **Resumen**

El presente texto considera cuestiones en torno a la continuidad y la discontinuidad entre la vida y la mente. Inicia examinando dichas cuestiones desde la perspectiva del principio de energía libre (PEL). El PEL se ha vuelto considerablemente influyente tanto en la neurociencia como en la ciencia cognitiva. Postula que los organismos actúan para conservarse a sí mismos en sus estados biológicos y cognitivos esperados, y que lo logran al minimizar su energía libre, dado que el promedio de energía libre a largo plazo es entropía. El texto, por lo tanto, argumenta que no existe una sola interpretación del PEL para pensar la relación entre la vida y la mente. Algunas formulaciones del PEL dan cuenta de lo que llamamos una perspectiva de independencia entre la vida y la mente. Una perspectiva de independencia es la perspectiva cognitivista del PEL, misma que depende del procesamiento de información con contenido semántico, y por ende, restringe el rango de sistemas capaces de exhibir mentalidad. Otras perspectivas de independencia ejemplifican lo que llamamos la demasiado generosa perspectiva no-cognitivista del PEL, que parecen ir en dirección opuesta: sugieren que la mentalidad se encuentra casi en cualquier lugar. El texto continúa argumentando que el PEL no-cognitivista y sus implicaciones para pensar la relación entre la vida y la mente puede ser útilmente delimitado por las recientes aproximaciones enactivas a la ciencia

---

<sup>1</sup> El presente capítulo es una traducción a cargo de Laura Rodríguez Benavidez (Facultad de Filosofía y Letras, UNAM) del texto original: Kirchhoff, M. D. and Froese, T. (2017). Where there is life there is mind: In support of a strong life-mind continuity thesis. *Entropy*, 19(4): 169. doi:10.3390/e19040169.

cognitiva. Concluimos que la versión más contundente de la relación vida-mente las considera fuertemente continuas, y esta continuidad se basa en conceptos particulares de vida (autopoiesis y adaptabilidad) y mente (básica y no-semántica).

**Palabras clave:** continuidad vida-mente; principio de energía libre; enactivismo radical; enactivismo autopoietico.

---

## 1. Introducción

¿Cómo es que la vida y la mente se caracterizan respectivamente, y cómo se conciben las relaciones entre ellas?

En este texto, iniciamos examinando esta pregunta desde la perspectiva del principio de energía libre (PEL). Se suele argumentar que el PEL estipula una base radical para el funcionamiento del cerebro; ofrece una teoría unificada de la percepción, la cognición y la acción (y demás capacidades psicológicas), y sugiere un marco a partir del cual entender la relación entre la vida y la mente [1–3]. Postula que los organismos actúan para conservarse a sí mismos en sus estados biológicos y cognitivos esperados, y que lo logran al minimizar su energía libre, dado que el promedio de energía libre a largo plazo es entropía [4, 5] (Por “estado” entendemos un estado en el espacio de estados de un sistema. Uno de los estados en los que un sistema espera encontrarse a sí mismo es el de “estar vivo”; por lo tanto, un sistema buscará reducir la probabilidad de encontrarse en un estado no-anticipado respecto a su modelo generativo. En otras palabras, al minimizar la energía libre, en promedio y a lo largo del tiempo, el sistema auto-organizará los parámetros de sus estados internos para ocupar un número limitado de estados en promedio y a lo largo del tiempo ([6], p. 180)). Luego entonces, minimizar la energía libre equivale a reducir el desorden, en el sentido de incertidumbre.

Después argumentaremos que no existe una única versión del PEL para pensar la relación entre la vida y la mente (O, mínimamente, no existe un acuerdo de cómo interpretar mejor las propiedades de la energía libre variacional para pensar la vida, la mente y la relación entre ellas). Estas diferentes perspectivas en torno a la relación vida-mente pueden apreciarse al considerar la respuesta que uno daría a la siguiente pregunta: “¿Los fenómenos mentales se restringen a los sistemas vivos?”

Algunas formulaciones de la energía libre responden negativamente a la pregunta. Llamamos a esta imagen de la relación vida-mente la *perspectiva de independencia entre la vida y la mente*. En el contexto del PEL, se presenta en, al menos, dos formulaciones, cada una con sus propias implicaciones para pensar la vida, la mente y la relación entre ambas:

- El principio de energía libre cognitivista [6–8].
- El principio de energía libre no-cognitivista y demasiado generoso [9, 10].

El principio de energía libre cognitivista (el PEL cognitivista) trata de contingente a la relación entre vida y mente. En ocasiones, se refiere a él como la hipótesis del cerebro auto-evidenciadora [7] o, simplemente, la mente predictiva [6]. Una perspectiva de independencia de este tipo puede aún sostener que algunos sistemas cognitivos son sistemas vivos, pero sostendrá a la vez que esta relación es puramente contingente, por ejemplo al asociar las mentes con procesos computacionales, con propiedades semánticas (i. e. llenas de contenido), o al considerar la posibilidad de que las mentes pueden realizarse por completo, independientemente de la vida con el tipo apropiado de sistema artificial de apoyo. Dichas mentes pueden decirse estando epistémicamente excluidas del mundo, constituidas por poderosos modelos generativos, que no pueden estar “necesariamente casados con órganos biológicos” ([8], p. 7).

El principio de energía libre no cognitivista (PEL no-cognitivista) toma un punto de partida muy distinto al del PEL cognitivista. Los orígenes del PEL se encuentran en la termodinámica, donde se han utilizado teoremas de no-equilibrio en la energía libre para explicar dinámicas auto-organizadoras en sistemas capaces de permanecer lejos del equilibrio termodinámico [4]. En este sentido, los orígenes del PEL no tienen, intrínsecamente, nada que ver con la vida y la mente, a pesar de que hoy en día éste se aplique de manera directa al explicar sistemas vivos y cognitivos [2, 9–11]. El PEL demasiado generoso, no-cognitivista es la postura que establece que todos los sistemas que conservan sus variables dentro de un rango limitado de valores pueden entenderse como poseedores de una forma de mentalidad o proto-mentalidad, dado que el PEL abarca cualquier sistema capaz de conservar su integridad estructural de cara a un ambiente fluctuante como dedicado a predecir sus propios estados futuros; lo cual significa que el retener integridad descansa en los procesos cuya función es maximizar la evidencia del modelo, es decir, que estos procesos exhiben dinámicas auto-evidenciadoras. Sin embargo, esta generosa perspectiva de auto-evidencia podría parecer dirigir a alguna forma de pansiquismo.

Otras formulaciones de la energía libre responden afirmativamente a la pregunta de si la mente se restringe a la vida. Dichas formulaciones son parte de una imagen más general de la relación vida-mente a la que nos referimos como la *perspectiva de dependencia entre la vida y la*

*mente*. Existen diferentes versiones de ésta en la literatura y lo común en todas ellas es que todas se adscriben a una postura mucho menos generosa pero aun así no-cognitivista del PEL. Etiquetamos, a continuación, las posibles versiones:

- PEL no-cognitivista + posturas tardías evolutivas de la mente (por ejemplo, [12])
- PEL no-cognitivista + posturas fuertes de la continuidad vida-mente ([13, 14] y el presente texto).

Las posturas tardías evolutivas de la mente enfatizan discontinuidades entre seres meramente vivos y aquéllos que también son cognitivos, de manera tal que las propiedades de la mente pueden pensarse como complejizaciones de las propiedades de la vida (ver [15] para mayor discusión de estas posturas, aunque en un contexto ligeramente distinto). Por ejemplo, la mentalidad, pero no la vida, requiere de la existencia de una maquinaria neural generativa sofisticada que no aparece en formas simples de vida como los organismos unicelulares [12]. Por lo tanto, bajo esta perspectiva, es posible estar vivo sin (necesariamente) ser cognitivo (A pesar de su defensa por una teoría representacional de la mente, orientada a la acción, situamos a Clark [12] en el bando del PEL no-cognitivista por su propugnación a la complementariedad del PEL y sus esquemas de procesamiento predictivo con trabajo en los paradigmas encorporeizados (*embodied*), extendidos (*extended*) y enactivos a la ciencia cognitiva. Hablaremos más sobre la postura de Clark a su debido tiempo).

En este texto, defenderemos al PEL no-cognitivista y a una postura fuerte de la continuidad vida-mente, basados en los desarrollos recientes de la ciencia cognitiva encorporeizada (*embodied*) y enactiva. Son las premisas clásicas del cognitivismo, especialmente el internalismo y el representacionalismo, las que aplican directamente los paradigmas enactivos a la problemática del PEL [1, 16]. No estamos solos en el desarrollo de formulaciones anti-cognitivistas del PEL desde la perspectiva del enactivismo. Los trabajos de Bruineberg et al. [13] y Kirchhoff [14, 17] prepararon el escenario. A continuación, una lista de los puntos en los que estamos de acuerdo con Bruineberg et al. [13]: (a) la concepción helmholtziana de la percepción como una inferencia inconsciente es inherente a las formulaciones cognitivistas del PEL; (b) hay buenas razones para creer que la concepción helmholtziana de la inferencia perceptual es incompatible con la inferencia bayesiana aproximativa bajo una formulación no cognitivista del PEL, y (c) una vez observado desde los lentes del enactivismo, el PEL puede abordar el tema de cómo la vida y la mente comparten el mismo conjunto de propiedades organizacionales básicas. La principal diferencia entre Bruineberg et al. [13] y este trabajo es que mientras Bruineberg et al. ([13]; y ver también [18]) buscan establecer que la función de los modelos generativos es mantener un sistema cerebro-cuerpo-nicho

robusto (ver [17] para una aproximación metafísica), nosotros directamente apuntamos hacia una tesis fuerte de la continuidad vida-mente, al seguir desarrollando los argumentos recientes de [14].

Argumentaremos que el PEL no-cognitivistista y sus implicaciones al pensar la relación entre vida y mente, pueden ser útilmente constreñidas y aumentadas por ideas clave en paradigmas enactivos recientemente radicales y autopoieticos de la ciencia cognitiva [14, 17, 19–22]. Nuestro argumento tiene dos pasos. El primero aborda la naturaleza de las *mentes básicas* como seleccionadas por evolución hacia una direccionalidad intencional sin contenido semántico [21]. El segundo paso basa el concepto de mentes básicas en el de vida básica abordado en términos de autopoiesis y adaptividad [19, 20]. Consecuentemente, llegamos a la visión fuerte de la continuidad vida-mente, mientras que al mismo tiempo, nos mantenemos lo suficientemente alejados del tipo de generalización de lo mental asociada con las interpretaciones demasiado generosas del PEL en torno al lugar de la mente en el mundo natural (Para otras articulaciones relacionadas pero distintas de la tesis fuerte de la continuidad vida-mente, ver [23–26], especialmente [27]).

## **2. El principio de energía libre (PEL)**

Lo que se llama principio de energía libre (PEL) es un imperativo para la auto-organización en sistemas dinámicos abiertos, que especifica que para que los sistemas vivos mantengan su integridad estructural y funcional, deben minimizar la “energía libre” en el contexto de inferencia activa: deben cambiar la relación que tienen con su nicho a fin de preservar su integridad [4, 10, 14]. El PEL es, por tanto, la afirmación de que todos los sistemas biológicos deben resistir activamente la tendencia natural al desorden. [13, 28].

La energía libre se define clásicamente en términos de principios termodinámicos, pero aquí sólo nos concierne la energía libre como energía libre variacional que se sigue de la teoría de la probabilidad y de las estadísticas Bayesianas, dado que ésta es la concepción de energía libre que concierne al PEL. De acuerdo a la teoría de la información, la energía libre es una cota superior de sorpresa (*surprisal*), donde la sorpresa se define como la diferencia entre las predicciones de un organismo (o sus anticipaciones) sobre sus entradas sensoriales y lo que recibe, en realidad, como entrada sensorial. Luego entonces, la sorpresa es una medida de improbabilidad que no debe confundirse con la noción psicológica de sorpresa (aunque en ocasiones ambas converjan). Los organismos que logran permanecer lejos de los límites de una fase terminal (y, por lo tanto, consiguen mantenerse con vida), señala el PEL, “lo hacen al minimizar la tendencia a entrar en este tipo particular de estados de sorpresa (esto es, no anticipados)” ([5] p. 1).

La relación entre la energía libre variacional y la entropía debería entenderse de la siguiente manera. La energía libre es una cota superior de sorpresa, y el promedio a largo plazo de sorpresa es la entropía. Para ver esto más claramente, considérese que un estado puede decirse poseer alta sorpresa si se considera poco probable de ocurrir relativo a un modelo generativo. La idea principal es que los organismos se convierten en modelos (aproximados) de sus nichos locales, ya que estos sistemas, en promedio y a lo largo del tiempo, destilan regularidades estadísticas de sus nichos y, por ende, tienden a encorporeizar estas regularidades en la forma de su cuerpo físico y en sus dinámicas globales internas [10, 11]. Si se considerara un estado en alta y constante sorpresa, éste sería un estado con entropía alta (con una alta improbabilidad estadística). De manera alterna, si un sistema es capaz de predecir las causas recibidas externa e internamente de su entrada sensorial, sería porque se encuentra en un estado de entropía baja, y por tanto de baja sorpresa. Esto no es más que decir que los estados esperados tienen una distribución de entropía baja. Y a la inversa, mientras más alto sea el número promedio de observaciones requeridas para describir la dispersión de estados para equis variable, mayor sería la entropía de la distribución de probabilidad de esa variable. En consecuencia, el PEL postula que los sistemas vivos pueden mantenerse a sí mismos dentro de límites entrópicos al buscar minimizar su energía libre.

A primera vista, al menos, lo que el PEL posibilita es la generación de una unidad profunda, subyacente que conecta “los procesos de adaptación, vida y mente” ([5], p. 1) catalogados en términos de una noción de teoría de la información de la energía libre. Todo esto equivale a decir que lo que permite a los sistemas vivos sobrevivir no es sino el mismo proceso que les permite percibir, actuar, pensar, etc. Luego entonces, el PEL ofrece la esperanza de proveer un único marco a través del cual unificar la teorización de la vida y la mente, y lo hace al apelar a un solo imperativo: la minimización de la energía libre [4, 10, 11]. Es así, como el PEL ofrece razones para ligar procesos de vida con procesos de mente a través de lo que es esencialmente una perspectiva de minimización de la incertidumbre de la vida y la mente. Sin embargo, en el contexto del PEL, una inspección minuciosa revela la tensión que existe entre diferentes concepciones en torno a cómo entender mejor las implicaciones de la minimización de la energía libre y las propiedades centrales que ésta conlleva.

### *2.1. La energía libre y la predicción cognitivista de minimización del error*

Algunas formulaciones de la energía libre añaden lo que se llama una restricción cognitivista, que tiene implicaciones en cómo una perspectiva semejante entiende la relación vida-mente. Por “restricción cognitivista” entendemos una restricción sobre la naturaleza del procesamiento de

información en cuestión; que sería pensado como procesamiento de representaciones con contenido semántico. Además de postular las representaciones mentales semánticas, el PEL cognitivista también concibe la minimización de energía libre a través de unos lentes epistemológicos particulares; a saber, que el teorema de la energía libre lleva a un escepticismo global en torno a la relación mente-mundo ([8], p. 2). A pesar de que hay mucho que decir respecto a las implicaciones del PEL cognitivista, nos enfocaremos en el problema de las representaciones mentales semánticas.

El PEL cognitivista se enmarca, casi exclusivamente, dentro de una interpretación particular de la inferencia bayesiana acumulativa que subyace en términos de “minimización de predicción del error” (MPE). Mientras que el PEL toma como punto de partida problemas relacionados con la auto-organización en termodinámica, sistemas fuera del equilibrio, y, por ende, puede aplicarse a un mayor número de fenómenos diferentes, las estrategias de minimización de predicción del error se han asociado más directamente con el funcionamiento cerebral ([6]; ver [1] para más referencias). El cerebro aquí se perfila como un modelo jerárquicamente generativo para minimizar la predicción de una cantidad de error que refleja la probabilidad de entrada sensomotriz referente a un modelo interno y basado-en-conocimiento. Es esta formulación basada-en-conocimiento la que agrega al supuesto cognitivista. Desglosemos lo que queremos decir con esto, dado que el PEL, generalmente, se toma por implicar que los estados internos están involucrados en la inferencia bayesiana, en la que estados tanto internos –que comprenden un modelo generativo de sistema— como activos pueden modelarse minimizando la energía libre de manera inferencial [29]. Así que, no es el problema de la inferencia o la posesión de un modelo generativo como tal lo que ejemplifica el supuesto cognitivista; sino que son las propiedades particulares asociadas con tales procesos inferenciales las que destacan un tipo particular, cognitivista de minimización de la energía libre.

El PEL cognitivista otorga especial importancia al procesamiento de información interno con propiedades semánticas (es decir, con contenido). Se trata de una postura semántica de la teoría computacional de la mente, en la que las representaciones internas se catalogan en términos de inferencias probabilísticas *top-down* de distribuciones de densidad de probabilidad. Semejante visión del funcionamiento cerebral usualmente se formula en el lenguaje de la psicología popular (el lenguaje de las creencias, los deseos, la atención y las razones) y procede de la premisa de que el procesamiento de información con propiedades semánticas es lo que constituye la cognición. La razón común para postular estados internos con propiedades semánticas es que a menos de que verdaderamente existan tales estados, éstos no serán cognitivos pues sistemas puramente físicos no serían capaces de representar el mundo más allá de sus estados internos. Dado que las mentes

generalmente se asumen manipulando representaciones y dado que la mayoría de los sistemas que ocurren naturalmente no se asumen manipulando tales cosas, se sigue que el PEL cognitivista dibuja una gruesa línea divisoria entre sistemas mentales y no mentales. Por ejemplo, Hohwy [3, 6, 7] arroja esos resultados cognitivistas del PEL (ver también [30]).

Trabajos de este tipo motivan a lo que nos referimos antes como la perspectiva de independencia entre la vida y la mente, que ubica los orígenes de la mente posteriores a los de la vida y que trata la relación entre lo vivo y lo mental de manera contingente. Sin embargo, debido a su marco funcionalista, este enfoque amenaza con introducir un vacío difícil de explicar entre formas de vida cognitivas, más complejas y el resto del mundo vivo, negando así, y definitivamente, la posibilidad de cualquier continuidad vida-mente. En una perspectiva cognitivista de la mente como ésta, sólo algunos sistemas vivos evolucionan la maquinaria neural capaz de realizar procesamiento de información que involucre propiedades semánticas.

## *2.2. Minimización no cognitivista de la energía libre*

Otra variante menos cognitivista o, incluso, anti-cognitivista llega a la conclusión de que uno puede respaldar el PEL sin respaldar una lectura cognitivista de la inferencia bayesiana aproximativa en el contexto de minimización de predicción del error. El PEL no-cognitivista presenta la minimización de la energía libre en sistemas físicos en términos de la “entropía” de Shannon en la teoría de la información [5]. Por lo tanto, es posible postular al PEL como el principio unificador de la vida y la mente, y, al mismo tiempo, negar que las características más básicas de la vida y la mente implican inferencias probabilísticas con contenido semántico, aún si la base de la vida y la mente implica en sí misma una inferencia probabilística (ver también [10, 14, 29]).

En algunas articulaciones generosas de esta perspectiva, se argumenta que la minimización de la energía libre no sólo ocurre en sistemas biológicos sino que también toma lugar en sistemas no vivos que abarcan desde la sincronización de relojes hasta el caldo primitivo y las redes sociales [9, 10]. A diferencia de PEL cognitivista, para quien la predicción de minimización del error es una función evolutiva de los cerebros que están en proceso de convertirse en modelos generativos jerárquicos, estas perspectivas “generosas” parecerían estar minando la continuidad y unidad de la vida y la mente. La razón de esto es que el principio que postulan y por medio del cual unifican vida y mente se aplica a sistemas que posiblemente no son ni vivos, ni mentales. Lo cual trae consigo algunos problemas. El primero es que si la mentalidad se realiza en procesos de minimización de la energía libre, y la minimización de la energía libre se aplica a todo desde seres humanos, relojes de péndulo hasta el caldo primordial, entonces la mentalidad resulta encontrarse casi en cualquier



lugar. Es difícil evaluar si un pansiquismo de esta forma es correcto. Sin un forma clara de separar la mentalidad de la no-mentalidad, la vida de la no-vida, cualquier sobre-generosa versión del PEL se vuelve demasiado general y pierde así su valor explicativo al abordar la relación entre la vida y la mente.

En las siguientes dos secciones, pasaremos a desarrollar estas dos formulaciones del PEL para la relación vida-mente, mientras tendremos en cuenta los preceptos principales de ambas versiones para llevarlas a sus respectivas conclusiones.

### **3. Del PEL Cognitivista a la Discontinuidad Vida-Mente**

Algunos psicólogos, neurocientíficos y científicos cognitivos adoptan la postura de que la mente puede explicarse en términos de computación y añaden que las mentes son computacionales. La mayor suposición de esta postura es que la computación no puede explicarse sin hacer referencia a contenido semántico y representación mental. A lo cual, el PEL cognitivista no se opone.

#### *3.1. Conceptos de computación*

Hohwy [3, 6, 7] se basa explícitamente en el PEL para desarrollar una teoría del cerebro como si se encontrara embebido en procesos de minimización de predicción del error con contenido semántico. Como hemos visto, el PEL sostiene que para que los organismos se mantengan lejos de la frontera de las fases terminales deben minimizar su energía libre. Hohwy afirma que esta formulación de la energía libre “corresponde al trabajo del cerebro de minimizar la predicción del error, de muestrear selectivamente datos sensoriales, de optimizar precisiones esperadas y de minimizar la complejidad de los modelos internos” y que estas descripciones de trabajo “remiten a la percepción, la acción, la atención y la selección de modelos, respectivamente” ([3], p. 1).

La MPE es un marco computacional. En él las funciones computacionales realizadas por el cerebro se proyectan en términos de minimización de predicción del error, que reflejan la probabilidad de entrada sensomotriz (o información) relativa a un modelo estadístico. Bajo esta perspectiva, el cerebro se concibe llevando a cabo los papeles funcionales requeridos para codificar un modelo bayesiano generativo jerárquico, no local, que abarca todo el cerebro [31]. Un modelo bayesiano es generativo dado que su función es “capturar la estructura estadística de algunos conjuntos de entradas observadas, al inferir una matriz causal, capaz de dar lugar a esta estructura” ([32], p. 21). Luego entonces, un modelo generativo es un modelo estadístico que mapea causas escondidas con consecuencias sensoriales, y, por ende, codifica previas creencias probabilísticas

acerca de cuáles efectos sensomotrices poseen tales causas provenientes del cuerpo y/o del mundo [33]. Lo que significa que en cada nivel de la jerarquía cortical del cerebro, las funciones de densidad de probabilidad se codifican de las señales de predicción del error que llegan del nivel cortical inferior [31].

La MPE combina esta imagen computacional del cerebro con la perspectiva de que el procesamiento de información es inferencial y dice que la inferencia debería entenderse como inferencia bayesiana aproximada relacionada con algo semejante al contraste de hipótesis. En este sentido, la inferencia estadística es una herramienta de la ciencia. Hohwy adopta esta perspectiva del cerebro al apuntar que “los científicos están en el negocio de minimizar el error en predicciones generadas de sus hipótesis” ([7], p. 3). Existen diferentes maneras a través de las cuales uno puede sumergirse en el contraste de hipótesis. O bien, uno puede ajustar ciertos parámetros de su propio modelo, o bien, uno puede intervenir en las muestras obtenidas a fin de tener un mejor acuerdo entre el modelo y las muestras que uno va recolectando. Según Hohwy, así como “sucede a la inferencia estadística, así le sucede al cerebro en la percepción [...] y en la acción” ([7], pp. 3–4). En la percepción inferencial, la percepción reduce la predicción del error en virtud de determinar predicciones previas (i.e., creencias probabilísticas que constituyen el modelo generativo). En la inferencia activa, la acción minimiza la predicción del error al trabajar en cambiar la entrada sensorial en sí misma a través de su movimiento en el mundo [7]. Percepción y acción, respectivamente.

La minimización de predicción del error puede, entonces, mostrarse como un mapeo a un mecanismo computacional, el modelo generativo realizado por el cerebro, que funciona para minimizar señales de error ([6], capítulo 2). Bajo tal perspectiva, los sistemas computacionales concretos son mecanismos funcionales de un tipo especial [34]. Y aunque la comunidad científica aún no se ha decidido respecto a los detalles de una implementación específica, la MPE se perfila como aquél que ofrece revelaciones nuevas a los mecanismos computacionales detrás de la percepción, la acción, la atención y otros procesos cognitivos. Mecanísticamente<sup>2</sup>, siguiendo a Hohwy, “esto se logra al suprimir la predicción del error en múltiples niveles de la jerarquía ordenada temporalmente” ([6], p. 85).

---

<sup>2</sup> N. del T. Del original en inglés “mechanistically”, término acuñado por el autor.

### *3.2. Conceptos de información: hacia una consideración semántica de la computación*

El PEL modela los estados internos embebiéndolos en la inferencia bayesiana aproximativa y donde ésta última es implementada en procesos de minimización de predicción del error: así es como convergen entre sí el PEL y la MPE. Sin embargo, la formulación cognitivista del PEL añade una dimensión semántica, explícitamente definida al perfil computacional de la MPE.

Para ver lo que esto implica, se necesita trabajar lo que la noción técnica de “información” podría involucrar en los procesos computacionales de reducción de predicción del error, en promedio y a lo largo del tiempo. Es posible distinguir diferentes conceptos de información; aquí consideraremos tres opciones.

La primera es que puede definirse termodinámicamente. Un sistema termodinámico cerrado, puede decirse, es uno que tiene contacto con su entorno circundante tan sólo en virtud del trabajo y del intercambio de calor. Termodinámicamente, la información equivale a la diferencia entre dos estados distinguibles (como son, el potencial gravitacional alto y bajo). Es común pensar esta discrepancia en términos de información y, de este modo, ligarla con las nociones termodinámicas de entropía y energía libre.

Segundo, uno también podría intentar definir información en términos de teoría de la información. Aquí la información es una medida del promedio de una distribución de probabilidad. Dicho de otra manera, la información es una medida de la media de la verosimilitud de que un mensaje sea transmitido entre una fuente y un receptor [35]. Friston [10] propone la minimización de la energía libre en términos de minimizar una cantidad de sorpresa, y dice que este concepto de información debería ser entendido en el sentido preciso de la información de Shannon [5]. Yendo un poco más allá, considérese lo que dicen Godfrey-Smith y Sterenly sobre la información de Shannon: “existe un tipo de ‘información’ a la que se remite en biología, el tipo originalmente descrito por Shannon, que no es problemática [...]. La información, en este sentido, existe siempre que haya contingencia y correlación” ([36], p. 4). Lo cual nos da una imagen en la que, incluso formas de vida simples (como las bacterias) pueden verse rastreando información y donde la información, en este sentido mínimo, puede mostrarse existente en sistemas complejos como los cerebros, y entre los organismos y su entorno. No obstante, es importante no sobre-intelectualizar esta noción. Como más adelante enfatizan Godfrey-Smith y Sterenly, si decimos que formas de vida simple pueden ser entendidas como instancias de la información de Shannon, o si decimos que los genes contienen información sobre las proteínas que producen, entonces “no estamos diciendo más de lo que decimos cuando decimos que existe una conexión de información entre el humo y el fuego, o entre los anillos de un árbol y su edad” ([36], p. 4). En efecto, de acuerdo a Godfrey-Smith

y Sterenly, la información de Shannon implica que “cualquier cosa es una *fente* de información si tiene un rango de estados posibles y una variable *lleva* información sobre otra al grado que sus estados se correlacionan físicamente” ([36], p. 1). Intuitivamente, mientras más se pueda inferir sobre el estado de una variable por el estado de otra, mayor información transmite la relación entre las dos variables. Esto también se conoce como información mutua. La covarianza es formalmente equivalente a la información mutua en el siguiente sentido: si dos variables yacen en una relación de covarianza, entonces la información sobre una reduce sorpresa (*surprisal*) acerca de la otra; lo que no es sino decir que la covarianza, como la información mutua, optimiza la evidencia del modelo.

Tanto la termodinámica como la información de Shannon no pueden respaldar la formulación cognitivista del PEL. Hohwy presenta al procesamiento predictivo jerárquico dando una imagen *representacionalista* de la naturaleza de la mente y la cognición [7] (Ver [30] como apoyo a la idea de que la MPE es una teoría representacional de la mente). Sin embargo, esto impide inmediatamente tanto a la termodinámica como a la información de Shannon jugar un papel importante al minimizar la sorpresa bajo la formulación cognitivista de la energía libre. La razón es simple. Aún si los anillos de un árbol son una fuente de información respecto a otros estados posibles (su edad), esto no significa que los anillos en sí mismos representen algo acerca de la edad del árbol: los dos estados son informacionalmente covariantes; entre ellos no existe una relación representacional.

La glosa representacional de Hohwy en torno a la mente es un acompañante más natural para un tercer concepto de información, a saber, una noción semántica de la información, y, de ahí, una perspectiva semántica de la computación. Una versión semántica/representacional de computación impone una restricción semántica. Piccinini presenta tal restricción como sigue: “Sólo los estados físicos que califican como representaciones pueden mapearse como descripciones computacionales, y por tanto calificar como estados computacionales” [[34], p. 9) Piccinini continúa diciendo que “la versión semántica es, probablemente, la más popular en la filosofía de la mente, porque parece cubrir mejor sus necesidades específicas que las otras. Dado que, generalmente, se asume que las mentes y las computadoras digitales manipulan [...] representaciones, resultan computar” ([34], p. 9). Así que, el PEL cognitivista postula que los cerebros computan información, y donde el tipo relevante de información es el tipo de información que usualmente nos parece importante para propósitos epistémicos.

Actualmente, se dice que el modelo generativo bayesiano, realizado en los circuitos neurales del cerebro tiene una arquitectura representacional doble. Esta arquitectura es “una que en cada nivel combina representaciones de entradas *bastante tradicionales* con representaciones de

error. De acuerdo a la propuesta doble, lo que queda fuera de la explicación es la señal de error, que (en estos modelos) figura como computada por ‘unidades de error’ dedicadas. Mismas que están ligadas, pero son distintas, a las llamadas unidades de representación encargadas de codificar las causas de las señales sensoriales” ([1], p. 187; las cursivas son mías). Así que, lo que el PEL cognitivista nos presenta es una perspectiva de la mente embebida en procesos computacionales de predicción semántica, codificados en representaciones jerárquicas de situaciones recibidas externa e internamente.

### 3.3. Problemas de integración

El PEL cognitivista eleva las demandas a favor de una integración de la información. En otras palabras, formas sofisticadas de cómputo, en términos de contenido semántico representacional, parecerían situar, para ponerse en marcha, una alta demanda del tipo de maquinaria requerida en la realización de tales modelos generativos jerárquicos y de amplio alcance. Se trata de una maquinaria que, probablemente, muchas formas de vida simple y otros organismos no tienen; dado que todo el trabajo de minimización de predicción del error, muestreo selectivo de datos, optimización de precisión, etcétera, se toma para ser “ordenado jerárquicamente en la corteza [cerebral]” ([7], p. 15). Clark [12] llega, precisamente, a esta conclusión. Al considerar formas de vida básica como organismos unicelulares capaces de quimiotaxis, Clark señala: “Semejante forma de vida puede responder a perturbaciones ambientales utilizando una variedad de trucos y estratagemas, *para ninguno de las cuales, se le requiere, estar embebida en un proceso* en el que simulaciones sensoriales recibidas se encuentran intentando generar la señal ‘de ‘arriba a abajo’” ([12], p. 4). Además, Clark indica que, incluso, si tales bacterias son auto-evidenciadoras en el sentido de optimizar un modelo de su entorno, tales formas de vida básica “necesitan no depender de predicciones *top-down* para estructurar e informar sus intercambios con el mundo. El procesamiento predictivo constituye entonces una teoría de proceso biológicamente plausible que puede o no puede ser implementada en cualquier sistema” ([12], p. 5). Una manera de leer a Clark [12] es creer que lo que dice es que mientras los organismos unicelulares están vivos, no son cognitivos, dado que carecen del tipo de arquitectura requerido por un sistema para instanciar un modelo generativo jerárquico por medio del cual se embeban en una inferencia probabilística. De hecho, si, en general los sistemas nerviosos centrales y más particularmente, la corteza cerebral son requeridos para que se dé la inferencia bayesiana, y es este tipo de proceso el que se requiere para la mentalidad, entonces esto empequeñece la continuidad entre la vida y la mente, al despojar ramas enteras del árbol de la vida de poseer propiedades mentales.

Esto implica que algunas formas de vida son sin mente, mientras que, al mismo tiempo, otras (más avanzadas neuralmente hablando) son de un tipo totalmente distinto: están vivas y tienen mente. Sin embargo, esto implica una separación profunda entre las mentes (con semántica) y el resto del mundo natural, vivo y no-vivo (sin semántica). Son de una índole distinta.

#### *3.4. Problemas de significado*

La dependencia del PEL cognitivista en el contenido semántico genera una serie de problemas sobre cómo explicar de manera naturalista propiedades semánticas de representaciones mentales, un problema que retomaremos más adelante. Primero, queremos prestarle atención a un problema diferente, pero relacionado. Todos los sistemas físicos pueden entenderse en términos de la información de Shannon; sin embargo, bajo esta concepción la información “es tan sólo información en términos de transferencia (insignificante) de energía y no debe confundirse con contenido intencional o semántico” ([13], p. 17). No obstante, generalmente, sólo algunos sistemas físicos se asumen capaces de manipular representaciones con contenido semántico. Una formulación cognitivista del PEL en la línea de la inferencia aproximativa bayesiana con propiedades semánticas es entonces consistente con una imagen que “conserva en sí a las mentes y a las computadoras, mientras deja fuera todo lo demás, vindicando así la teoría computacional de la cognición como una teoría sólida y no trivial” ([37], p. 9).

A pesar de que una teoría de la mente semejante es atractiva para los que defienden el PEL cognitivista, de inmediato enfrenta dificultades. Hohwy [3] hace uso del PEL para llegar a una versión cognitivista de la mente. Pero tomar como punto de partida el PEL, que inicia con consideraciones termodinámicas acerca de la auto-organización en sistemas dinámicos aleatorios, no da licencia para inferir una formulación cognitivista de la minimización de la energía libre. La razón es que el PEL presenta la auto-organización en términos de la información de Shannon, esto es, en términos de covarianza, correlación física y sincronía generalizada. ¿Qué sigue de esto?

Primero, en una formulación no-semántica del PEL, donde la información se entiende en el sentido de Shannon, la arquitectura de las formas básicas de cognición (por ejemplo, adaptarse a y actuar en un ambiente dinámico de una manera más que meramente dispuesta) no involucra, necesariamente, la adquisición y manipulación de estados internos con contenido semántico [21]. Luego entonces, las versiones cognitivistas del PEL ponen un estándar innecesariamente alto respecto a qué tipo de actividades calificarían de mentales. De hecho, existen argumentos que muestran (si son correctos) que la información-como-covarianza no es ningún tipo de contenido semántico [21]. En la sección 5, profundizaremos en estos puntos.

Segundo, las mentes han evolucionado “para tener las cosas hechas en tiempo real” ([38], p. 5). Por lo tanto, las mentes (y los cerebros) han evolucionado primariamente para la acción y sólo secundariamente para el pensamiento, como tradicionalmente se había concebido. Las formulaciones cognitivistas del PEL restan importancia al papel protagónico de la acción a favor de sus estrellas computacionales: la información semántica y las representaciones *top-down*, realizadas por los modelos generativos, neuralmente codificados. No obstante, se ha de considerar que a los sistemas sensoriales actuales “no les concierne la verdad ni la precisión como tales, sino más bien, la acción y la necesidad de mantener la estabilidad funcional de los organismos en los que éstos se sitúan” ([39], p. 90). Esto se ajusta difícilmente a la epistemología del PEL cognitivista que supone estados internos con condiciones de verdad y/o precisión que ligen tales estados internos con situaciones en el mundo. Considérese (nuevamente) el caso de la quimiotaxis en organismos unicelulares como la bacteria *E. coli*. La quimiotaxis es un ejemplo de un mecanismo de control empleado por los organismos para conservar el estado de homeostasis y, gracias a ello, prolongar la probabilidad de su supervivencia. En este caso, es la búsqueda de comida (como los azúcares) basada en los gradientes de concentración de repelentes y atrayentes químicos en el entorno fluido de una bacteria ([40], p. 311). Siguiendo a Auletta [40], la quimiotaxis puede modelarse consistiendo de las llamadas predicciones homeostáticas o las expectativas previas dotadas por selección natural, mismas que aproximan los estados externos preferidos por los organismos — llamémosles preferencias intrínsecas de los organismos— (en la última sección, retomaremos el desglose de esta noción). De acuerdo a Auletta, “las creencias previas dotadas genéticamente, implicadas por los estados internos y la configuración del organismo... especifican lo que es innatamente sorprendente y permiten actuar para enfrentar desviaciones no predichas de estados esperados” ([40], p. 315). En esta versión, el modelo generativo encorporeizado de la *E. coli* comprende creencias previas que especifican que debe luchar para encontrarse en gradientes altos de azúcar y se opone a lo contrario. Así que, la *E. coli* puede modelarse buscando estos estados a través de la inferencia activa. Lo cual sitúa a la acción y a las prioridades en pos de la acción en la base de la homeostasis y la vida. No hay razón para creer que la quimiotaxis que busca mejorar la vida está desprovista de mentalidad, dado que nada le impide tratar a los nutrientes que recibe como datos sensoriales, la membrana celular como una manta de Markov y el comportamiento constantemente escurridizo de la bacteria como casos de inferencia activa.

Esta diferencia comprometida con la teoría de la información —entre información semántica y no semántica— abre una brecha entre la vida y la mente. La vida tiene un lado metabólico. El metabolismo es una de las características principales de la vida [41]. Los sistemas auto-

organizativos pueden conservar su organización a pesar de toparse con tendencias termodinámicas hacia el desorden entrópico. Lo que se necesita es que los sistemas vivos sean capaces de localizar los estados correctos, en la totalidad de su espacio de estados, que les permitan mantenerse dentro de límites viables; como en el caso de la quimiotaxis. Por ejemplo, la “temperatura ideal de un humano está determinada por su corporalidad: a los 37°C las enzimas que regulan el metabolismo funcionan de manera óptima, mientras el costo metabólico de mantener la temperatura del cuerpo sea asequible bajo ciertas condiciones ambientales” ([13], p. 6). Mayor temperatura equivale a mayores niveles de sorpresa y viceversa. El gradiente de temperatura es una fuente de información, que se encuentra sistemáticamente relacionada con una distribución de probabilidad desconocida, en el sentido de que la sorpresa no puede ser directamente evaluada por un organismo. Sin embargo, esto no es sino decir que la información en cuestión es una covariación. No se trata de un estado semántico representacional que trae consigo información sobre otras situaciones. Luego entonces, formas básicas de procesos de auto-organización en sistemas vivos no tienen nada que ver con los requerimientos de mentalidad postulados por el PEL cognitivista.

Hohwy [3] asume que una lectura cognitivista de la minimización de predicción del error es co-extensiva con el PEL. Hemos argumentado que la concepción técnica de información utilizada en el PEL (información de Shannon) difiere sustancialmente de la información semántica. Además argumentamos que la primera no implica la última noción de información, lo que explica, en parte, por qué una formulación cognitivista del PEL termina separando de tajo la vida de la mente. En la sección siguiente, consideraremos más a detalle formulaciones no cognitivistas del PEL y exploraremos las posibles implicaciones derivadas de una tesis de continuidad vida-mente, basada en esta perspectiva no-cognitivista del PEL.

#### **4. De la minimización de energía libre a una perspectiva vida-mente demasiado generosa**

El PEL es un imperativo para la auto-organización de sistemas dinámicos. Dado que la vida existe, el PEL afirma que los sistemas vivos deben manifestar las propiedades siguientes. Primero, los sistemas biológicos deben tener un modelo de cómo se genera su entrada sensorial y de los tipos de estado en los que esperan encontrarse. En otras palabras, el PEL concibe a los sistemas biológicos próximos a los modelos estadísticos óptimos de sus nichos. Como Friston dice: “el principio de energía-libre toma la existencia de los agentes como su punto de partida y concluye que cada fenotipo o agente *encorporeiza* (*embodies*) un modelo óptimo de su nicho. Esta optimización se logra al minimizar la energía-libre, misma que delimita la evidencia por cada



agente (o modelo), producto de las interacciones sensoriales con el mundo” ([11], p. 89). La idea de que un sistema biológico encorporeiza un modelo requiere de un desglose cuidadoso.

Una manera de leer esta afirmación es que la evolución ha situado sistemas biológicos tales que, en promedio y a lo largo del tiempo, destilan (esto es, extraen) regularidades estadísticas de sus nichos, y, por lo tanto, las encorporeizan en su modelo. Quizá sea esto lo que lleve a los defensores del PEL cognitivista decir lo siguiente: “La mente puede entonces entenderse en términos internalistas y solipsistas, dejando de lado el cuerpo, el mundo y otras personas” ([7], p. 7). Una vez que el sistema biológico ha destilado las regularidades estadísticas de su nicho, puede deshacerse de él y, en cambio, depender de su modelo interno del mundo. Es ésta la implicación representacionista del PEL cognitivista.

#### *La ergodicidad y la Manta de Markov*

Ésta no es la manera como se entiende el concepto del modelo en el PEL según Friston (A menos de que explícitamente añadamos “cognitivista” al PEL, en lo que sigue, tendremos en mente una perspectiva no-cognitivista del PEL). La idea aquí es que el sistema biológico *es* un modelo de su nicho [11]. Como señala Friston: un sistema biológico “carece de un modelo de su mundo: él es un modelo. En otras palabras, la forma, la estructura y los estados de nuestros cerebros encorporeizados no contienen un modelo de lo sensorio: ellos mismos son el modelo” ([10], p. 32). En este sentido, cada fenotipo o sistema vivo “encorporeiza un modelo óptimo de su nicho” ([11], p. 89). Cabe añadir, que no sólo el sistema biológico es el que puede decirse que encorporeiza (*embodies*) el medio en el que se sitúa (*embedded*), sino que “el medio encorporeiza al agente” ([11], p. 89). En segundo lugar, por lo tanto, puede mostrarse a los modelos encorporeizados dependientes de expectativas previas respecto a cómo los estados del entorno han de desenvolverse en el tiempo. Esta característica es también conocida como ergodicidad. Decir que un sistema es ergódico es simplemente decir que sus estados tenderán a visitar a algún miembro de su conjunto atractivo a lo largo del tiempo, o que éste regresará, una y otra vez, al mismo vecindario de su espacio de estados durante el transcurso de su vida. En otras palabras, los sistemas ergódicos “[ocuparán] un pequeño número de estados de alta probabilidad y [evitarán] un gran número de otros estados” [42].

Los puntos principales que explican cómo es que los sistemas vivos perduran en el tiempo aplican del mismo modo a bacterias, plantas y a organismos como nosotros. Esto plantea las preguntas en torno a las continuidades y dis-continuidades que queremos explorar en el PEL. Las

respuestas a estas preguntas aún no son del todo trabajadas en la literatura del PEL y tienden a oscilar entre resultados radicalmente distintos.

En algunos pasajes, la formulación de Friston del PEL parece ofrecernos una visión profundamente continua de la relación vida-mente; lo que llamamos, en la introducción, el PEL no-cognitivista + una fuerte continuidad vida-mente. Friston enfatiza que existe más alrededor de la vida que la mera auto-organización. Los sistemas vivos también “negocian [...] un ambiente cambiante de manera que les permita perdurar largos periodos de tiempo” ([4], p. 422). En otras palabras, se encuentran entregados al comportamiento adaptativo. Bruineberg et al. [13] caracteriza esta habilidad adaptativa como “una instancia de lo que en la fenomenología filosófica se describe como ser-movido para tender hacia un agarre óptimo” ([13], p. 10; ver también [14, 22] para observaciones similares).

Sin embargo, en otros pasajes, Friston ofrece lo que para nosotros parece ser una perspectiva diferente y demasiado generosa del PEL no-cognitivista. Considérese, por ejemplo, su siguiente cita: el PEL “se aplica a cualquier... sistema que resiste una tendencia al desorden; desde los organismos unicelulares hasta las redes sociales” ([9], p. 293). O, como Friston dice aquí: “La motivación para minimizar la energía libre ha utilizado, hasta ahora, el siguiente tipo de argumento: los sistemas que no minimizan energía libre no pueden existir, porque la entropía de sus estados sensoriales no estaría limitada y aumentaría indefinidamente; según el teorema de fluctuación” ([10], p. 2). Estrictamente hablando, lo que Friston [10] dice aquí es que para que cualquier sistema exista, debe hacer trabajo para minimizar la energía libre. Lo que lo compromete con alguna de las siguientes tres implicaciones. Primero, si la minimización de la energía libre es suficiente para que exista mentalidad, entonces todo sistema tiene una mente, incluso aunque no todos los sistemas estén vivos. Segundo, si la minimización de la energía libre es suficiente para la vida y la mente, entonces todos los sistemas que existen están vivos y son mentales. Finalmente, los sistemas biológicos, como todos los otros que existen, necesitan hacer trabajo para minimizar la energía libre. La última opción señala que la minimización de energía libre no es una propiedad característica de los sistemas vivos, y, como tal, da lugar a alguna de las siguientes dos implicaciones. O bien (opción número uno) el PEL sitúa la mentalidad en una clase de sistemas que incluye pero no queda restringida a los sistemas vivos y, por lo tanto, vira hacia alguna forma de pansiquismo; o bien (opción número dos) el PEL equipara la continuidad vida-mente con una perspectiva que ve a la vida y a la mente en todas partes.

Algunos científicos y filósofos han abrazado la primera opción para sus teorías de la mente (por ejemplo, [43, 44]). Sin embargo, a excepción de las implicaciones pansiquistas o

pancomputacionales, se sigue que la minimización de la energía libre en sí misma no puede ser suficiente para la vida y la mentalidad, incluso tratándose de una propiedad necesaria para todo sistema vivo y cognitivo, como medio para establecer un límite máximo al desorden. ¿Qué más se requeriría entonces? La ergodicidad no parece ser suficiente ni para la vida ni para la mente dado que cualquier sistema dinámico aleatorio manifestará comportamiento ergódico porque tal sistema evolucionará en el tiempo hacia lo que se llama su atractor global aleatorio [45]. Considérese que arrojar una moneda al aire no sólo una vez, sino muchas es un ejemplo de un proceso ergódico, dado que la probabilidad de observar “caras”<sup>3</sup> convergerá como función del tiempo al mismo valor.

Además de la ergodicidad, Friston apela a la idea de que la manta de Markov delimita a los sistemas. Una manta de Markov está compuesta de estados sensoriales y activos, que separan estados internos de externos (en el sentido estadístico de independencia condicional). En una célula, digamos, la superficie de una célula constituye una manta de Markov al separar estados intracelulares (internos) de los extracelulares (externos). Una manta de Markov puede a su vez dividirse a sí misma en estados perceptuales/sensoriales y estados activos [10, 29]. Si asumimos que uno puede interpretar los estados internos parametrizando “una densidad arbitraria (variacional) o algunas creencias bayesianas... sobre estados externos, entonces la dinámica de estados internos y activos puede describirse como un descenso del gradiente en energía libre variacional” ([29], p. 3). En otras palabras, uno puede describir un sistema en virtud de procesar una manta de Markov, como si maximizara una evidencia del modelo a través de la inferencia bayesiana aproximativa. Por lo tanto, puede decirse que los estados internos de la manta de Markov infieren las causas ocultas de su entrada sensorial por medio de inferencia perceptual o activa.

Cabe mencionar que no todos los fenómenos poseen una manta de Markov; lo que muestra, desde la perspectiva del PEL, que la mera auto-organización no es suficiente ni para la vida ni para la mente. La flama de una vela es un caso concreto. A pesar de que puede parecer entregada a procesos de auto-manutención, no lo hace recursivamente [46], y, por lo tanto, no puede asegurar que los estados que la constituyen y sus interdependencias cambien lentamente; lo que se requiere para que una célula o un sistema ergódico posea una manta de Markov [10]. A pesar de estas consideraciones acerca de las mantas de Markov, las preocupaciones que planteamos aquí también se aplican a sistemas que poseen una manta de Markov, dado que los relojes de péndulo pueden modelarse constituidos por una manta de Markov, embebidos en una inferencia bayesiana y

---

<sup>3</sup> N. del T. En el caso de las monedas, usualmente se usan las parejas cara-cruz (España), cara-sello (Colombia), cara-seca (Argentina) un poco más relacionadas con el inglés *head-tail*. Para el lector mexicano, se trata tan sólo del equivalente águila-sol.

mostrando cierto grado de ergodicidad [10]. En el caso de los relojes de péndulo, Bruineberg et al. [13] sostienen que la sincronización de sistemas caóticos puede explicarse, más parsimoniosamente, apelando a una sincronía generalizada que a la idea de que cada reloj infiere el estado de otro, en tanto que cada uno puede entenderse como modelo de la dinámica interna del otro. Según Bruineberg et al. [13] aunque es completamente posible interpretar este caso en términos de la dinámica de la manta de Markov, no es necesario insistir en una lectura inferencial, sino que bastaría apelar a una sincronía generalizada. No nos interesa esta discusión en particular, lo que nos interesa es señalar que en otras publicaciones, Friston formula la sincronía generalizada como consistente con estados internos que se embeben en la inferencia bayesiana aproximativa [47]. Nuestro punto es que dado que los conceptos principales del PEL no-cognitivista —la inferencia bayesiana aproximativa, la ergodicidad, las mantas de Markov, etc.— pueden aplicarse, por un lado, a sistemas vivos y cognitivos, por otro lado, existe el riesgo evidente de sobreexplotar la aplicación de estos conceptos, misma que resultaría en o bien concebir a la vida y la mente casi en todo lugar, o bien, en despojar al PEL de su poder explicativo cuando se trata de abordar la naturaleza de la vida y la mente y la relación entre ambas (Supongamos que el primer punto puede defenderse, entonces el PEL no-cognitivista demasiado generoso no ejemplificaría una perspectiva de independencia entre la vida y la mente, en tanto que la vida y la mente esencialmente estarían presentes en cualquier lugar).

## **5. Restringiendo el PEL no-cognitivista con la REC<sup>4</sup>: una consideración de la mente evolutiva y tardía**

Nuestra estrategia consistirá, ahora, en mostrar que las implicaciones de la perspectiva demasiado generosa del PEL no-cognitivista pueden restringirse. En lugar de situar a la mentalidad fuera del dominio de la vida, o incluso pensar que todo lo que existe es semejante a la vida y entonces semejante a la mente, debemos sostener que tales implicaciones pueden constreñirse útilmente con el trabajo del enactivismo encorporeizado radical (REC) en la filosofía de la cognición naturalista [21, 48]. Esta restricción nos ayudará a tomar los primeros pasos hacia el desarrollo de una perspectiva no-cognitivista del PEL que evite cualquier tipo de afiliación con un panpsiquismo y en

---

<sup>4</sup> N. del T. REC (del inglés “radical enactive cognition” = cognición enactiva radical) es un término hoy día utilizado como nombre propio por los autores del enactivismo encorporeizado radical (radical embodied enactivism) y no como simple abreviatura como el caso de PEL (o FEP en inglés, “free energy principle”).

la que la presencia de contenido mental semántico (esto es, estados internos con condiciones de corrección) marque una transición en el reino de lo mental, en oposición a que señale un camino de lo no-mental a lo mental.

Los principales compromisos de la REC pueden articularse en dos principios básicos. El primero, es una negación de la perspectiva “usual” de las ciencias cognitivas y su filosofía, que señala que la cognición, en un sentido fundamental, involucra representación con contenido mental. Según la REC, “el vasto mar de posibilidades de lo que los seres humanos [y otros organismos] hacen y experimentan se entiende mejor al apelar al desenvolvimiento dinámico, las interacciones situadas y corporizadas y los compromisos con lo que el mundo nos da” ([48], p. 1). El segundo principio afirma que este tipo de actividades básicas cognitivas se dan en la actividad corpórea, entregada-al-mundo [21]. Ahora proseguiremos a destacar dos puntos de convergencia entre la REC y el PEL no-cognitivista.

Primero, ambos enfatizan la co-dependencia entre lo interno y lo externo. Recordemos que el PEL toma al sistema biológico como modelo de su econicho y agrega, además, que el econicho es un modelo del sistema biológico. Por ejemplo, mientras que la morfología y la posibilidad de acción de la araña reflejan su nicho, la telaraña y el entorno circundante reflejan el tipo de organismo que lo habita. En definitiva, el PEL realmente postula una separación entre los estados internos y externos: la manta de Markov implica dicha separación [28]. Sin embargo, a diferencia del PEL cognitivista, esto no significa que tal separación indique una separación *epistémica*. Así que, la mera existencia de una manta de Markov no implica que los estados internos de un organismo tengan que representar estados más allá de la manta en virtud de construir estados internos con contenido semántico sobre aquellos estados externos; sino más bien, sugiere, de manera similar a como lo hace el enactivismo, que “el organismo y el entorno están vinculados en una especificación y selección recíproca” ([49], p. 174: citado en [32], p. 289). Allen y Friston enfatizan esta mutualidad entre lo interno y lo externo al decir: “El punto es que la frontera en sí misma se constituye, más bien, por un intercambio dinámico ergódico entre ‘interno’ y ‘externo’, y no por una ejecución cognitivista de procesamiento interno” ([28], p. 16). La REC transforma esta mutualidad entre lo interno y lo externo en forma de covarianza, lo que no es sino expresión de cómo dos variables cambian (o no) juntas. La covarianza es formalmente equivalente a la emergencia de la sincronía generalizada que resulta de la inferencia activa —un corolario del PEL (ver [47])— y expresa un acoplamiento dinámico de dos o más sistemas dinámicos aleatorios. En este sentido, tanto la REC como el PEL formulan, a partir de principios básicos, por qué uno puede esperar una disminución de entropía cuando se da una sincronía generalizada o covarianza en

acoplamiento agente-entorno. De manera similar a la inferencia activa en el PEL no-cognitivista, la REC entiende el ensamblaje y la orquestación de tales acoplamientos dinámicos como resultado de la actividad encorporeizada (embodied).

Segundo, tanto el PEL no-cognitivista como la REC conciben el tipo de información disponible para un organismo biológico del tipo de la variedad de Shannon. Sin embargo, donde el PEL asume que este tipo de información da cuenta de la auto-organización en los sistemas biológicos [5], la REC introduce una implicación adicional que pretende plantear un problema serio para todas las teorías cognitivistas de la mente, incluyendo el PEL cognitivista [21]. El PEL cognitivista asume que un tipo de interacción inteligente con el mundo demanda contenido semántico. No obstante, esta suposición choca contra lo que Hutto y Myin [21] llaman el *difícil problema del contenido*, cuando dicen que “postular contenido informacional es incompatible con el naturalismo explicativo. El problema de raíz es que la Covarianza no Constituye Contenido” ([21], p. xv). La idea aquí es que el contenido semántico no existe independiente de ciertas prácticas socio-culturales, ergo, no constituye una propiedad inherente ni de los sistemas biológicos, ni, por ende, de la vida. ([21], p. xv).

No diremos nada sobre la relación entre el contenido y las prácticas sociales (pero ver [50]). En vez de eso, puliremos la siguiente implicación: si asumimos que la información-como-covarianza no puede dar lugar a la información-como-contenido y si asumimos, además, que responder adaptativamente a la información-como-covarianza es una propiedad esencial de los sistemas vivos y cognitivos, entonces, la mentalidad no es, en sus formas más básicas, un asunto de procesar cualquier tipo de contenido. Es, precisamente, por esta razón que sostuvimos que el PEL cognitivista conduce a una posición de independencia o de no-continuidad respecto a la relación vida-mente. Afirmamos esto, al observar que el PEL cognitivista no es compatible con el tipo de información en su base (como ésta se aplica a la auto-organización en los sistemas biológicos, esto es, información-como-covarianza), para dar cuenta de procesos que involucran información semántica, y, por lo tanto, esto lo obliga a excluir la mayoría de los procesos en biología del ámbito de lo mental.

La REC va más allá de sólo problematizar teorías de la mente cognitivista. Desarrolla una imagen positiva de actividad cognitiva sin contenido mental. De acuerdo a la REC si uno niega que cualquier tipo de interacción orgánica con el mundo debe, necesariamente, implicar contenido, no se sigue que este tipo de interacción sea, necesariamente no-mental. Bajo esta perspectiva, es posible que los organismos sean tanto *intencionalmente dirigidos*, activos, involucrados en compromisos dirigidos-al-mundo como responsivos informacionalmente, sin que esta dirección esté

mediada por estados internos con contenido semántico. Como Hutto y Myin señalan: “Las formas más simples de vida son capaces de una respuesta intencionalmente dirigida” ([21], p. x), pero esta respuesta no es un asunto de interpretación, entendimiento o cualquier otro tipo de actividad representacional sofisticada.

La REC desarrolla su propia explicación de la direccionalidad intencional sin contenido a través de una versión modificada del teleofuncionalismo. En términos simples, el teleofuncionalismo es la perspectiva de que lo que hace que algo sea una actividad mental tiene que ver con la función que desempeña en un organismo. La noción de función se concibe en términos de *funciones (biológicas) propias*, y tales funciones comúnmente se entienden a la luz de las condiciones evolutivas y/o históricas bajo las cuales las respectivas funciones fueron seleccionadas y/o adquiridas. Por lo tanto, el teleofuncionalismo explica por qué un organismo tiene ciertas capacidades biológicas y mentales, apelando al tipo de operaciones para las cuales tales funciones fueron seleccionadas, a fin de aumentar la probabilidad de lucha y supervivencia.

Muchos teleofuncionalistas son representacionistas con respecto a la mentalidad. La REC rechaza este compromiso del teleofuncionalismo. Según Hutto y Myin [21] la direccionalidad intencional de las formas básicas de la actividad cognitiva (por ejemplo, cuando una rana usa su lengua como un látigo para atrapar a una mosca) están constituidas por su historia natural de selección. Llegamos a esta misma conclusión cuando discutimos arriba la quimiotaxis en organismos unicelulares; lo cual implica que la direccionalidad intencional tiene una dimensión normativa. Una forma de actividad intencionalmente dirigida (del tipo no-semántico) “busca provocar ciertos tipos de respuesta orgánicas a ciertas cosas (o tipos de cosas) y no a otras” ([51], p. 142). Un bonito resultado de la REC es que implica una perspectiva de la mente dentro de la cual una transición de actividades carentes de contenido a otras llenas de contenido no indica una transición de la no-mente a la mente. La REC es, por lo tanto, una teoría transformativa de la mente en el ámbito de lo mental.

No todos los sistemas que ocurren naturalmente satisfacen la condición de la REC para la mentalidad. Esto es, no todas las cosas que existen exhiben direccionalidad intencional; por ejemplo, un termómetro. No porque un dispositivo como éste carezca de funciones sistémicas del tipo definido por el papel que desempeñan. Según la REC, semejante maquinaria carece del tipo de historia de selección natural asociada con las funciones biológicas apropiadas, que se requiere para que un sistema manifieste direccionalidad intencional. Si asumimos que la propuesta de la REC está en lo correcto, desarrollar el PEL no-cognitivist bajo los auspicios del teleofuncionalismo de la REC implicaría entonces la siguiente restricción importante: que sólo un subconjunto de sistemas

minimizadores de energía libre son intencionalmente dirigidos hacia ciertos motivos de su entorno local, y, por lo tanto, sólo un subconjunto de tales subsistemas es mental.

Este problema requiere mucha mayor discusión que la que aquí le daremos. No obstante, la distancia entre las versiones no-cognitivistas del PEL y el REC no es mucha, al menos no una vez que se reconoce un llamado a la historia de la selección. Así pues, leemos: “de acuerdo al RPP [i.e. procesamiento predictivo radical (*radical predictive processing*) basado en el principio de energía libre] la predicción de error desemboca no sólo en una escala ontogenética, sino también en una filogenética; si el cerebro (y el cuerpo) constituyen un modelo generativo, entonces aquellos [modelos generativos] que mejor se adecúan a su nicho ambiental serán seleccionados por evolución” ([28], p. 9). Esto es revelador. Allen y Friston [28] reconocen aquí la necesidad de desarrollar una perspectiva de modelos generativos, acoplamiento dinámico y actividad corporal en la línea sugerida por la REC. Al reconocer este traslapo entre el PEL no-cognitivista y la REC, llegamos a la conclusión de que es posible desarrollar una perspectiva no-semántica de organismos que minimizan la energía libre, sin que esto nos lleve a una generalización de lo mental sin restricciones.

Sin embargo, argumentaremos ahora que la REC se queda corta al defender una versión fuerte de la tesis de la continuidad vida-mente. (Nótese que esto no implica un dilema fatal para la REC al tener que dar cuenta de los orígenes del contenido en términos naturalistas (ver [52] para mayor discusión)). Por extensión, si el PEL no cognitivista tuviera que desarrollarse dentro de los confines de la REC, también se quedaría corto en llegar a tal perspectiva fuertemente unificadora de la vida y la mente. La razón de esto es que una vez que el REC se añade al PEL, esta adición resulta en una perspectiva evolutiva tardía de la mente. REC sitúa los orígenes de la vida previos a la emergencia de la mente, dada su adopción al teleofuncionalismo. Se sigue que el REC puede caracterizarse por aceptar los siguientes tres supuestos: el primero, que existe la vida pero no la cognición; el segundo, que existe la vida y que la vida y la mente convergen cuando los sistemas vivos se hace capaces de formas sin-contenido de la direccionalidad intencional, y, finalmente, que ciertos tipos de sistemas vivos (los seres humanos, por ejemplo) son capaces de comprometerse con actividad cognitiva con contenido semántico (como escribir o hacer matemáticas). No abordamos esta última implicación del REC, sino que nos enfocamos y nos hemos enfocado en una segunda afirmación, que es aquella según la cual organismos REC exhiben mentalidad —definida como acciones intencionales dirigidas— dada una historia de selección. Por lo tanto, aunado al PEL, el REC reduce el alcance del PEL no-cognitivista, dado que su apelación a una historia de selección



podría implicar que las primeras formas de vida no tienen propiedades mentales en tanto que carecerían (al menos todavía) de una historia natural de selección.

Estamos en desacuerdo con la perspectiva de la relación vida-mente que la REC conlleva, especialmente cuando resulta poco claro el trabajo explicativo que se está haciendo al apelar a la selección. Considérese, por ejemplo, cuando el primer sistema vivo empieza a dividirse. Este proceso resulta en un número de clones del primer sistema individual. Aunque, dadas las fluctuaciones ambientales, no todos los clones sean capaces de sobrevivir. Lo que nos lleva a la pregunta: ¿los sobrevivientes han sido “seleccionados naturalmente”? Si respondemos afirmativamente, entonces los sobrevivientes tienen direccionalidad intencional y los otros no. No obstante, es posible que no exista una diferencia funcional notable entre el individuo original y sus clones –independientemente de que sobrevivan o mueran. La implicación de esto sería entonces que organismos psicológicamente idénticos pueden diferenciarse en términos de mentalidad, lo que lleva a una preocupación filosófica similar a la que evoca el experimento pensado del “hombre del pantano” o la noción de zombies filosóficos (libres de energía).

Y aún más, la apelación a la “selección” arrastra problemas adicionales. Por ejemplo, ¿en qué punto podemos decir que una función fue “seleccionada”? ¿Se trata acaso del momento en el que una función contribuye por vez primera a niveles relativamente mayores de reproducción de un organismo, comparados con los de sus congéneres? ¿O acaso se requiere de varias generaciones de individuos por encima de la aptitud (*fitness*) promedio? ¿Y cómo es que esta particular contribución de la función a la aptitud puede ser desenredada no-trivialmente del resto de factores en juego? Lo que es seguro es que sin duda, a lo largo de las generaciones, la selección natural puede cambiar la manera particular en la que los organismos se dirigen intencionalmente a su entorno, pero aún no es claro si en primer lugar este proceso evolutivo debería tomarse constitutivo de esta direccionalidad intencional como tal (y por ende, mentalidad básica).

## **6. De la REC al EA y el PEL no-cognitivista: una fuerte continuidad vida-mente**

Finalmente, sugeriremos una manera a partir de la cual rehabilitar efectivamente el problema arriba mencionado; y lo haremos mostrando que el PEL no-cognitivista también se encuentra a la par del enactivismo autopoietico (EA), permitiéndonos tejer así conjuntamente los orígenes de la vida y los orígenes de la mente (Esta versión constituye un desarrollo más a la línea de argumentación perseguida en [14]).

La REC toma como punto de partida que existe una población de individuos que puede tomar forma por las fuerzas de la selección natural. Sin embargo, ¿qué es lo que es precisamente un individuo? (No estamos en posición de discutir esto a profundidad aquí –una tarea para otra ocasión. Para una discusión inicial detallada ver [53] y el resto de los artículos en *Biology and Philosophy*). Si este concepto es demasiado amplio, entonces dificulta delimitar apropiadamente la noción de direccionalidad intencional, dado que muchos tipos de sistemas pasan por historias de cierto tipo de selección respecto a su funcionamiento, como un termostato. Además es innegable que el funcionamiento de un termostato puede ser descrito normativamente, esto es, en términos de si regula correctamente o no la temperatura. Sin embargo, dado que sus diseñadores y usuarios especifican externamente las condiciones de “corrección”, su normatividad es tan sólo un tipo derivado de normatividad. Éste no es el caso de los sistemas vivos. Por ejemplo, nuestros cuerpos están regulando continuamente la temperatura para mantenerla dentro de unos límites específicos. Según el PEL, esto se logra por la minimización de la energía libre, en promedio y a lo largo del tiempo. Esta función se encuentra también sujeta a condiciones de éxito y fracaso, pero aquí las condiciones son intrínsecas al cuerpo en lugar de ser definidas de manera externa, por la simple razón de que están determinadas por un rango de viabilidad, que al final constituyen una cuestión de existencia, de vida y muerte [28].

Siendo justos, incluso un termostato descompuesto puede desintegrarse eventualmente por sobrecalentamiento. No obstante, esto no es suficiente para atribuirle una forma intrínseca de normatividad. La diferencia principal es que nuestro cuerpo, o cualquier sistema vivo en este caso, y en contraste con el termostato, es un sistema físicamente auto-productivo (autopoiesis). Vivir es un proceso que se mantiene a sí mismo lejos de condiciones termodinámicas de equilibrio al invertir trabajo en prevenir su propia desintegración. De esta manera, debe definir una frontera entre lo que le pertenece a sí mismo y lo que le pertenece a su entorno. Esta frontera lo define como un individuo autónomo. Según el PEL no-cognitivista, un sistema semejante es un sistema adaptivo capaz de una inferencia activa con una manta de Markov, donde la manta de Markov no conlleva nada parecido a una frontera epistémica que separa la mente del mundo. Mientras que el modo de ser de un termostato se caracteriza por una persistencia pasiva, un organismo se caracteriza por una existencia activa (o inferencia activa). Sin mencionar que, dado que el ser de un organismo es su hacer, su existencia como individuo es inherentemente precaria y aun así los seres vivos generalmente hacen todo lo que pueden para mantener su forma de vida contra todo pronóstico. Esto es, a pesar de tener la posibilidad de dejarse simplemente sucumbir ante las fuerzas de decaimiento y desintegración, consistentemente regulan sus fronteras para evadir este destino. En

otras palabras, los seres vivos no pasan simplemente por perturbaciones como los sistemas no-vivos, sino que responden selectivamente, y el éxito de esta respuesta está normativamente ligado a la preservación de su forma de vida.

Por lo que, el enactivismo autopoietico (EA) nos da los inicios de una versión de cómo el origen de la vida es también el origen de una individualidad autónoma y de una normatividad intrínseca (los orígenes de sistemas que mantienen los estados internos de sus mantas de Markov, vía inferencia activa). Sin embargo, al principio esta normatividad no es específica de una función particular sino que concierne a un individuo como un todo: toda la actividad que no mata al individuo tienen un valor positivo para el sistema vivo. La actividad que mata al individuo tendría un valor negativo excepto que ya no puede ser así para el sistema (que ya no está). A fin de poder ser capaz de funcionar de una manera más eficiente, es necesario que el sistema vivo sea capaz de responder de manera diferencial y, particularmente, de detectar y evitar interacciones potencialmente letales antes de que éstas se vuelvan letales. Es poco posible que este tipo de complejidad pueda emerger espontáneamente en el origen de la vida, en tanto requiere una historia de selección, concebida en términos generales como el resultado de la interacción, el desarrollo y la evolución. En otras palabras, es aquí donde el EA se conecta con las preocupaciones de la REC y la direccionalidad intencional entra en escena como una manera de responder adaptativamente a características relevantes del medio interno y externo del organismo. Mientras que el EA ayuda a la REC a constreñir apropiadamente su apelar a una historia de selección; en respuesta, la REC puede ayudar al EA a crear un puente a la “brecha cognitiva” [26] entre mentes básicas y mentes no-básicas.

Aproximarse a la direccionalidad intencional de esta manera tiene la ventaja de llegar al contexto de un individuo autónomo del que ya se encuentra activamente distinguido, al mismo tiempo que relacionado con, lo que es otro (su entorno) y cuya actividad ya está intrínsecamente caracterizada por una forma básica de normatividad holística. Además no hay necesidad de preocuparse por el estatuto mental de artefactos con funciones seleccionadas artificialmente, dado que carecen del criterio de vida básica (autopoiesis y adaptividad, ver [54]). La direccionalidad intencional es en efecto, bajo esta perspectiva, un refinamiento y elaboración de un perfil existencial básico por medio de historia selectiva. Las condiciones de lo que la REC considera una mente básica, es decir, un individuo autónomo que tiene normatividad intrínseca, ya estarían dadas. De ser así, podemos decir que tanto la mente básica como la vida básica emergen juntas en el origen de la vida, un evento de capital importancia, que después debería reconceptualizarse como el origen de la-vida-y-la-mente –como los orígenes de organismos con mantas de Markov capaces de mantener

sus propios estados internos (o procesos) a través de una inferencia activa. (En este artículo, hemos dado pasos hacia el desarrollo de una aproximación a la tesis fuerte de la continuidad vida-mente que puede juntar revelaciones clave del principio de energía libre con los últimos trabajos en la ciencia cognitiva enactiva. No obstante, esta tarea está lejos de estar concluida. Sospechamos que trabajos posteriores de este proyecto buscarán evaluar el PEL no sólo en relación con los marcos de la REC y el EA, sino también en términos de otras versiones de dinámica predictiva y anticipativa. Para trabajos apuntando en esta dirección, ver [55-57]).

## **7. Conclusiones**

Si algo como el PEL es lo suficientemente apropiado para abordar la relación vida-mente, entonces la vida y la mente co-emergen dentro de la amplia clase de procesos que le conciernen a la minimización de incertidumbre. Sin embargo, no existe ninguna versión singular de esta clase general para pensar la relación entre vida y mente. En este texto, hemos explorado dos perspectivas diferentes del PEL en torno a la relación vida-mente. En primer lugar, argumentamos que el PEL cognitivista implica que no existe continuidad entre la vida y la mente. Esencialmente, el apelar del cognitivismo al contenido semántico como signo de lo mental sitúa una alta barrera que sólo puede ser cruzada por un pequeño número de especies recientes y, potencialmente por sus artefactos. En segundo lugar, argumentamos que formulaciones demasiado generosas del PEL no-cognitivista amenazan con implicar que la mentalidad se encuentra prácticamente en cualquier lugar.

Para evitar estas implicaciones, posteriormente argumentamos que una versión no-cognitivista del PEL puede restringirse útilmente al integrarse a ideas clave de la ciencia cognitiva enactiva acerca de la individualidad y la normatividad. Finalmente, concluimos que la versión más contundente de la relación vida-mente las trata de fuertemente continuas y que esta continuidad se basa en conceptos particulares de vida básica (autopoiesis y adaptividad) y mente básica (intencionalmente dirigida aunque no semántica) en el contexto del principio de energía libre.

## **Referencias**

1. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 2013, *36*, 181 – 253.
2. Friston, K.J. The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* 2010, *11*, 127 – 138.

3. Hohwy, J. The Neural Organ Explains the Mind. Disponible en línea: [http://open-mind.net/papers/the-neural-organ-explains-the-mind/at\\_download/paperPDF](http://open-mind.net/papers/the-neural-organ-explains-the-mind/at_download/paperPDF).
4. Friston, K.J.; Stephan, K.E. Free-energy and the brain. *Synthese* 2007, *159*, 417–458.
5. Friston, K.; Thornton, C.; Clark, A. Free-energy minimization and the dark-room problem. *Front. Psychol.* 2012, *3*, 130.
6. Hohwy, J. *The Predictive Mind*; Oxford University Press: Oxford, UK, 2013.
7. Hohwy, J. The self-evidencing brain. *Noûs* 2014, *50*, 259–285.
8. Hohwy, J. How to Entrain Your Evil Demon. Disponible en línea: [http://predictive-mind.net/papers/how-to-entrain-your-evil-demon/at\\_download/paperPDF](http://predictive-mind.net/papers/how-to-entrain-your-evil-demon/at_download/paperPDF).
9. Friston, K.J. The free-energy principle: A rough guide to the brain? *Trends Cogn. Sci.* 2009, *13*, 293–301.
10. Friston, K.J. Life as we know it. *J. R. Soc. Interface* 2003, *10*.
11. Friston, K.J. Embodied inference: Or ‘I Think Therefore I am, If I Am What I Think’. In *The Implications of Embodiment (Cognition and Communication)*; Tschacher, W., Bergomi, C., Eds.; Imprint Academic: Exeter, UK, 2011; pp. 89–125.
12. Clark, A. How to Knit Your Own Markov Blanket: Resisting the Second Law with Metamorphic Minds. Disponible en línea: <http://www.x-spect.org/uploads/9/8/1/5/98154170/knittingmarkov8.pdf>.
13. Bruineberg, J.; Kiverstein, J.; Rietveld, E. The anticipating brain is not a scientist: The free-energy principle from an ecological-enactive perspective. *Synthese* 2016.
14. Kirchhoff, M.D. Autopoiesis, free energy, and the life-mind continuity thesis. *Synthese* 2016.
15. Godfrey-Smith, P. Mind, matter, and metabolism. *J. Philos.* 2017.
16. Froese, T.; Ikegami, T. The brain is not an isolated “black box,” nor is its goal to become one. *Behav. Brain Sci.* 2013, *36*, 213–214.
17. Kirchhoff, M.D. Species of realization and the free energy principle. *Australas. J. Philos.* 2015, *93*, 706–723.
18. Bruineberg, J.; Rietveld, E. Self-organization, free energy minimization, and optimal grip on a field of affordances. *Front. Hum. Neurosci.* 2014.
19. Di Paolo, E. Extended life. *Topoi* 2009, *28*, 9–21.
20. Froese, T.; di Paolo, E.A. The enactive approach: Theoretical sketches from cell to society. *Pragmat. Cogn.* 2011, *19*, 1–36.
21. Hutto, D.D.; Myin, E. *Radicalizing Enactivism*; The MIT Press: Cambridge, MA, USA, 2013.

22. Kirchhoff, M.D. Experiential fantasies, prediction, and enactive minds. *J. Conscious. Stud.* 2014, 22, 68–92.
23. Stewart, J. Life = Cognition: The Epistemological and Ontological Significance of Artificial Life. En *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*; Varela, F.J., Bourgine, P. Eds.; MIT Press: Cambridge, MA, USA, 1992; pp. 475–483.
24. Stewart, J. Cognition = life: Implications for higher-level cognition. *Behav. Process.* 1996, 35, 311–326.
25. Bitbol, M.; Luisi, P.L. Autopoiesis with or without cognition: Defining life at its edge. *J. R. Soc. Interface* 2004, 1, 99–107.
26. Froese, T.; di Paolo, E. Sociality and the life-mind continuity thesis. *Phenomenol. Cogn. Sci.* 2009, 8, 439–463.
27. Thompson, E. *Mind in Life*; MIT Press: Cambridge, MA, USA, 2007.
28. Allen, M.; Friston, K.J. From cognitivism to autopoiesis: Towards a computational framework for the embodied mind. *Synthese* 2016.
29. Friston, K.J.; Levin, M.; Sengupta, B.; Pezzulo, G. Knowing one's place: A free energy approach to pattern regulation. *J. R. Soc. Interface* 2015, 12, 20141383.
30. Gladziejewski, P. Predictive coding and representationalism. *Synthese* 2015.
31. Hohwy, J. Attention and conscious perception in the hypothesis testing brain. *Front. Psychol.* 2012, 3, 74–87.
32. Clark, A. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*; Oxford University Press: Oxford, UK, 2016.
33. Colombo, M.; Wright, C. Explanatory pluralism: An unrewarding prediction error for free energy theorists. *Brain Cogn.* 2016.
34. Piccini, G. Computation in Physical Systems. Disponible en línea: <https://plato.stanford.edu/entries/computation-physicalsystems>.
35. Shannon, C. A mathematical theory of communication. *Bell Syst. Tech. J.* 1948, 27, 279–423.
36. Godfrey-Smith, P.; Sterenly, K. Biological Information. Disponible en línea: <https://plato.stanford.edu/entries/information-biological/>.
37. Piccinini, G. *Physical Computation: A Mechanistic Account*; Oxford University Press: Oxford, UK, 2015.
38. Wilson, R.; Foglia, L. Embodied Cognition. Disponible en línea: <https://plato.stanford.edu/entries/embodied-cognition/>.

39. Shani, I. Intentional directedness. *Cybern. Hum. Know.* 2006, 13, 87–110.
40. Auletta, G. Information and metabolism in bacterial chemotaxis. *Entropy* 2013, 15, 311–326.
41. Boden, M.A. Is metabolism necessary? *Br. J. Philos. Sci.* 1999, 50, 231–248.
42. Rabinovich, M.I.; Friston, K.J.; Varona, P. (Eds.) *Principles of Brain Dynamics: Global State Interactions*; The MIT Press: Cambridge, MA, USA, 2012.
43. Chalmers, D. Panpsychism and panprotopsychism. In *Consciousness in the Physical World: Perspectives on Russellian Monism*; Alter, T. Nagasawa, Y. Eds.; Oxford University Press: Oxford, UK, 2015; pp. 246–276.
44. Tononi, G.; Koch, C. Consciousness: Here, there and everywhere? *Philos. Trans. R. Soc. B* 2015, 370.
45. Crauel, H. Global random attractors are uniquely determined by attracting deterministic compact sets. *Annali di Matematica Pura ed Applicata* 1999, 4, 57–72.
46. Bickhard, M.H. Interactivism: A manifesto. *New Ideas Psychol.* 2008, 27, 85–95.
47. Friston, K.J.; Frith, C. A duet for one. *Conscious. Cogn.* 2015, 36, 390–405.
48. Hutto, D.D.; Kirchhoff, M.D.; Myin, E. Extensive enactivism: Why keep it all in? *Front. Hum. Neurosci.* 2014.
49. Varela, F.; Thompson, E.; Rosch, E. *The Embodied Mind*; MIT Press: Cambridge, MA, USA, 1991.
50. Hutto, D.D. *Folk-Psychological Narratives*; The MIT Press: Cambridge, MA, USA, 2008.
51. Hutto, D.D. Against passive intellectualism: Reply to Crane. In *Radical Enactivism*; Menary, R. Ed.; John Benjamins Publishing Company: London, UK, 2006.
52. Hutto, D.D. Against passive intellectualism: Reply to Crane. In *Radical Enactivism*; Menary, R., Ed.; John Benjamins Publishing Company: London, UK, 2006.
53. Pradeu, T. The many faces of biological individuality. *Biol. Philos.* 2016, 31, 761–773.
54. Froese, T.; Ziemke, T. Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artif. Intell.* 2009, 173, 366–500.
55. Bickhard, M.H. The anticipatory brain: Two approaches. In *Fundamental Issues of Artificial Intelligence*; Müller, V.C., Ed.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 259–281.
56. Rosen, R. *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations*; Springer: Berlin/Heidelberg, Germany, 2012.
57. Nasuto, S.J.; Hayashi, Y. Anticipation: Beyond synthetic biology and cognitive robotics. *BioSystems* 2016, 148, 22–31.