UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE FILOSOFÍA

TESIS DOCTORAL

Conciencia, proceso y autotransformatividad

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Enrique Aramendia Muneta

Directores

Óscar Lucas González Castán
Mariano Luis Rodríguez González
Miguel Ángel Sebastián González

Madrid

© Enrique Aramendia Muneta, 2023

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE FILOSOFÍA

TESIS DOCTORAL

Conciencia, proceso y autotransformatividad

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR
Aramendia Muneta, Enrique

DIRECTORES
González Castán, Óscar Lucas
Rodríguez González, Mariano Luis
Sebastián González, Miguel Ángel

Madrid, 2022

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE FILOSOFÍA

TESIS DOCTORAL

Conciencia, proceso y autotransformatividad

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

Programa: D9AC - Doctorado en Filosofía

PRESENTADA POR
Aramendia Muneta, Enrique

DIRECTORES
González Castán, Óscar Lucas
Rodríguez González, Mariano Luis
Sebastián González, Miguel Ángel

Madrid, 2022

Agradecimientos

Muchísimas gracias a Óscar, Mariano y Miguel Ángel. A Pedro Chacón. A Carlos Sierra-Lechuga y el grupo *Realidad y Proceso*. A Johanna Seibt e Idoya Zazpe. A Asier Arias. A Juan Hermoso Durán. A Raúl Jauregi. A Mariola Elia. Aitari. Amari. Susanari. Leari. Esker anitz.

Sobre TEFLONX

Teflon X(cc0 1.0(documentación) MIT(código))es una plantilla de LaTeX creada por David Pacios Izquierdo con fecha de Enero de 2018. Con atribuciones de uso CC0.

Esta plantilla fue desarrollada para facilitar la creación de documentación profesional para Trabajos de Fin de Grado, Trabajos de Fin de Máster o Doctorados. La versión usada es la X

V:X Overleaf V2 with XeLaTeX, margin 1in, bib

Contacto
Autor: David Pacios Izquiero
Correo: dpacios@ucm.es
ASCII: ascii@ucm.es
Despacho 110 - Facultad de Informática

Índice general

Resumen iv

Abstract vi

Exordio viii

Introducción 1

1. Conciencia 5
1.1. Introducción: conciencia, del qué es al cómo es . . . 5
1.2. Cómo es la conciencia . . . 8
1.2.1. Propiedades fenoménicas . . . 9
1.2.2. Propiedades funcionales . . . 18
1.2.3. Propiedades dinámico-estructurales . . . 23
1.2.4. Propiedades temporales . . . 30
1.2.5. Conclusiones . . . 35
1.3. Cómo se ha explicado la conciencia . . . 36
1.3.1. Tipos de teorías de la conciencia . . . 38
1.3.2. Teorías representacionales . . . 38
1.3.3. Teorías informacionales . . . 45
1.3.4. Teorías biológicas . . . 57
1.3.5. Teorías físicas . . . 61
1.3.6. Teorías entrelazantes . . . 63
1.4. Conclusiones . . . 64

2. Proceso 67
2.1. Introducción: un marco procesual para la conciencia . . . 67
2.2. Qué es un proceso . . . 68
2.2.1. Proceso, proceso general y proceso transformativo . . . 70
2.2.2. Tipos de procesos según la forma mereológica . . . 72
2.3. Cómo analizar los procesos . . . 74
2.3.1. Mereologías clásicas para el espacio y el tiempo . . . 75
2.3.2. Una mereología no clásica para los procesos . . . 79
2.4. Autotransformatividad . . . 87
2.4.1. Interferencia y clausura . . . 88
2.4.2. Modulación y clausura . . . 93
2.4.3. Clausura de interferencia y clausura de modulación . . . 105
2.4.4. Caracterización completa de los procesos autotransformativos . . . 112

2.5. Conclusiones 112

3. Conciencia y proceso **115**
3.1. Introducción: por qué un enfoque de procesos 115
3.2. Teorías de la conciencia y proceso 116
3.2.1. Autotransformatividad en la Teoría del Espacio de Trabajo Global 116
3.2.2. Autotransformatividad en la Teoría de la Información Integrada . 120
3.2.3. Autotransformatividad en la Teoría de la Doble Microgénesis . . . 128
3.2.4. Autotransformatividad en la Teoría Autorrepresentacional 133
3.2.5. Conclusiones 141
3.3. Propiedades de la conciencia y proceso 141
3.3.1. Argumento de las propiedades no fenoménicas 142
3.3.2. Argumento de la intimidad 149
3.4. El enfoque de procesos en dos problemas concretos 160
3.4.1. Introducción: la conciencia como proceso en el espacio y en el tiempo160
3.4.2. La conciencia de la certeza en la percepción 164
3.4.3. La experiencia *punta-de-la-lengua* 172
3.5. Conclusiones 180

Conclusiones **182**

Índice de figuras **185**

Índice de cuadros **186**

Bibliografía **187**

Conciencia, proceso y autotransformatividad

Resumen

Introducción
La conciencia es un hecho cotidiano. Todos experimentamos conciencia a cada momento. Al mismo tiempo, la conciencia es un hecho complejo. Se presenta bajo muchas formas y aspectos, y su explicación resulta extraordinariamente elusiva y está plagada de grandes problemas que la ciencia y la filosofía han señalado convenientemente. Para describir en qué consiste la conciencia a menudo se han venido empleando términos como "sustancia", "estado" o "inmediatez". Por desgracia, esta terminología resulta confusa y no refleja de manera adecuada la riqueza y la naturaleza dinámica de la conciencia. Por contraste a ese tipo de enfoques, en esta tesis voy a usar la noción de proceso para explicar el hecho de la conciencia.

Síntesis
El objetivo de esta tesis es explicar la conciencia a través de la noción de proceso. Este enfoque procesual lo llevaré a efecto en tres grandes pasos. Primero, voy a tratar de definir cómo es y cómo se ha explicado la conciencia. El "cómo es" nos llevará a las propiedades de la conciencia, mientras que el "cómo se ha explicado" nos llevará a las teorías de la conciencia. Segundo, voy a presentar un marco teórico basado en la filosofía de procesos, en el cual se introducirán sucesivamente la noción de proceso transformativo y la noción de proceso autotransformativo. Tercero, voy a usar el marco teórico anterior para resolver algunos de los problemas que rodean al hecho de la conciencia.

El término "conciencia" es polisémico. Con "conciencia" aquí me refiero al hecho de tener experiencias conscientes. Es difícil decir qué es una experiencia consciente. No obstante, es posible describir cómo es una experiencia consciente. Las experiencias conscientes tienen siempre un carácter fenoménico, bien sea subjetivo o cualitativo. Además, típicamente tienen cierto contenido, el cual puede ser directa o indirectamente accesible por el sujeto para hablar, razonar o actuar de acuerdo a él. Las experiencias conscientes tienen típicamente cierta composición, es decir, tienen partes experienciales, y, pese a tener composición, se vive como dotada de unidad. Para finalizar, las experiencias conscientes suceden siempre en el tiempo. Tienen un principio y un final, y consumen energía.

Las teorías de la conciencia es el modo en que se ha tratado de explicar el hecho de la experiencia consciente. Hay muchas teorías de la conciencia, aunque se pueden dividir en cuatro tipos fundamentales: teorías representacionales, funcionales, biológicas y físicas. Aquí me centraré en cuatro teorías de la conciencia que, por su impacto o grado de determinación, resultan especialmente relevantes: la Teoría de la Información Integrada, la Teoría del Espacio de Trabajo Global, la Teoría de la Doble Microgénesis y la Teoría Autorrepresentacional.

Para tratar algo como la experiencia consciente es necesario echar mano de herramientas específicas. Una de esas herramientas es la noción de proceso. Los procesos suceden en el tiempo, tienen un principio y tienen un final. Además, están compuestos, es decir, tienen partes. Esas partes se describen de manera funcional (no de manera espacio-temporal) y, a menudo, el todo es algo más que la mera suma de las partes. Hay un tipo de proceso que resulta especialmente relevante: el proceso autotransformativo. Los procesos autotransformativos se describen a través de dos propiedades: (1) tienen una composición basada en interferencia mutua y (2) se describen a través de variables que se relacionan entre sí a través de modulaciones con clausura.

Cuando se aplica la noción de proceso a la experiencia consciente se llega a la conclusión de que tiene precisamente la forma de un proceso autotransformativo. Esto se puede demostrar de dos formas. Primero, las propiedades de la conciencia conducen a afirmar que las experiencias conscientes son procesos y que el proceso autotransformativo es precisamente el tipo de proceso que puede soportar estas propiedades. Segundo, las teorías de la conciencia (y, entre ellas, las cuatro teorías antes mencionadas), cuando se les exige un grado alto de determinación en sus mecanismos internos, acaban por introducir aspectos propios de los procesos autotransformativos, y esa introducción no es anecdótica, sino fundamental en ellas.

Conclusiones

Las conclusiones principales de esta tesis son dos. La primera es que la noción de proceso es imprescindible para entender y explicar el hecho de la experiencia consciente. Los enfoques no procesuales, por el contrario, conducen a callejones sin salida. La segunda es que la noción de proceso autotransformativo es una herramienta fundamental que, como mínimo, contribuye a simplificar muchos de los nudos explicativos que rodean al hecho de la experiencia conciencia y su aspecto específico.

Consciousness, process and self-transformativity

Abstract

Introduction
Consciousness is an ordinary fact. Everybody is conscious and everybody has conscious experiences. Simultaneously, consciousness is an extremely complex fact. It appears under manifold appearances, and it is very hard to explain, as science and philosophy have appositely pointed out. Terms like "substance", "state" and "immediacy" are often used to refer to conscious experiences. However, this terminology does not reflect the dynamic nature of consciousness. Unlike former approaches, in this thesis I am going to use the notion of process to explain the fact of consciousness.

Summary
The aim of this thesis is to use the notion of process to explain the fact of consciousness. This process-based approach is composed by three main steps. First, I try to define how consciousness is and how has been explained. The former question is linked to the properties of consciousness while the latter is linked to the theories of consciousness. Second, I present a theoretical framework, based on process philosophy, where I introduce two very relevant notions, namely, the transformative processes and the self-transformative processes. Third, I apply this framework to solve some problems that surround the fact of consciousness.

"Consciousness" is a polysemous term. By "consciousness" here I mean the fact that we have conscious experiences. It is hard to define what a conscious experience is. Nevertheless, we can try to describe it by its properties. Conscious experiences always have a phenomenal character which should be subjective or qualitative (one of both, at least). Conscious experiences have typically a content that can be accessed directly or indirectly to report, reason or act according to it. Conscious experiences are typically composed, that is, they have experiential parts. However, we live them as if they had some kind of unity. Finally, conscious experiences always occur through time. They take us time and energy.

The theories of consciousness constitute the way consciousness has been explained. There

are mainly four types of theories of consciousness: representational, informational, biological, and physical theories. Here I will focus on four specific theories, which I select because they offer a high determination on their internal mechanisms. Those theories are Integrated Information Theory, Global Workspace Theory, Self-Representational Theory and Double Microgenesis Theory.

To explain consciousness, we do need specific tools. One of these tools, maybe a neglected one, is the notion of process. Processes occur through time, have a beginning and an end. They have also a composition where parts are defined functionally (not merely spatio-temporally) and the whole can be more than the sum of its parts. There is a special kind of process that is particularly noteworthy: the self-transformative processes. Self-transformative processes are characterized by two properties: (1) a specific mutually interfering composition and (2) a loop-like modulation between the variables that describe the process.

When the process-based approach is applied to consciousness, the conclusion is that conscious experience has the aspect of a self-transformative process. We can show this by two different strategies. First, the properties of conscious experience lead to stipulate self-transformative processes as the kind of processes that underpin conscious experiences. Second, many theories of consciousness (among them, the aforementioned), if stressed by demanding a great determination on their internal mechanisms, introduce self-transformative aspects at some stage, and these aspects are not trivial, but relevant ones.

Conclusions

The conclusions of this approach are mainly two. First, we need the notion of process to explain consciousness. Second, the notion of self-transformative processes is very important, indispensable, and helps to simplify some of the big problems that surround the fact of consciousness and its specific nature.

Exordio

1.

Hace ya muchos años, en el contexto de un trabajo de fin de máster acerca del Argumento de la Habitación China, recopilamos y clasificamos las réplicas que se habían propuesto en su contra en nueve grupos. (Aunque solo sea por un momento, digo "recopilamos", en plural, para referirme a quienes empezaron este largo viaje, quienes se quedaron en el camino y quienes se incorporaron más tarde y me han acompañado hasta la meta.) Dos de estos grupos los apuntamos bajo los epígrafes de *speed does matter* y *complexity does matter*. Después, ofrecimos una nueva réplica aprovechando un juego de palabras que no era original pero sí muy efectivo: *thought does matter*. La réplica consistía en defender la idea de que es imposible que la persona que está dentro de la habitación no acabe por adquirir ciertos conocimientos de chino, y no solo conocimientos sintácticos (contra lo que Searle sostenía), sino también semánticos, a no ser que se retiren de la conversación, por así decirlo, por decreto, determinadas cuestiones metaconversacionales. Algunas de estas cuestiones son tan simples y cotidianas como las siguientes: ¿Puedes dejar de hablar durante dos segundos? ¿Puedes enviarme de nuevo el signo anterior? ¿Puedes enviarme el signo de hace dos veces y después el de hace tres?

La nueva réplica no resolvía la cuestión de fondo en uno u otro sentido, es decir, en favor de la inteligencia artificial fuerte o en favor del naturalismo biológico. Pero tenía la virtud de romper la cáscara de la nuez, es decir, de neutralizar el argumento de Searle como mecanismo definitivo para cerrar la cuestión. La respuesta, seguramente, habría de venir de la ciencia y no de la mano de experimentos mentales que, por más necesarios que resulten, no son definitivos.

Este fue el punto de partida. Sin embargo, con el tiempo, empezaron a interesarme cada vez más otras cuestiones relacionadas con la conciencia y dejé aparcado de manera provisional el problema de la inteligencia artificial. En qué consiste la conciencia, cuáles son sus propiedades, cuáles sus tipos, cómo puede emerger del cuerpo. Esta última cuestión parecía especialmente recalcitrante y elusiva. ¿Qué evidencias tenemos al respecto en nuestra pequeña experiencia cotidiana y en la amplia literatura de la psicofísica y la neurociencia? A la cabeza me viene ahora el eco de algunas de esas evidencias en la forma de nombres propios: Gall, Spurzheim, La Mettrie, Willis, Phineas Gage, Mademoiselle X, Wertheimer, Stigler, Pieron, Henry Molaison, Brenda Milner, Clive Wearing, DB, Penfield, Libet, Logothetis, Ramachandran, Duffau... "Emerger" es un término algo ambicioso. "Describir" o "entender" quizás no lo son tanto.

Una de las propiedades que describen la conciencia me resultaba especialmente inquietante: la conciencia se cambia a sí misma, y lo hace de una manera distinta a lo que sucede en otros procesos tales como el digestivo, el metabólico o el respiratorio. Es un cambiarse constante y continuo, sobre la marcha, que le confiere una sorprendente capacidad para adaptarse y una gran plasticidad. Cada sensación se parece en algo a la anterior, pero, al mismo tiempo, se vuelve algo sobre lo que conversar. Siempre es nueva materia.

No obstante, seguía estando ahí el problema mente-cuerpo.

Si no podía resolver este problema, al menos podía aprender a vivir con ello. Para ello, era imprescindible revisar las nociones de materia y propiedades. Siempre existe la tentación de entender la materia y las propiedades como cosas independientes. Parecía urgente apartar ese tipo de pensamientos de la mente y admitir que no se puede entender la materia al margen de sus propiedades ni la conciencia al margen de sus propiedades, como no se puede entender la idea de corpúsculo sin que tenga una forma u ocupe un espacio, como no se puede entender una mente sin los procesos que la sostienen, como no se puede entender un concepto sin una dinámica subyacente. Así surgió un concepto revitalizado de materia que contenía lo que normalmente se ha entendido por sustancia material y, a la vez, sus propiedades. Nada nuevo bajo el sol, pero me permitía circular de una manera novedosa y eficaz entre algunos de los problemas "clásicos". Materia y propiedades no se pueden entender por separado. Si quitas una cosa, no te queda la otra. Y quien intenta hacer eso pronto se encontrará en serios aprietos. El dualismo, por ejemplo, se podía entender como la disociación antinatural de un tipo de materia, el sistema nervioso central y un tipo de propiedades, aquellas que se corresponden con los procesos conscientes. Las soluciones reduccionistas eran la otra cara de la misma moneda disociativa, al separar artificialmente la materia de un cierto tipo de propiedades como la conciencia, la cual quedaba relegada o, en última instancia, negada. La nueva noción de materia permitía también incluir en ella los dos grandes sentidos del término: lo tangible y lo pensable, lo externo y lo interno, lo objetivo y lo subjetivo. Por un momento, tuve la tentación de llamar a todo esto "materismo", expresión que tomé prestada de una corriente pictórica. Pronto abandoné esa idea, ante el temor de sustituir una teoría por otra y un problema por otro. Mi intención no era sustituir unas teorías por otras sino precisamente lo contrario: hacer que convivan las teorías, resaltar los aspectos neutrales de las mismas y contraponerlas únicamente en los aspectos no neutrales. A menudo, a la hora de la verdad, a la hora de poner todo el dinero en el rojo o en el negro, las teorías que antes parecían completamente opuestas empiezan a mostrar parecidos razonables. Me vienen a la mente ahora las posiciones de Searle y Dennett con respecto a la posibilidad de que se creen máquinas conscientes en el futuro y también algunas de las posturas de Chalmers con respecto al epifenomenalismo.

Llamaré a esto el "principio de neutralidad mutua" y lo usaré una vez más sin nombrarlo.

2.

Esto fue solo el principio. Pero, después, la cosa se fue enredando. Lo que sucedió se puede ilustrar mediante la parábola del "teléfono etiópico". Imaginemos un plano corto donde aparece un teléfono móvil abandonado en una carretera. El teléfono suena sin parar y está lleno de polvo. Se ve también que está bajo el sol del mediodía. Uno se pregunta una serie de cosas: quién es el dueño del teléfono, quién llama y con qué intenciones, cómo es posible que haya cobertura, cuánto tiempo lleva ese teléfono abandonado sin que nadie se acerque a cogerlo. Ahora, el plano se aleja y el escenario se ve mucho mejor. En la carretera hay una larga fila de mujeres y niños caminando. Llevan a cuestas todo lo que pueden, como si huyeran. Nadie parece darse cuenta de ese teléfono. Las primeras preguntas dejan de tener interés. Pasan, por así decirlo, a un discreto segundo plano.

Creo que el efecto teléfono etiópico sucede varias veces en el curso de una tesis. Poco a poco, los problemas iniciales pierden importancia y quedan eclipsados por nuevos problemas que van surgiendo en el camino. Mi prioridad ha sido siempre saber cómo funciona la conciencia. Argumentos como la habitación china, el espectro invertido, los *qualia* danzantes, la habitación de Mary y otros de ese tipo son herramientas auxiliares y normativas, que ayudan hacia un objetivo mayor, pero que no deberían acaparar todo el discurso. Quien lea esta tesis encontrará argumentos así, pero siempre de manera auxiliar y no en la centralidad del asunto.

Mientras estudiaba acerca de la conciencia, la misma idea aparecía una y otra vez. Era la idea de que la conciencia es algo capaz de moverse y cambiarse a sí mismo. A esto lo llamé de manera vaga “automodulación”. Al principio solo era una idea modesta.

Pero era una idea.

Más tarde descubrí que no era posible entender la automodulación a través de una metafísica de la sustancia, y que, para poder avanzar, era necesario introducir en el discurso la noción de proceso.

Y ya eran dos ideas.

3.

Hablemos un momento de música. Si uno inserta en la segunda parte de *MGMT* la letra de *Urroztarra, Fuerza y Garra*, todo encaja perfectamente. La métrica es la misma. Sin embargo, que la métrica encaje no quiere decir que dos cosas sean la misma. La noción de automodulación encaja también en la métrica de la conciencia: la modulación está en la base de la biología y en la neurología, en la sinapsis y en la sincronía. Pero no basta con esa coincidencia en la métrica. Para que sea aceptada como un concepto válido para entender la conciencia tiene que haber algo más. Tiene que hacer más sencillos algunos de sus problemas y tiene que ayudar en su resolución. Esto es el meollo de la cuestión y de esto va la parte principal de la tesis. De usar esta nueva noción y de ver si, con eso,

se simplifican las cosas.

4.

A menudo, por la mañana, después del ritual de la ducha y el desayuno, siempre en el mismo instante, que es ese instante justo después de cerrar la puerta de casa y llamar el ascensor, me asalta una sensación que tira de la mano hacia atrás. Esta sensación parece vacía, carente de contenido concreto, pero me da dos datos importantes. Uno, me dice que hay algo que se me olvida. La sensación no da información acerca de qué es lo que se me olvida. Pueden ser la tarjeta de la oficina, la fiambrera con la comida, unos papeles... Pero me dice que hay algo que no marcha bien. Dos, me dice si esa cosa que se me olvida es importante. Cuanto más importante es la cosa, más intensa parece ser la sensación.

Hablando de la experiencia consciente, hay gente a la que les interesa especialmente sensaciones como ver el rojo del tomate o tener un orgasmo. A otro tipo de autores, en cambio, les interesa más aquello que sucede en los nueve meses posteriores. Thomas Nagel, Michael Tye y Frank Jackson son tres autores del primer tipo, que tienden a usar este tipo de ejemplos: sensaciones más o menos estáticas e intensas. Lo que les interesa explicar es el "efecto Coca-Cola" de la conciencia. La chispa que la hace irresistible y única.

William James y Bruce Mangan, en cambio, son autores del segundo tipo. A veces se les reconoce porque están interesados en experiencias conscientes menos paradigmáticas, más periféricas, más marginales, tales como la sensación ir tras la pista correcta, o la sensación de tener algo *en-la-punta-de-la-lengua*. Estas sensaciones también son importantes. Quizás tanto como la sensación de ver un tomate rojo o tener un orgasmo.

Más allá de las filosofías de tomates y orgasmos, más allá de las filosofías de TOKs y FOKs, esta tesis aspira a moverse entre las experiencias paradigmáticas y los difusos marcos de la experiencia consciente. Y todo ello sin olvidar la evidencia empírica.

5.

Esto no es una metáfora: esta tesis es el resultado de subir una y otra vez la misma montaña. Es una montaña que está en Navarra, y que está rodeada de un bosque compuesto por dos capas: una capa baja formada por bojes y otra capa alta y tupida formada por hayas. Dos bosques en uno.

Subiendo la montaña se aprenden muchas cosas. Se aprende, para empezar, a atender aquellas cosas que ya sabíamos y hace tiempo dejamos de atender. Se aprende que, lo que desde abajo parece una subida monótona, desde la cima no es sino el primero de los riscos. Que hace falta subir hasta lo alto para poder evaluar bien el camino, pero que hacer falta haber hecho primero el camino para poder identificarlo desde arriba. Que el camino no es una senda señalizada y balizada, sino que tiene multitud de recovecos, de rincones, de atajos y que, a veces, da vueltas en círculo.

No todas las ideas que voy a defender han nacido en esa montaña. Pero muchas de ellas se han forjado allí tal y como son. Y, sobre todo, allí he adquirido la determinación para defenderlas.

6.

La idea de subir siempre la misma montaña tiene, además, algo spinoziano: uno se conoce en relación con el sitio en el que está. Subiendo la montaña se experimentan de primera mano varias evidencias. La primera es que, al mover la vista, las imágenes del bosque se vuelven menos claras, más difusas. El procesamiento no es gratis. Cuanto más procesamiento, más energía se gasta y más oxígeno se requiere en el cerebro. En segundo lugar, dos o tres veces he percibido una pequeña raya en la zona visual derecha, oblicua, ligeramente más inclinada que la barra del teclado del ordenador. La raya era independiente del paisaje alrededor, se movía con los ojos y cesaba al parar y beber un trago de agua. Probablemente era una diminuta hipoxia en el lóbulo occipital izquierdo, quizás en el córtex estriado.

Estas dos experiencias me han mantenido siempre alerta contra todo intento de explicar la conciencia al margen de la aquello que la sustenta, esto es, el cerebro y el cuerpo donde se aloja ese cerebro. No hay atajos o pasos intermedios hacia la conciencia que no pasen por la neurociencia y el conocimiento de sus soportes físicos. Incluso aunque no determinen completamente la experiencia consciente, deben formar parte de la explicación.

7.

Voy a describir dos experiencias más que son relevantes y, en cierto aspecto, originarias. La primera es la experiencia de una esfera de luz en el exterior. Esa luz no se diferencia del interior. Dentro y fuera se funden. No hay una idea de dentro y una de fuera. Dentro y fuera son coordenadas inexistentes. No hay una distinción entre sujeto y objeto. Ser es (se reduce a) ser estrella en un océano de oscuridad. Pero esa pequeña estrella alberga un halo palpitante.

La segunda es una sensación asfixiante en la que el barro se me acumula sobre la frente y me impide respirar. El barro no pasa de manera continua, como tampoco pasa de manera continua el agua del río por su cauce. Su color marrón no tiene nada que ver con una sensación de homogeneidad. Está marcado por estrías luminosas. Es opresivo, pero también es translúcido. Es angustioso, pero, al mismo tiempo, está lleno de luz.

Estos dos recuerdos uterinos son experiencias balbuceantes de una conciencia en ciernes. Son otros dos tipos de experiencia consciente que también hay que explicar.

8.

No hay, que yo sepa, una definición estándar de la conciencia. Dewey (1967: 8) negó explícitamente que la conciencia pudiera ser definida o descrita, puesto que es precisamente aquello que se necesita en psicología para describir el resto de las cosas. Entre los autores

contemporáneos, una de las definiciones más utilizadas es la de Searle:

> By consciousness I simply mean those subjective states of awareness or sentience that begin when one wakes in the morning and continue throughout the period that one is awake until one falls into a dreamless sleep, into a coma, or dies or is otherwise, as they say, unconscious. (Searle 1999: 40)

De esta definición, confesadamente pragmática y no esencial, se puede decir lo mismo que decía Sklovski acerca de la definición del Volga como el río que pasa por Yaroslavl, Kostromá, Kineshma y Gorki: no encierra mentira o error, pero solo explica una pequeña parte de lo que refiere. No es extraño que Crick o Block prefieran pasar página rápidamente y entrar en otras cuestiones.

9.

Uno de los retroacrónimos de Gary Busey cobra, en este momento, más valor que nunca: NOW (*No Other Way*). En conciencia, NOW quiere decir que todo empieza y todo acaba. Exactamente igual que comienza y acaba un texto. Igual que empieza y acaba la sombra de una farola en el camino de vuelta a casa.

Introducción

En pocas palabras, la tesis que voy a sostener aquí es la siguiente: ser consciente es tener experiencias conscientes y tener una experiencia consciente es un hecho específico y complejo de explicar. Para entender este hecho es necesario pensarlo como un proceso: que sucede en el tiempo, que tiene partes, como cualquier proceso, y que tiene las limitaciones propias de cualquier proceso. Pero no es un proceso cualquiera, sino uno de un tipo especial: un proceso autotransformativo. Es decir, la conciencia es un proceso cuyas partes interfieren mutuamente entre sí para formar algo que es más que la mera adición de ellas.

Esta tesis la voy a justificar en tres grandes pasos. El primero consiste en describir cómo es la conciencia, es decir, identificar a qué nos referimos exactamente cuando usamos términos como "conciencia" o "experiencia consciente". La cuestión no es trivial, porque bajo el mismo término hay un hecho diverso que se manifiesta de muchas formas distintas. El segundo paso es establecer un marco conceptual para la noción de proceso. Este marco debe estar hecho a medida, para que encajen en él cierto tipo de procesos especiales como son los procesos autotransformativos Los procesos autotransformativos son muy comunes en la naturaleza. Pero no todo proceso en la naturaleza es autotransformativo. Un arco tensado es un proceso autotransformativo, esto es, sus partes (el cuerpo y la cuerda) interfieren mutuamente para formar un todo que es más que la mera adición de partes. Una cazuela llena de agua hirviendo sobre el fuego, en cambio, no lo es, al menos de manera significativa, puesto que la interferencia entre sus partes (el fuego, la cazuela, el agua) no es mutua. El tercer y último paso consiste en justificar que la experiencia consciente y los procesos autotransformativos están relacionados. Esto son realmente dos tareas distintas: justificar que las teorías de la conciencia necesitan de aspectos autotransformativos, por lo menos si se les exige un nivel de determinación alto, y justificar que las propiedades que describen la conciencia se explican mejor y de manera más sencilla si se entiende la conciencia como un proceso autotransformativo.

El capítulo 1 se ocupa del primer paso. Allí se trata de dar respuesta a dos grandes cuestiones: qué es la experiencia consciente y cómo se ha explicado la experiencia consciente. La cuestión del "qué es" deriva en la cuestión del "cómo es", esto es, en describir la conciencia a través de sus propiedades más fundamentales y típicas. Algunas propiedades son fundamentales. Por ejemplo, la experiencia consciente viene siempre acompañada de algo que es el "cómo es" estar en esa experiencia para quien la tiene (propiedad fenoménica). O, por ejemplo, la experiencia consciente acerca de algo siempre viene acompañada de cierta accesibilidad a ese algo. En otras palabras, el sujeto puede reportar, actuar o razonar

de acuerdo con ese contenido, incluso aunque el contenido no sea propiamente atendido (propiedad funcional). Otras propiedades son típicas, es decir, acompañan a la conciencia virtualmente siempre o casi siempre, que se sepa. Por ejemplo, la experiencia consciente tiene cierta composición y, a pesar de ello, siempre se presenta como dotada de cierta unidad (propiedad estructural). La experiencia consciente, además, es lenta, es decir, lleva su tiempo (propiedad temporal). Esto sugiere que es un hecho específico, que requiere un cierto tiempo de procesamiento, y que consume unos recursos que no son ilimitados.

Con respecto a la segunda cuestión, las teorías de la conciencia constituyen el modo "como se ha explicado" la experiencia consciente. Las teorías de la conciencia se pueden dividir en cuatro grandes grupos, según la clase de hechos que usan para la explicación: representacionales, informacionales, biológicas y físicas. Entre ellas, hay teorías de orientación más filosófica y otras de orientación más neurofisiológica. Por su capacidad explicativa, así como por criterios de diversidad, he escogido centrarme en cuatro teorías especialmente relevantes: la Teoría Autorrepresentacional, la Teoría del Espacio de Trabajo Global, La Teoría de la Información Integrada y la Teoría de la Doble Microgénesis.

El segundo paso en esta tesis consiste en ofrecer un marco conceptual para la articulación de la noción de proceso. La filosofía de procesos nos ofrece ese marco y de eso trata, fundamentalmente, el capítulo 2. Un proceso es "una familia organizada de ocurrencias que están sistemáticamente relacionadas entre sí, bien sea causalmente, bien sea funcionalmente".[1] Por tanto, un aspecto crucial de los procesos consiste en determinar cómo se relacionan estas ocurrencias entre sí y cómo se relacionan con el proceso del cual forman parte. En consecuencia, el marco conceptual de proceso deberá incluir una mereología, esto es, una teoría acerca del todo y las partes que defina los principios de acuerdo con los cuales los procesos se componen y se descomponen. Esta mereología, además, debe estar especialmente preparada para permitir composiciones donde el todo es algo más que la suma de las partes y en las cuales las partes interfieren entre sí.

Dentro de este marco conceptual y mereológico es posible definir un tipo especial de procesos, a saber, los procesos autotransformativos, que se caracterizan por algunos rasgos concretos. Por un lado, son complejos, sus partes interfieren entre sí y estas interferencias tienen cierta clausura, es decir, estas interferencias son mutuas. Por otro lado, esta clausura de interferencia se refleja en una clausura de modulación. "Clausura de modulación" quiere decir, básicamente, tres cosas. Una, que el proceso puede ser descrito a través de variables cuantitativa o cualitativamente determinadas. Dos, que esas variables son dinámicas, es decir, dependen del tiempo. Y tres, que el modo en el que varían en el tiempo está determinado por su propio valor en cada momento, es decir, que en cada momento la variación está determinada de una manera clausurada.

Una vez que se ha descrito cómo es la conciencia y cómo es posible entender algo a través de la noción de procesos, el tercer paso consiste en relacionar ambos conceptos. De esto trata el capítulo 3. Esta tarea consiste en responder a dos cuestiones: (1) qué razones hay para explicar la conciencia como un proceso autotransformativo y (2) en qué problemas concretos de la conciencia se puede aplicar este enfoque de procesos. Empezando por la

[1] Rescher (1996: 38)

primera pregunta, las razones para pensar que la conciencia es un proceso autotransformativo son, a su vez, dos. Por un lado, las teorías de la conciencia, cuando se les exige un nivel de determinación alto acerca de los mecanismos internos de la conciencia, acaban recurriendo a aspectos autotransformativos. Esto lo veremos en las cuatro teorías antes mencionadas. La Teoría del Espacio de Trabajo Global requiere de procesos autotransformativos para explicar la paradoja del umbral y, de hecho, cuando tiene que dar cuenta de mecanismos neurofisiológicos, la teoría echa mano de procesamientos recurrentes de larga distancia para explicar la conciencia de acceso. La Teoría de la Información Integrada, por su parte, explica la conciencia a través de sistemas dinámicos en los cuales es necesario introducir la clausura de interferencia y de modulación propias de los procesos autotransformativos. La Teoría de la Doble Microgénesis define específicamente procesos de reentrada que explican la modulación que los sistemas no específicos producen en los sistemas específicos. Por último, La Teoría Autorrepresentacional, pese a no tener un enfoque procesual, obliga a introducir aspectos autotransformativos para justificar que en la conciencia existe genuina autorrepresentación y no mera representación entre partes de los estados mentales.

La segunda razón para pensar la conciencia como un proceso autotransformativo proviene de las propiedades que describen la conciencia. Aquí, la argumentación consta de dos pasos. En primer lugar, las propiedades no fenoménicas (esto es, funcionales, estructurales y temporales) obligan a introducir el análisis de procesos por las siguientes razones:

- Los procesos están compuestos, suceden en el tiempo y sus partes tienen relaciones funcionales entre sí.
- La conciencia tiene composición, es decir, sucede en el tiempo de manera no trivial y sus partes tienen relaciones funcionales relevantes.
- Por consiguiente, la conciencia es un proceso y debe ser analizada como tal.

En segundo lugar, la explicación de la propiedad fenoménica de la intimidad requiere estipular procesos intrínsecos (en el sentido de no relacionales con el contenido primario de la experiencia consciente) los cuales, a su vez, reflejan un aspecto autotransformativo del análisis de procesos. La base de esta parte es el denominado argumento de intrinsicidad. El argumento de intrinsicidad relaciona la existencia de partes de la experiencia consciente que cumplen dos criterios (independencia y particularidad) con los aspectos autotransformativos de los procesos que dan cuenta de esa experiencia consciente. Además, como corolario, es lógico pensar que en estos procesos existe modulación cualitativa, que puede ser de tipo agónico (si es destructiva) o de tipo no agónico (si no es destructiva).

La segunda cuestión es identificar problemas donde es útil aplicar el nuevo enfoque de procesos. Las aplicaciones son numerosas, pero me centraré en la resolución de dos problemas concretos, la experiencia consciente de certeza en la percepción y la experiencia consciente de tener algo *en-la-punta-de-la-lengua*. El argumento de intrinsicidad es aplicable en ambos casos.

La experiencia de certeza en una percepción visual, esto es, la experiencia del sujeto de seguridad acerca de que está viendo una "U" (y no, por ejemplo, una "O") en una revisión oftalmológica, es independiente del contenido primario de la experiencia consciente (esto es, es independiente de que la letra que se está percibiendo sea una u otra) y, al mismo tiempo, es particular de esa experiencia (esto es, es particular de esta experiencia de ver esa letra en concreto y no otra cosa). Por tanto, tiene procesos específicos, que deben entenderse como procesos intrínsecos que son los propios de un proceso autotransformativo. En este caso, la modulación de cualidad resultará ser de tipo no agónico.

Con respecto a los estados de tener algo *en-la-punta-de-la-lengua*, en la literatura se encuentran, fundamentalmente, dos teorías explicativas: la teoría de acceso directo y la teoría heurístico-metacognitiva. Hay evidencias que favorecen a la segunda teoría con respecto a la primera. En esta teoría, la experiencia consciente de sentir que el ítem que se intenta recuperar está a punto de ser recordado tiene una función cognitiva relevante. Para explicarla es necesario aceptar que median procesos intrínsecos que no dependen del ítem que se quiere recordar pero que son particulares del intento de recordar ese ítem y no otro.

En conclusión, la noción de proceso autotransformativo es una nueva herramienta que tenemos a nuestra disposición para analizar qué es y cómo se puede entender la experiencia consciente. De manera obvia, no todo proceso autotransformativo es un proceso consciente. Quizás esta herramienta conceptual no sea suficiente para comprender completamente el hecho de la conciencia. Pero es necesaria para entender algunos de sus aspectos más fundamentales, tales como la intimidad o el papel de la experiencia en la metacognición.

Algunas cuestiones formales antes de entrar en materia. El formato de referencias y citación que he usado en este texto es APA (*American Psychological Association*). Las traducciones que aparecen en las citas son mías, salvo que se indique lo contrario, en cuyo caso, junto a la referencia de la fuente original, aparecerá la referencia de la traducción. Por otro lado, para hablar de la experiencia consciente (en inglés, *consciousness*) he optado por el término castellano "conciencia", en detrimento del término "consciencia". Lo he hecho por criterios de simplicidad y porque la ambigüedad con la noción de conciencia moral no parece problemática en este contexto. De igual manera, se emplearán expresiones como "ser consciente de" o "darse cuenta de" para referirse a la cualidad de la conciencia de ser acerca de algo que, a menudo, es exterior (esto es, aquello a lo que se refiere el término inglés "*awareness*").

Por lo que se refiere a la redacción del texto, he empleado a menudo, quizás más allá de lo prudente, los paréntesis. Este uso puede ser criticable, pero es premeditado. Se justifica doblemente. Por un lado, desde el punto de vista del autor, el paréntesis es como parar en medio de la ascensión para echar un trago de agua. Permite al caminante detenerse, tomar una bocanada de aire y retomar el camino. Por otro lado, desde el punto de vista del lector, el paréntesis permite escoger en cada punto entre esas dos opciones de lectura que tiene todo lector frente a su tarea: la lectura rápida de comprensión general y la lectura de detalle, atenta y recurrente.

Capítulo 1

Conciencia

1.1. Introducción: conciencia, del qué es al cómo es

La conciencia es un hecho elusivo. Todo el mundo sabe que es consciente, pero definir en qué consiste exactamente ese "ser consciente" no es tan fácil. Searle (1998) sostiene que la conciencia es lo ontológicamente subjetivo y epistemológicamente objetivo. Sin embargo, un partidario del fisicalismo diría que existen procesos físicos que objetivamente constituyen la conciencia, mientras que un diseñador de experimentos sabe lo importante que puede llegar a ser el reporte personal para determinar si efectivamente hay conciencia o no. Brentano (1874) y James (1931) sostienen que los estados conscientes tratan sobre algo. Sin embargo, la ansiedad y la tristeza son estados conscientes que pueden suceder sin contenido y quienes han vivido una experiencia mística insisten en que se trata de un estado sin ningún contenido al que, irónicamente, denominan "conciencia pura" (Forman 1990a). Baars (1988) sostiene que la conciencia es la atención. Sin embargo, multitud de experiencias que se califican como conscientes, como la sensación de olvidar algo al salir de casa o la sensación de estar tras la pista buena, suceden discretamente en un segundo plano de la conciencia.

La etimología no ayuda a mitigar esta elusividad. El término "conciencia" proviene del latín *conscientia* que, a su vez, se asocia con el término griego συνείδησις (Zucker 1928; Jung 1933; Chacón 2000; Thiel 2011). Las primeras veces que se emplea el término se hace en un sentido moral (Cfr. Zucker 1928). Demócrito lo utiliza para referirse a la capacidad de discernir entre lo bueno y lo malo.[2] No es hasta Crísipo y Hierocles que se le otorga un sentido biológico, como autoconciencia, familiarización o capacidad para discernir lo

[2] ἔνιοι θνητῆς φύσεως διάλυσιν οὐκ εἰδότες ἄνθρωποι, συνειδήσει δὲ τῆς ἐν τῶι βίωι κακοπραγμοσύνης, τὸν τῆς βιοτῆς χρόνον ἐν ταραχαῖς καὶ φόβοις ταλαιπωρέουσι, ψεύδεα περὶ τοῦ μετὰ τὴν τελευτὴν μυθοπλαστέοντες χρόνου. (Diels 1903: 459) "Algunos hombres, sin saber de la disolución de su naturaleza mortal, se vuelven conscientes de las desgracias de su vida y al sufrir por ello se engañan acerca de su vida inventando fábulas sobre el más allá."

propio de lo que no lo es.[3] [4]

En el sentido más contemporáneo el uso del término "conciencia" no es unívoco. Con conciencia nos referimos a diferentes cosas. A menudo nos referimos a ciertas experiencias sensoriales paradigmáticas, tales como qué es lo que se siente al ver el color rojo vivo de un objeto, al percibir un dolor agudo o al saborear la carne de una fruta. Este tipo de ejemplos son sobreabundantes en el argumentario filosófico (Locke 1825; Wittgenstein 1953; Block y Fodor 1972; Nagel 1974; Jackson 1982), quizás porque son los ejemplos que mejor reflejan la grieta explicativa que hay entre lo físico y lo fenoménico. No obstante, ahí no se agota la diversidad de las experiencias conscientes. Igualmente conscientes son otros estados tan comunes como el deseo o la emoción. Y también lo son otros más sutiles pero en absoluto anecdóticos, tales como la leve sensación de haberse dejado olvidada una cosa, el efecto de tener algo en la punta de la lengua, la experiencia ¡ajá!, el bucle fonológico, el efecto Zeigarnik, la sensación de ir por el buen camino o tras una pista falsa, la experiencia de *déjà vu*, la sensación de estar divagando, el ASMR (*autonomous sensory meridian response*), etc.

Tampoco los términos con los que se acompaña a la conciencia están libres de confusión. Se habla continuamente de estados conscientes, quizás por influencia de la noción de conciencia como estado genérico de un individuo o criatura, que diferencia entre el estado consciente propio de la vigilia (y de cuando se duerme con sueños) y el estado no consciente que sucede cuando se duerme sin sueños o cuando se pierde la conciencia. No obstante, la realidad cambiante de la conciencia dista mucho de ser algo que quede completamente descrito bajo un concepto de estaticidad, sino que más bien se parece a procesos que se suceden dentro de aquello que James (1931) llamó el "flujo de la conciencia".

Incluso la subjetividad que parece acompañar permanentemente a la conciencia es escurridiza. La noción cartesiana de yo, como sustancia portadora de propiedades ha dejado paso a una progresiva disolución de ese yo. El yo se entiende como conjunto de percepciones (Hume 1888 1739), como centro de gravitación narrativa (Dennett 1991), como perspectividad que se da desde la dimensión experiencial (Zahavi 2005) o incluso como mero ser en el ahora (Windt 2015). El "yo" ha dejado de ser el pegamento de la conciencia. Más bien, parece sugerirse lo contrario.

¿Qué se puede decir, entonces, acerca de la conciencia? En primer lugar, se pueden establecer ciertos ejes distintivos para categorizar aquello a lo que nos referimos cuando hablamos de conciencia. Kriegel (2004) recoge, al menos, cuatro ejes para establecer otros

[3] Τὴν δὲ πρώτην ὁρμήν φασι τὸ ξῷον ἴσχειν ἐπὶ τὸ τηρεῖν ἑαυτό, οἰκειούσης αὐτῷ τῆς φύσεως ἀπ' ἀρχῆς· καθά φησιν ὁ Χρύσιππος ἐν τῷ πρώτῳ περὶ Τελῶν, πρῶτον οἰκεῖον λέγων εἶναι παντὶ ξῴῳ τὴν αὑτοῦ σύστασιν καὶ τὴν ταύτης συνείδησιν (Arnim y Adler 1964: 43) "Pues los animales son dados desde el primer momento a detenerse y mirarse a sí mismos, y les es familiar su propia naturaleza desde el inicio. Lo mismo dijo Crísipo en Acerca de los Fines, que lo primeramente propio de todos los animales es su propia constitución y la conciencia que tienen de ella."

[4] Hierocles (2014: I, 38; IV, 3-60; II, 1) recoge la noción de οἰκείωσις o familiarización, que permite a un animal diferenciar entre lo que es propio (por ejemplo, su pata derecha) de lo que no lo es (por ejemplo, una posible presa).

tantos sentidos del término "conciencia". El primer eje es la distinción entre conciencia como estado biológico de una criatura y conciencia como estado concreto. La conciencia como estado biológico es el estado de "ser consciente" en general, que se predica del individuo (o la criatura) como un todo, que se tiene si está despierto o durmiendo con sueños y que no se tiene se está durmiendo sin sueños o cuando se pierde la conciencia.[5] Por otro lado, la conciencia como estado concreto es la experiencia consciente, que se vive de una manera específica y tiene una forma y contenido específicos. El segundo eje es la distinción entre conciencia transitiva y conciencia intransitiva. En la conciencia transitiva los estados conscientes tienen cierto contenido, es decir, son relacionales, son acerca de algo. En la conciencia intransitiva, por el contrario, los estados conscientes no tratan acerca de nada en especial. El tercer eje es la distinción entre conciencia sobre el exterior y autoconciencia (a veces, también llamado eje conciencia no reflexiva/conciencia reflexiva). La conciencia trata sobre lo exterior a uno mismo o, cuanto menos, sobre algo exterior a su estado consciente. La autoconciencia, sin embargo, trata sobre uno mismo o sobre el propio estado consciente. El último eje distingue entre los estados cuyo contenido está en el centro de atención, a saber, la conciencia focal, y los estados sobre contenidos que permanecen en la periferia, a saber, la conciencia periférica. Estos ejes pueden combinarse hasta cierto punto. Lo que mide el test de Glasgow es el estado biológico de conciencia de un individuo. Una sensación de ansiedad es una experiencia consciente, quizás intransitiva y normalmente periférica. Percibir el rojo de un tomate, en cambio, es una experiencia consciente, transitiva y focal que trata acerca del mundo exterior. El modo en que un corredor de fondo siente y reconoce el cansancio de su cuerpo durante la carrera es una experiencia transitiva y autoconsciente. La música que se escucha mientras uno estudia es una experiencia consciente, transitiva y periférica.

En segundo lugar, se puede tratar de describir la conciencia o, al menos, de describir cómo es en sus formas más habituales, es decir como una experiencia consciente que generalmente trata sobre algo. Windt (2015) propone una distinción fundamental entre el *aquí fenoménico* y el *ahora fenoménico*. El *aquí fenoménico* está en las formas más habituales de conciencia, esto es, en aquellas que suceden con contenido o en la periferia de cierto contenido o como emoción y que distinguen entre el yo y lo otro, es decir, conllevan una cierta perspectividad. El *ahora fenoménico*, en cambio, es la experiencia consciente sin contenido ni emoción que además no requiere de esa distinción entre el yo y lo otro. Esta diferenciación es ortogonal a la noción de conciencia temporal: la conciencia de la duración o de sucesión o de que algo se haga presente forman parte del *aquí fenoménico*. Si mi intención fuese dar una caracterización esencialista de la conciencia, sería preciso centrarme en qué es común a esas dos esas dos experiencias conscientes y lo diferencia de aquello que no es consciente. El resultado, entonces, sería que no habría analizado el hecho de la conciencia en su complejidad, sino una parte del mismo. Por otra parte, la inmensa mayoría de las evidencias empíricas a nuestra disposición son acerca de *aquí fenoménicos* y en concreto acerca de la conciencia visual. La conciencia que voy a describir va a ser de este tipo, es decir, una conciencia en torno a un contenido y que se vive desde perspectiva, y trataré de mantener un ojo en ella y otro ojo en la posibilidad de que exista también otro tipo de experiencia consciente.

[5] Aunque Thompson (2015) y Windt (2015) defienden que hay una forma de conciencia incluso durante el sueño profundo.

La manera en que se describe cómo es la conciencia propia del *aquí fenoménico* es a través de sus propiedades. A esta parte se dedica la sección 1.2. Primero, analizaré cuatro grandes grupos de propiedades. Las propiedades fenoménicas son las más conocidas en la literatura y describen aquello que es lo más propio y característico de la conciencia, esto es, el cómo es para mí estar en una experiencia consciente. Las propiedades funcionales describen qué papel tiene la conciencia para el reporte, la acción o el razonamiento. Las propiedades estructurales describen qué aspecto tiene la conciencia, si está compuesta y si tiene forma de unidad. Por último, las propiedades temporales describen cómo se crea la experiencia consciente a lo largo del tiempo y tienen una importancia crucial para explicar por qué es como es. Después trataré de mostrar cómo se han explicado esas propiedades de la conciencia a través de las teorías de la conciencia. Esa será la parte principal de la sección 1.3, donde me centraré especialmente en cuatro teorías de la conciencia, a saber, la Teoría Autorrepresentacional (Kriegel 2009), la Teoría del Espacio de Trabajo Global (Baars 1988; Dehaene et al. 2001, 2006, 2011), la Teoría de la Información Integrada (Tononi 2008; Balduzzi y Tononi 2008; Oizumi et al. 2014) y la Teoría de la Doble Microgénesis (Bachmann 2000; Bachmann y Hudetz 2014).

1.2. Cómo es la conciencia

La experiencia consciente se puede describir de diversas maneras. Se puede describir de acuerdo a "cómo es" estar en esa experiencia para mí, que es la forma más habitual. Pero también se puede describir de acuerdo a cómo es su estructura (si es compleja o simple, si se presenta como algo unificado), de acuerdo a cuál es su función (si es que la tiene) o de acuerdo a cómo se comporta a través del tiempo. A todas estas formas diferentes de describir cómo es la conciencia se las puede denominar propiedades de la conciencia.

El término "propiedad", no obstante, no es filosóficamente neutro. Muy al contrario, destila cierto aroma a metafísica de la sustancia. Por lo menos desde Aristóteles, la sustancia se ha entendido como aquello que es portador de propiedades.[6] Así, al hablar de propiedades de la conciencia, parece sugerirse que existe algo que yace debajo de todas esas propiedades y que es esencial o más fundamental que las mismas. Esta posición es lícita, pero no es la que aquí se toma. El término propiedad en relación con la conciencia ha de entenderse simplemente como lo que describe cómo es la experiencia consciente y que nos permite reducir su indeterminación. Si se insiste en ese término "propiedad" es solo por razones prácticas y porque las alternativas no siempre resultan funcionales.[7]

La historia de las propiedades de la conciencia es larga. De forma muy simplificada, se puede decir que las propiedades de la conciencia aparecen en la literatura en tres grandes momentos o generaciones. La primera propiedad en aparecer es la unidad, que tiene su origen en la propia etimología de los términos *conscientia* o συνείδησις. Las propiedades

[6]"El término filosófico 'sustancia' se corresponde con el griego *ousia*, que significa 'ser', y se transmite a través del latín *substantia*, que significa 'aquello que yace bajo las cosas o las fundamenta'. [...] Las sustancias se han tipificado como [...] los sujetos paradigmáticos de predicación y portadores de propiedades." (Robinson 2021)

[7]Zubiri (1989), por ejemplo, propone hablar de "notas de". Llevado a la conciencia, las propiedades de la conciencia serían notas de la conciencia, esto es, formas en las que la conciencia *se hace notar*.

de segunda generación surgen cuando la conciencia comienza a tratarse como un hecho en sí. Esto sucede especialmente a partir de James (1931) y su célebre listado de cinco propiedades, que incluye la unidad (tanto sincrónica como diacrónica), pero también la subjetividad, la diversidad, la relación con el exterior y la selectividad (o atención). Por último, las propiedades de tercera generación son aquellas que se descubren de manera explícita a finales del siglo XX, en buena medida gracias a los avances científicos en neurociencia y neurofisiología. Estas propiedades son quizás las más escurridizas desde el punto de vista filosófico. Entre ellas se encuentran las propiedades temporales de la conciencia.

Por motivos de claridad, para la presentación de las propiedades de la conciencia no usaré una clasificación cronológica, sino una clasificación de tipo conceptual en cuatro grandes grupos de propiedades: fenoménicas, funcionales, dinámico-estructurales y temporales. Comenzaré con las propiedades fenoménicas, ya que son aquellas que se toman, tradicionalmente, como las más características de la experiencia consciente.

1.2.1. Propiedades fenoménicas

Cada vez que miro un tomate rojo, me pincho el dedo con una aguja o tengo un orgasmo (por solo mencionar tres experiencias conscientes que se repiten una y otra vez en la literatura), junto con los procesos físicos que acompañan a cada una de estas experiencias sucede algo más. Ese algo es el hecho de cómo es para mí estar en el percibir ese tomate junto a su color y forma, en el sentir ese dolor particular o en el vivir ese placer en concreto. Ese "cómo es" estar en la experiencia consciente de percibir, sentir y vivir para mí (desde mi perspectiva, como mío) es el aspecto fenoménico de la conciencia, que Block (1995) ha llamado también la conciencia fenoménica.[8]

El aspecto fenoménico de la conciencia tiene dos formas que se han identificado como "caracteres" de la conciencia. La primera es la subjetividad, que James asociaba con una intimidad y calidez en la experiencia, similar a la que siente el sujeto al despertarse cada mañana, en el estado hipnopómpico donde "el yo llega a reconocerse a sí mismo, o a ser realmente percibido, con calidez e intimidad" (James 1931: 333). Nagel lo denominó de manera genérica el carácter subjetivo de la experiencia y afirmó que "un organismo tiene estados mentales conscientes" si y solo si se da ese *carácter subjetivo*, esto es, si y solo si "hay algo que es *cómo se es* ese organismo [...] *para* el organismo" (Nagel 1974: 436).

Sin embargo, existe cierto acuerdo en que ese carácter subjetivo no agota el aspecto fenoménico de la experiencia consciente. Husserl distingue entre dos orientaciones que se dan en toda vivencia o estado consciente del yo: la orientación objetiva que "mira" al

[8]Esta forma de referirse a las propiedades fenoménicas no deja de ser una manera ostensiva y ciertamente insatisfactoria de definición. En su defensa solo diré dos cosas: (1) que la conciencia es algo señalable puesto que es algo que todos tenemos y hemos experimentado alguna vez y (2) que, por lo menos, esta forma de definición soslaya en parte la autorreferencialidad evidente en la que se incurre cuando se emplean otras expresiones tales como el "cómo se siente" cuando se tiene una experiencia consciente.

objeto del que somos conscientes y la orientación subjetiva que "mira" al estado mental o vivencia de la cual el sujeto es consciente.[9] Crane nos recuerda que la expresión "cómo aparece algo ante ti" tienen dos sentidos, a saber, "cómo ese algo se aparece" y "cómo es para ti el hecho de que ese algo se aparezca"(Crane 2000: 185). Kriegel (2009, 2011) también expresa el mismo aspecto de la conciencia al distinguir entre dos componentes de la experiencia de contemplar un cielo azul: el componente azulado de la experiencia y el componente para-mí (*for-me-ness*) de la experiencia. Y Zahavi (2005) abunda en una distinción similar, en este caso asociando el carácter subjetivo a una especie de carácter cualitativo de la experiencia y no del objeto de la experiencia:

> Repito: la cuestión acerca del "cómo es ser" consta de dos aspectos: "cómo es el objeto para el sujeto" y "cómo es la experiencia del objeto para el sujeto". No obstante, aunque estos dos aspectos se pueden distinguir conceptualmente, no se pueden separar. No es posible desagregar alguno de estos dos aspectos de la experiencia fenoménica. No es posible encontrar uno en ausencia del otro.(Zahavi 2005: 123)

En definitiva, el aspecto fenoménico de la experiencia no tiene uno sino dos caracteres:

- carácter cualitativo: es el *cómo es* estar en una experiencia consciente particular
- carácter subjetivo: es el cómo es *para mí* estar en esa experiencia consciente.

Con respecto al estatus de esos dos caracteres cabe hacer dos preguntas. Primero, si son dos formas de entender el mismo hecho, es decir el aspecto fenoménico de la conciencia, de tal manera que no es posible desagregar uno del otro, al menos en la mayoría de las experiencias conscientes. La respuesta no es incontrovertida y, si bien algunos autores como Kriegel (2009) parecen inclinarse por esa opción, también parecen identificar la explicación de ambos caracteres como dos trabajos conceptualmente bien diferenciados.[10]

En segundo lugar, cabe preguntarse si ambos caracteres son necesarios para la existencia de una experiencia consciente y en qué medida. De nuevo, no hay acuerdo. Si solo se atiende a las teorías representacionales, ya hay posiciones divergentes.[11] Kriegel (2009) sostiene que el carácter subjetivo es el que hace que una experiencia sea consciente, mientras que el carácter cualitativo hace la experiencia consciente tal y como es, esto es, lo que conforma su "cómo es". Sin embargo, los autores representacionalistas de primer orden (Dretske 1995; Tye 1995) tienden a dar más importancia al carácter cualitativo. Para ser neutrales en este asunto, aquí se mantendrá la tesis de que la existencia de alguno de esos dos caracteres (uno de ellos o ambos) es necesaria para que pueda existir experiencia consciente.

A continuación voy a analizar algunas de las propiedades de ambos caracteres y voy a sostener que el carácter cualitativo depende en parte del objeto primario de la experiencia

[9] Cfr. González-Castán (2008: 200): "Cada vivencia intencional tiene necesariamente dos lados u orientaciones: una objetiva que 'mira' al objeto del que somos conscientes, y la otra subjetiva que 'mira' al sujeto que es consciente de ella y de su objeto."

[10] Esa es la Tesis de la División de la Labor Conceptual defendida por Kriegel (2009).

[11] Como luego veremos, las teorías representacionales son aquellas que otorgan a la noción de representación un papel relevante para la explicación de la conciencia.

consciente, si bien puede tener cierto aspecto intrínseco, en el sentido de no relacional con ese objeto primario. Sostendré también que el carácter subjetivo está relacionado con cierta perspectiva, y que no tiene una fenomenología homogénea.

1.2.1.1. Carácter cualitativo

El carácter cualitativo de una experiencia lo constituye el "cómo es" estar en una experiencia consciente, es decir, tal y como es, en toda su extensión fenoménica. A menudo, "carácter cualitativo" y *qualia* (del latín, en singular, *quale*) se han tomado como sinónimos. Sin embargo, algunos autores prefieren evitar el término "*qualia*" por su excesiva carga semántica.[12] En general, el término "*qualia*" se asocia a una experiencia privada, inefable, intrínseca, incorregible e inmediata, como la experiencia única de ver el color rojo de un tomate. Por desgracia, es complejo y muy discutible establecer en qué consisten todas y cada una de esas propiedades acerca de una experiencia consciente.[13] Por el contrario, es posible que solo algunas formas de ellas sean constitutivas del carácter cualitativo de la experiencia consciente. Por ese motivo, aquí evitaré el término "*qualia*".

El carácter cualitativo distingue la experiencia de ver un tomate rojo y la experiencia de ver una manzana verde. Ambos objetos no se perciben como la misma cosa: la experiencia de percibir cada uno de ellos es distinta. Pero el carácter cualitativo también distingue la experiencia de ver un tomate rojo ahora (que lo veo por segunda vez) de la experiencia de ver el mismo tomate rojo ayer (cuando lo vi por primera vez), ya que ambas experiencias no son exactamente la misma, pese a que sean muy parecidas. Esto apunta a un primer rasgo en el carácter cualitativo de la experiencia consciente. Por un lado, el carácter cualitativo tiene cierto componente que está relacionado con el mundo exterior, es decir, que está determinado por aquello que se percibe. Por otro lado, tiene cierto componente que no está completamente determinado por el mundo exterior, que se puede denominar intrínseco, en el sentido de no relacionado con el exterior. Ambos componentes, y no uno solo de ellos, son los que constituyen el carácter cualitativo de la experiencia consciente tal y como es.

Sin embargo, no todas las experiencias conscientes tienen contenido, en el sentido de objeto (intencional). Ese es el caso de la experiencia de ansiedad, de felicidad no orientada a nada especial o los eventos de conciencia pura propios del misticismo (*pure consciousness event*, PCE).[14] Y el carácter cualitativo está presente también en estas experiencias conscientes sin contenido alguno. Por consiguiente, la explicación del carácter cualitativo no puede reducirse únicamente a una relacionalidad o intencionalidad, al menos si no se quiere negar la existencia de este tipo de experiencias.

A continuación voy a presentar la noción de relacionalidad de la experiencia consciente, para después tratar de defender que la relacionalidad con el mundo exterior no es suficiente para determinar completamente el carácter cualitativo de la experiencia consciente.

[12] Cfr. Dennett (1988); Sebastián (2018)
[13] Ver, por ejemplo, la crítica de Dennett (1988) a la noción de *qualia*.
[14] Forman (1990a)

Para ello expondré la postura contraria, esto es, la postura de transparencia, de acuerdo a la cual la relacionalidad con el exterior determina por sí sola el carácter cualitativo de la experiencia consciente para, a continuación, exponer varias objeciones a esa postura. La conclusión es que el carácter cualitativo de la experiencia consciente con contenido está determinado en parte por su relacionalidad con el mundo exterior y en parte por cierta intrinsicidad de la experiencia consciente. La noción de intrinsicidad es ambigua y tiene una carga filosófica importante. Por ese motivo, me detendré en definir en qué sentido se habla aquí de intrinsicidad. Por último, me ocuparé de las experiencias conscientes sin contenido alguno, y trataré de defender su posibilidad y la necesidad de que ellas tengan también un carácter cualitativo.

Acerca de los estados conscientes con contenido

Supongamos un estado mental con cierto contenido. Sea consciente o no, la relación de ese estado con su contenido, esto es, su direccionalidad, su "ser acerca de" es lo que se ha denominado tradicionalmente "intencionalidad". Del mismo modo, el contenido al cual se dirige el estado mental se ha denominado "objeto intencional" o, más brevemente, "objeto". La intencionalidad, esto es, la relación entre un estado mental y su objeto (intencional), puede caracterizarse de diferentes modos. En primer lugar, puede caracterizarse de manera representacional, esto es, mediante representación (Dretske 1995) o no representacional, esto es, sin representación (Searle 1983). Si es representacional, a su vez, puede caracterizarse de distintas maneras. Dos de las más habituales son a través del acceso actual a ese contenido (Baars 1988) o a través de la disposición de ese contenido para ser accedido (Kriegel 2009).

Sea cual sea la caracterización de ese "ser acerca de" de los estados mentales, cabe preguntarse si, en los casos en que hay conciencia, la intencionalidad determina o no el carácter cualitativo del estado consciente. La respuesta a esta cuestión no es única, pero es muy común aceptar que el carácter cualitativo de una experiencia de percepción consciente está determinado, *al menos en parte*, por el objeto que se percibe, y por consiguiente, que el carácter cualitativo tiene un cierto aspecto relacional.[15] Esto se debe a que afirmar que la intencionalidad no determina de ninguna forma el carácter cualitativo nos aboca a una fenomenología plana, que es siempre la misma, o bien a una fenomenología azarosa, completamente al albur de procesos contingentes, donde la forma en que "se siente" la percepción de un tomate hoy y la forma en que "se siente" la percepción del mismo tomate ayer no tienen ningún parecido.

Aceptado, pues, que existe cierta determinación de la relacionalidad en el carácter cualitativo, la cuestión reside en establecer hasta qué punto llega esa determinación.

Desde el funcionalismo, algunos autores (Harman 1990; Tye 1995) han defendido que esta determinación es completa, es decir, que el carácter cualitativo de la experiencia consciente está completamente determinado por su relación con el objeto que es su contenido. Esta posición radical ha llevado a postular una postura de transparencia sobre la

[15] Sebastián (2018)

experiencia consciente,[16] de acuerdo a la cual toda experiencia consciente es transparente puesto que el contenido comparece ante el sujeto tal y como es, es decir, sin ninguna veladura.[17]

Frente a la postura de transparencia, no obstante, caben varias objeciones. La primera proviene de la misma fenomenología. La postura radical de transparencia implica que el carácter cualitativo de la experiencia consciente no tiene ninguna variabilidad en lo que se refiere a cómo el sujeto vive esa experiencia *desde su perspectiva*. En consecuencia, el hecho de la vivencia personal no tendría trascendencia para conformar la experiencia consciente, al margen del mero hecho de que esa experiencia sea consciente. Efectivamente, hay posiciones que pueden vivir con esta limitación,[18] pero no es una postura habitual y parece ir en contra de aquello que uno siente como individuo cada vez que tiene una experiencia consciente.

La segunda objeción proviene de la experiencia de percepción. Brentano, por ejemplo, distingue entre objeto primario y objeto secundario de la experiencia consciente. En la percepción de un sonido, el objeto primario es aquello que se oye, mientras que el objeto secundario es el "oír mismo" (Brentano 1874: 167). La idea que sostendré aquí es que ese objeto secundario, ese oír mismo, no se puede entender sino como el proceso de ese oír concreto, el proceso de esa experiencia consciente de oír. Propondré aquí un ejemplo para ilustrar este hecho. En la percepción se llama fenómeno entópico a aquel fenómeno visual cuyo origen no proviene del exterior sino de los propios ojos. Ese es el caso de los fosfenos o manchas luminosas después de frotarse los ojos. Los fosfenos son fenómenos no relacionados con el objeto que se percibe, sino con los órganos de percepción y, por tanto, que surgen en una etapa temprana del proceso de percepción. No obstante, resulta ciertamente paradójico aceptar la existencia de fenómenos intrínsecos a las etapas tempranas del proceso de percepción (como los entópicos) y, al mismo tiempo, negar completamente la posibilidad de que existan fenómenos intrínsecos en etapas más tardías del proceso de percepción, tales como aquellas en las que se forma la experiencia consciente. Esos fenómenos intrínsecos, en la medida en que forman parte del objeto de la experiencia, constituyen su objeto secundario.

La tercera objeción proviene de la posibilidad de existencia de experiencias conscientes sin contenido específico, esto es, no dirigidas a nada en particular, tales como la sensación de ansiedad, la sensación de depresión o aquellas que surgen en la práctica de algunos tipos de meditación, como la PCE o evento de conciencia pura (Forman 1990a). Si, como parece, este tipo de experiencias son posibles y si el carácter cualitativo está completa-

[16] El término "transparencia" tiene cierta ambivalencia. Literalmente, *trans-pārēre* hace referencia al hecho de aparecer (dejarse ver, manifestarse) algo a través de otra cosa, la cual no es percibida o se percibe en mucho menor grado. Por consiguiente, toda transparencia consta de dos aspectos: (a) el aparecer $algo_1$ a través de $algo_2$ y (b) el desaparecer (o disolverse) $algo_2$ en $algo_1$.

[17] Tye (1995: 31) extiende la noción de transparencia y la aplica tanto a estados mentales claramente intencionales, tales como la sensación de ver un color, como a estados mentales no tan claramente intencionales, tales como la sensación de dolor en la pierna. Para Tye, en ambos casos la experiencia consciente se presenta al sujeto de manera transparente: por mucho que el sujeto haga introspección y trate de concentrarse en su propia experiencia de percibir el color o sentir el dolor, no hay nada nuevo que se pueda desvelar al sujeto acerca de esas experiencias, puesto que todas ellas son manifiestas.

[18] Cfr. Kriegel (2009)

mente determinado por el objeto del contenido, entonces ¿cómo es posible que tengan carácter cualitativo?

La conclusión entonces es que el carácter cualitativo está determinado por la relacionalidad de la experiencia consciente con su objeto primario, pero no solo está determinado por ella. Deben existir ciertos aspectos intrínsecos de la experiencia consciente que determinen su carácter cualitativo. ¿Cuáles son estos aspectos? A esto dedicaré el punto siguiente.

La intrinsicidad de la experiencia consciente

Una vez establecido que el carácter cualitativo de la experiencia consciente tiene un componente aparte del componente relacional con el objeto primario de esa experiencia, la cuestión es: ¿en qué consiste entonces ese aspecto? A veces, se le ha llamado intrinsicidad de la experiencia consciente (Carruthers 2005). Sin embargo, la intrinsicidad es una noción que posee una alta carga filosófica, y es necesario precisar en qué sentido se usa.

Lewis (1983: 197) define la distinción entre intrínseco y extrínseco en base a cuatro nociones: adscripción, esencia, dependencia y persistencia. De acuerdo con Lewis, una propiedad p es intrínseca a una cosa c si (1) p es enteramente acerca de c y no es acerca de nada más; (2) c posee p en virtud de sí mismo y nada más; (3) p depende de c y nada más y (4) p se mantiene para todo duplicado perfecto de c. De acuerdo con Lewis, la masa de un objeto sería una propiedad intrínseca que cumple (1), (2), (3) y (4) mientras que el peso (en el sentido estar sujeto a fuerza gravitatoria), no.[19]

Este concepto de intrinsicidad no está exento de problemas. Lewis (1983) comienza apuntando una primera gran dificultad, a saber, cómo es posible identificar lo intrínseco con soledad y oponerlo a la identificación de lo extrínseco con compañía. De hecho, la propiedad de la soledad es, de por sí, una propiedad extrínseca, como lo sería la propiedad de no tener hermanos o no tener patria. Todas ellas dependen de lo demás: que no haya otros y que los otros no sean hermanos y que haya patrias y que las patrias no sean propias.

Existe, sin embargo, otra forma de entender lo intrínseco de manera más débil: como lo no relacional con respecto a algo. Ese algo puede ser diferentes cosas.[20] En nuestro caso, ese algo será lo relevante al propio proceso de percepción, a saber, el objeto primario de la experiencia consciente. Es precisamente en este sentido que se afirma aquí que el carácter cualitativo tiene un componente intrínseco.

[19]Que la masa sea una propiedad completamente intrínseca, sin embargo, dista mucho de ser completamente incontrovertible, como así lo sugiere el hecho de que en la masa participan las cuatro fuerzas del modelo estándar de partículas y la posible existencia del bosón de Higgs.

[20]Por ejemplo, Galilei (1623: ap. 48), Boyle (1666: ap. VI) y Locke (1825), distinguen entre cualidades primarias y secundarias de acuerdo a su no relacionalidad con el observador. Igualmente, Searle (1983: 27) y Searle (2002: 62), distingue entre intencionalidad intrínseca o derivada según su relación con el observador. Ojo, el uso de intrinsicidad que se hace aquí no es por la no relación al observador, sino por la no relación con el contenido de la experiencia consciente.

Ese componente puede ser muy diverso. Puede incluir la subjetividad, si es que se acepta que existe una parte cualitativa en el carácter subjetivo de la experiencia consciente.[21] Puede incluir aspectos metacognitivos de la experiencia, por ejemplo, una conciencia periférica de haber percibido algo de la misma manera antes, o una conciencia periférica de que algo no encaja en la manera en que se ha percibido ese contenido. Por último, puede incluir caracteres contingentes, tales como el efecto de una droga en la experiencia consciente.

Aunque se denominen a todas esas formas "componentes intrínsecos de la experiencia consciente", esto es, aunque sean distintos del componente relacionado con el objeto primario de la experiencia, esto no implica que no sean relacionales *de alguna forma*. En concreto, son relacionales con la propia experiencia, es decir con los procesos que dan cuenta de la dicha experiencia o con la forma en que se vive esa experiencia. La explicación de en qué puede consistir esa forma de intrinsicidad es algo complejo, y para evitar caer en regresiones o círculos viciosos es necesario tener en consideración el hecho propio de los procesos que acompañan a la experiencia consciente.[22]

Acerca de los estados conscientes sin contenido

Del mismo modo que existen experiencias conscientes con contenido, es posible que existan experiencias conscientes sin contenido alguno, es decir, que no son acerca de nada. En general, este tipo de estados conscientes no son muy habituales, no forman parte de los ejemplos paradigmáticos de experiencia consciente, y, con frecuencia, o bien no les ha prestado atención suficiente o bien son ignorados en la literatura. Sin embargo, esto no quiere decir que no existan. De hecho, hay muchos autores que defienden la existencia de tales estados conscientes. Forman, por ejemplo, defiende la existencia de experiencias conscientes de tipo místico, que él denomina los eventos de conciencia pura, y que se definen como experiencias conscientes que suceden "en vigilia" pero "en ausencia de contenido", es decir, sin "intencionalidad" (Forman 1990a: 8).

De existir, se pueden decir dos cosas acerca de dichos estados. En primer lugar, que deben tener cierto carácter cualitativo, es decir, que hay algo que es el "como es" estar en una experiencia consciente sin contenido alguno. En segundo lugar, que estos estados deben ser explicados a través de aspectos diferentes de los anteriormente mencionados y en concreto, al margen de la intencionalidad. La única opción es acudir a procesos intrínsecos de la experiencia consciente, en el sentido de diferentes del componente relacional con cualquier objeto de la experiencia consciente, sea este primario o secundario.

[21] Esto es, si se acepta que es posible una zona de solapamiento entre el carácter cualitativo y el carácter subjetivo.

[22] A esa cuestión me dedicaré en el capítulo 3.

1.2.1.2. Carácter subjetivo

Hacer la digestión, regular el nivel de calcio mediante la segregación de PTH o disparar una neurona cortical son actividades que suceden en mí sin ser vividas o sentidas como *para mí*. Esto las distingue de otras actividades tales como percibir la ventana y su paisaje, sentir una caricia en la mano o sufrir un dolor de espalda. Las segundas suceden en mí y las vivo como *para mí*.

El carácter subjetivo de una experiencia consciente es el "cómo es" estar en esa experiencia *para mí* o *desde mi perspectiva*. ¿En qué consiste ese "cómo es" estar *para mí*? James (1931) describió esa subjetividad como una cierta "intimidad". La percepción sensible, la sensación placer o dolor se experimentan con cierta intimidad. Al dormir sin sueños esa intimidad parece desvanecerse, mientras que, al despertarnos cada mañana, esa intimidad vuelve a estar junto a nosotros. Incluso aceptando el término como descripción fenomenológica, es necesario indicar que esa intimidad puede adoptar distintas formas en las diferentes experiencias y, de hecho, no todos los autores la interpretan de la misma manera. Veamos primero qué cosas hay en común en toda experiencia de intimidad en la experiencia consciente y qué se puede decir acerca de cómo es.

La perspectividad de primera persona

Lo primero que se puede decir es que la experiencia consciente de un objeto se vive como desde cierta perspectividad o con cierto escorzo.[23] Esa perspectiva no es una cualquiera, sino que es una perspectividad de primera mano: es en primera persona. Pero, al mismo tiempo, no depende completamente del contenido de la experiencia. Primero, no se reduce a la perspectiva visual que se tiene al contemplar un valle desde el pico de una montaña o una gota de agua desde el microscopio. La perspectividad de primera persona sucede igualmente en otras modalidades como la percepción háptica o auditiva y en otras formas de conciencia como las emociones o los pensamientos. Segundo, es independiente también de si la conciencia es autoconciencia o no (es decir, de si es reflexiva o no).[24] En otras palabras, esa perspectividad existe igualmente en experiencias conscientes de contenidos dispares, como leer un prospecto, contemplar un cuadro, sentir un dolor o concentrarme introspectivamente en mi sensación corporal. Tercero, es independiente también de si mi experiencia consciente es real o es ilusoria. Esto es, existe perspectiva igualmente tanto cuando se percibe un tomate que realmente está delante de nosotros como cuando suceden efectos de ilusión óptica tales como el fenómeno del retardo de destello[25] o la ceguera inducida por movimiento[26]. En otras palabras, el sujeto puede estar equivocado en aquello que cree percibir, pero no puede estar equivocado en cuanto al hecho de que lo está percibiendo.[27]

[23] Cfr. Husserl (1913); Zahavi (2005); González-Castán (2008); Kriegel (2009); Sebastián (2020).
[24] Farrell y McClelland (2017)
[25] Watanabe (2022)
[26] Ceurstemont (2022)
[27] Inspirado en una distinción de Wittgenstein, Shoemaker (1968) introduce la noción de inmunidad al error de identificación de la primera persona (IEM, *immune to error of misidentification of first-person*), que aplica a los estados conscientes.

La riqueza de la perspectividad

Existe cierto consenso acerca de la existencia de una forma de perspectividad que constituye la intimidad que acompaña a la experiencia consciente. Sin embargo, no hay tanto consenso a la hora de caracterizar en qué consiste exactamente esa forma de perspectividad, como tampoco lo hay para establecer si es necesaria o no para la existencia de la experiencia consciente.[28] Para empezar, la cuestión de la perspectividad depende de la noción de sujeto que se acepte: no es lo mismo aceptar un yo sustancial (Descartes) o una noción de yo como constructo (Dennett 1991; Zahavi 2005; Metzinger 2009).[29] Dejando a un lado este interesante debate, y centrándonos en el aspecto estrictamente fenoménico de la perspectividad, tampoco existe consenso.

Con respecto a la diversidad fenoménica de la perspectividad, caben dos posiciones: que la perspectividad se experimente de manera plana, es decir, que no haya matiz fenoménico alguno en la intimidad de la experiencia consciente, o que esa intimidad se experimente con cierta variabilidad, es decir, con cierta distinción cualitativa entre las diferentes experiencias conscientes.

Con respecto a su forma fenoménica, caben también diferentes opciones. Algunos autores la han caracterizado como una *para-mí-idad* (Nagel 1974; Kriegel 2009), otros como una *mí-idad* (Zahavi 2005).

Guillot (2017) y Farrell y McClelland (2017) distinguen tres formas principales de articular la perspectividad propia del carácter subjetivo de la experiencia consciente:

- Como la experiencia apareciendo al sujeto (*para-mí-idad*)
- Como el sujeto apareciendo ante sí mismo a través de la experiencia (*yo-idad*)
- Como el sujeto apareciendo como poseedor de la experiencia (*mí-idad*)

La estipulación de esos tres tipos de perspectividad viene dada por el hecho de que no todos los casos son reducibles entre sí ni reducibles a uno solo de ellos, así como por la necesidad de explicar fenómenos específicos de trastorno de la subjetividad (como el Síndrome de Cotard y el fenómeno del pensamiento insertado).[30] Pero también porque existe una variabilidad fenoménica entre sí. En concreto, existe una cierta gradualidad entre pasividad y actividad que va del primero (de la *para-mí-idad*) al tercero (a la *mí-idad*).

[28] Por recordar, la tesis de cautela que se adopta aquí es que alguno de los dos caracteres (cualitativo, subjetivo o ambos) es necesario para la existencia de experiencia consciente.

[29] Afecta, por ejemplo, al debate acerca de la necesidad del carácter subjetivo para la experiencia consciente. Desde el yo sustancial, por ejemplo, resulta difícil sostener estados conscientes sin carácter subjetivo.

[30] El Síndrome de Cotard (Debruyne et al. 2009) es un trastorno neuropsiquiátrico muy raro en el cual el paciente niega la existencia de su propio cuerpo, con riesgo de conducir a la inanición. El fenómeno de inserción de pensamiento sucede cuando el paciente niega que un pensamiento concreto sea suyo y afirma que alguien (una voz, por ejemplo) lo ha puesto ahí. Ambos se asocian con trastornos en el carácter subjetivo de la conciencia (Farrell y McClelland 2017).

Por último con respecto a su universalidad, es decir con respecto a la pregunta de sí toda experiencia consciente viene con alguna de esas perspectividades, caben tres posibilidades: que la perspectividad sea constitutiva de toda experiencia consciente (universalismo), que se constitutiva de la mayoría de las experiencias conscientes, con la excepción de algunos casos (tipicalismo) o que no sea constitutiva de ninguna experiencia consciente (absentismo).[31]

La posición con respecto al carácter subjetivo de la conciencia depende de la posición que se adopte acerca de esos tres respectos, pero es importante tener en cuenta que la diversidad permite combinar los diferentes aspectos para dar como resultado una riqueza de la intimidad que acompaña a la experiencia consciente. Esto es, si se acepta diversidad combinada con cierto tipicalismo aplicable a las tres formas de perspectividad, el resultado es que el carácter subjetivo tiene cierta riqueza fenoménica de la que habrá de dar cuenta.

1.2.2. Propiedades funcionales

Cada vez que percibo conscientemente desde mi ventana que está lloviendo, solo por ese mero hecho, soy capaz de hacer varias cosas: puedo avisar a alguien de que está lloviendo, puedo recordar dónde guardé el paraguas o puedo correr a recoger la ropa tendida. Si, por el contrario, percibo periféricamente y de manera difusa que algo está volando en mi habitación, quizás no pueda asegurar si es una mosca u otra cosa, pero algo puedo hacer: sé que puedo dirigir la mirada hacia cierto sitio para obtener esa información si fuese preciso. Por último, cuando aprendo a tocar el piano, la experiencia consciente de la postura mano sobre el teclado es crucial, al menos en un primer momento. Después esa conciencia empieza a entorpecer y la postura se vuelve automática, esto es, no consciente. Hasta que, por fin, un día me doy cuenta de que he adquirido un vicio de postura de la mano. Entonces, para corregirlo, tengo que ser de nuevo consciente de esa postura errónea.

Estos tres ejemplos de conciencia focal, conciencia periférica y aprendizaje mediante una macrodinámica de la conciencia son fiel reflejo de la diversidad de funciones que se asocia con la experiencia consciente: la conciencia implica cierta accesibilidad a un contenido que no siempre tiene por qué ser directa, y su existencia tiene funciones positivas o pros (para acometer algunas tareas) y otras negativas o contras (para acometer otras tareas).[32]

El ejemplo de la lluvia refleja lo que se ha llamado conciencia de acceso (Block 1995). La conciencia de acceso es la propiedad de los estados conscientes de acuerdo a la cual

[31] Farrell y McClelland (2017: 4)

[32] La funcionalidad negativa de la conciencia no es una idea nueva. En su célebre pasaje del lector absorto Plotino (1989: I, 4, 10) nos recuerda que mientras el lector está sumido en su lectura, no es consciente de cada una de las palabras que está leyendo. Si lo fuera, la lectura se volvería trabajosa y dispersa: no podría concentrarse. También aparece en Schrödinger (1958) —en los ejemplos de atarse los zapatos, apagar la luz, o desvestirse al anochecer— y de manera sistemática en Baars (1988). Hay, por supuesto, controversia acerca de la función de la experiencia consciente (Block 1995; Rosenthal 2008; Block 2011; Cohen y Dennett 2011). Por el momento aquí se toma la postura más prudente, aceptando que existen propiedades funcionales de algún tipo.

el contenido de la conciencia es directamente accesible para ser reportado, para razonar a partir de él o para actuar de acuerdo a él. La conciencia de acceso se define, de esa manera, por contraste con la conciencia fenoménica, esto es, la propiedad de los estados conscientes de tener algo que es el "cómo es" estar en una experiencia consciente para el sujeto. Aunque conciencia de acceso y conciencia fenoménica parecen aludir a dos hechos distintos, es más lógico pensar en que son dos propiedades de un mismo hecho, a saber, la conciencia.

Por su parte, el ejemplo de la conciencia periférica de la mosca refleja cierta conciencia de acceso, pero se trata de una conciencia de acceso más atenuada, donde se alerta de una cierta accesibilidad indirecta, esto es, una disponibilidad de determinado contenido (la mosca) para ser accedido mediante un mecanismo intermedio (dirigir la mirada hacia cierto lado). Es importante distinguir entre estas dos formas de acceso, puesto que tienen rasgos funcionales diferentes. Sin embargo, ambas describen cómo es la experiencia consciente desde el punto de vista funcional. A continuación voy a caracterizar estas dos formas de conciencia de acceso por separado. Para la primera me basaré en los rasgos funcionales definidos por Baars (1988). Para la segunda estableceré algunos rasgos funcionales específicos y me apoyaré en Mangan (1993).

1.2.2.1. Acceso directo

La estipulación de la accesibilidad directa como una propiedad de toda experiencia consciente se recoge en la hipótesis de acceso. Esta hipótesis establece lo siguiente acerca de cualquier estado S que pertenece a un sujeto y tiene un contenido p:

> Hipótesis de acceso: S es un estado consciente solo si su contenido p es 'directamente' accesible por el sujeto (es decir, si es utilizable sin necesidad de procesamiento adicional) para llevar a cabo un amplio repertorio de tareas cognitivas tales como reportar p, razonar de acuerdo con p o actuar en base a p. (Phillips 2018: 1)

La hipótesis de acceso dice que la accesibilidad es necesaria para la conciencia. Expresado en una forma lógica, dice que si no hay accesibilidad A, no hay conciencia C, esto es, $\neg A \rightarrow \neg C$, y por tanto, $C \rightarrow A$. Si no hay accesibilidad, no hay estado consciente, mientras que, si la hay, puede haber estado consciente o puede no haberlo (esto es, puede que también exista cierta accesibilidad para algunos estados mentales no conscientes). Por otro lado, esta hipótesis solo exige accesibilidad directa, pero no especifica si se refiere a una disposición del contenido para ser accedido directamente o bien se refiere a ser de hecho (actualmente) accedido. Por último, la hipótesis establece tres maneras de accesibilidad, a saber, reporte, razonamiento o acción de acuerdo al contenido p.

Esta hipótesis está asociada con la atención y la memoria de trabajo. Baars (1988) es quien ha definido de manera más sistemática las características funcionales propias de la conciencia cuyo contenido es directamente accesible. Estas características se pueden dividir en tres subgrupos: ineficiencia computacional, orquestación y estructura unitaria (véase cuadro 1.1).

Procesos conscientes	Procesos no conscientes
Ineficiencia computacional: abundancia de errores, baja velocidad, probabilidad de interferencia	Eficiencia computacional: especialización, escasez de errores, velocidad alta, baja interferencia entre procesos
Orquestación: gran rango de contenidos diversos en el tiempo, gran capacidad de relacionar contenidos entre sí, gran capacidad de relacionar contenidos conscientes con contextos no conscientes	Modularidad: procesos específicos de rango limitado, funcionamiento relativamente autónomo de cada proceso
Estructura unitaria: consistencia interna, serialidad, capacidad limitada	Estructura dispersa: diversidad, se procesa en paralelo, de alta capacidad

Cuadro 1.1: Comparación de propiedades funcionales entre procesos conscientes y procesos no conscientes

Existen numerosas evidencias de las características anteriores. La ineficiencia computacional, por ejemplo, refleja las limitaciones propias del procesamiento consciente.[33] Si uno trata de atender a cada una de las teclas al tocar al piano, acabará tocando más lentamente y cometerá más errores. Igualmente, si uno trata de contar los coches que pasan por una calle mientras mantiene una conversación sobre otra cosa, es fácil que acabe perdiendo la cuenta. La ineficiencia computacional de los procesos conscientes significa que la conciencia puede interferir y ralentizar los procesos cognitivos y que, por tanto, no resultaría ventajoso que todos los procesos cognitivos fueran procesos conscientes.

En efecto, no parece que la conciencia aporte gran cosa en aquellos procesos que se llevan a cabo de manera habitual y rutinaria.[34] En cambio, en lo que se refiere a procesos cognitivos novedosos e informativamente relevantes, la conciencia puede ser de gran utilidad. Así sucede en la detección de sistemas incoherentes tales como el "tridente del diablo", la resolución de ambigüedades semánticas, sintácticas o visuales, la ruptura de contextos o esquemas preestablecidos, la creación de nuevos contenidos o de relaciones nuevas entre contenidos existentes, el aprendizaje de nuevos procesos o la adaptación a contextos o escenarios novedosos. Las características funcionales de la conciencia en este tipo de casos son tres. Primero, acepta un amplio rango de contenidos. Por ejemplo, un sujeto puede ser consciente de objetos de manera modal —visual, auditiva, olfativa, etc.— o intermodal —como una unidad en los diferentes modos—, pero también puede ser consciente de conceptos abstractos (como números u operaciones), de recuerdos, de deseos o, incluso, de otras experiencias conscientes. En otras palabras, en la conciencia pueden ser potencialmente accesibles muchas formas diferentes de contenido. Segundo, a través de la conciencia es posible relacionar contenidos diversos mediante la incorporación de los diferentes contenidos conscientes en procesos cognitivos comunes. Tres ejemplos de ello son las mencionadas capacidades de reportar, razonar y actuar de acuerdo al contenido

[33]Cfr. Baars (1988: 27) para ver más ejemplos de ineficiencia computacional en eventos conscientes, en comparación con la eficiencia en el caso de eventos inconscientes, en tareas lingüísticas, aritméticas, perceptivas.

[34]Véase los ejemplos de Baars (1988: 27), así como lo de Schrödinger (1958).

consciente. Tercero, la experiencia consciente puede ser modificada por un contexto o jerarquía de procesos no conscientes cooperantes. [35] Por ejemplo, los procesos conscientes permiten al sujeto moderar su lenguaje en un contexto arisco o radicalizarlo en un contexto amigable sin necesidad de ser consciente de esa situación.

Por último, los procesos conscientes tienen una estructura funcionalmente unitaria. Por un lado, no aceptan la inconsistencia lógica. Por ejemplo, un sujeto no puede ser consciente al mismo tiempo de que alguien está y no está en la habitación. Por otro lado, hay contenidos que compiten entre sí. Por ejemplo, el sujeto no puede ver al mismo tiempo las dos perspectivas del cubo de Necker. Denominaré a esto la serialidad de la conciencia. Por último, tienen una capacidad limitada. Por ejemplo, el sujeto puede percibir de un golpe de vista la cantidad de objetos de una imagen —en un fenómeno que se denomina *subitización*—, al menos hasta cierto número que ronda los siete objetos. Esto es, se puede percibir de un golpe de vista que hay dos o tres objetos en una imagen, pero no se puede percibir de un golpe de vista que en una imagen hay once, veintitrés o doscientos cuarenta y tres objetos.

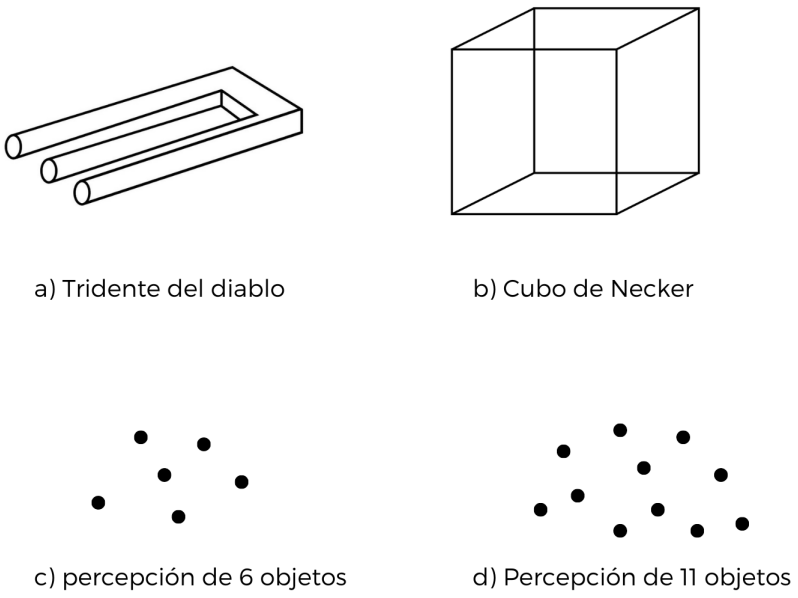


Figura 1.1: El tridente del diablo (a) y el cubo de Necker (b) son dos figuras ambiguas. Al contemplarlas se reconocen dos propiedades funcionales de la experiencia consciente visual: la consistencia interna y la serialidad. No es posible ser consciente al mismo tiempo de dos figuras incompatibles entre sí, pero es posible pasar de una interpretación a otra al cabo de unos segundos. Por otro lado, las figuras c y d muestran cómo la *subitización*, esto es, la capacidad de ver de un golpe de vista la cantidad de elementos de una imagen, está limitada fundamentalmente por el número de elementos y por la disposición de los mismos.

[35] El contexto es aquello que condiciona y limita la experiencia consciente sin ser consciente. (Baars 1988: 138).

1.2.2.2. Acceso indirecto

Las características funcionales directamente accesibles constituyen la parte más notoria de la experiencia consciente, pero ahí no se agota toda la diversidad funcional de la experiencia consciente. Por el contrario, se pueden encontrar experiencias conscientes marginales que carecen de las propiedades funcionales de acceso directo al contenido pero que, sin embargo, todavía mantienen ciertas propiedades funcionales, en el sentido de mediar para que ese acceso sea posible. Una de las experiencias de ese tipo más evidente es la sensación de "ir tras la pista buena".[36] No obstante, existen muchas sensaciones similares, como la sensación de tener algo *en-la-punta-de-la-lengua*, la sensación de haber visto fugazmente algo o la sensación de que algo no tiene sentido.

Para establecer un marco en el cual la existencia de este tipo de propiedades funcionales sea posible es necesario redefinir la hipótesis de acceso:

> Nueva versión de la hipótesis de acceso: S es una experiencia consciente solo si existe un contenido p asociado a S (1) que es directamente accesible por el sujeto para llevar a cabo un amplio repertorio de tareas cognitivas tales como reportar p, razonar de acuerdo con p o actuar en base a p o bien (2) para la accesibilidad del cual S intercede.

Las experiencias marginales de este tipo se caracterizan de la siguiente manera (Mangan 1993):

- son tenues, al menos si se comparan con los estados nucleares de la conciencia (por ejemplo, compárese la sensación de que un lugar nos resulta familiar con la experiencia de contemplarlo)[37]
- son escurridizos, en el sentido de que el sujeto no puede focalizar con facilidad la atención en ellas, al contrario de lo que sucede con los estados nucleares
- dependen del contexto, especialmente de los objetivos e inclinaciones del sujeto[38]
- relacionan contenidos de los estados de conciencia entre sí al tiempo que les otorgan una valencia positiva, por así llamarla, que permite favorecer unos contenidos sobre otros

El margen es tenue, escurridizo, contextual y relacional. Además, posee una función de mediación en la recuperación de información (*retrieval-mediation*).

En definitiva, la conciencia siempre viene acompañada de cierta accesibilidad a su contenido, que se manifiesta de dos formas. Con respecto a las experiencias conscientes acerca de un contenido, la conciencia viene acompañada de accesibilidad directa para reportar, razonar o actuar de acuerdo a ese contenido, y esa accesibilidad viene acompañada de ineficiencia computacional, capacidad para la orquestación y una estructura unitaria en

[36] Mangan (1993).
[37] James (1931) distingue entre estados nucleares y estados marginales de la experiencia consciente.
[38] Se usa en término "contexto" en el sentido de contexto recogido en Baars (1988).

el sentido de ser internamente consistente, secuencial y limitada en capacidades. Por otro lado, con respecto a las experiencias conscientes marginales, la conciencia viene acompañada de una capacidad para interceder en la accesibilidad de cierto contenido.

1.2.3. Propiedades dinámico-estructurales

La experiencia consciente de contemplar un paisaje y escuchar una sinfonía (o la de ambas cosas al mismo tiempo) se presenta ante el sujeto con dos aspectos que, aparentemente, es necesario reconciliar. Por un lado, la experiencia está dotada de una extremada riqueza, cargada de matices y de cierta composición interna, bien sea composición espacial o temporal. Al mismo tiempo, y a pesar de esa riqueza, la experiencia se vive con el aspecto de una única experiencia en el espacio y en el tiempo, muy distinta de la suma de las experiencias que tendría el sujeto si viera cada elemento del paisaje por separado o si escuchase cada nota de la sinfonía por separado. La experiencia consciente, entonces, se manifiesta ante el sujeto como dotada de cierta estructura. Tiene cierta composición y cierta unidad, y esa estructura forma parte de la experiencia consciente, es decir, forma parte del "cómo es" estar en esa experiencia concreta. El problema de reconciliar la composicionalidad y la unidad de la experiencia consciente tiene diferentes vertientes y diversos niveles de solución. En parte es un problema cognitivo, que consiste en explicar cómo todos los rasgos que forman parte del paisaje (o de la sinfonía) se integran en formas, éstas en objetos y los objetos, a su vez, en un único paisaje. A menudo a esto se le ha llamado el problema del enlace.[39] En parte, es un problema experiencial, que consiste en explicar cómo diferentes partes experienciales forman una única experiencia la cual no es la mera suma de sus partes, sino otra experiencia en sí, una experiencia con unidad. Esto se le ha llamado el problema de la unidad en la conciencia.[40]

La estructura de la conciencia es una cuestión compleja, a la cual la literatura ha dedicado mucho tiempo.[41]. En general se acepta que la conciencia viene siempre o de manera típica con cierto aspecto de unidad.[42] Sin embargo, por lo que respecta a la composicionalidad, la controversia es mayor. Hay autores que niegan o limitan a la mínima expresión la composicionalidad de la experiencia consciente (Tye 1995). No obstante, la mayoría la aceptan (Searle 1992; Koch 2004; Kriegel 2009), aunque no siempre estén completamente de acuerdo en cuanto a cómo caracterizar esa composicionalidad.

[39] *Binding problem.* Cfr. von der Malsburg (1994) y Koch (2004).

[40] *Diachronic/synchronic unity.* Cfr. Rashbrook (2013).

[41] Cfr. Revonsuo (1999); Tye (2003); Cleeremans (2003); Dainton (2000, 2018); Rashbrook (2013)

[42] Op. cit. El trastorno de cerebro escindido (Gazzaniga 2009) supone un caso especial que se ha propuesto, a veces, como contraejemplo de esta idea de unidad. En este trastorno, pacientes sometidos a callosotomía son incapaces de acometer determinadas tareas relacionadas con la contralateralidad de los sistemas perceptuales vs. la lateralidad de las áreas del habla. Por ejemplo, no son capaces de nombrar un objeto que se palpa con la mano izquierda sin verlo. El trastorno de cerebro escindido, sin embargo, no tiene por qué suponer necesariamente un contraejemplo de la unidad de la experiencia consciente. "Es posible que en los casos de cerebro escindido exista una especie de ruptura en la unidad de acceso. [...] Pero [...] una ruptura en la unidad de acceso no implica una ruptura de la unidad fenoménica. Es posible que los sujetos con cerebro escindido mantengan un campo fenoménico unificado, con algún tipo de fenomenología conjunta que subsuma cada uno de los contenidos separados." (Bayne y Chalmers 2003)

Quienes aceptan cierta composicionalidad se enfrentan al problema de explicar cómo es posible tener una unidad a partir de esa composicionalidad de la conciencia. Una estrategia para resolver un problema es dividirlo en problemas más pequeños, a sabiendas de que las soluciones podrían tener o no factores comunes. A continuación voy a seguir esa estrategia. Primero expondré y defenderé la idea de que la conciencia tiene cierta composición, y que esa composición no solo opera a nivel cognitivo, sino también a nivel experiencial, de tal manera que es necesario hablar propiamente de partes experienciales. Después expondré la idea de que la conciencia aparece de manera típica con cierta forma de unidad que es necesario explicar. El problema de explicar esta unidad se ha dividido habitualmente en dos subtipos: el problema de explicar la unidad de las partes experienciales que suceden al mismo tiempo en un instante o periodo de tiempo determinado (unidad sincrónica) y el problema de explicar la unidad de las diferentes experiencias a través del tiempo, esto es, como un único flujo de conciencia (unidad diacrónica).[43]

El resultado de aceptar estas dos propiedades estructurales de la conciencia (composicionalidad y unidad) es que surge un problema mereológico, a saber, el problema de explicar cómo es posible la experiencia de unidad más allá de la simple suma de las partes experienciales que componen la experiencia. Este hecho será una de las razones por las cuales se hace necesario un enfoque dinámico y procesual para explicar la conciencia. Por ese motivo, a las propiedades estructurales las he denominado propiedades dinámico-estructurales.

1.2.3.1. Composicionalidad

Afirmar que algo tiene composicionalidad es afirmar que está formado por partes o componentes, mientras que negar que algo tenga composicionalidad es negar que tenga ningún componente en absoluto. En el caso de los estados mentales conscientes, la cuestión de la composicionalidad se puede dividir en otras dos cuestiones, a saber, si existe una composicionalidad cognitiva en los estados mentales conscientes, y si existe una composicionalidad experiencial en la experiencia consciente que acompaña a esos estados. La negación de una composicionalidad cognitiva sería una postura radical que haría imposible explicar siquiera cómo puede ser que existan diferencias conceptuales dentro de los estados mentales. Por ese motivo, los autores que niegan que exista una composicionalidad en la experiencia consciente no lo hacen porque nieguen una composicionalidad a los estados mentales en cuanto a mentales, sino porque niegan que esa composicionalidad se deba traducir necesariamente en una composición en la manera en "cómo es" para el sujeto estar en la experiencia consciente asociada a ese estado mental.

Las posiciones con respecto a la composicionalidad de la experiencia consciente, entonces, se pueden dividir en dos grupos.[44] Por un lado, el enfoque de ausencia de partes experienciales (a veces llamado *no experiential parts view*, no-EPV) sostiene que no hay nada en la experiencia consciente que se pueda reconocer como una parte experiencial, ya que la experiencia es diversa pero completamente homogénea. Tye (1995, 2003) y su teoría PANIC de la conciencia —donde los estados conscientes son abstractos y no conceptuales—

[43] Esta es también la estrategia de muchos autores, entre ellos Searle (1992) y Rashbrook (2013).
[44] Bayne y Chalmers (2003).

es un ejemplo de una postura que defiende la ausencia de partes experienciales:

> De acuerdo con lo anterior, en realidad en nuestra conciencia habitual y normal no hay entidades tales como las experiencias puramente visuales o las experiencias puramente auditivas o las experiencias puramente olfativas. Allí donde existe unidad fenomenológica de las modalidades sensoriales, las experiencias específicas de cada sentido no existen. Solo son los "figmentos" de los filósofos y los psicólogos. Y por consiguiente, no hay algo como un problema de unidad de esas experiencias. No hay experiencias que unificar. [...] Solo hay una única experiencia multimodal, que se puede describir de manera más o menos profusa. (Tye 2003: 28)

Por otro lado, el enfoque de partes experienciales (también llamado *experiential parts view*, EPV) sostiene que la experiencia consciente en efecto está formada por partes experienciales que son manifiestas para el sujeto. A este enfoque se refiere Bayne y Chalmers cuando dicen que "en un momento dado, un sujeto posee una multiplicidad de experiencias conscientes" (Bayne y Chalmers 2003: 23).

Como sucedía con otras propiedades de la conciencia, estas dos posturas pueden ser matizadas. En concreto, el enfoque de partes experienciales puede variar entre quienes lo aplican a toda experiencia consciente (universalismo), a la gran mayoría de experiencias conscientes salvo casos especiales o clínicos (tipicalismo) o bien a ninguna (absentismo). Quienes aceptan el enfoque de partes experienciales, sea universal o típicamente, además, tienen que hacer frente a un problema adicional, a saber, el problema de explicar cómo es posible que las diferentes partes experienciales se sientan como una única experiencia. Y, sin embargo, esta postura es la más habitual y es la que se acepta aquí.[45] Las razones son las siguientes.

La primera y fundamental razón consiste en que rechazar absolutamente la existencia de partes experienciales equivale prácticamente a anular cualquier intento de hacer fenomenología de la experiencia consciente, lo que va directamente en contra de la misma sensación de experiencia consciente, esto es, de "cómo es" para el sujeto estar en esa experiencia.

En segundo lugar, aceptar la distinción entre carácter fenoménico y carácter subjetivo de la conciencia (sobre la cual ya se ha argumentado) implica aceptar que existe cierta composicionalidad en la experiencia consciente, puesto que ambos caracteres son, efectivamente, partes constituyentes de la experiencia fenoménica.

En tercer lugar, la existencia de un flujo de conciencia (James 1931) compuesto de estados marginales y estados nucleares que se solapan entre sí también implica la convivencia de partes experienciales diferentes dentro de la experiencia consciente. Si se niega que existe solapamiento de estados marginales y estados nucleares, entonces es necesario articular una explicación del flujo de la conciencia sin solapamiento de estados o bien una explicación del flujo basada en experiencia sin ningún anclaje con la intencionalidad.

[45] Searle (1992), Crane (2000), Bayne y Chalmers (2003) o Kriegel (2009) han abrazado de una manera u otra la postura de partes experienciales.

En cuarto lugar, muchos autores aceptan que vivimos el mundo tal y como es, al menos hasta cierto punto.[46] De acuerdo con estos autores, la complejidad del mundo exterior tiene un reflejo en la complejidad de la experiencia consciente: la experiencia de ver un coche "es vivida como" experiencia de coche, la experiencia de ver un gato "es vivida como" experiencia de gato. Esta es la postura habitual entre quienes aceptan que la experiencia consciente tiene cierta relación con la intencionalidad.[47] Esta posición es incompatible con la ausencia de partes experienciales en la experiencia consciente de percepción, ya que la intencionalidad puede tener diferentes contenidos al mismo tiempo. Por ejemplo, el sujeto puede ser consciente de ver un gato y escuchar un gato al mismo tiempo, y esa experiencia es distinta de la que el sujeto tendría si solo viera el gato o solo escuchase el gato.

1.2.3.2. Unidad

La experiencia consciente está formada de partes diversas que, sin embargo, se viven con un aspecto de unidad. Cuando yo observo un triángulo rojo, mi experiencia consciente se vive con cierto aspecto de unidad que es distinto de la mera adición de la experiencia de percibir un triángulo, por un lado, y la experiencia de percibir un color rojo, por otro lado. Igualmente, cuando yo escucho tres notas sucesivas, mi experiencia consciente es vivida con cierto aspecto de unidad que la distingue de la mera adición de las experiencias consistentes en escuchar cada nota por separado. La experiencia consciente se presenta ante el sujeto con cierto aspecto de unidad, y esa unidad se vive tanto en experiencias simultáneas como en experiencias sucesivas, es decir, tanto en el tiempo como *a través* del tiempo.

La aceptación de la composicionalidad de la experiencia consciente, esto es, la aceptación de que la experiencia consciente consta de partes experienciales, conduce al problema de cómo se produce la experiencia de unidad acerca de esas partes experienciales. Ese es el denominado problema de la unidad,[48] el cual se divide, a su vez, en dos problemas:[49]

- Unidad sincrónica: es el problema de explicar cómo se produce una experiencia de unidad a partir de partes experienciales simultáneas en el tiempo. Este problema se puede formular en términos de cómo es posible explicar una conciencia de unidad que es distinta de la mera suma de las partes experienciales simultáneas que la componen.
- Unidad diacrónica: es el problema de explicar cómo se produce una experiencia de unidad a partir de experiencias distintas que se suceden en el tiempo. Este problema se puede formular en términos de cómo es posible explicar una conciencia de unidad que es distinta de la mera suma de las experiencias que se suceden en el tiempo.

[46] Searle (1992) es un representante notable de esta posturas, aunque también es una postura habitual entre los representacionalistas.

[47] Un poco más arriba se ha defendido una postura acerca del carácter cualitativo de acuerdo a la cual la intencionalidad determina, al menos en gran parte, la experiencia consciente.

[48] Searle (1992); Bayne y Chalmers (2003); Rashbrook (2013).

[49] Rashbrook (2013: 466)

La distinción entre estos dos problemas no implica necesariamente que la solución a ambos no pueda tener aspectos comunes. De hecho algunos autores apuntan a la posibilidad de solapamiento entre experiencias conscientes para explicar la unidad diacrónica, de tal modo que el problema de la unidad diacrónica podría ser un tipo especial del problema de la unidad sincrónica.[50] No obstante, distinguir ambos problemas permite no tener que asumir *a priori* enfoques reductivistas y hace posible diferenciar conceptualmente cada ámbito concreto del problema.

Ahora voy a analizar las diferentes formas del problema de la unidad para cada uno de los dos tipos.

Unidad sincrónica

El problema de la unidad sincrónica de la experiencia consciente consiste en explicar cómo es posible que la experiencia consciente, a pesar de estar compuesta por partes experienciales simultáneas, se presente al sujeto con cierto aspecto de unidad. Revonsuo (1999) recomienda distinguir entre dos formas del problema que, a su vez, no tienen por qué estar completamente compartimentadas. Por un lado, el problema del enlace (*binding problem*) es un problema específico de la percepción visual que alude a cómo es posible que la integración de los distintos rasgos de grano fino de los estímulos visuales (orientación, contorno, color, intensidad, etc.) en partes y de las distintas partes en objetos durante la percepción consciente. Este problema puede tener cierto aspecto fenoménico, pero es fundamentalmente un problema cognitivo y neurofisiológico. Por otro lado, el problema de la unidad de la experiencia es el problema específico de explicar cómo las distintas partes experienciales forman una experiencia de unidad diferente de la mera adición de esas partes.

Como sucedía con la distinción entre unidad sincrónica y diacrónica, la distinción de estos problemas no significa que su solución no pueda tener cierto solapamiento, sino que únicamente significa que conceptualmente requieren un enfoque distinto. En el cuadro 1.2 se muestran algunos ejemplos de problemas de la unidad que van desde el problema del enlace hasta el problema de la unidad de la experiencia.

Este último, el problema de la unidad de la experiencia fenoménica, es genuinamente el problema de la unidad de la conciencia, y tiene algunas características específicas que es preciso destacar.

La primera característica de este problema es que tiene un aspecto mereológico, pues consiste precisamente en determinar cómo se explica que la suma de las experiencias no sea igual que la experiencia de la suma. Supongamos que un sujeto escucha cantar a un pájaro que está sobre sus manos mientras lo observa y lo acaricia. El sujeto vive esa experiencia e como siendo, al mismo tiempo, la experiencia de ver al pájaro e_1, la experiencia de escucharlo e_2 y la experiencia de tocarlo e_3. ¿Cómo se relaciona la experiencia unificada e que el sujeto tiene con respecto al pájaro con esas tres experiencias distintas e_1, e_2 y e_3? La mera suma de partes experienciales no es idéntica a la experiencia unificada de e_1, e_2

[50] Cfr. James (1931), Dainton (2000) o Thiel (2011: 216)

y e_3. En concreto, la experiencia unificada de e_1, e_2 y e_3 implica la experiencia individual de e_1, la experiencia individual de e_2 y la experiencia individual de e_3, mientras que la mera suma de esas tres experiencias no implica la experiencia unificada de ellas.[51]

La segunda característica tiene que ver con evitar la regresión que supone la unión de todas las experiencias. Se puede exponer de la siguiente manera.[52] La experiencia $e(e_1, e_2, e_3...)$, por ser una experiencia de unidad diferente de la suma de las experiencias que unifica, debe ser también unificada con el resto de experiencias, puesto que la conciencia es única. Por consiguiente, es necesario una nueva experiencia $e'(e, e_1, e_2, e_3...)$, que será diferente de las experiencias que unifica. A su vez, está experiencia necesita de una experiencia $e''(e', e, e_1, e_2, e_3...)$, y así sucesivamente.

La solución del problema de la unidad de la conciencia debe tener en cuenta ambos problemas.

Unidad diacrónica

La forma en que el sujeto vive la experiencia de escuchar un trueno de súbito no es como puro trueno, sino como *trueno-que-rompe-el-silencio-y-contrasta-con-él* (James 1931: 240). Del mismo modo, la forma en que el sujeto vive cada una de las campanadas de la iglesia no es la de una campanada aislada, sino la de una sucesión de campanadas. El problema de la unidad diacrónica es, por tanto, el problema de explicar cómo puede la experiencia tener un aspecto de unidad a través del tiempo pese a estar compuesta por diferentes partes experienciales que se suceden unas a las otras. Este problema se ha denominado también a veces el problema de la continuidad de la conciencia (James 1931: 237), si bien es necesario hacer una distinción entre dos formas de ese problema. Por un lado, el problema de la continuidad de la conciencia es el problema de explicar la continuidad de la experiencia consciente, que sucede como un hilo, es decir, sin cese ni ruptura, al menos mientras el sujeto permanece consciente o despierto. Por otro lado, a veces se ha llamado el problema de la continuidad al problema de la identidad personal, esto es, cómo es posible seguir siendo idéntico tras las interrupciones del estado consciente tales como el sueño, la anestesia total o la pérdida de conocimiento.[53] La unidad diacrónica se refiere al primer problema, esto es, al problema de explicar cómo tiene la conciencia un aspecto de unidad a través de experiencias conscientes que se suceden en el tiempo con un espacio corto de tiempo entre ellas.

De manera muy general, las soluciones a este problema son de dos tipos: los modelos

[51] Este problema ha aparecido en distintos sitios y bajo distintas apariencias. Solo algunas de ellas: la disputa entre Collins y Clarke (Uzgalis 2011) acerca de la unidad de la conciencia, la exposición de la regresión de la unidad en Tye (2003) o la noción de unidad subsuntiva en Bayne y Chalmers (2003): "Cuando [dos estados conscientes] A y B están fenoménicamente unificados, no solo ocurre que hay algo que es como tener cada estado por separado: hay que es como tener A y B juntos. Y la fenomenología de estar en A y B juntos conlleva la fenomenología de estar en A y la fenomenología de estar en B". (Bayne y Chalmers 2003: 33)

[52] Siguiendo a Tye (2003)

[53] Cfr. Thiel (2011).

Problema de unidad	Referencia	Descripción	Ejemplos
Enlace de propiedades	*property binding* (Revonsuo 1999)	Cómo se integran las diferentes propiedades de un contenido perceptual	La percepción de una persona de "las Ts de color rojo" (Treisman y Gelade 1980)
Enlace de partes	*part binding* (Revonsuo 1999) y *object unity* (Tye 2003; Bayne y Chalmers 2003)	Cómo se integran las diferentes partes de un objeto que es el contenido perceptual. Cómo se reconocen objetos, relacionado con el flujo ventral (Goodale y Milner 1992)	Cubo de Necker. La percepción de un rostro
Agrupación espacial	*spatial grouping* (Revonsuo 1999), *Gestalt unity* (Tye 2003)	Cómo se produce el agrupamiento espacial de acuerdo con los principios de la Gestalt, a saber, proximidad, similitud, simetría, continuidad, clausura y destino común.	Percibir un único gato a través del marco de la ventana
Enlace de ubicación	*spatial unity* (Tye 2003; Bayne y Chalmers 2003), *location binding* (Revonsuo 1999)	Cómo se perciben los objetos unos con respecto a los otros (de acuerdo con su posición) y con respecto a un único espacio de orden superior (egocéntrico o alocéntrico)	Experiencia egocéntrica con respecto a los objetos que están en mi habitación. Distinción entre la experiencia "estar a la derecha" y "estar a tu derecha"
Enlace semántico-conceptual	*semantic-conceptual binding* (Revonsuo 1999)	Cómo se unifican estados que pueden pertenecer al ámbito de lo consciente y de lo no consciente.	Los prejuicios, los sesgos y las referencias culturales comunes
Sentido de presencia espacial y temporal	*sense of presence* (Revonsuo 1999)	Cómo puede el sujeto siente experiencia de ser uno dentro de un único universo	Conciencia introspectiva
Unidad corpórea	*phenomenal unity* (Tye 2003)	Cómo puede el sujeto tener la experiencia de unidad corpórea de toda experiencia consciente	Sentir un dolor como algo propio y dentro de un cuerpo (familiarización).Trastorno de la Identidad de la Integridad Corporal
Unidad de la experiencia fenoménica	*phenomenal unity* (Tye 2003)	Cómo se unifican distintas experiencias que le suceden al sujeto en un momento	Un sujeto mira a un pájaro y, al mismo tiempo, escucha cómo canta y le acaricia el plumaje con los dedos.

Cuadro 1.2: En este cuadro se exponen las distintas formas en que se ha presentado el problema de unidad de la conciencia en distintos autores. Estas formas van desde el nivel cognitivo hasta el nivel experiencial, y desde el problema del enlace hasta el problema de la unidad de la experiencia fenoménica.

extensionales de solapamiento y los modelos retencionales.[54] Por un lado, los modelos extensionales de solapamiento proponen que las experiencias conscientes se suceden con cierto solapamiento entre ellas. James (1931), por ejemplo, sostiene que la experiencia consciente está formada por partes sustantivas y partes transitivas. Las partes sustantivas de la corriente de conciencia constituyen el núcleo de la conciencia. Por ejemplo, la sucesión de caras al recorrer con la mirada una fotografía de grupo constituyen las partes sustantivas de la experiencia consciente de observar la fotografía. Sin embargo, la sucesión entre dos estados nucleares de conciencia se produce con solapamiento. Este solapamiento entre estados nucleares constituye lo que se llama halo, penumbra o margen de la conciencia (*fringe*). Por consiguiente, este modelo de solapamiento asocia la continuidad a una forma especial de unidad sincrónica.

Por otro lado, los modelos retencionales proponen que la experiencia se vive en cada instante con la sombra de la experiencia anterior. Husserl (1916), por ejemplo, defiende que el presente, como en James, no se puede entender como un punto entre el pasado y el futuro, sino como una duración, pero distingue tres partes en esa duración:[55]

- La presentación originaria, que se corresponde con la fase del ahora en el objeto presentado[56]
- La retención, que se corresponde con la fase recién transcurrida del objeto
- La protención, que se corresponde con la fase del objeto que se anticipa y está a punto de ocurrir

En resumen, la experiencia consciente se manifiesta típicamente ante el sujeto con cierta estructura, es decir, como dotada de una composición, esto es, con partes y, al mismo tiempo como dotada de cierta unidad, tanto en el momento (unidad sincrónica) como en el transcurso de esos momentos (unidad diacrónica). Uno de los aspectos claves de la explicación de la conciencia consiste precisamente en comprender cómo es posible esta experiencia de unidad entre las diferentes partes de la experiencia consciente, la cual es algo más que la mera adición de dichas partes.

1.2.4. Propiedades temporales

Las propiedades fenoménicas y las propiedades estructurales son propiedades manifiestas en el sentido de que forman parte del "cómo es" estar en la experiencia consciente para el sujeto. Sin embargo, existen otras propiedades que, pese a no ser tan manifiestas, son igualmente importantes para explicar la experiencia consciente. Para determinar y establecer estas propiedades es necesario inferir su existencia a través de las evidencias empíricas o a partir de la observación de fenómenos concretos que, a veces, son muy habituales y, otras veces, no lo son tanto. Este era, en parte, el caso de las propiedades

[54] Cfr. Dainton (2018).

[55] Husserl y Serrano de Haro (2002) y también Zahavi (2005: 56).

[56] También impresión originaria (*Urimpression*) Zahavi (2005) lo traduce como *primal-impression* o *primal presentation*.

funcionales. Y es el caso, también, de las propiedades temporales.

Las propiedades temporales de la experiencia consciente son aquellas propiedades relacionadas con el hecho específico de la conciencia como proceso: la conciencia es algo que sucede en un marco espacio-temporal dado, conlleva procesos neurofisiológicos, lleva tiempo y cuesta energía. En otras palabras, no sucede *en el acto*, sino que es una *acción* que tiene una dinámica propia.[57] Hay algunos fenómenos que nos sugieren la importancia de estas propiedades. El efecto de retardo del destello, por ejemplo, es un fenómeno que sucede cuando un estímulo presentado como destello se percibe con cierto retraso con respecto a otro presentado en movimiento continuo, pese a que ambos sean simultáneos.[58] El efecto de fusión de dos estímulos en uno sucede cuando dos estímulos se presentan brevemente y de manera sucesiva. Por ejemplo, un disco rojo seguido de un disco verde, ambos presentados solo durante 10 msg, son percibidos como uno solo, a saber, como un destello amarillo.[59]

La mayoría de las evidencias que se tienen acerca de las propiedades temporales de la conciencia son de carácter empírico y provienen de la manipulación de la percepción visual a través de mecanismos de enmascaramiento.[60] Este tipo de experimentos permiten establecer dos propiedades principales desde el punto de vista temporal en la dinámica de la conciencia: (1) que la experiencia consciente es lenta y tiene un tiempo de ignición y (2) que esa lentitud es variable y puede acelerarse bajo determinadas circunstancias. A continuación voy a presentar estas dos propiedades incluyendo referencias a las evidencias empíricas que las soportan.

1.2.4.1. Lentitud y tiempo de ignición

La percepción visual, como el resto de tipos de percepción, no es inmediata. Pero ¿qué significa esto? En filosofía, la expresión "inmediato" se ha tomado al menos en tres sentidos: (1) lo inmediato como sin necesidad de ideas intermedias de la experiencia,[61] (2) lo inmediato como sin necesidad de mecanismos o procesos intermedios y (3) lo inmediato como lo instantáneo (esto es, sin transcurrir de tiempo). Se puede afirmar que, como mínimo, la percepción visual no es inmediata en los últimos dos sentidos. Contra (2), la percepción visual requiere de ciertos procesos neurofisiológicos en los que no solo forma

[57] Koch (2004: 249).

[58] Flash-lag Effect: Visual Illusion 3D (Watanabe 2022), Watanabe et al. (2010) y Murakami (2001: 3101).

[59] Efron (1973) y Koch (2004: 270).

[60] En los experimentos de enmascaramiento se presentan dos estímulos, S_1 y S_2. Uno de ellos es el estímulo objeto y el otro es el estímulo enmascarador. El estímulo enmascarador oculta al estímulo objetivo, que en su ausencia sería percibido sin problemas. El tiempo transcurrido entre el momento de inicio de S_1 y S_2 se denomina SOA (*stimulus onset asynchrony*, tiempo entre el inicio de los estímulos). Cfr. Bachmann y Francis (2014).

[61] En la Dióptrica (1637), por ejemplo, Descartes sostenía que las seis cualidades básicas de la visión (luz, color, forma, distancia, tamaño y posición) eran inmediatas con respecto a no necesitar de la experiencia, pero mediatas en cuanto a requerir de una geometría y una física innata. (Berkeley 1732), por su parte, sostenía que sólo la luz y el color se podían percibir inmediatamente, es decir, sin experiencia, mientras que el resto de cualidades necesitaban de ideas de la experiencia intermedias.

parte la retina, sino también otras partes del cerebro como el núcleo geniculado lateral talámico, el córtex estriado —área V1—, áreas de procesamiento visual superior —V2, V3, V4, MT—, partes del lóbulo temporal y, quizás, otras regiones en el lóbulo frontal y parietal.[62] Además, contra (3), la percepción visual requiere de cierto tiempo, que, en general, se ha estimado como no menor a 250 msg, de los cuales unos 150 msg se invierten en propagar la señal del estímulo hacia las áreas de procesamiento superior y al menos 100 msg en hacer accesible esa información al resto de sistemas cognitivos.[63]

Al margen de que cálculos anteriores sean más o menos precisos, hay indicios más que razonables de que los procesos asociados con la percepción visual conllevan cierta lentitud, y de que esa lentitud no es únicamente debida a la propagación de las señales propias del estímulo, sino también a la activación de procesos específicos los cuales forman parte, propiamente, de la conciencia. Las evidencias más importantes provienen de los experimentos de enmascaramiento por metacontraste.

Enmascaramiento por metacontraste

Los experimentos de enmascaramiento se basan en el efecto de interferencia en la percepción de un estímulo, llamado "estímulo objetivo", por parte de otro, llamado "estímulo enmascarador" que se presenta inmediatamente antes (en el caso de *forwarding masking*, enmascaramiento anterógrado o hacia delante) o inmediatamente después (en el caso de *backwarding masking*, enmascaramiento retrógrado o hacia atrás). El estímulo enmascarador puede ubicarse en la misma zona visual que el estímulo objetivo o en una zona adyacente. En el primer caso se habla de enmascaramiento local, o enmascaramiento a secas, mientras que en el segundo caso se habla de metacontraste.[64]

Los experimentos de metacontraste resultan especialmente interesantes para el estudio de la dinámica de la conciencia. En ellos, el enmascaramiento de un estímulo (S_1, estímulo objetivo) por parte de otro posterior (S_2, estímulo enmascarador) se produce incluso a pesar de que el segundo no se presenta en la misma zona visual del sujeto, sino en una zona adyacente. En estos casos, ese efecto de enmascaramiento se puede manipular con solo variar el tiempo entre el inicio de ambos estímulos (denominado SOA o *stimulus onset asynchrony*) y manteniendo constantes en el resto de parámetros, incluida la duración de los estímulos S_1 y S_2. Lo interesante es que, al presentarse el estímulo enmascarador en zonas adyacentes y no en la misma zona que el estímulo objetivo, el efecto no se puede explicar como una interferencia en el procesamiento relacionado con la zona visual. Esto indica que existe una actividad asociada específicamente al procesamiento consciente. Es decir, una dinámica específica de la conciencia.

[62] Crick (1994), Koch (2004), Dehaene (2014).

[63] Koch (2004: 273)

[64] Para una historia acerca de los primeros experimentos de metacontraste, véase Alpern (1953), donde se recopilan diferentes fenómenos de distorsión de la percepción de la luminosidad del primer estímulo S_1 por parte de estímulos posteriores S_2 en zonas adyacentes. Los experimentos de metacontraste más modernos se enfocan en el reconocimiento de caracteres S_1 y la interferencia de ese reconocimiento causada por el estímulo S_2. Para hacerse una idea de la complejidad de los últimos experimentos de enmascaramiento y metacontraste, véase Dehaene et al. (2001), Del Cul et al. (2007) o Bachmann y Francis (2014).

Los experimentos basados en enmascaramiento son comunes y fácilmente reproducibles, están bien medidos y metaestudiados (Bachmann 2000; Bachmann y Francis 2014). De estos metaestudios se conluye que estos experimentos no siempre dan lugar a funciones monótonas. Por el contrario, bajo determinadas condiciones dan como resultado funciones de enmascaramiento no monótonas, es decir, funciones en las cuales no se cumple siempre que a mayor SOA, menor enmascaramiento.[65] Las funciones no monótonas propias del metacontraste tienen forma de U o J. En ellas se comprueba que el pico de enmascaramiento del estímulo enmascarador S_2 sobre el estímulo objetivo S_1 (esto es, el valle de reconocimiento consciente del estímulo objetivo S_1) se produce entre los 40/50 y los 100 msg.[66]

Bachmann (2000) ofrece una interesante interpretación de los resultados de los experimentos de metacontraste. De acuerdo con Bachmann, en la percepción visual consciente concurren dos procesos de generación o microgénesis: el proceso de generación de la representación (*representational microgenesis*, RM) y el proceso de generación del percepto consciente (*phenomenal microgenesis*, PM). El proceso RM es de tipo modal y se produce en sistemas específicos, mientras que el proceso PM es de tipo amodal y se produce en sistemas no específicos.[67] Puesto que el metaconstraste está basado en un enmascaramiento en la zona visual adyacente, la interferencia propia del efecto de enmascaramiento solo puede entenderse como interferencia relacionada con los sistemas no específicos, esto es, con la actividad propia del proceso PM, y no como interferencia de los sistemas específicos propios del RM. Para SOAs muy cortos, el proceso PM del primer estímulo todavía no ha comenzado cuando comienza el proceso PM del segundo. Por consiguiente, el efecto de enmascaramiento es débil. Por otro lado, para SOAs muy largos, el proceso PM del primer estímulo prácticamente ha finalizado cuando comienza el proceso PM del segundo. En consecuencia, tampoco hay apenas efecto de enmascaramiento. Por último, para SOAs intermedios los procesos PM de ambos estímulos coinciden en el tiempo y se produce el enmascaramiento. La conclusión de Bachmann es que la duración del proceso de generación de la conciencia puede oscilar entre los 40 y los 100 msgs.

Por su parte, Del Cul et al. (2007) han medido el tiempo de procesamiento requerido para que el reconocimiento de estímulo visual sea consciente mediante enmascaramiento por metacontraste de un estímulo enmascarador (de unos 16 msg de duración) con diferentes SOAs (de 16 a 100 msg). Coincidiendo con las aproximaciones de Bachmann, los autores concluyeron que la curva de interferencia no es lineal y que la interferencia en la percepción

[65]Suceden más habitualmente "en metacontraste que en enmascaramiento de ruido/luz y cuando la energía (duración y/o luminancia) o contraste del estímulo presentado primero es mayor que la de los estímulos posteriores, o, al menos, igual a ellos". (Bachmann y Francis 2014: 6)

[66]Para obtener estos datos reflejados en la figura 1.2, Bachmann (2000: 99) echa mano de más de 138 funciones de enmascaramiento publicadas durante años, y se vale de técnicas tales como la normalización, el reescalado, la interpolación y la ponderación de funciones de acuerdo al número de mediciones.

[67] Aquí se emplea la distinción entre sistemas específicos y no específicos, habitual en los manuales de anatomía funcional (Gonzalo et al. 1993: 197). Así, por ejemplo, el núcleo geniculado lateral o el núcleo geniculado medial son partes del tálamo que, a su vez, forman parte de los sistemas específicos, puesto que están relacionados con modalidades de percepción específicas, a saber, la sensibilidad visual y auditiva, respectivamente. Otras partes del tálamo, por el contrario, no se relacionan específicamente con una modalidad concreta de percepción.

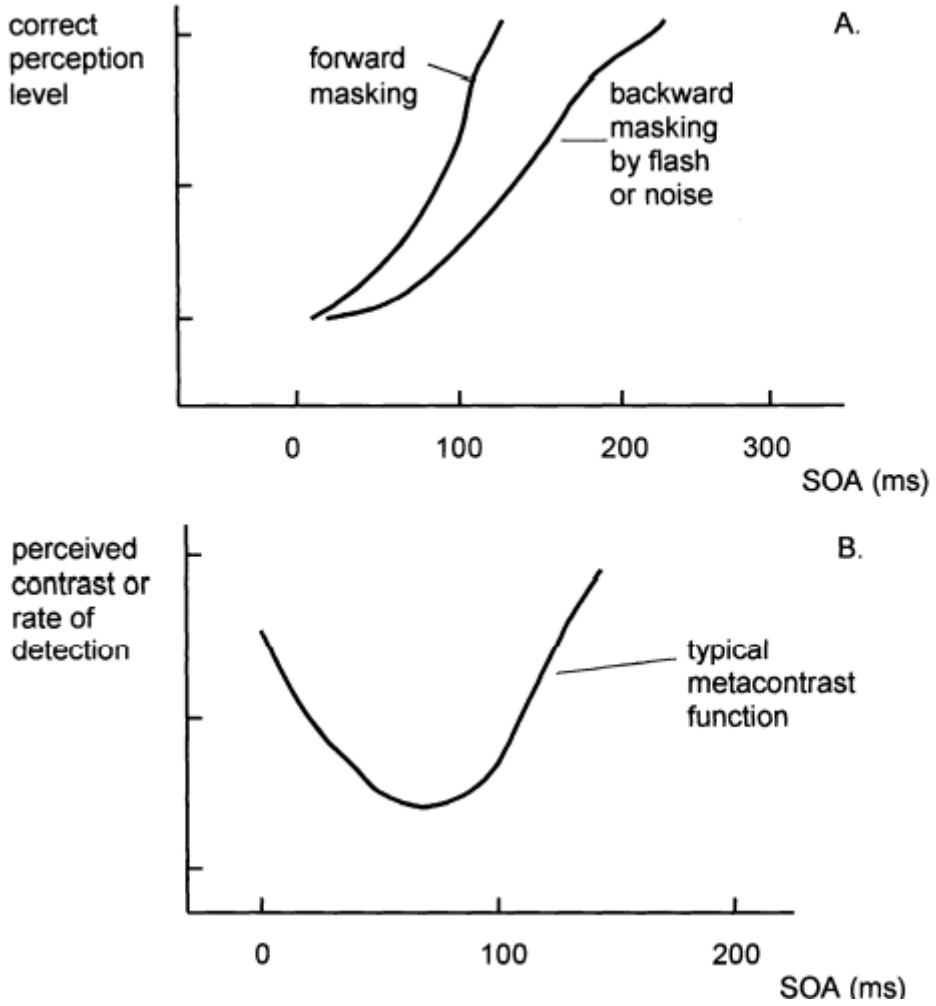


Figura 1.2: Representación de las funciones de enmascaramiento más habituales, recogido de un metaestudio de Bachmann (2000: 100). La gráfica A muestra el nivel de percepción correcta con respecto al SOA, es decir, al tiempo transcurrido entre el comienzo del estímulo objetivo y el comienzo del estímulo enmascarador (para enmascaramiento directo y retroactivo). La gráfica B muestra una función típica de enmascaramiento por metacontraste, donde la variación del nivel de percepción correcta depende del SOA de manera no monótona. Las funciones en este caso tienen forma de J o U. El efecto del enmascaramiento es máximo para SOAs entre 50 y 100 ms, y a partir de esos valores comienza a disminuir.

consciente disminuye significativamente entre los 50 y los 100 msgs.[68]

1.2.4.2. La aceleración de la conciencia

El tiempo de ignición o procesamiento puede acelerarse cuando se cumplen determinadas condiciones. Este es el caso de la aceleración mediante enmascaramiento y es el fenómeno que explica el retardo del destello. En la aceleración mediante enmascaramiento (Bachmann 2000: 111) se percibe un efecto de aceleración de la percepción consciente del estímulo enmascarador (en este caso, S_2) con respecto a la percepción consciente de ese mismo estímulo de manera aislada. Dicho con otras palabras, la percepción del estímulo parece ser facilitada cuando se produce un estímulo previo enmascarado (S_1), como si el primer estímulo S_1 preparase de antemano parte del proceso consciente para el segundo S_2. En el retardo del destello un estímulo de tipo destello se percibe con retraso frente a otro estímulo de tipo continuo (en movimiento lineal) pese al hecho de que ambos se produzcan de manera simultánea.[69] En este caso la explicación más plausible es de nuevo que los estímulos previos preparan y, por consiguiente, facilitan de algún modo la percepción del estímulo continuo, cosa que no sucede con el estímulo de destello, que ha de ser generado desde cero.

1.2.5. Conclusiones

La experiencia consciente es un hecho complejo de describir y lleno de matices. Si algo caracteriza a la experiencia consciente es su fenomenología, esto es, el "cómo es" estar en esa experiencia consciente particular para el sujeto. No obstante, la experiencia consciente tiene otros rasgos que no son tan manifiestos pero que igualmente la acompañan. No se puede explicar la conciencia tomando solo unas propiedades y dejando de lado otras. La idea de que es posible explicar una experiencia consciente de rojo al margen de sus propiedades temporales es, simplemente, dejarse demasiados pelos en la gatera. A menudo, quienes separan completamente las propiedades fenoménicas del resto de propiedades aluden a ese tipo de ejemplos de experiencia consciente, ya sea la percepción de un color, la sensación de dolor o placer. Por suerte o por desgracia, la conciencia no se agota ahí. Por el contrario, resulta difícil explicar la percepción de movimiento de un objeto sin tener en cuenta el tiempo de ignición de la conciencia, como también resulta difícil explicar la experiencia consciente de tener algo e*n-la-punta-de-la-lengua* o *ir-tras-la-pista-correcta* al margen de la función que cumple.

He descrito la conciencia a través de cuatro grandes grupos de propiedades. Las propiedades fenoménicas y estructurales son manifiestas. Las propiedades funcionales y temporales no son siempre manifiestas y, a veces, es necesario acudir a la evidencia empírica para

[68] Posteriormente analizaron mediante EPR la dinámica cerebral subyacente y obtuvieron una activación temprana occipito-temporal no correlada con el acceso consciente y una activación posterior fronto-parieto-temporal distribuida que sí estaría correlada con el acceso consciente.

[69] Flash-lag Effect: Visual Illusion 3D (Watanabe 2022). Murakami (2001: 3101), Watanabe et al. (2010). Ver también Bachmann (2000: 112, 136)

poder identificarlas. Fenoménicamente hablando, la experiencia consciente tiene dos caracteres: el carácter cualitativo, es decir, el "cómo es" estar en esa experiencia consciente *particular* y el carácter subjetivo, es decir, el "cómo es" estar en esa experiencia *para el sujeto*. Funcionalmente hablando, la experiencia consciente viene típicamente acompañada de cierto acceso a un contenido, ya sea directo, cuando el contenido de la conciencia es focal, ya sea indirecto, cuando el contenido de la conciencia es periférico o marginal. Estructuralmente, la experiencia consciente tiene típicamente cierta composicionalidad, esto es, está compuesta de partes experienciales las cuales, además, se manifiestan ante el sujeto como dotadas de cierta unidad. Temporalmente, la experiencia consciente es un hecho específico que sucede en el espacio y en el tiempo, que requiere procesos específicos diferenciados de los procesos representacionales y que tiene un tiempo de ignición que puede ser más o menos lento.

Del análisis de estas propiedades se pueden obtener dos conclusiones provisionales. La primera es que la conciencia es un hecho específico, con procesos específicos, que sucede en un tiempo concreto en unas condiciones dadas, con funciones cognitivas positivas y con funciones cognitivas negativas. A veces se ha denominado a esto como la necesidad de entender la conciencia como un hecho en bruto.[70]

La segunda es que para explicar la experiencia consciente con todas sus propiedades es necesario entenderla como un proceso dinámico y no como un estado. La experiencia consciente sucede en el tiempo, tiene aspectos intrínsecos, es algo más que la mera adición de sus partes experienciales y responde a una dinámica propia. Esto justifica el análisis de la conciencia a través de la noción de procesos que se introducirá en el siguiente capítulo.

1.3. Cómo se ha explicado la conciencia

La experiencia consciente es un hecho específico que se describe fundamentalmente por su carácter fenoménico, como aquello que es el "cómo es" estar en esa experiencia para quien la posee. La experiencia consciente también requiere de otros aspectos para ser completamente descrita, tales como los aspectos funcionales, a saber, su accesibilidad por parte de los sistemas cognitivos, los estructurales, es decir, su composicionalidad y su unidad y los temporales, esto es, cómo sucede a lo largo del tiempo. Explicar este hecho específico de la conciencia es explicar cómo es la conciencia, que se describe a través de sus propiedades. Esto es aquello de lo que se ocupan las teorías de la conciencia.

¿Qué se considera aquí una teoría de la conciencia? En la explicación científica es conocida la distinción entre *explanandum* (el fenómeno que se quiere explicar) y *explanans* (aquello que explica o da cuenta de ese fenómeno).[71] Entonces, un fenómeno A se puede

[70] Koch (2004: 7).

[71] "De acuerdo con el modelo nomológico-deductivo, una explicación científica consiste en dos 'constituyentes' básicos: un *explanandum*, que es una sentencia 'que describe el fenómeno que se va a explicar' y un *explanans*, que es 'la clase de sentencias a las que se atribuye el dar cuenta del fenómeno'." (Woodward y Ross 2021).

asociar explicativamente con otro fenómeno B de cuatro maneras distintas. En primer lugar, A puede ser el *explanandum*, es decir, aquello que se explica a través del fenómeno B. En segundo lugar, A puede ser el *explanans*, es decir, aquello que sirve para explicar el fenómeno B. En tercer lugar, A y B pueden ser ambos *explanandum*, en cuyo caso deberá existir otro fenómeno C que los explique. Por último A y B pueden ser ambos *explanans* y *explanandum* al mismo tiempo, en la medida en que estén entrelazados en una explicación mutua, la cual requiere de ambos fenómenos para ser completa.

En sentido amplio, aquí se entenderá una teoría de la conciencia como cualquier teoría que tome la experiencia consciente (descrita a través de sus propiedades fenoménicas, funcionales, estructurales y temporales) fundamentalmente como *explanandum*. Así, se excluyen dos tipos de teorías que involucran a la conciencia: (1) aquellas teorías en las que el hecho de la conciencia es parte fundamental de la teoría pero es tratado únicamente como *explanans*, y (2) aquellas teorías "de todo" en las que el hecho de la conciencia es un *explanandum* más y, por tanto, no forma parte fundamental de la teoría.[72]

De ese modo, es posible hacer ya una primera distinción entre las teorías de la conciencia. Por un lado, las teorías explicativas de la conciencia, o teorías de la conciencia en el sentido clásico, tratan de explicar las propiedades que describen la conciencia. Por otro lado, las teorías entrelazantes de la conciencia no tratan tanto de explicar la conciencia como de incluirla como un elemento más de una red explicativa, de la cual forma parte fundamental como *explanans* y *explanandum* al mismo tiempo.

En función de la clase de hechos que aparecen como *explanans*, las teorías explicativas de la conciencia se pueden dividir en cuatro grandes tipos: teorías representacionales, informacionales, biológicas y físicas. Las teorías representacionales usan la noción de representación o representación adecuada para explicar la conciencia. Las teorías informacionales, por su parte, emplean la noción de información o la difusión de la misma para explicar la conciencia, y pueden hacerlo aceptando o no la noción de representación. Las teorías biológicas se centran en procesos neurofisiológicos para dar cuenta de la conciencia. Por último, las teorías físicas se centran en procesos puramente físicos y generalmente de naturaleza cuántica. Todas estas teorías, sin embargo, también son dependientes del *explanandum* en un sentido muy relevante: cada teoría es profundamente deudora de las propiedades de la conciencia sobre las que pone el foco de atención. Dicho de otro modo, hay teorías explicativas de la conciencia mejor dotadas para explicar las propiedades fenoménicas, mientras que otras están mejor dotadas para explicar las propiedades funcionales, estructurales o temporales. En la sección siguiente mostraré en qué se diferencian estos tipos de teorías y presentaré con mayor detalle cuatro de esas teorías que resultan especialmente significativas en atención a su poder explicativo.

Las teorías entrelazantes de la conciencia, por su parte, ofrecen una nueva perspectiva acerca de la cuestión. Su mayor virtud es abrir nuevas cuestiones acerca del problema de

[72]Por ejemplo, Wigner (1961) usa la conciencia para explicar el problema de la medida, por lo que su teoría pertenece al grupo de teorías que emplean la conciencia como *explanans*. Igualmente, las teorías del campo unificado tratan de explicar la realidad y, dentro de ella, la conciencia, mediante la estipulación de un ente básico y unificador. Esta teoría pertenecería al grupo de teorías "de todo".

la conciencia. Un ejemplo de esas teorías es la teoría de las Múltiples Mentes (Lockwood 1996). Por estar en la mayoría de los casos aún en ciernes, tan solo incluiré una breve reseña acerca de ellas.

1.3.1. Tipos de teorías de la conciencia

De acuerdo con su *explanans*, las teorías explicativas de la conciencia se pueden dividir en cuatro grandes tipos: representacionales, informacionales, biológicas y físicas.[73] A continuación describiré los rasgos comunes de cada uno de estos tipos de teorías, incluyendo también un tipo de teoría no propiamente explicativa, sino entrelazante. Dado que las teorías físicas no están adecuadamente soportadas empíricamente, me centraré especialmente en cuatro teorías que son especialmente representativas de los tipos representacional, informacional y biológico, a saber, la Teoría Autorrepresentacional, la Teoría del Espacio de Trabajo Global, la Teoría de la Información Integrada y la Teoría de la Doble Microgénesis.

¿Por qué escoger estas teorías y no otras? Si revisamos la literatura a fecha de hoy, resulta difícil negar que la Teoría del Espacio de Trabajo Global y la Teoría de la Información Integrada son dos de las teorías más mencionadas y tenidas en cuenta.[74] La tercera teoría que se menciona más habitualmente es la Teoría Representacional de Orden Superior. Sin embargo, esta última teoría adolece de ciertos problemas que Kriegel (2009) ha señalado y que se pueden resumir en la idea de que no ofrece una explicación adecuada para la unidad de la experiencia consciente a partir de la diversidad de estados representacionales que estipula. Por ese motivo, en lugar de escoger la Teoría Representacional de Orden Superior, se escoge una variante, a saber, la Teoría Autorrepresentacional, la cual recoge la esencia de las teorías de nivel superior, esto es la representación de un estado por otro estado, solo que, en este caso, resulta ser el mismo. Esta teoría, además, tiene un nivel de determinación de sus mecanismos internos muy alto. Por último, y para dibujar un panorama completo de las teorías de la conciencia, se escoge una teoría biológica, a saber, la Teoría de la Doble Microgénesis. Si bien la Teoría de la Doble Microgénesis no es una de las teorías más extendidas, es, sin embargo, de entre las biológicas, una de las teorías con mayor nivel de determinación y mayor nivel explicativo.

1.3.2. Teorías representacionales

Las teorías representacionales son aquellas teorías que tratan de explicar las propiedades fenoménicas de los estados conscientes a través de su naturaleza representacional.

La representacionalidad está presente en toda la tradición epistemológica desde Descartes hasta Hume, si bien es Brentano quien ofrece una versión más moderna, asociada en este caso a los fenómenos psíquicos o mentales:

[73] Véase figura 1.3

[74] Seth (2022), Niikawa (2020), Del Pin et al. (2021). Véase también Block (2009) y Boly et al. (2013).

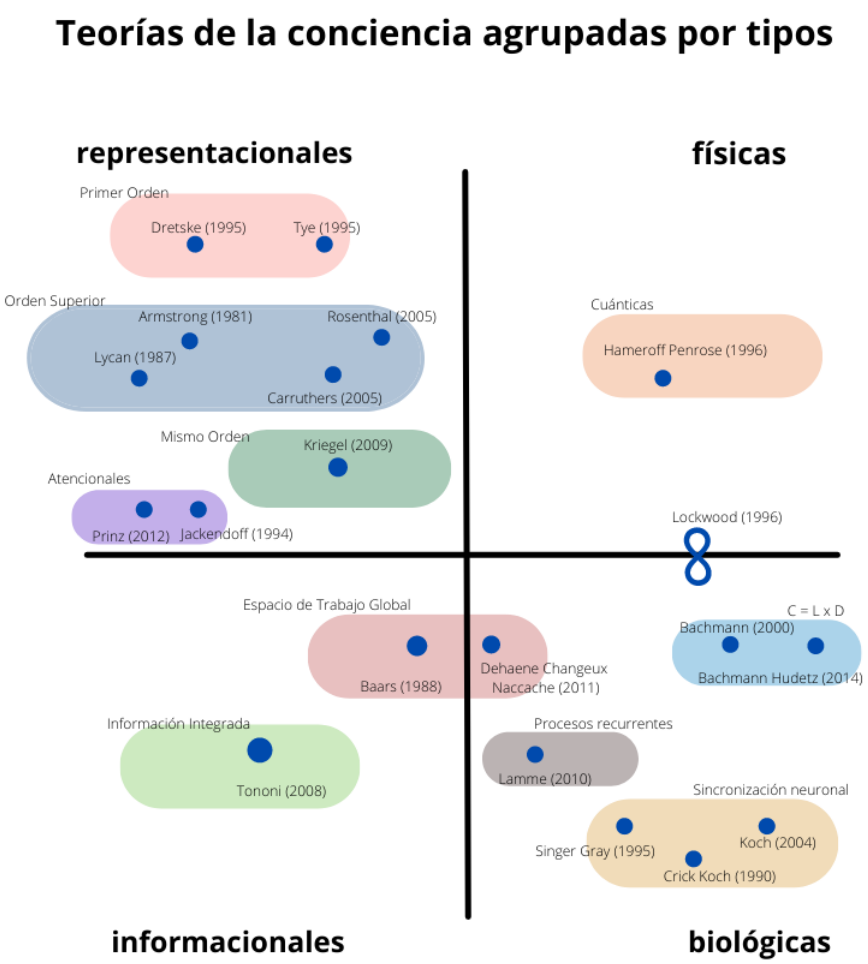


Figura 1.3: Representación de las principales teorías de la conciencia de acuerdo con su ubicación dentro de cuatro grandes tipos de teorías: representacionales, informacionales, biológicas y físicas.

> ...designábamos con el nombre de fenómenos psíquicos, tanto a las representaciones como a todos aquellos fenómenos cuyo fundamento está formado por representaciones. Apenas necesitamos advertir que una vez más entendemos por representación, no lo representado, sino el acto de representarlo. (Brentano 1874: 104/libro II, cap 1, §3) (Benítez 2005)

La base de esta noción "moderna" de representación en Brentano se resume en una frase: la representación no es copia sino acto, acto de representar. Esta idea permanece también en Dretske, quien define de este modo la tesis que subyace a las teorías representacionales:

> [La] Tesis Representacional [...] consiste (con alguna excepción menor) en que (1) todos los hechos mentales son hechos representacionales y (2) todos los hechos representacionales son hechos acerca de funciones informacionales. [...]
> La Tesis Representacional es suficientemente plausible para actitudes proposicionales tales como creencias, pensamientos, juicios, etc. [...] Es menos plausible para hechos sensoriales y para los aspectos cualitativos o fenoménicos de nuestra vida mental. Sin embargo, [...] me concentraré en la experiencia perceptual. [...]
> Además, hay otras experiencias, como, por ejemplo, la sensación general de depresión, acerca de las cuales no sé muy bien qué decir. Ese es el motivo de incluir la expresión "con alguna excepción menor" en la formulación de la Tesis Representacional. (Dretske 1995: xii-xv)

La Tesis Representacional es la base de las teorías representacionales de la conciencia, puesto que relaciona de manera fundamental los estados representacionales y los estados mentales. Sin embargo, no relacionan de manera fundamental los estados representacionales y los estados conscientes. Y no todo estado representacional tiene por qué ser consciente. Por ese motivo, las teorías representacionales de la conciencia imponen restricciones adicionales a la representación tales que, si se cumplen, habrá conciencia y, si no se cumplen, no la habrá. Estas restricciones pueden ser de tipo representacional o de tipo atencional. Dos ejemplos de restricciones de tipo atencional son la teoría de la mente computacional de Jackendoff (1994) o la teoría AIR de Prinz (2012).[75] Las teorías con restricciones puramente representacionales se dividen en las teorías de primer orden, teorías de orden superior y teorías de mismo orden.

Las teorías de primer orden establecen que los estados conscientes son estados representacionales sobre objetos y propiedades de los objetos. Para que estos estados representacionales sean conscientes, deben ser de tipo natural y determinado de acuerdo a su relación con el contexto (Dretske 1995), o bien deben estar preparados para la producción de deseos y creencias, y ser abstractos, no conceptuales e intencionales (Tye 1995).

En las teorías de primer orden el sujeto no aparece o bien lo hace de manera derivada (desplazada, se dice), a saber, como un objeto secundario de los estados representacionales.[76] Sin embargo, todo aspecto fenoménico de la conciencia parece tener adherido

[75] Representaciones de nivel intermedio atendidas (*Attended Intermediate-level Representations*, AIR).

[76] Como "percepción desplazada" (Dretske 1995), es decir, como algo que aparece únicamente a lo largo de la representación de un contenido y de una manera secundaria.

cierto carácter subjetivo: el sujeto, en situación normal, siente los estados conscientes como propios. Por su parte, las teorías de orden superior surgen con la intención de explicar el carácter subjetivo de la experiencia consciente y, para ello, añaden el requisito de un orden jerárquico a los estados representacionales conscientes. En estas teorías, los estados conscientes son aquellos estados representacionales que se dirigen a otros estados representacionales.[77] Estas teorías se dividen entre las que entienden esta representación superior como percepción introspectiva (Armstrong 1981; Lycan 1987) y las que entienden esta representación superior como dotada de cierta conceptualidad, bien sea en acto (Rosenthal 2005), bien sea como disposición para ser usada (Carruthers 2005).[78]

La teoría de Rosenthal (2005) es una de las expresiones más claras de las teorías de orden superior. Muy resumidamente se puede caracterizar en tres puntos:

- Principio de representacionalidad: todo estado mental es un estado intencional, esto es, que es acerca de algo, o bien tiene propiedades fenoménicas.[79]
- Principio de transitividad: un estado mental es consciente solo si hay un sujeto que es consciente de ese estado.[80].
- Noción de conceptualidad: esa relación entre el sujeto y su estado consciente tiene cierto carácter conceptual, esto es, tiene una manera de representar fundada en el pensamiento en oposición a lo sensorial o perceptual.[81]

Las teorías de orden superior —como la de Rosenthal— ponen al sujeto en el centro de la conciencia y, de ese modo, dan una explicación al carácter subjetivo de la experiencia

[77]No es ésta una postura original. Descartes, por ejemplo, dice en la correspondencia con Burman: *Conscium esse est quidem cogitare et reflectere supra suam cogitationem* (Adam y Tannery 1903: 283). Leibniz, por su parte, afirma lo siguiente: *conscientia est reflexio in actionem, seu memoria actionis nostrae, ita ut cogitemus, nostram esse. Involvit hoc ipsam substantiam veram seu* τὸ εγο (Couturat 1903: 495). Nota: cfr. Thiel (2011) y especialmente Hennig (2006) para algunos problemas acerca de la noción de conciencia en Descartes.

[78]Dretske (1995), Gennaro (2004a) y Carruthers (2005) tienen su propia versión de esta clasificación. A menudo, las teorías de orden superior basadas en la percepción se denominan también teorías HOP, (*higher-order perception theories*), mientras que las basadas en conceptualidad se denominan teorías HOT, (*higher-order thought theories*). Las teorías de orden superior se denominan genéricamente HOR (*higher-order representationalism*), en oposición a las teorías de primer orden o teorías FOR (*first-order representationalism*) y a las teorías de mismo orden o teorías SOR (*same-order representationalism*).

[79]"Todo estado mental, sea cual sea, exhibe propiedades de uno de estos dos tipos: propiedades intencionales y propiedades fenoménicas o sensoriales. Algo tiene una propiedad intencional si tiene un contenido proposicional o es acerca de algo. Las propiedades sensoriales, por contra, son menos homogéneas. [...] Algunos estados mentales pueden tener ambos tipos de propiedades. Pero sea cual sea la verdad de los estados mentales, es claro que no se puede considerar un estado como mental a no ser que tenga alguna propiedad intencional o alguna propiedad fenoménica." (Rosenthal 2005: 23). Que Rosenthal incluya las propiedades fenoménicas junto a las intencionales en su versión del principio de representacionalidad no supone ninguna novedad. Recuérdese la expresión "con alguna excepción menor" (*plus minus a bit*) en la formulación original de Dretske (1995).

[80]"Llamaré principio de transitividad al hecho de que los estados mentales solo son conscientes si uno es consciente de ellos de alguno modo" (Rosenthal 2005: 4)

[81]"Como no sentimos ni percibimos nuestros estados conscientes, solo es posible que seamos conscientes de ellos teniendo pensamientos acerca de ellos. Y como esos pensamientos son acerca de otros estados mentales, les podemos llamar pensamientos de nivel superior" (*higher-order thoughts*, HOTs). (Rosenthal 2005: 5).

consciente.[82] Sin embargo, aunque la estipulación de un nuevo estado de orden superior resuelve algunos problemas, lo cierto es que deja otros abiertos (Kriegel 2009). Por un lado, no se acaba de aclarar por qué esa representación de nivel superior es suficiente para explicar la conciencia. Si una representación de primer orden no garantiza que haya conciencia, ¿por qué lo iba a hacer una representación de segundo nivel? Por otro lado, queda abierta la cuestión de si el nuevo estado de orden superior es, a su vez, consciente. Si se defiende que el sujeto nunca es consciente de ese estado de orden superior, entonces ¿cómo es posible que el sujeto se percate de ese estado? Si, por el contrario, el estado superior es siempre consciente, ¿acaso no aparece una regresión infinita de estados que explican la conciencia de los otros?[83] Por último, si se defiende que el estado de orden superior solo es consciente hasta cierto nivel, a lo sumo, hasta el tercer nivel (Rosenthal 2005: 9), entonces, ¿cómo es posible tener una experiencia consciente unificada a partir de más de un estado mental de los cuales, a su vez, el sujeto es consciente?

La necesidad de dar solución a los problemas anteriores nos conduce a una nueva forma de teorías representacionales, a saber, las teorías del mismo orden. Para ellas, los estados representacionales conscientes son aquellos que tienen una representación adecuada. Y la representación adecuada para un estado es la capacidad de autorrepresentarse o representarse a sí mismo. La Teoría Autorrepresentacional de Kriegel (2009) es el conspicuo ejemplo de esas teorías.

1.3.2.1. Teoría autorrepresentacional

La Teoría Autorrepresentacional (Kriegel 2009) es deudora de la tradición que Brentano inició más de un siglo antes. Esta teoría parte de la idea de que aquello que hace consciente a un estado mental es su carácter subjetivo, que es el hecho de que esos estados son vividos como *para mí*. Kriegel sostiene que para que un estado mental tenga ese carácter subjetivo es necesario que tenga un tipo adecuado de representación, y sostiene que ese tipo adecuado de representación solo puede ser la autorrepresentación. El núcleo de esta teoría es el denominado argumento maestro, que concluye la necesidad de que un estado sea autorrepresentado para que tenga carácter subjetivo y, en consecuencia, para ser consciente.

De manera resumida, el argumento maestro consta de los siguientes pasos:[84]

1. Los estados que tienen carácter subjetivo son estados que se presentan para el sujeto, y, por consiguiente, deben ser estados representados. En esquema:
   Para cualquier estado mental M y cualquier sujeto S, M tiene carácter subjetivo

[82] Involucra al yo (*self-involving*), y no al estado mental (*mental state-involving*). (Sebastián 2012)

[83] Aristóteles ya era conocedor de este problema, y dice en *Acerca del alma*: "Dado que sentimos que vemos y oímos, el acto de ver habrá de sentirse forzosamente o con la vista o con algún otro sentido. Ahora bien, en este último supuesto el mismo sentido captaría la vista y el color, objeto de esta. Luego habrá dos sentidos que capten el mismo objeto a no ser que el mismo sentido se capte a sí mismo. Pero es que, además, si fuera otro el sentido encargado de captar la vista o bien habrá una serie infinita o bien habrá, el último término, algún sentido que se capte a sí mismo." (Aristóteles 1988: 217–218; III,2,10-19)

[84] Para una presentación completa, ver Kriegel (2009: 156).

solo si hay un estado mental $M*$ tal que (i) S está en $M*$ y (ii) $M*$ representa a M.

2. Los estados que tienen carácter subjetivo son estados representados, pero no de cualquier manera, sino de una manera adecuada, a saber, representados conscientemente. En esquema:
   Para cualquier estado mental M y cualquier sujeto S, M tiene carácter subjetivo solo si hay un estado mental $M*$ tal que (i) S está en $M*$; (ii) $M*$ representa a M y (iii) esa representación es adecuada (o sea, $M*$ es un estado consciente).

3. La representación adecuada propia de la representación consciente solo puede ser autorrepresentación. En esquema:
   Para cualquier estado mental M y cualquier sujeto S, M tiene carácter subjetivo solo si hay un estado mental $M*$ tal que (i) S está en $M*$ y (ii) $M*$ representa a M y (iii) $M* = M$

4. Por consiguiente, todo estado tiene carácter subjetivo (y es consciente) solo si es autorrepresentado.

El argumento maestro es un argumento basado en la necesidad, al cual se le van añadiendo y refinando las condiciones de esa necesidad en cada uno de los pasos. En el paso 1 se concluye que todo estado con carácter subjetivo debe ser un estado representado (Kriegel 2009: 101-113). Kriegel da este paso en virtud del carácter manifiestamente fenomenológico de la experiencia consciente, oponiéndose a dos posibles soluciones alternativas, a saber:

(a) que el estado con carácter subjetivo lo sea sin necesidad de que el sujeto se percate de él, debido a una especie de propiedad intrínseca, inexplicable y no estructurada como en la forma de un "resplandor intrínseco"[85] o

(b) que el estado con carácter subjetivo sea percatado pero no representado, es decir, que sea el producto de una certeza íntima de conocer (*acquaintance*) no necesariamente representacional.[86]

Contra (a), Kriegel argumenta que si el estado con carácter subjetivo lo fuese a través de una propiedad intrínseca, inexplicable y no estructurada como un "resplandor intrínseco" —*intrinsic glow*—, entonces el sujeto no podría percatarse de ese estado. Sin embargo, el sujeto se percata de los estados conscientes, esto es, estos estados le son manifiestamente accesibles y, por tanto tienen cierta estructura para él.

Contra (b), Kriegel (2009: 113) sostiene que la certeza íntima de conocer (*acquaintance*) no es un concepto aceptable, puesto que es factivo. En otras palabras, presupone aquello que es, a saber, el grado de intimidad de los estados con carácter subjetivo.

Una vez descartadas esas dos opciones, a saber, que los estados con carácter subjetivo puedan serlo en atención a una propiedad intrínseca, inexplicable y no estructurada

[85] Kriegel (2009: 101) atribuye una variante de esta opción a Husserl (1964 1928).
[86] Kriegel (2009: 107) atribuye esta variante a Levine (2001).

(*intrinsic glow*) y que puedan ser en atención a una certeza íntima de conocer (*acquaintance*), la única opción es que los estados mentales tengan carácter subjetivo en atención a ser estados representados.

El paso 1 todavía no se diferencia de lo que podría ser una postura representacional de orden superior. Los pasos 2 y 3 constituyen lo que Kriegel denomina el argumento epistemológico, y van distinguir meridianamente la postura de mismo nivel de las posiciones representacionales de orden superior. El paso 2 señala que la representación del estado mental con carácter subjetivo (esto es, $M*$) no puede ser una representación cualquiera, puesto que el sujeto tiene evidencia fenoménica directa de ese estado. Esto solo puede ser debido a que la representación de ese estado es adecuada, esto es, representación consciente. Ahora se pueden distinguir dos propiedades con respecto al objeto representado: propiedades *schmalitativas* y propiedades cualitativas. Las propiedades *schmalitativas* son las propias de una representación no consciente, y hacia ellas el sujeto no tiene un nivel de intimidad.[87] Estas propiedades pasan desapercibidas. El sujeto no se percata de ellas. Las propiedades cualitativas, por el contrario, son las que forman la parte esencial del denominado carácter cualitativo de la conciencia fenoménica y el sujeto tiene un alto grado de intimidad acerca de ellas.

> Debemos distinguir dos tipos de propiedades: las propiedades representadas y las propiedades constituidas por la representación. A las primeras las podemos llamar *schmalitativas*, mientras que a las segundas las podemos llamar cualitativas. De la misma manera que existe una diferencia entre medir 2 metros y ser representado como midiendo 2 metros, también existe una diferencia entre ser *schmalitativamente* azulado y ser representado como siendo *schmalitativamente* azulado.[...]
> El carácter cualitativo es carácter *schmalitativo* apropiadamente representado.[. . .]
> Las propiedades *schmalitativas* no forman parte del carácter fenoménico de la experiencia. De hecho, no son en absoluto fenomenológicamente manifiestas. Las únicas propiedades fenomenológicamente manifiestas, las únicas que forman parte del carácter fenoménico, son las propiedades cualitativas. (Kriegel 2009: 109-110)

El paso 3 se demuestra simplemente por la imposibilidad de una regresión infinita de estados conscientes. Si todo estado consciente es representado conscientemente, entonces existen tres alternativas: o bien es representado (conscientemente) por el mismo estado mental o bien es representado (conscientemente) por un estado mental diferente del estado mental propio de la representación o bien sucede una mezcla de ambos. Sin embargo, las dos últimas opciones conducen a regresiones infinitas de estados conscientes en las que siempre es necesario un nuevo estado para justificar el adjetivo "consciente" aplicado al anterior o a estados entrelazados, donde el primero representa al segundo y viceversa. La única salida, por tanto, es una autorrepresentación que evite la regresión infinita.

En consecuencia, un estado mental tiene carácter subjetivo (y, en consecuencia, es consciente) solo si se representa a sí mismo.

[87] *Schmalitative* es un neologismo que Kriegel introduce en oposición a *qualitative*, no sin cierta ironía. Refleja el carácter desapercibido y no fenoménico que caracteriza a estas propiedades frente a las propiedades cualitativas (*qualitative*, relativas a los *qualia*) propias de los estados conscientes y del carácter fenoménico de los mismos.

De acuerdo con el argumento maestro, todo estado consciente es una autorrepresentación. Pero no toda autorrepresentación (una fotografía de un espejo, por ejemplo) es, *per se*, consciente. La conciencia fenoménica subjetiva tiene otras propiedades específicas: es periférica, frente a la noción de conciencia focal y da información acerca del interior, en oposición al estado consciente acerca de objetos externos. Kriegel (2004, 2009) la define como un "tener una consciencia interior periférica" (*an inner peripheral awareness*).

¿Y qué significa entonces que el sujeto tenga una consciencia interior periférica? Que sea interior significa que la conciencia implica que el sujeto es consciente del interior a través de la representación exterior, es decir, a través de las propiedades cualitativas que constituyen esa representación, las cuales provienen de la autorrepresentación propia de los estados conscientes. Que la consciencia sea periférica significa que su función es indicar al sujeto que hay información disponible acerca de los estados mentales en caso de que requiera de ella. Por consiguiente, la consciencia interior periférica proporciona al sujeto "información suficiente sobre su propia experiencia como para poder recuperar más información de manera rápida y sin esfuerzo si fuera preciso" (Kriegel 2009: 259).

En conclusión, la Teoría Autorrepresentacional de la conciencia permite explicar las propiedades fenoménicas de la experiencia consciente y, en concreto, permite explicar su carácter subjetivo a través de la autorrepresentación, que se vuelve una noción necesaria (aunque quizás no suficiente). También explica, al menos parcialmente, cuál podría ser una de las funciones de la conciencia. Sin embargo, no explica completamente de las propiedades funcionales, temporales y estructurales de la conciencia.

1.3.3. Teorías informacionales

La representación va asociada a la información, tal y como se pone de manifiesto en el segundo punto de la Tesis Representacional (Dretske 1995). En ese sentido, se puede afirmar que las teorías representacionales tienen cierto aspecto informacional. Sin embargo, no todas las teorías informacionales son teorías representacionales o, al menos, no todas las teorías informacionales ponen el peso explicativo en la noción de representación. Ese es el caso de dos de las teorías de la conciencia más populares: la Teoría del Espacio de Trabajo Global y la Teoría de la Información Integrada.

1.3.3.1. Teoría del Espacio de Trabajo Global

La Teoría del Espacio de Trabajo Global (Baars 1988) emplea elementos informacionales, representacionales y neurofisiológicos para explicar la conciencia y, especialmente, sus propiedades funcionales. De acuerdo con ella, la experiencia consciente es experiencia de un contenido atendido, el cual se procesa de manera local y se difunde de manera amplia por zonas diversas del cerebro, en lo que se denomina el espacio de trabajo global. Este contenido está dotado de representacionalidad, en el sentido de información, aunque esta

teoría no pone el peso explicativo en el acto de representar, sino en el proceso de difusión de la representación, es decir, de la información, para que sea disponible por los sistemas cognitivos. Esta teoría tiene una versión más determinada neuronalmente, la Teoría del Espacio de Trabajo Global Neuronal(Dehaene y Naccache 2001; Changeux y Dehaene 2008; Dehaene et al. 2011), en la cual se asocia esa difusión con actividades concretas que tienen lugar en zonas amplias del cerebro que incluyen desde el flujo ventral hasta el córtex frontal y parietal. Ahora voy a exponer los aspectos básicos de esta teoría, tal y como Baars (1988) los presentó. Después recopilaré las evidencias neurofisiológicas relacionadas con la versión neuronal de la teoría. Para finalizar, evaluaré breve y críticamente los logros y defectos de esta teoría.

La teoría de Baars (1988) parte de dos ideas principales:

1. El cerebro es una máquina de procesamiento masivo en paralelo que está jerarquizada en unidades de procesamiento que, anatómicamente, se agrupan en columnas[88] y que, funcionalmente, se diferencian en módulos orientados a tareas específicas. Estos módulos coinciden con columnas corticales (de entre 500 a 1000 micras) que a su vez son agrupaciones de minicolumnas de unas 30 micras y decenas de células, a las que Baars llama procesadores específicos.[89] Computacionalmente, los procesadores específicos son muy eficientes para el desempeño de tareas concretas.
2. Se acepta una especie de representacionalismo en el siguiente sentido: igual que en un programa informático hay variables locales y variables globales, en el procesamiento del cerebro existen variables locales, que solo adquieren sentido dentro de uno o unos pocos procesadores, y variables globales, que adquieren sentido para todos los procesadores. Estas últimas son el tipo de información que se difunde en el espacio de trabajo global, e inducen a pensar en una *lingua franca* común a todo el sistema nervioso.[90]

El espacio de trabajo global ofrece a cada uno de los procesadores específicos del cerebro la posibilidad de tener acceso a información proveniente de otros procesadores específicos (por ejemplo, aquellos que procesan los estímulos externos) a través de la difusión amplia de dicha información. La conciencia sucede cuando se produce esa difusión amplia. Baars

[88]Mountcastle (1978)

[89]Nótese que, por un lado, hay procesamiento consciente y, por otro lado, procesadores no conscientes tales como los específicos. Los procesos conscientes no se llevan a cabo en procesadores específicos, sino que se producen al difundir de manera global la información proveniente de estos procesadores específicos.

[90]Para diferenciar correctamente entre la *lingua franca* de Baars (1988: 66) y la *lingua mentis* (*Sprache des Geistes* o *language of thought*) de Fodor (1975), nótese que para Baars las variables globales de la *lingua franca* conviven con las variables locales correspondientes a las *linguae non francae*, esto es, los códigos propios que se restringen a los procesadores específicos, mientras que en el lenguaje del pensamiento esta diferenciación no existe. Para Fodor no hay representacionalidad particular. Para Baars sí la hay, pero es precisamente el carácter global de esa representacionalidad lo que hace posible la conciencia. La *lingua franca* para Baars (1988: 67) se define por su globalidad: son códigos susceptibles de ser difundidos ampliamente a cualquier parte del sistema nervioso central y periférico, y suponen una carga mayor para el sistema global ya que deben ser serializados, debido precisamente a la amplia difusión de la que son objeto. De la existencia de este tipo de códigos dan noticia dos datos empíricos: actos aparentemente involuntarios como el ritmo cardiaco, los movimientos peristálticos o la transpiración se pueden controlar (1) mediante bioretroalimentación y (2) mediante imágenes mentales evocadoras.

propone entender el espacio de trabajo global como la pantalla gigante en un congreso, desde la cual cada uno de los ponentes difunde al resto cierta información relevante y significativa, esto es, globalmente relevante y globalmente significativa. En su propio ámbito, cada procesador específico puede desempeñar determinada tarea de manera autónoma y eficiente. Sin embargo, cuando se requiere trabajo cooperativo y, simplemente, no se sabe quién puede ser el destinatario adecuado para esa información, lo mejor es difundir la información en el espacio de trabajo global. En este caso los procesos se hacen más lentos y se vuelven más propensos a errores, pero, a cambio, se consigue un desempeño mejor frente a tareas novedosas o que requieran gran capacidad relacional.

A grandes rasgos, la Teoría del Espacio de Trabajo Global se puede describir del siguiente modo:

- Es una teoría de la conciencia enfocada en explicar las propiedades funcionales de la conciencia.
- No obstante, es una teoría que trata de explicar también aquellos aspectos o propiedades que constituyen el hecho específico de la conciencia, tales como el hecho de que la conciencia tenga un tiempo de procesamiento o que su contenido se presente de forma serial (esto es, que sus contenidos no sean compatibles entre sí).
- Se interesa por hacer compatibles dos aspectos aparentemente contradictorios en los enfoques psicológicos y neurológicos: la serialidad propia de la conciencia, por un lado, y la estructura paralela y masiva de procesamiento cerebral, por otro.[91]

La clave de esta teoría es la noción de espacio de trabajo global. El espacio de trabajo global es el proceso de difusión, en el cual toman parte los sistemas emisores, los sistemas receptores y el propio mensaje. A estos elementos se suma, además, el contexto, que es definido como aquellos procesos no conscientes que, sin formar parte propiamente del proceso de la difusión, tienen capacidad de influir en él. A continuación voy a presentar cada uno de estos elementos con su problemática asociada.

Emisores: ¿cómo se lleva a cabo la competición por el espacio de trabajo global?

El espacio de trabajo global es el proceso de difusión amplia de un mensaje por parte de un emisor. El espacio de trabajo global es un recurso limitado y su acceso ha de producirse con serialidad: es posible que dos procesadores específicos procesen de manera local contenidos incompatibles entre sí, pero no es posible que esos contenidos incompatibles entre sí ocupen simultáneamente este espacio global. Esto da lugar al problema de explicar cómo se produce la competición entre los diferentes emisores por ese espacio de trabajo. Si se difunde siempre toda la información, el espacio de trabajo global estará continuamente saturado con información poco relevante. Por otro lado, si la información no se difunde, no será posible determinar si es relevante, puesto que la relevancia o significatividad de la información viene dada por su significatividad para cada procesador específico y, por

[91] Esto la distingue de la Teoría de la Información Integrada, que se verá después.

consiguiente, no se puede conocer hasta que no se ha producido una cierta difusión. A este aparente contrasentido Baars lo denomina "paradoja del umbral" de conciencia (*threshold paradox*). Baars (1988: 98) ofrece dos posibles soluciones para esta paradoja. La primera consiste en que cada emisor candidato haga difusiones en entornos pequeños para que esas difusiones se hagan paulatinamente más amplias de acuerdo con la respuesta de los receptores más cercanos. La segunda consiste en que cada emisor candidato haga una difusión amplia durante un breve periodo de tiempo, el cual es insuficiente para producir conciencia, pero suficiente para evaluar la respuesta de los receptores de manera global. En ambos casos la retroalimentación por parte de los receptores es necesaria para decidir si se seguirá adelante con difusión amplia o no.

Receptores: ¿cómo de amplia es la difusión en el espacio de trabajo global?

Otro de los aspectos importantes dentro de la Teoría del Espacio de Trabajo Global consiste en determinar el alcance de la difusión o, dicho de otro modo, en determinar si todos y cada uno de los procesadores del sistema nervioso central y periférico son accesibles por la información difundida en el espacio global o si, por el contrario, hay procesadores específicos completamente no conscientes. Baars sostiene que la difusión es potencialmente amplia, esto es, que el espacio de trabajo global puede influir en prácticamente cualquier procesador específico relacionado con el sistema nervioso central y periférico, si bien algunos de estos procesadores están biológicamente preparados para ser más estables y autónomos. Para que nos hagamos una idea de hasta qué grado es global esa difusión, Baars (1988: 100) expone tres ejemplos. El ritmo cardíaco, que se controla desde el cerebro posterior (*hindbrain*), puede ser modificado por las experiencias conscientes, incluso aunque solo sea de una manera paulatina y moderada. Igualmente, es posible modificar la actividad de una unidad motora individual de la médula espinal como, por ejemplo, la que dirige la activación de determinado músculo de la mano, hasta el punto de que, ayudado por la biorretroalimentación adecuada, el sujeto es capaz de producir secuencias de actividad artificiales (como redobles de tambor) en esa unidad motora (Basmajian y John V. 1979). Por último, es posible alterar a través de la experiencia consciente mecanismos aparentemente automáticos tales como la dilatación de los vasos sanguíneos, la respuesta galvánica de la piel,[92] el movimiento ocular, el incremento del ritmo cardíaco o la dilatación de pupilas.

Mensaje: ¿qué tipo de información se puede difundir en el espacio de trabajo global?

La caracterización del mensaje que se difunde es uno de los aspectos más complejos y, quizás, más problemáticos dentro de la Teoría de Espacio de Trabajo Global. Aun así, se pueden decir algunas cosas acerca del mensaje y su contenido. Para empezar, el contenido significativo para cada procesador específico tiene un rango limitado. Por ejemplo, algunos procesadores específicos solo pueden trabajar para integrar aspectos concretos de

[92]La respuesta galvánica de la piel (*Galvanic Skin Response*, GSR) refleja la respuesta emocional del sujeto a través de la variación del nivel de sudoración en la piel medido mediante electrodos ubicados generalmente en las manos. Es uno de los parámetros que se miden en las máquinas detectoras de mentiras.

la percepción visual (colores, figuras o movimientos concretos). En cambio, el contenido que se puede difundir a través del espacio de trabajo global es de rango amplio, es decir, puede ser de naturaleza muy diversa (sensorial, perceptual, conceptual, emocional). Un requisito es que esa información difundida sea internamente consistente.[93] No es posible, por ejemplo, ser consciente de las dos posiciones del cubo de Necker al mismo tiempo.[94] Tampoco es posible ser consciente del significado de las palabras "*turn*", "*look*" o "*book*" al mismo tiempo. Esta noción de consistencia explica junto con la limitación de recursos, la serialidad propia de la conciencia.

Contexto: ¿cómo influyen los procesadores específicos en el mensaje del espacio de trabajo global?

El contexto es un concepto que ha adquirido una importancia capital en la Teoría del Espacio de Trabajo Global. El contexto se define como coaliciones entre procesadores específicos[95] que, sin pasar a formar parte del espacio global, lo condicionan y restringen de manera continuada y significativa. Por ejemplo, dentro del ámbito de los procesos cognitivos relacionados con la comprensión lectora, términos como "crisis", "recuperación" o "desaceleración" tienen una interpretación distinta en un contexto económico, médico, ético o físico, sin que en la mayoría de los casos sea necesaria una conciencia específica de esa distinción. Se dice entonces que los conceptos detrás de estos términos forman parte del contexto (de un contexto conceptual en este caso), y que la interpretación de los mismos se realiza de manera automática, es decir, en función del contexto.

El contexto conceptual no es único. Puede haber contextos perceptuales, conceptuales o basados en objetivos y esos contextos pueden formar, a su vez, estructuras jerárquicas.[96] Por otro lado, los contextos no son estados de cosas inmóviles, sino que son procesos jerárquicos que, a su vez, pueden ser alterados por el propio proceso de difusión del mensaje. En otras palabras, las coaliciones de procesadores específicos pueden alterarse a través del espacio de trabajo global. La contextualización consiste en la constitución de un contexto a través del espacio de trabajo global, a lo largo del cual procesos conscientes pasan a ser no conscientes (es decir, procesos de difusión global pasan a formar coaliciones nuevas entre procesadores específicos). Por ejemplo, el zumbido continuo de un frigorífico deja de ser advertido conscientemente por el sujeto al cabo de un rato. Por otro lado, la descontextualización consiste en la alteración de determinados contextos que pasan a ser conscientemente representados (es decir, coaliciones de procesadores específicos pasan a ser difundidos en el espacio de trabajo global). Por ejemplo, el cese del zumbido del frigorífico hace que el sujeto sea consciente de la extinción de ese zumbido que, un instante antes, pasaba inadvertido.

Si el espacio de trabajo global es la parte fundamental de esta teoría, los procesos neurofisiológicos que se han propuesto para dar cuenta de este espacio han sido tales que

[93] Baars (1988: 82).

[94] Ver figura 1.1

[95] Coalición aquí tiene el mismo sentido que la noción de asamblea de células (*cell assemblies*) en Hebb (1949).

[96] Baars (1988: 137).

fuesen adecuados para ofrecer una explicación de las diferentes cuestiones planteadas. En general, se han propuesto mecanismos de procesamiento recurrente, donde la difusión propia del espacio de trabajo depende de procesos de realimentación en los cuales los procesadores específicos toman parte. Baars sostiene que la difusión se produce mediante sistemas complejos de activación en los que participan el tálamo, la formación reticular mesencefálica y el córtex cerebral.[97] Dehaene (Dehaene y Naccache 2001; Changeux y Dehaene 2008; Dehaene et al. 2011) ha desarrollado la Teoría del Espacio de Trabajo Global Neuronal donde los procesos de multidifusión se asocian con la activación estable sostenida en el tiempo de determinadas áreas corticales. Esta peculiar activación deja señales inequívocas que se pueden rastrear a través de las tecnologías disponibles hoy en día para la detección de la actividad cerebral. Dehaene (2014: 115) denomina a estas señales "marcadores característicos de la conciencia"[98] e identifica cuatro de ellos. En primer lugar, mediante fRMI[99] se puede identificar cierta actividad asociada solo con los procesos conscientes en zonas tan distantes como el córtex parietal y el córtex prefrontal. Esta actividad, además, presenta el aspecto de una ignición: no es lineal, sino que pasa de no darse en absoluto a darse completamente en los procesos conscientes. En segundo lugar, a través de EEG[100] se puede detectar una onda lenta de tipo P3 (o P300) que sucede alrededor de 300 msg después del estímulo, y que correla de manera convincente con los casos de percepción consciente y no con los casos de percepción subliminal.[101] Las ondas de este tipo, además, nunca se presentan seguidas, sino que existe un periodo refractario entre ellas (de ahí el nombre de ondas "lentas"). La onda P3 refleja dos propiedades de conciencia: (a) que procesos asociados a la conciencia son lentos y se retrasan con respecto al mundo real (una cosa es cuándo sucede el estímulo y otra cuándo el sujeto es consciente del mismo y, por consiguiente, no hay percepción inmediata) y (b) que existe una competición por acceder a la conciencia, esto es, por el acceso al espacio de trabajo global (puesto que existe ese periodo refractario). Por último, mediante la implantación de electrodos en el cerebro profundo de pacientes sometidos a craneotomías[102] se pudieron identificar dos nuevos marcadores característicos de la conciencia. Por un lado, una ráfaga de oscilaciones de frecuencia alta, correspondientes a la banda gamma (más de 30 Hz), que se detectaba de manera generalizada en el cerebro (como es propio de una actividad multidifundida). Por otro lado, una actividad sincronizada en zonas distantes del cerebro (como es propio de una actividad global).

[97] Sistema ERTAS (*Extended Reticular-Thalamic Activating System* (Baars 1988: 119).

[98] Dehaene (2014: 142) distingue entre marcadores característicos (*neural signatures of consciousness*) y correlatos neuronales (*neural correlates of consciousness*). Los marcadores característicos tienen una determinación mayor. Son señas de identidad de la conciencia, no meras correlaciones.

[99] La imagen por resonancia magnética funcional (*functional magnetic resonance imaging*, fRMI) permite obtener imágenes del cerebro a través del contraste dependiente del abasto de oxígeno en sangre en cada zona. fRMI se emplea para identificar espacialmente qué zonas cerebrales tienen mayor actividad durante determinado desempeño. Esta técnica no tiene precisión temporal.

[100] El electroencefalograma (EEG) mide la actividad eléctrica del cerebro a través del cuero cabelludo. Los potenciales evocados medidos mediante EEG se pueden ubicar temporalmente con gran exactitud. Esos potenciales evocados se pueden asociar a potenciales relacionados con eventos cognitivos (ERPs, *event-related potentials*). P3 es un tipo de ERP que se asocia con la percepción consciente.

[101] La cuestión acerca de cuáles son los marcadores característicos de la conciencia es controvertida. Algunos autores (Tononi y Koch 2015), por ejemplo, defienden la existencia de casos en los cuales puede haber conciencia sin P3.

[102] Las técnicas fRMI y EEG no son invasivas, pues se miden sin necesidad de acceder al cerebro del paciente. La implantación de electrodos en el cerebro profundo, en cambio, requiere intervención oportunista y deslocalizada durante la cirugía transcraneal realizada a pacientes con tumores o epilepsia severa.

En conclusión, la Teoría del Espacio de Trabajo Global explica propiedades funcionales y temporales de la conciencia tales como la propensión a cometer errores, la lentitud, la serialidad de los contenidos de la conciencia y la capacidad de relación y la sensibilidad al contexto. Además, esta teoría es quizás una de las que mayor evidencia empírica ha recopilado en torno a los procesos que pueden dar cuenta de la conciencia, demostrando una gran capacidad para ser contrastada, en especial mediante experimentos de enmascaramiento en percepción visual. A pesar de ello, esta teoría se enfrenta a numerosos problemas. Por un lado, existen evidencias empíricas que parecen disociar las nociones de atención y conciencia.[103] Además, esta teoría no ofrece respuestas concretas para las propiedades fenoménicas de la conciencia. En concreto, carece de una herramienta convincente para explicar el carácter subjetivo de la conciencia. Tampoco presenta una caracterización completa e indudable de qué puede ser aquello que se denomina "el mensaje" en la difusión global. Como respuesta a este último problema surgieron otras teorías como la Teoría de la Información Integrada.

1.3.3.2. Teoría de la Información Integrada

Existen teorías informacionales que no requieren de la noción de representación. La Teoría de la Información Integrada (*Integrated Information Theory*, IIT)[104] tiene cierta orientación matemática inspirada en Shannon (1948), aunque la noción de información que usa no está basada en comunicación sino en estructura. Partiendo de la idea de que la experiencia consciente es un hecho específico, diferente en cada momento, con unidad, y que debe ser explicado desde posicionamientos no dualistas, esta teoría establece que la conciencia está asociada a una *in-formatividad*, en el sentido de diversidad de estados con poder de causa-efecto que hacen la diferencia, los cuales, además, están dotados de integración, en el sentido de dependencia mutua entre los diferentes elementos que componen esa informatividad. La Teoría de la Información Integrada identifica la experiencia consciente con información integrada y propone que las zonas del cerebro que dan cuenta de la experiencia consciente deben ser tales que puedan hacer posible esta integración de la información. Esta teoría se presentó en su versión 1.0 en Tononi (2004) y ha sido actualizada con el tiempo en una versión 2.0 (Tononi 2008; Balduzzi y Tononi 2008) y otra versión 3.0 (Tononi 2012; Oizumi et al. 2014). La versión 3.0 se distingue de las anteriores en diversos aspectos: establece explícitamente los cinco axiomas básicos de la experiencia consciente, da mayor peso a la idea de jerarquía en esa información integrada, relaciona el estado actual del sistema con el mecanismo causal de ese sistema como causa (pasado) y como efecto (futuro) e identifica la experiencia consciente con una estructura conceptual maximalmente irreducible. Además, matemáticamente también tiene algunas diferencias.[105] Aquí voy a presentar los cinco axiomas de la versión 3.0. Sin embargo, para la exposición del corpus matemático me basaré en la versión 2.0, por resultar más

[103]Como, por ejemplo, el experimento de supresión basada en destello continuo (CFS o *continuous flask suppression*) (Koch y Tsuchiya 2007)

[104]Tononi (2004); Balduzzi y Tononi (2008); Oizumi et al. (2014)

[105]La versión 3.0 de esta teoría identifica la experiencia consciente con la estructura causal mayor del sistema que no puede reducirse a otras estructuras menores. Para más acerca de las diferencias entre cada versión de la IIT, consultar el texto de apoyo S1 en Oizumi et al. (2014).

accesible y sencilla, al tiempo que contiene conceptualmente la noción de integración de información. Primero presentaré las nociones más relevantes de la teoría. Después recordaré algunas de las evidencias en su favor. Para finalizar, expondré algunos de los problemas y retos a los que se enfrenta esta teoría.

Esta teoría parte de tres asunciones iniciales:[106]

- Asunción física: la explicación de la conciencia solo requiere de procesos físicos. Como consecuencia, se niega el dualismo mente/cuerpo y se acepta que es posible falsear las teorías de la conciencia mediante la evidencia científica.
- Asunción evolutiva: la conciencia es producto de la selección natural. Es muy posible que existan formas de conciencia primaria en los animales. Toda teoría de la conciencia debe explicar esa gradualidad.[107]
- Asunción de *qualia*: la conciencia tiene un carácter subjetivo y no transmisible. La conciencia depende del sujeto y es diferente en cada momento (no hay dos estados conscientes exactamente iguales).

Además de estas tres asunciones, la Teoría de la Información Integrada subraya especialmente cinco propiedades básicas de la conciencia (Oizumi et al. 2014). Estas propiedades son también llamadas "axiomas" de la teoría, por su carácter irrenunciable, inmediato e irreducible entre ellas:

- Existencia intrínseca: la conciencia existe como hecho en bruto que hay que explicar y además existe de modo intrínseco, en el sentido de que existe independientemente del observador. Mi experiencia consciente de ver (o imaginar) una manzana roja no depende de otra cosa que no sea un sistema complejo específico (tal como mi cerebro).
- Composición: la conciencia posee una estructura, esto es, una diferenciación compositiva. Mi experiencia consciente de ver una manzana roja se engloba en un paisaje visual compuesto por la mesa, el mantel y la manzana encima de todo. Incluso si solo aparece una manzana, mi experiencia consciente está compuesta por las partes de esa manzana. Incluso cuando contemplo un punto en un papel blanco, mi experiencia consciente tiene cierta estructura: posee la orientación del punto con respecto al suelo, al cielo o a mí mismo.
- Información: dos estados conscientes no son nunca exactamente iguales, es decir, cada estado posee siempre una diferencia informacional: es específica. Mi experiencia consciente de ver una manzana roja nunca es la misma. Si la veo por segunda vez, por ejemplo, la experiencia consciente está teñida del carácter de lo repetido. Si no recuerdo haberla visto antes, está teñida del carácter de novedad.

[106] Edelman y Tononi (2000)

[107] La conciencia primaria reduce sus objetos a lo presente, y es posible encontrarla tanto en el ser humano como en los animales. La conciencia primaria es la conciencia básica a partir de la cual se desarrolla la conciencia del ser humano, la cual, a diferencia de la primaria, puede tener contenidos acerca del pasado o del futuro.

- Integración: a pesar de la composición y la información, la experiencia consciente está unificada y, por consiguiente, no es reducible a partes independientes. Mi experiencia consciente de ver una manzana se integra con la mesa y el mantel. No es reducible a las experiencias conscientes resultantes de ver la manzana sola más la experiencia de ver la mesa sola, más la experiencia de ver el mantel solo.
- Exclusión: en cada momento, mi experiencia consciente excluye a las demás, ya que no hay algo como una superposición de estados conscientes, sino un estado consciente completo. Mi experiencia consciente de ver una manzana roja es una única experiencia compleja que no admite otras experiencias simultáneas pero diferentes. Esta propiedad es la contrapartida de la integración.

Las tres asunciones anteriores junto con estas cinco propiedades en forma de axiomas constituyen la base de la Teoría de la Información Integrada. Sin ellas, esta teoría no se puede interpretar correctamente. Para la Teoría de la Información Integrada, medir el grado de integración y diferenciación de la información en un sistema equivale a medir el grado de conciencia de ese sistema. Con el objeto de medir el grado de integración y diferenciación de la información, Tononi introduce un nuevo concepto matemático: Φ o la integración de la información. Resulta imposible valorar una teoría como la IIT sin tener en cuenta las nociones matemático-informacionales en las que se basa. Por ese motivo, a continuación se presentará un resumen de esta parte (basado en la versión 2.0).[108]

La estructura matemática de la noción de información integrada tiene similitudes con la teoría clásica de la información de Shannon (1948). Siguiendo a Shannon, la información es la entropía de un sistema, la cual mide el grado de incertidumbre de dicho sistema. La información, por tanto, requiere de incertidumbre: solo puede existir información en un contexto de reducción del grado de incertidumbre en caso de comunicación exitosa entre la fuente de información y el destino. Un fanal, por ejemplo, que emite siempre la misma luz con la misma frecuencia, no posee más información que la de marcar la cercanía de la costa. Es predecible y su entropía es muy pequeña. La señal emitida por un telégrafo, en cambio, tiene un valor de entropía mayor. No es (tan) predecible a partir de sus estados anteriores.

Como es conocido, la entropía H de una distribución de probabilidad $p = p_1, ...p_n$ se mide a través de la fórmula siguiente:

$$H(p) = -\sum_{i=1}^{n} p_i \log_2(p_i)$$

donde la entropía (y el grado de información) es mayor cuanto más iguales y menores sean las probabilidades p_i. Por ejemplo, la comunicación de un bit con probabilidad idéntica de 0 o 1 arroja un valor de entropía $H = -\log_2(1/2) = 1$, mientras que la comunicación de 8 bits (un byte) con idénticas probabilidades arroja una entropía mayor, $H = -\log_2(1/8) = 3$. La comunicación de 0 bits arroja un valor de entropía $H = 0$, esto es, la mayor certidumbre posible y en consecuencia la menor información posible.[109]

[108] Balduzzi y Tononi (2008)

[109] La formulación de entropía presupone el teorema de Bayes y por consiguiente está expuesta a las críticas del bayesianismo provenientes de la mecánica cuántica y el teorema de Bell.

La Teoría de la Información Integrada hace uso de la estructura matemática propia de la información en Shannon para definir la noción de información integrada Φ de un modelo de referencia o sistema X.[110] El sistema X está formado por una serie de elementos cada uno de los cuales puede tener diferentes estados. El conjunto de esos estados en un momento dado i conforma el estado del sistema en ese momento (x_i). Además, los elementos del sistema X se relacionan entre sí de acuerdo a una mecánica causal de conexiones ($mech$). El objetivo de la teoría es medir el grado de integración de los diferentes elementos de ese sistema. Para ello se define la noción de información integrada o Φ en relación con el estado actual del sistema (x_1). La información integrada del sistema en relación con un estado actual se define como la divergencia entre dos distribuciones de probabilidad. La primera es la distribución de probabilidad para el repertorio de posibles estados del sistema antes de llegar al estado actual, esto es, $p(X_0(mech, x_1))$. La segunda distribución de probabilidad resulta de dividir el sistema X en dos o más particiones con diferenciación mínima y obtener el producto de las distribuciones de probabilidad de cada partición calculadas por separado, suponiendo las conexiones intraparticionales como parte de los mecanismos causales de la partición, y las conexiones interparticionales como entradas aleatorias, esto es, como si no formaran parte de los mecanismos causales internos. Para calcular la divergencia se emplea la definición de entropía relativa de Kullback-Leibler, de acuerdo con la cual la divergencia entre dos distribuciones de probabilidad p y q se calcula así:

$$D_{KL}(p|q) = H[p|q] = \sum_{i=1}^{n} p_i \log_2(p_i/q_i)$$

La información integrada Φ es el resultado de aplicar la divergencia anterior a las dos distribuciones de probabilidad del sistema antes indicadas.[111] Esto es:

$$\Phi(X(mech, x_1)) = H[p(X_0(mech, x_1))| \prod p({}^kM_0(mech, \mu_1))]$$

donde: ${}^kM_0 \in$ MIP; MIP es la partición de información mínima, es decir, la partición para la cual las partes son más independientes (o, en otras palabras, están menos integradas entre sí)[112] y $\prod p({}^kM_0(mech, \mu_1))$ es el producto del repertorio actual de las partes ${}^kM_0$

[110]Que haya semejanzas en la estructura matemática no significa que la noción de información de esta teoría sea la misma de Shannon. En el texto de apoyo S3 de Oizumi et al. (2014) se resaltan estas diferencias, y se denomina información intrínseca (y enfocada a estructura) a la noción de información de la IIT, en contraste con la información extrínseca y enfocada a comunicación de Shannon. Esta distinción es notable. Mientras que un defensor de la Teoría del Espacio Global sostendría que la conciencia es la difusión global de un mensaje, un defensor de la Teoría de la Información Integrada respondería: ¿Qué mensaje? No existe mensaje, la información es estructura causal, es *in-formación*.

[111]Así es en la versión 2.0 de esta teoría. Sin embargo, para la comparación de distribuciones de probabilidad la versión 3.0 no usa la divergencia de Kullback-Leibler, sino la métrica de Wasserstein (o EMD, *earth mover's distance*). Si $D_{KL}(p|q)$ mide la divergencia entre dos distribuciones de probabilidad p y q, EMD(p, q) mide la distancia entre esas dos distribuciones de probabilidad. EMD es más exacta que la divergencia de Kullback-Leibler y cumple simetría, es decir, EMD(p, q)=EMD(q, p). Por contra, EMD es más compleja de calcular y requiere siempre de computación. La idea básica es que EMD(p, q) mide el trabajo mínimo necesario para pasar de una forma de un montón de tierra (la distribución de probabilidad p, con valor total 1) a otra forma de montón de tierra con cantidad total idéntica (la distribución de probabilidad q, igualmente con valor total 1).

[112]Intuitivamente, si debo realizar la partición del sistema formado por la rama de un naranjo y su naranja en dos partes, lo más sencillo es partirlo por el punto más débil (y de menor integración), esto

tomadas por separado.[113]

La variable Φ refleja el grado de integración dentro del sistema de referencia: dice cuánto diverge el sistema si se toma en conjunto con respecto a si se tomara de manera particionada. Un modelo compuesto por n elementos no conectados entre sí tendrá obviamente $\Phi = 0$, dado que es un sistema estático, en el cual no hay divergencia alguna entre la distribución de probabilidad del conjunto y la de las partes. Un modelo compuesto por n elementos donde los elementos solo se conectan de dos en dos, tal y como una cámara de fotodiodos (Tononi 2008), tendrá igualmente $\Phi = 0$, puesto que no hay divergencia entre la distribución de probabilidad del sistema tomado como un conjunto y la resultante de particionarlo en dos conjuntos, cada uno de ellos formado por dos elementos. Por el contrario, un sistema rico en conexiones de reentrada entre los elementos presentará $\Phi > 0$.

Además de explicar cuantitativamente la experiencia consciente, el sistema complejo permite también explicarla cualitativamente: "Igual que la cantidad de conciencia generada por un complejo de elementos está determinada por la cantidad de información integrada que éste genera, la cualidad de la conciencia está determinada por el conjunto de todas las relaciones informacionales generadas por sus mecanismos." (Tononi 2008: 234). En definitiva, la estructura causal del sistema complejo en un determinado momento explica cualitativamente la experiencia consciente, mientras que la noción de información integrada describe cuantitativamente el grado de esa experiencia consciente.

La Teoría de la Información Integrada tiene una fuerte determinación neurofisiológica. Tononi (2008) destaca las siguientes evidencias empíricas a su favor:

- Las simulaciones del sistema corticotalámico se corresponden con sistemas con valores altos de Φ, esto es, en los que se combinan especialización funcional (modularidad) con integración funcional. En concreto así es la estructura columnar y microcolumnar del córtex. La conciencia no depende de ninguna región cortical concreta.
- Los casos de callosotomía (propios del fenómeno de cerebro escindido) suponen una reducción drástica de Φ, pero no la hacen cero, ya que existen conexiones indirectas entre los dos hemisferios. Esto explica la existencia de conciencia unificada en esos casos.
- Por otro lado, los estudios acerca de la arquitectura de conexión del cerebelo (Bower 2002) indican poca integración. Pese a la profusión y gran informatividad de las conexiones cerebelares, en términos de sinapsis esta conectividad es sencilla, al menos, si se compara con lo que sucede en el córtex. Esto encaja con las evidencias

es, por el tallo que las une, en lugar de hacerlo partiendo la naranja en dos o partiendo la rama en dos. Esta intuición aplica para la definición de la partición de información mínima. El cálculo exacto de la MIP o partición de información mínima está establecido y en la versión 2.0 de IIT requiere de una normalización. Véase Balduzzi y Tononi (2008).

[113]La manera de hallar el solapamiento de dos distribuciones de probabilidad independientes es el producto, de igual manera que, para dos sucesos independientes A y B, la probabilidad de $P(A \cap B) = P(A)P(B)$. Este es el caso de las distribuciones de probabilidad resultantes de cada una de las partes de la partición.

de que el cerebelo no contribuye esencialmente en los procesos conscientes (Yu et al. 2014).

- Por lo que sabemos, los caminos aferentes no participan de las complejas conexiones del sistema, lo que explicaría que, aun siendo primordiales en el disparo de la conciencia, no sean habilitadores. Ese es el caso de la zona V1 (Crick y Koch 1995).
- La asociación de información integrada con conciencia es compatible con la idea de un núcleo dinámico (Edelman y Tononi 2000) y, por consiguiente, con la idea de que la base habilitadora de la conciencia no está siempre formada por los mismos elementos, sino que es variable.
- Los informes acerca de sujetos despertados durante periodos NREM[114] de sueño indican falta de estados conscientes, lo que coincide según las medidas EEG con periodos de actividad cerebral (niveles bajos de acetilcolina) asociados a una pérdida de la integración de la información.
- La conciencia requiere un tiempo de procesamiento o ignición (Bachmann 2000; Del Cul et al. 2007), lo que es compatible con la dependencia de procesos de integración de la información.

La Teoría de la Información Integrada está especialmente preparada para explicar las propiedades estructurales y funcionales de la conciencia tales como la diversidad informacional y la integración y unidad de ese contenido. Además, ofrece una explicación de los *qualia* propios de la experiencia consciente. Sin embargo, adolece de ciertos problemas. El primero es que los cinco axiomas no cumplen adecuadamente con el propósito de recoger adecuadamente las propiedades de la conciencia (Bayne 2018) y, en concreto, que no incluyen el carácter subjetivo propio de la experiencia consciente ni su lentitud y aceleración (de acuerdo con el listado de propiedades presentado antes). Otro problema es que esta teoría abre la posibilidad a que cualquier sistema con un Φ mayor que cero, esté formado por cuatro diodos o por millones de neuronas, presente cierta conciencia. Aún más problemático resulta que se pueden imaginar sistemas complejos de elementos con valores altos de Φ no correlacionados con una mayor conciencia (Aaronson 2014). El "sistema Vandermonde", basado en una mecánica descrita por la matriz de Vandermonde, base de códigos de corrección de errores tales como Reed-Solomon, que se usan para la lectura de CDs o para la transmisión en comunicaciones, presenta valores de Φ altos sin que, por ello, se pueda afirmar que el lector de CDs disfrute de conciencia.[115] Tononi y Koch (2015), por su parte, han respondido a estas objeciones de distintas formas. Primero, han defendido que una gradualidad de la conciencia en el mundo físico permite resolver problemas tales como la posibilidad de existencia de cierta conciencia en animales. Segundo, han recordado que los computadores que procesan este tipo de mecánicas funcionales no tienen realmente mecánicas con información integrada alta a nivel físico, puesto que son máquinas de Von Neumann con procesamiento mayormente en serie. El debate sigue abierto.

[114] *Non-rapid eye movement.*

[115] Se pueden imaginar otros sistemas ficticios basados por ejemplo en la encriptación Rijndael de AES (*Advanced Encryption Standard*), donde dada la confusión y difusión de la información inicial en cada una de las etapas (*rounds*) cabe esperar valores altos de Φ.

1.3.4. Teorías biológicas

Si las teorías representacionales explican la conciencia a través de la noción de representación y las teorías informacionales a través de la noción de información, las teorías biológicas de la conciencia lo hacen a través de la noción de vida. En la línea del naturalismo biológico defendido por Searle (1984, 1992), las teorías biológicas explican la conciencia como un proceso específico de los organismos vivos que es similar a —o, al menos, no muy diferente de— otros procesos tales como la respiración, la digestión o la reproducción. Por tanto, al contrario de lo que sucede en otros tipos de teorías, las biológicas marcan una acusada distinción entre organismos con vida y sistemas sin vida. Los primeros son susceptibles de tener conciencia, mientras que los segundos no.

Existen distintos tipos de teorías biológicas.[116] Las más conocidas son las teorías de tipo microgenético, esto es, aquellas que explican los procesos neurofisiológicos que dan cuenta de la formación de la experiencia consciente. Se pueden destacar dos tipos de teorías microgéneticas de la conciencia: las teorías de la actividad sincronizada (Singer y Gray 1995; Koch 2004) y la Teoría de la Doble Microgénesis (Bachmann 2000).

1.3.4.1. Teorías de la actividad sincronizada

Las teorías de la actividad sincronizada responden al problema del enlace expuesto por von der Malsburg (1994). El problema del enlace es el problema de explicar cómo es posible la percepción unificada de cada objeto en una imagen, por ejemplo, cómo es posible percibir un triángulo rojo y un cuadrado azul (y no al revés, por ejemplo, el triángulo como azul y el cuadrado como rojo) en la misma imagen pese a que la figura y el color se procesan en zonas distintas del cerebro. Malsburg aportó dos grandes e influyentes ideas con respecto al problema del enlace. En primer lugar, identificó la correlación entre la actividad neuronal en distintas partes del cerebro como una de las grandes cuestiones a esclarecer en relación con la función cerebral en el ámbito de la conciencia. En segundo lugar, en contra de las antiguas respuestas basadas en la localización de la conciencia en una parte del cerebro (localización de neuronas de integración), propuso soluciones basadas en la correlación temporal.

Posteriormente, apoyándose en experimentos propios basados en la implantación de electrodos en el córtex visual primario (zona V1) de gatos sometidos a estimulación visual así como haciéndose eco de otras evidencias empíricas,[117] Singer y Gray hicieron suya la hipótesis de correlación temporal y defendieron que el enlace entre las diferentes zonas del cerebro, necesario para resolver el problema del enlace, debe llevarse a cabo mediante una actividad de disparo sincronizado de las neuronas pertenecientes a esas zonas dispersas (Singer y Gray 1995). Este disparo sincronizado se produciría en el orden de los

[116] Hay teorías microgenéticas, que tratan de esclarecer cuáles son los mecanismos biológicos específicos que explican al experiencia consciente. Hay también teorías biológicas de tipo filogenético (Damasio 2010; Humphrey 1992), que tratan de explicar la emergencia de la conciencia a través de la evolución.

[117] Logothetis y Schall (1989) y Leopold y Logothetis (1996).

milisegundos.

Igualmente, Crick y Koch se atrevieron a formular la hipótesis de que la conciencia se asocia con la memoria a corto plazo y con patrones de actividad neuronal sincrónica alrededor de la banda de frecuencias gamma, es decir, en torno a los 40 Hz (Crick y Koch 1990). Posteriormente Koch (2004) ha sofisticado la noción de correlatos neuronales de la conciencia, estableciendo una diferencia entre correlatos habilitadores de la conciencia (NCCe, *enabling*) y correlatos específicos de la conciencia (NCCs, *specific*). Los correlatos habilitadores son aquellos sin los cuales no sería posible que existiera conciencia, tales como el abasto de sangre al cerebro, la formación reticular mesencefálica,[118] las zonas no específicas del tálamo o determinados neuromoduladores tales como la acetilcolina y la norepinefrina (o noradrenalina). Los correlatos específicos, por el contrario, conforman el aspecto particular de cada experiencia consciente, tienen su sede principal en el córtex cerebral y no se reducen a zonas primarias ni, en particular, a la zona visual primaria (V1).

El tiempo ha atemperado las expectativas iniciales que se crearon por las teorías de la actividad sincronizada. Tanto Crick (1994) (1994) como Koch (2004) han matizado y suavizado su hipótesis inicial. Koch (2004, 2012) ya solo defiende que únicamente es necesaria la memoria icónica para tener conciencia y, recientemente, ha mostrado públicamente su simpatía por la Teoría de la Información Integrada.[119]

1.3.4.2. Teoría de la Doble Microgénesis

La Teoría de la Doble Microgénesis da respuesta a las evidencias empíricas acerca de la temporalidad de la experiencia consciente en los casos de percepción visual. Bachmann (2000) realiza un exhaustivo metaestudio acerca de los experimentos de metacontraste que le lleva a concluir que la experiencia consciente es el resultado de dos procesos sucesivos: la microgénesis de representación y la microgénesis fenoménica.[120] La teoría de Bachman ha sido refinada a lo largo del tiempo. A continuación me detendré en su versión inicial (Bachmann 2000) y en una versión actualizada (Bachmann y Hudetz 2014). Para finalizar, señalaré algunos de los aspectos positivos y negativos de esta teoría.

Bachmann desarrolla su primera hipótesis sobre la génesis de la conciencia en lo que denominó la Teoría del Recontacto Perceptual.[121] La teoría de Bachmann se apoya en los siguientes conceptos:

- Microgénesis. La conciencia se ha de comprender como una pequeña génesis, es decir, como un proceso que se sostiene en el tiempo durante un breve periodo de procesamiento.

[118] Moruzzi y Magoun (1949).
[119] Cfr. Tononi y Koch (2015).
[120] Ver sección de propiedades temporales 1.2.4 para un resumen de las conclusiones fundamentales de ese estudio.
[121] *Perceptual Retouch Theory*. Bachmann (2000: 195).

- Recapitulación. Bachmann (2000: 38) recupera la difunta teoría de recapitulación de Haeckel, según la cual la ontogenia resume los pasos fundamentales de la filogenia. En la interpretación de Bachmann, la microgénesis de la conciencia resume los pasos fundamentales de la ontogenia y la filogenia, esto es, excitabilidad, sensitividad, perceptividad, intelección y conciencia.
- Sustracción y acumulación. Los procesos de generación de la experiencia consciente se acumulan, es decir, unos procesos necesitan de otros para llevarse a cabo. Es posible medir el tiempo de procesamiento de esos procesos mediante la sustracción o adición de estímulos y la medición del tiempo de reacción en esos casos. En la teoría de Bachmann, sustracción y acumulación son la base para aceptar los experimentos de enmascaramiento mediante metacontraste. El enmascaramiento es, de hecho, el mecanismo empírico que permite accionar (y manipular) la sustracción de la génesis de la experiencia consciente. La variable más importante en estos experimentos es el tiempo entre estímulos.[122] Esta variable permite calcular el retardo de procesamiento de la percepción consciente y es la base para inferir cuál es el tiempo de procesamiento propio de la microgénesis de la conciencia.
- Representación y conciencia. En relación con la acumulación y la sustracción, se diferencian dos microgénesis: microgénesis representacional y microgénesis fenoménica. En la primera se forma una representación mental activa (en el sentido de que puede condicionar los procesos cognitivos), al margen de si es consciente o no. En la segunda, la representación mental activa se vuelve consciente. La segunda, consecuentemente, requiere de la primera.[123]

Sobre estas bases Bachmann concibe la conciencia como un doble contacto, más en concreto, como un contacto que se produce por la acción de sistemas no específicos tras el primer contacto que es el efecto del estímulo en los sistemas específicos el cerebro.[124] Este doble contacto podría entenderse como la repetición del efecto del estímulo en otra zona del sistema nervioso, es decir, un segundo contacto en el que se reproduce el contacto físico inicial, a saber, el de las membranas del individuo con el exterior. El segundo contacto se asocia con la etapa fenoménica, esto es, con la microgénesis fenoménica.

Con la noción de doble acción (sistemas específicos y sistemas no específicos) Bachmann quiere decir que la caracterización de la conciencia y su contenido viene dada por dos procesos. En primer lugar, por un proceso que sucede en un brevísimo lapso de tiempo y en el que intervienen sistemas específicos del cerebro. En segundo lugar, por un proceso cuyo tiempo de procesamiento es, probablemente, algo más largo (hasta 100 msg), que es común a todas las formas de conciencia y en el cual toman parte zonas no específicas del

[122]SOA, *Stimulus-onset asynchrony*: es el tiempo transcurrido entre el inicio del primer estímulo S_1 y el inicio del segundo estímulo S_2. SOA $= t(S_2) - t(S_1)$.

[123]La Teoría del Recontacto Perceptual, en ese sentido, parece aceptar la necesidad representacional. La representación es necesaria para la experiencia consciente. Sin embargo, las versiones posteriores de esa teoría permiten concebir conciencia sin contenido representacional.

[124]El tálamo es el gran concentrador por donde pasan la gran mayoría de señales provenientes de los sentidos (Baars y Gage 2007: 42). Los sistemas específicos (*specific systems*, SP) del tálamo son las partes del tálamo que recogen esas señales, ya sean visuales, auditivas o táctiles. Los sistemas no específicos (*non-specific systems*, NSP) son las partes del tálamo no asociadas con un modo sensorial concreto. Esta distinción se puede extender al córtex. Véase también Gonzalo et al. (1993).

cerebro y en particular algunas zonas talámicas no específicas. Para entender este doble proceso se puede usar la metáfora de un montacargas. En el primer proceso, llamado específico, los agentes específicos colocan en el montacargas la carga que se ha seleccionado, mientras que, en el segundo proceso, llamado no específico, los agentes no específicos accionan el montacargas y toda su carga se eleva "hacia la conciencia", al margen de lo que haya dentro del mismo. En otras palabras, la selección del contenido de la conciencia se hace, de hecho, en la microgénesis de la conciencia.

La Teoría del Recontacto Perceptual ha evolucionado a lo largo del tiempo, pero siempre manteniendo la idea de que la conciencia necesita de dos procesos, una microgénesis representacional y una microgénesis fenoménica. En su versión de Bachmann y Hudetz (2014) se formuló como la combinación de dos procesos diferenciados: el nivel crítico de actividad neuromodulatoria (denominado sistema L, level), de carácter no específico, y los contenidos de la conciencia provenientes de los sistemas de representación (denominado sistema D, data), de carácter específico. La conciencia es el resultado de la combinación (que no la suma) de estos dos procesos. Bachmann y Hudetz simbolizan esta combinación de la siguiente manera:

$$C = L \times D$$

Sin abandonar esta idea básica de la modulación de un proceso sobre otro, Bachmann ha ido añadiendo diferentes descubrimientos con respecto a las posibles bases fisiológicas que dan cuenta de ambos sistemas. En el mencionado trabajo de Bachmann y Hudetz (2014) se sugiere que el sistema D estaría formado por los nervios aferentes (provenientes de los sistemas receptores) que finalizan en las células piramidales de la capa 5 del córtex visual primario. El sistema L, por su lado, se basaría en diversos procesos relacionados con la conexión intracortical que se ven favorecidos por el nivel de los neuromoduladores y, en concreto, por una activación particular del pico neuronal, que se basa en las puertas de calcio y que se produce en sentido inverso al sentido de propagación habitual del potencial de acción. Este mecanismo se denomina activación BAC (*backpropagation activated calcium spike firing*) y tiene la virtud de producir una activación más sostenida en el tiempo y, por tanto, más adecuada para aplicar sobre ella la modulación rápida que corresponde al sistema D.

La Teoría de la Doble Microgénesis es, quizás, una de las que mejor explica las propiedades temporales relacionadas con la percepción visual consciente e integra en ella una gran cantidad de evidencias empíricas provenientes de los experimentos de metacontraste, tales como la lentitud del tiempo de procesamiento de la percepción visual y la posibilidad de acelerar esos procesos. La distinción entre base consciente (sistema L) y contenido consciente (sistema D) introduce un matiz conceptual para diferenciar el carácter subjetivo del carácter cualitativo de la experiencia consciente, así como para explicar cómo es posible tener experiencias conscientes con diferentes contenidos, (o, incluso, sin contenido alguno) y por qué la experiencia consciente puede ser vivida como dotada de cierta unidad. No obstante, a pesar de estos aspectos positivos, la Teoría de la Doble Microgénesis es un proyecto en progreso que todavía está lejos de tener un aspecto finalizado. De nuevo no explica la subjetividad propia de la experiencia consciente y está todavía parcialmente indeterminada en cuanto a los procesos exactos que dan lugar a la experiencia consciente.

1.3.5. Teorías físicas

Las teorías biológicas de la conciencia son, evidentemente, también teorías físicas, en la medida en que tratan de explicar los mecanismos físicos sobre los que sucede la experiencia consciente. Sin embargo, priorizan los mecanismos neuronales y biológicos sobre otros mecanismos puramente físicos. Por el contrario, existen cierto grupo de teorías que tratan de explicar la experiencia consciente a través de fenómenos exclusivamente físicos y no (esencialmente) biológicos. Estas son las denominadas teorías físicas de la conciencia. La Teoría de la Conciencia Cuántica (Penrose 1994; Hameroff y Penrose 1996) es quizás la que más eco ha tenido hasta el momento. A continuación revisaré brevemente en qué consiste y cuáles son sus argumentos para, seguidamente, mencionar algunos de los problemas que hacen que, hoy por hoy, esta teoría no sea una alternativa aceptable.

Los argumentos en favor de esta teoría son fundamentalmente dos. El primero surge de una interpretación especial del teorema de Gödel (Lucas 1961; Penrose 1994). Este teorema indica que todo sistema consistente S será un sistema incompleto, puesto que siempre cabe establecer una proposición (la proposición de Gödel) tal como "esta proposición no es demostrable en S", cuya validez conlleva su falsedad (y a la inversa) dentro del sistema S. Lucas (1961) sostiene que la proposición de Gödel es la prueba de que la mente humana no puede nunca asimilarse (contra los defensores de la inteligencia artificial fuerte) a una máquina de Turing, puesto que la mente puede resolver el problema expuesto por Gödel y comprender sin dificultad la verdad de su proposición, mientras que una máquina de Turing (o, en general, cualquier máquina de computación) nunca podría resolverlo. Penrose, por su parte, va un paso más allá al sostener que la proposición de Gödel demuestra que la mente no es asimilable a la computación clásica, y que para su explicación es necesario introducir elementos no computacionales.[125] A juicio de Penrose, la mecánica cuántica es el único campo que ofrece respuestas de este tipo.

El segundo argumento está relacionado con la temporalidad especial de la conciencia que se arroja de los experimentos de Libet (Libet et al. 1979), de acuerdo con los cuales existe una actividad cerebral previa a la decisión consciente, lo que parece sugerir que no es posible compatibilizar un monismo materialista con la idea de libre voluntad. Para Penrose, esta compatibilidad es posible siempre y cuando se explique el periodo preconsciente en términos de computación cuántica, en lugar de en términos de computación clásica.

Penrose y Hameroff sugirieron que la conciencia podía estar causada por un fenómeno puramente cuántico que denominaron "reducción objetiva" (OR, *objective reduction*). La reducción objetiva es un colapso de la función de onda debido a causas intrínsecas. El

[125]La conciencia es el producto de una acción física específica del cerebro, pero esta acción física nunca puede ser simulada computacionalmente de manera apropiada (Penrose 2002: 227). Esta afirmación de Penrose se opone a la idea de que la mente se puede replicar con una máquina de Turing (esto es, la denominada hipótesis de inteligencia artificial fuerte) y a la idea de que la mente puede ser simulada computacionalmente (esto es, la denominada hipótesis de inteligencia artificial débil).

colapso de la función de onda supondría la transición de un periodo preconsciente al momento de conciencia. Por otro lado, debido a los estados de superposición cuántica previos al colapso, el autocolapso tendría ciertas características especiales: es no computable, es múltiple en el espacio y en el tiempo y está orquestado.

Dando un paso más allá, Penrose y Hameroff especularon con la posibilidad de que la reducción objetiva fuese una reducción de los estados de masa (asociados a las posiciones del electrón) de las tubulinas, que son proteínas de dos partes que componen los microtúbulos de los que, a su vez, se componen las células neuronales.[126] Estos estados de masa, que se asocian con propiedades piezoeléctricas y ferroeléctricas específicas, se encontrarían en superposición en el momento preconsciente, y colapsarían en el momento de la conciencia. Sin embargo, el colapso de los estados de una única tubulina no es suficiente para explicar un hecho macrofísico (y macrofísicamente unificado) como la conciencia. De ahí que Penrose y Hameroff propongan en su lugar estados de coherencia masiva donde, a través de las conexiones entre los microtúbulos (a saber, las proteínas asociadas a los microtúbulos o MAPs, *micro-tubule associated proteins*), se producía, a su vez, una reducción objetiva orquestada de todos los estados de las tubulinas.

Con respecto a los problemas de implementación, la cuestión es, si cabe, más compleja. Se trata de establecer las condiciones bajo las cuales es posible esa secuencia de fenómenos cuánticos que explica la conciencia. En otras palabras, cómo y dónde se implementa la reducción objetiva orquestada a partir de los estados de coherencia masiva de los estados de masa de las tubulinas propios de los estados preconscientes. Penrose y Hameroff (1996: 461-462) aventuraron que estos estados podrían justificarse a partir de la coherencia de Fröhlich, que únicamente se podría dar en determinadas condiciones ambientales. [127]

Por desgracia, las condiciones ambientales bajo las cuales puede darse la coherencia de Fröhlich son tan restrictivas que resulta impensable su aplicación para el caso particular de los microtúbulos de las células nerviosas. Los estudios de Reimers acerca de la coherencia de Fröhlich (Reimers et al. 2009) indican que estos estados de coherencia solo se podrían dar en situaciones de muy baja temperatura (y muy baja energía) que son totalmente incompatibles con la existencia de vida.

En la actualidad, la Teoría de la Conciencia Cuántica permanece en punto muerto, hasta el punto de que el propio Ned Block la excluye de las denominadas "teorías de la concien-

[126] Hameroff y Penrose (1996)

[127] Herbert Fröhlich [...] también predijo el fenómeno de la coherencia cuántica en células vivas. [...] Fröhlich teorizó que los conjuntos de dipolos proteínicos sometidos a un campo electromagnético común (e.g. proteínas dentro de una membrana polarizada o subunidades dentro de un polímero electreto como los microtúbulos) sufren una excitación conformacional coherente cuando se les suministra energía. Fröhlich postuló que esa energía puede provenir de la energía bioquímica y térmica del "baño caliente" del entorno. De acuerdo con Fröhlich, es posible que emerjan procesos organizados y cooperativos conducentes a excitaciones coherentes, debido a la coherencia estructural de los dipolos hidrofóbicos sometidos a un gradiente de voltaje común. Fröhlich dedujo las frecuencias de excitación coherente [...] (que están en la región espectral de las microondas, de orden de los GHz) y los denominó transiciones acusto-conformacionales o fonones coherentes. En física cuántica, estos estados de coherencia se denominan condensados de Bose-Einstein y se ha sugerido [...] que pueden crear los estados cuánticos macroscópicos que soportan el enlace unitario propio de la conciencia.(Hameroff y Penrose 1996: 461-462)

cia que los neurocientíficos se toman en serio" (Block 2009: 1111).

1.3.6. Teorías entrelazantes

Al principio de esta sección se delimitó el ámbito de lo que se denominan teorías de la conciencia tomando como base la distinción entre *explanans* y *explanandum*. Las teorías de la conciencia son aquellas en las que el hecho de la conciencia forma parte esencialmente como *explanandum*, mientras que otros hechos (tales como la representación, la información, los procesos físicos o los procesos biológicos) forman parte esencialmente como *explanans*.

De ese modo, se retiraron del escenario las teorías que emplean la conciencia como *explanans*, tales como la teoría de Wigner (1961), según la cual el hecho de la conciencia explica algunos fenómenos de la mecánica cuántica y, en concreto, el problema de la medida (*explanandum*).[128] También se retiraron del escenario las llamadas teorías "acerca de todo" en las cuales la conciencia forma parte de una serie de *explananda* y, por consiguiente, no forma parte esencial de la teoría. Tal es el caso, por ejemplo, de la teoría trascendente del campo unificado u otras teorías y pseudoteorías que explican todo el resto de fenómenos a partir de una unidad prevalente de la cual todo depende de manera sistémica.

No obstante, se dejó abierta la posibilidad a una serie de teorías que se podrían llamar "de segundo nivel" en las que el hecho de la conciencia se entrelaza con otros fenómenos en redes de relaciones *explanandum*/*explanans*. Este tipo de teorías no constituyen la parte principal de las denominadas teorías de la conciencia, pero merecen una breve mención en la medida en que han de adquirir importancia cuando se vayan esclareciendo algunos de los subproblemas y misterios que rodean el hecho de la conciencia.

Entre el conjunto de teorías entrelazantes destacan los trabajos seminales de Albert y Loewer (Albert y Loewer 1988) y también de Lockwood (1989, 1996).[129] En ambos casos se exploran las inquietantes consecuencias que se derivan de la articulación del hecho de la conciencia en el contexto de la física cuántica y dentro de la interpretación de los muchos mundos, frente a otras interpretaciones tales como la de Copenhague que sigue Penrose,[130] o la teoría basada en variables ocultas de Bohm. Lockwood (1989) defiende que el colapso de la función de onda colapsa es un *deus ex machina* que se introduce en la cadena de mediciones sin base científica alguna.

Si la función de onda no colapsa, entonces lo más plausible[131] es que el estado macros-

[128]Para ser justos, se podría alegar, con cierta razón, que la posición de Wigner (1961) no es exclusivamente de este tipo, puesto que expone no una sino dos tesis: 1) que la conciencia solo surge en determinadas condiciones físico-químicas y 2) que la conciencia provoca el colapso de la función de onda.

[129]Véase también Chalmers y McQueen (2021) y Okon y Sebastián (2018).

[130]Hameroff y Penrose (1996). Chalmers y McQueen (2021) también avanzan hacia una posible interpretación de la conciencia como causa del colapso de la función de onda, en la línea de Wigner (1961), y en el ámbito de la interpretación de Copenhague.

[131]*Pace* Bohm

cópico del fenómeno cuántico observado (el gato de Schrödinger observado tras abrir la caja) sea solo el estado de uno de muchos mundos que discurren en paralelo. Esta es la interpretación de los múltiples mundos de DeWitt (1970). De acuerdo con ella, todas y cada una de las posibilidades en superposición forman parte de la descripción completa del mundo físico.

En la interpretación de los múltiples mundos el problema de la conciencia se convierte en el problema de cómo hacer compatibles los siguientes postulados:

1. Los estados físicos son de naturaleza mecánico-cuántica y todos siguen la ecuación de Schrödinger.[132] Por consiguiente, no sucede algo así como un colapso de la función de onda.
2. No hay estados en el mundo que no sean físicos, es decir, no existe algo como un dualismo de substancias. Los estados mentales deben sobrevenir de acuerdo a estados físicos.
3. Una explicación del mundo nunca será completa a menos que reconcilie la apariencia y la realidad.
4. Y, sin embargo, una mente no puede percibir los estados superpuestos. Por consiguiente, percibe los estados como si no hubiera coherencia.

La modulación de cada uno de estos postulados ofrece diferentes soluciones al problema de la compatibilización del hecho de la conciencia y los fenómenos físicos. Lockwood (1996), por ejemplo, defiende escrupulosamente el postulado (2), a saber, que los estados mentales sobrevienen de acuerdo a los estados físicos, y propone que existe un estado compuesto de la mente (que denomina MENTE o *MIND*, en mayúsculas), en el cual se encuentran a su vez las diferentes instancias de mente (*mind*, en minúsculas) correspondientes a los diferentes mundos. La explicación de la MENTE depende, por consiguiente, de fenómenos cuánticos, pero los fenómenos cuánticos no son, por sí mismos, suficientes para explicar la MENTE en la medida en que no lo son para explicar las diferentes instancias de mente que se encuadran en cada mundo posible. En otras palabras, cada instancia de mente es un estado ortogonal puro, independiente del resto de instancias mentales correspondientes a otros mundos, si bien la MENTE depende de todas esas instancias. En esto consiste precisamente la naturaleza entrelazante de la teoría de Lockwood.

1.4. Conclusiones

La conciencia es un hecho específico, que tiene propiedades específicas y que difícilmente se puede reducir a otros conceptos. Al mismo tiempo, la conciencia es un hecho complejo, que tiene muchas formas y aspectos. Cuando uno trata de reducir la conciencia a una de esas formas, la conciencia resulta ser especialmente elusiva y escurridiza. Se ha tratado de reducir la conciencia a representación, pero eso implica negar la posibilidad de que

[132] Albert y Loewer (1988: 203).

exista una experiencia de conciencia pura, esto es, arrepresentacional, y no todo el mundo está dispuesto a aceptar tal cosa. Se ha tratado de reducir la conciencia a atención, sin embargo, existen numerosas evidencias de disociación entre atención y conciencia. Se ha tratado de reducir la conciencia a procesos físicos o fisiológicos, pero la experiencia consciente parece tener un reino experiencial y una subjetividad propias que difícilmente se pueden explicar si empleamos únicamente nociones del ámbito de los procesos físicos. Se ha tratado de reducir la conciencia a una dimensión experiencial, pero resulta que esa dimensión depende fuertemente de aspectos tan físicos y mundanos como el tiempo de presentación de los estímulos (tal y como se deduce del efecto de retardo de destello) o el consumo de drogas y medicamentos.

La conciencia es un hecho específico y es difícil de definir. Pero se puede describir, al menos típicamente, a través de sus propiedades. La propiedad principal es que en la experiencia consciente hay algo que es el "cómo es" estar esa experiencia para el sujeto que la tiene. Además, si la conciencia tiene contenido, ese contenido es accesible, es decir, existe la posibilidad de decir, razonar o actuar con respecto a él, sea de manera directa o indirecta. De hecho, si no hubiese accesibilidad de ningún tipo, nada se podría saber acerca de la conciencia. Por otro lado, la experiencia consciente tiene típicamente cierta composición, es decir, está compuesta de partes experienciales. Si no tuviera cierta composición, no podría existir algo como la fenomenología. Paradójicamente, esa composición se vive siempre "como dotada" de cierta unidad. La conciencia se vive como una para el sujeto en cada momento y a lo largo del tiempo. Por último, la conciencia sucede en el tiempo: tiene un principio y tiene un final.

Las teorías de la conciencia es el modo en que tratamos de explicar la experiencia consciente. Estas teorías ofrecen diferentes explicaciones, que pueden apoyarse fundamentalmente en otros fenómenos de tipo representacional, informacional, biológico o físico. Aquí he puesto el foco en cuatro teorías diversas que, por su capacidad para integrar la evidencia empírica o por su poder explicativo, disfrutan de una posición de privilegio entre las demás. Pero el hecho específico de la conciencia es complejo, y no todas las teorías explican lo mismo. Es decir, las teorías se distinguen también por aquello que tratan de explicar. La Teoría Autorrepresentacional se especializa en explicar el carácter subjetivo de la conciencia, es decir, el hecho de "cómo es" *para mí* estar en una experiencia consciente. La Teoría del Espacio de Trabajo Global se especializa en explicar la experiencia consciente que tiene contenidos que son atendidos, y que sucede tras cierto tiempo, es secuencial, limitada en recursos y propensa a errores, pero también tiene una gran capacidad relacional y es muy sensible al contexto. La Teoría de la Información Integrada, por su parte, explica especialmente bien cómo es posible que la experiencia consciente sea tan diversa y, al mismo tiempo, se presente con cierta forma de unidad. Por último, la Teoría de la Doble Microgénesis está enfocada a explicar cómo sucede la experiencia consciente a través del tiempo, por qué cuesta el tiempo que cuesta y por qué ese tiempo no siempre es el mismo, sino que puede variar.

El resultado de esta enorme variabilidad teórica en cuanto a *explanans* y en cuanto a *explananda* es que, a menudo, resulta complejo comparar y contraponer las teorías entre sí. Esto no quiere decir que las teorías sean compatibles entre ellas, puesto que proponen

principios incompatibles entre sí. Solo dos ejemplos. La Teoría de la Doble Microgénesis defiende una estructura procesual que no tiene reflejo en la Teoría de la Información Integrada. Igualmente, las nociones de representación, así como los soporte neuronales propuestos en el ámbito de la Teoría Autorrepresentacional y la Teoría del Espacio de Trabajo Global son divergentes y, quizás, irreconciliables.

Sin embargo, sería injusto decir que las teorías carecen absolutamente de aspectos compatibles o, por así decirlo, aspectos con respecto a los cuales son mutuamente neutras.[133] Uno de estos aspectos es que muchas de las teorías (y, más en concreto, las cuatro teorías aquí presentadas) conllevan, de manera explícita o implícita, la idea de que la conciencia sucede a través del tiempo, como si fuera un proceso. Pero existen más. Dedicaré el capítulo 3 enteramente a dar razones en favor de esta idea. Pero, antes que nada, es necesario establecer un marco conceptual para la articulación de la noción de proceso.

[133] A este hecho lo denomino el "principio de neutralidad mutua" entre las teorías. Un grupo de teorías cumplen el principio de neutralidad mutua con respecto a un aspecto cuando las teorías son compatibles entre sí en relación a este aspecto determinado.

Capítulo 2

Proceso

2.1. Introducción: un marco procesual para la conciencia

Cuando uno se enfrenta a un hecho tan complejo y elusivo como la conciencia y, con cierta impotencia, comprueba una y otra vez que las dificultades no parecen resolverse, sino que, más bien al contrario, se multiplican y son cada vez mayores, es legítimo cuestionarse si las herramientas de análisis que se vienen empleando son las adecuadas. Tradicionalmente, el análisis de la conciencia está poblado de conceptos propios de la metafísica de la sustancia, tales como la idea de un yo sustancial o *res cogitans*, la idea de sujeto, la idea de estado (estados mentales, estados representacionales, estados conscientes) y la idea de inmediatez (percepción inmediata, aprehensión directa). Sin embargo, la experiencia consciente tiene algunas propiedades que resultan desconcertantes, por lo menos a la luz de los conceptos anteriores. La experiencia consciente sucede a lo largo del tiempo. Es lenta y esta lentitud no es algo puramente anecdótico, sino fundamental. Tiene cierta estructura y, sin embargo, dentro de esa estructura posee también cierto aspecto de unidad. Fluye y cambia. Y, al mismo tiempo, parece cambiarse a sí misma en el curso de ese fluir.

Aquí voy a proponer que es muy recomendable sustituir algunos de estos antiguos conceptos por otros nuevos. Para ello es imprescindible volver la mirada hacia la noción de proceso y la idea de dinamismo. El objetivo de este capítulo es ofrecer un marco conceptual acerca de la idea de proceso que pueda servir de firme base para el análisis de la conciencia.

Para ello, en primer lugar, introduciré de manera general qué es la filosofía de procesos. Mi objetivo no es hacer una defensa cerrada y dogmática de esta filosofía, sino tan solo poner el énfasis en su parte más positiva, es decir, en su aportación heurística. A esto se dedicará el apartado 2.2 Después definiré aquello que se entiende por proceso transformativo y por proceso autotransformativo, para lo cual voy a recurrir a una mereología

no clásica —inspirada en la mereología de niveles LEM de Seibt (2015)— que sea compatible con una composición dinámica concreta, a saber, de tipo bucle. Esto constituirá el núcleo del apartado 2.3 y sentará las bases para, a partir de ahí, articular una analítica de procesos con la cual se pueda entender mejor en qué consiste la experiencia consciente como proceso.[134] A continuación, en el apartado 2.4, expondré la idea de modulación y clausura, tomando como base las nociones de proceso transformativo y proceso autotransformativo. Como la noción de clausura de modulación no tiene precedentes en filosofía, aprovecharé el marco matemático que ofrece la dinámica de sistemas para poder caracterizar esta noción de manera sólida. Para acabar, pondré en relación la noción de clausura en la modulación y los procesos autotransformativos

El objetivo final es establecer un marco conceptual para el análisis procesual de la experiencia consciente de acuerdo con sus propiedades principales (las propiedades fenoménicas, funcionales, estructurales y temporales) en el cual no sea necesaria la idea de sustancia ni dar por hecho la existencia de algo así como un yo sustancial.

2.2. Qué es un proceso

Desde que Aristóteles incluyese la sustancia (οὐσία) como una de sus categorías, la ontología ha estado dominada por la distinción entre sustancia y propiedades, con especial énfasis en las propiedades de ubicación temporal y espacial.[135] Esta tendencia mayoritaria, que algunos (Rescher 1996, 2000; Seibt 2020) han venido a denominar genéricamente "metafísica de la sustancia", ha priorizado las cosas con toda su complejidad (mesa, silla, sol, yo, persona, sociedad, mundo...) sobre los procesos y, si bien no supone la negación de la posibilidad del cambio, a menudo ha supeditado su explicación a la sustancia —aquello que permanece— y sus propiedades —aquello que cambia—, siendo el movimiento (en tanto que modificación de las propiedades de lugar a lo largo del tiempo) una de sus categorías principales. La metafísica de la sustancia también ha tenido su reflejo en filosofía de la mente, especialmente en la idea cartesiana del yo como sustancia.[136]

Pese a su posición hegemónica, la metafísica de la sustancia no está exenta de polémica. Principalmente se expone a dos problemas. El primero se refiere a su alcance y responde a la cuestión de si existe para todas las cosas algo permanente (la sustancia) que sea el sustrato de cambio. Por el contrario, algunos fenómenos difícilmente se pueden entender

[134]Hay una influyente tradición en filosofía de procesos abanderada por Whitehead (1978 1929) que no es analítica y que necesita de otros conceptos tales como las "entidad actual", "objeto eterno", "prehensión", "creatividad" y "concrescencia". La analítica de procesos que se presenta aquí no es whiteheadinana, sino que se basa en Rescher (1996, 2000).

[135]Lo que no quiere decir que haya que entender necesariamente a Aristóteles como el gran exponente de la metafísica de la sustancia. En Aristóteles, la sustancia de un ser natural se puede entender como el conjunto de propiedades esenciales, no accidentales, y limitadas en número que son necesarias y suficientes para que un individuo pertenezca a una especie determinada.

[136]Hasta qué punto esta idea de yo sustancial está en Descartes es, también, una cuestión no exenta de controversia. Se podría alegar que detrás del yo "sustancial" de Descartes no están sino los *cogitos* y nada más que ellos.

como asociados a un sujeto que permanece (y, por extensión, a una sustancia que permanece). Este es el caso de una tormenta de nieve, las mareas o un torbellino de viento (Rescher 1996, 2000; Seibt 2020). Todos esos fenómenos existen, igual que existen las cosas o los animales y, sin embargo, carecen de un elemento que permanezca durante el cambio. Por el contrario, consisten en el mismo cambio, al margen de que los elementos que forman parte de ese cambio permanezcan o no. ¿Cómo es posible entenderlos desde la noción de sustancia?

El segundo problema se refiere a su coherencia y responde a la cuestión de si es posible una distinción real entre sustancia y propiedades, esto es, si puede darse sustancia sin propiedades. Si así fuera, la sustancia sería aquello que queda después de eliminar todas las propiedades. Sin embargo, no hay nada que quede después de eliminar de algo *todas* sus propiedades (Zubiri 1989), de igual manera que no queda alcachofa después de retirar todas sus hojas.

Por ese motivo, desde Heráclito hasta Whitehead (1978 1929) han surgido excepciones e impugnaciones a esta forma de hacer metafísica. El primer giro de Heráclito se basó en tres afirmaciones complementarias que resultan claves para comprender y contextualizar la posterior revuelta contra la idea de sustancia:[137]

1. El concepto de cambio debe pasar de ser aquello que hay que explicar (*explanandum*) a ser algo que tiene también valor explicativo (*explanans*).
2. El concepto de cambio no implica necesariamente desorganización o caos. Por el contrario, puede constituir unidad organizacional.
3. Es necesario distinguir de manera genérica dos formas de dinamismo: la dinámica transicional, por un lado, y la dinámica de permanencia, por otro lado.

En este sentido, una de las aportaciones de Heráclito para explicar la dinámica de permanencia es la introducción de la noción de conexión por tensionamiento mutuo.[138] El tensionamiento mutuo consiste en el efecto de tensión equilibrada entre las partes de un elemento y explica cómo la lira puede emitir sonidos o cómo el arco puede impulsar flechas a largas distancias. En estos casos, es precisamente la dinámica de tensiones, y no otra forma preestablecida, lo que asegura la permanencia del objeto y aquello que explica su función.

Posteriormente, Leibniz y, especialmente, la filosofía del organismo de Whitehead constituyen dos ejemplos conspicuos de sistematización que, cuando menos, cuestiona la metafísica de la sustancia desde la priorización de la categoría de proceso.[139]

En general, los argumentos en favor de una filosofía enfocada al proceso se pueden resumir

[137] Ver Seibt (2020). Seibt (2009) ha usado la expresión "el mito de la sustancia".
[138] *Palíntonos harmoniē*.
[139] Para un detalle exhaustivo de la historia de las filosofías del proceso, ver Rescher (1996) o Seibt (2020).

en tres afirmaciones básicas (Seibt 2020):

- Primero, la aceptación de una sustancia como sujeto que subyace o como poseedor de propiedades (que constituye la base de la denominada metafísica de la sustancia) no es una condición indispensable de las leyes del pensamiento. Al contrario, se puede prescindir de ella.
- Segundo, todas las categorías familiares a la metafísica de la sustancia (entre ellas, sustancia, propiedades, estados...) se pueden reformular en términos de filosofía del proceso.
- Tercero, ciertos hechos no pueden ser completamente explicados en el marco de una metafísica de la sustancia, sino que, por el contrario, solo pueden ser completamente explicados desde la noción de proceso.

Ni la primera ni la segunda afirmación están libres de cierta controversia. Ambas forman parte del debate que Rescher (1996, 2000) ha denominado con cierta ironía la "revuelta contra la filosofía del proceso". Su discusión con Strawson (1959) constituye un buen ejemplo de ello. Aquí, sin embargo, voy a obviar esa interesante polémica y voy a centrarme en la tercera afirmación, a saber, que existen hechos en la realidad que nos obligan a modificar nuestro marco categorial para atender a los procesos en lugar de a las sustancias. La tesis que defenderé es que uno de esos hechos es la experiencia consciente.

2.2.1. Proceso, proceso general y proceso transformativo

A continuación voy a introducir de manera breve un marco conceptual que permita emplear con propiedad la noción de proceso. La introducción de este marco conceptual es oportunista y se orienta al fin que se busca aquí, esto es, la explicación de la conciencia.

¿Qué es proceso? Por proceso se entiende "un grupo de cambios coordinados en la complexión de la realidad" y "una familia organizada de ocurrencias que están sistemáticamente enlazadas entre sí, bien causalmente, bien funcionalmente" (Rescher 1996: 38). Quizás esta definición no sea la mejor posible, pero es innegable que ofrece dos claves para el análisis de procesos: (1) los procesos son entidades complejas (familias, grupos de cambios) y, por consiguiente, para explicarlos es necesaria una mereología de procesos y (2) los procesos se determinan de manera funcional y causal, lo que quiere decir que no se pueden determinar únicamente por su posición en el espacio-tiempo.

Seibt (2009: 484), por su parte, introduce la noción de proceso general por oposición a la noción de proceso particular.[140] Los procesos particulares son aquellos que (1) están

[140]La definición de proceso general dada por Seibt (2009: 484) es muy técnica y puede resultar difícil de seguir. En términos de Seibt, un proceso general es un proceso no particular (en el sentido de no completamente determinado espacio-temporalmente), independiente, individual y concreto (en el sentido de determinado funcionalmente), con extensión espacio-temporal y dinámico. De acuerdo con esta definición, un tornado, por ejemplo, es un proceso general. Es general, puesto que no se puede determinar

determinados por una única posición en el espacio-tiempo y (2) su posición es necesaria para la individuación de la entidad referida. Por el contrario, los procesos generales son procesos que no están determinados por una referencia espacio-temporal, sino por su función. Este es el caso de procesos puramente dinámicos tales como los referidos por las expresiones "está nevando", "este tornado" o "esta canción". "La individualidad [de un proceso] está unida a la *especificidad-en-su-función*" dice Seibt (2009: 486). La sospecha de Seibt es que este tipo de procesos son más comunes y fundamentales de lo que parecería a primera vista.

¿Qué tipos de procesos cabe encontrar en la realidad? De manera muy general, Rescher (1996) distingue entre procesos productivos y procesos transformativos. Estos dos tipos de procesos se distinguen por su mereología, esto es, por las relaciones que tienen sus partes entre sí y con el todo.[141] La noción de proceso productivo no amenaza a la metafísica de la sustancia: un proceso productivo es aquel encaminado a la producción de un producto que puede ser entendido como una sustancia. En una fábrica de automóviles, por ejemplo, existen multitud de procesos productivos cuya suma resulta en la producción en serie de automóviles. La mereología de estos procesos consiste, fundamentalmente, en el ensamblaje de las piezas de las cuales están compuestos: el todo consiste en la suma de sus partes. En contraposición, los procesos transformativos son aquellos "que únicamente transforman estados de hechos en general y hacen posible que sucedan otros procesos ulteriores sin que resulten en cosas particulares" (Rescher 1996: 41). La dinámica del motor de un coche arrancado, por ejemplo, puede ser considerada como un proceso transformativo que se sostiene en el tiempo. La mereología, en este caso, es diferente. El todo no consiste únicamente en la suma de sus partes. Es algo más que eso.

Las denominadas "filosofías del proceso" no constituyen algo homogéneo. Una filosofía del proceso puede distinguirse de otra en su grado de rechazo de la sustancia y, por extensión, en su grado de rechazo de los procesos productivos, pudiendo existir un amplio abanico desde las posturas más radicales a las más atemperadas. No obstante, toda filosofía de proceso, para que se pueda denominar como tal, se caracteriza por prestar una especial atención a los procesos transformativos y acepta la idea de que este tipo de procesos están a la base o son fundamentales para la explicación de la realidad. Rescher dice literalmente que la filosofía del proceso "se caracteriza por su insistencia en la fundamentalidad de los procesos transformativos, y su diferenciación potencial de las cosas sustanciales" (Rescher 2000: 28).[142]

¿Cómo se pueden analizar los procesos transformativos? Para empezar, un proceso transformativo es algo que sucede a lo largo del tiempo y tiene una ocurrencia compleja.

de acuerdo a una referencia espacio-temporal (la referencia cambia en cada instante) y, sin embargo, se extiende en el espacio-tiempo, es individual (es ese tornado y no otro) y es dinámico.

[141] Para una definición completa de mereología, véase Varzi (2019).

[142] Zubiri, por su parte, distingue entre dinamismo de variación (que varía la sustantividad —que no sustancia— sin que deje de ser lo que es) y dinamismo de alteración (que hace a la sustantividad un otro, un *alter*) y afirma que las variaciones son "el dinamismo básico sobre el que están apoyados o montados todos los demás dinamismos" (Zubiri 1989: 127). Esta afirmación de Zubiri encierra gran similitud con la de Rescher.

Tener una ocurrencia compleja quiere decir que un proceso transformativo sucede entre elementos, esto es, consta de partes. De acuerdo con Rescher (1996: 23) un proceso es "mereológicamente homogéneo", en el sentido de que las partes de los procesos son siempre, a su vez, procesos.[143] De hecho, la relación *es-parte-de* constituye una de las categorías fundamentales de la filosofía del proceso (Seibt 2009). Por consiguiente, las partes de un proceso transformativo son, a su vez, procesos.

2.2.2. Tipos de procesos según la forma mereológica

Tal y como hemos visto más arriba, dentro de la polémica entre metafísica de la sustancia y filosofía del proceso, los defensores de esta última argumentan que todas las categorías familiares de la metafísica de la sustancia se pueden articular desde la filosofía del proceso, mientras que no sucede a la inversa y que, por consiguiente, la filosofía del proceso es ontológicamente superior.

Para apoyar esta superioridad ontológica, Seibt (Seibt 2009; Seibt et al. 2017) propone una taxonomía aplicable a los procesos donde caben desde las cosas y materiales más discernibles ("una mesa", "una silla", "un gato", "agua", "oro") hasta los entes más escurridizos desde el punto de vista de la metafísica de la sustancia tales como las actividades y los desarrollos ("una tormenta", "un torbellino de aire", "una explosión", "la formación de un volcán" o "el choque de dos placas tectónicas"). Nuestra tarea posterior consistirá en ver si la conciencia se acerca más al primero o al segundo tipo de cosas.

La taxonomía propuesta por Seibt está basada en dos nociones fundamentales: automeridad y homomeridad. Estas nociones sirven como dos ejes para distinguir los diferentes elementos que forman parte de la realidad.

2.2.2.1. Homomeridad

La homomeridad mide la capacidad que tiene un proceso de tipo T para que sus partes sean del mismo tipo T. Por tanto, la homomeridad es una noción que se predica acerca de un proceso p y acerca de un tipo T. La homomeridad puede ser temporal o espacial. Un proceso es altamente homomérico con respecto a lo temporal si se puede tomar cualquier parte temporal de su ocurrencia y esa parte sigue siendo del mismo tipo T que el proceso original. Por ejemplo, un gato es un tipo animal durante toda su vida sin importar la parte temporal de su existencia que se escoja, esto es, sin importar si se escogen los tres primeros años de su vida, los tres últimos o cinco minutos cualesquiera de su existencia. De igual manera, un proceso es escasamente homomérico con respecto a lo temporal si las partes o subregiones temporales de su ocurrencia no son del mismo tipo que el proceso original. Por ejemplo, la fase acuática de la rana siendo todavía renacuajo no es del mismo

[143] En otras palabras, un proceso es homomérico con respecto a la característica (al tipo) de "ser proceso". Ver en la sección 2.2.2 la definición de homomeridad en el sentido de Seibt (2009).

tipo que el proceso de ser rana tomado en su totalidad.

Con respecto a lo espacial, un proceso es altamente homomérico si se puede tomar cualquier parte espacial y esa parte es del mismo tipo que el proceso total. Por ejemplo, una línea recta sigue siendo del tipo "recto" si es dividida por la mitad, por un cuarto o por un octavo. Por el contrario, una línea en zigzag admite una división (a saber, cada uno de los segmentos rectos que la conforman) para la cual su homomeridad espacial será mínima, puesto que cada segmento será del tipo recto mientras que el proceso completo será del tipo no recto.

2.2.2.2. Automeridad

La automeridad es una noción más fuerte que mide la capacidad de un proceso para ocurrir dentro de sus partes espacio-temporales. Un proceso es máximamente automérico con respecto a lo temporal si el proceso ocurre completamente en cualquier parte dentro de su región temporal de ocurrencia. Este es el caso de procesos del tipo "actividad", como, por ejemplo, "está nevando" o "la leña arde". En cada pequeño instante se cumple que está nevando, está ardiendo. Por otro lado, un proceso es mínimamente automérico con respecto a lo temporal si es necesario tomar el intervalo de tiempo completo en su región temporal de ocurrencia para que ocurra. Este es el caso de procesos de tipo desarrollativo (de dinámica transicional, se podría decir), también llamados "eventos", tales como "convertirse en mariposa", "explosionar" o "formar una familia". En este tipo de procesos se puede hablar propiamente de fases del proceso.

La noción de automeridad también se aplica a lo espacial. Un proceso es máximamente automérico con respecto a lo espacial si en toda subregión dentro de su región espacial de ocurrencia el proceso ocurre completamente. El agua contenida en un vaso, por ejemplo, tiene automeridad espacial muy alta, que llega hasta el nivel atómico, mientras que un coche tiene una automeridad espacial baja, casi mínima, puesto que no hay partes dentro de él que sean un coche. La noción de automeridad se puede aplicar también a los procesos de tipo actividad. Una actividad como la referida por la expresión "el coro de la iglesia está cantando" puede tener automeridad media-baja en lo espacial, mientras que otra actividad como la referida por la expresión "la leña arde" puede tener automeridad alta en lo espacial.[144]

2.2.2.3. Tipos de procesos

Cada proceso que ocurre tiene una forma mereológica concreta, de acuerdo con su homomeridad y su automeridad. Esto permite a Seibt (2018) proponer una serie de tipos

[144]"El coro de la iglesia está cantando" tiene una automeridad espacial media-baja. Es posible tomar medio coro y que la actividad siga ocurriendo. Pero no es posible tomar una sola persona del coro (o la pierna de esa persona) y que siga ocurriendo esa actividad. Por el contrario, la actividad de arder tiene una automeridad espacial alta. Se puede coger cada parte del trozo de leña que está ardiendo y decir propiamente que ocurre en él la actividad, esto es, que "arde".

Seibt (2018)	AE	AT	HE	HT	Ejemplos
tipo 1- actividades		alta			una nevada, una tormenta, un torbellino de aire
tipo 2- materiales	normal	alta			agua, aceite
tipo 3- cantidad de material	baja	alta	normal		un litro de agua
tipo 4- desarrollos		baja		baja	una explosión, la formación de un volcán, la metamorfosis de una oruga en mariposa
tipo 5- cosas	baja	alta	baja		una taza, una mesa, un gato

Cuadro 2.1: Clasificación de procesos de acuerdo a su automeridad y a su homomeridad (Seibt 2018). La noción de proceso parece ser capaz de incluir cualquier elemento de la realidad que sucede en el espacio-tiempo. Nota: AE: automeridad espacial, AT: automeridad temporal, HE: homomeridad espacial, HT: homomeridad temporal.

de procesos o dinámicas, algunas de las cuales ya han ido apareciendo en los ejemplos anteriores. En concreto:[145]

- El tipo 1 de las dinámicas es el denominado de actividad, que se caracteriza por tener una automeridad alta en lo temporal, mientras que en lo espacial su automeridad puede ser alta, normal o baja. Sus paradigmas son fenómenos meteorológicos tales como un torbellino de aire, una tormenta o una nevada.
- El tipo 4 de las dinámicas es el desarrollativo, que se caracteriza por tener una automeridad y una homomeridad baja en lo temporal, mientras que en lo espacial su automeridad (y su homomeridad) puede ser baja o normal. Sus paradigmas son una explosión, la formación de un volcán o la metamorfosis de una oruga en mariposa.
- El tipo 5 de dinámicas es el de las cosas, esto es, los elementos discernibles e idénticos a sí mismos predilectos de la metafísica de la sustancia, tales como una taza, un gato o una mesa. Las cosas son altamente automéricas en lo temporal y bajamente automéricas en lo espacial.

2.3. Cómo analizar los procesos

Los procesos se definen por la composición que tienen: son "grupos de cambios" o "familias de ocurrencias".[146] En otras palabras, los procesos se descomponen en otros procesos. Esto quiere decir que, para comprender los procesos, es necesario tener a mano alguna herramienta de análisis y descomposición.

En este apartado voy a definir un marco teórico que permita descomponer correctamente los procesos transformativos. Este marco viene dado por la mereología. Primero expondré

[145]Para simplificar la búsqueda y comparación en la bibliografía, mantengo aquí la notación original (y, ciertamente, arbitraria) que propone Seibt (2018) como tipo 1, 2, 3, etc.
[146]Rescher (1996: 38)

la mereología clásica, que puede emplearse en procesos para definir sus partes espaciales y temporales. Esta mereología, no obstante, presenta serias deficiencias cuando se quiere aplicar en la descripción de partes funcionales o causales tales como las propias de los procesos, lo que obliga a optar por otro tipo de mereologías, denominadas no clásicas. Por ese motivo, se presentarán cuáles son esas deficiencias y se optará finalmente por una mereología alternativa, de corte no clásico e inspirada en la mereología de niveles de Seibt (2015).

2.3.1. Mereologías clásicas para el espacio y el tiempo

La mereología es la teoría que estudia las relaciones entre las partes que se corresponden con un todo y las relaciones de las partes con el todo (Varzi 2019). La homomeridad y automeridad, por ejemplo, son propiedades que se definen entre las partes temporales y el todo y entre las partes espaciales y el todo. Este tipo de partes espacialmente o temporalmente extensionales son los ingredientes paradigmáticos de las mereologías de corte clásico. A continuación voy a mostrar cuáles son las propiedades y principios de descomposición de este tipo de mereologías, para después señalar algunos de los problemas que nos encontramos cuando este tipo de mereologías se aplican en el análisis de partes no extensionales, tales como las de los procesos.

2.3.1.1. Propiedades y principios de descomposición

En las mereologías clásicas se define la relación *es-propiamente-parte-de* (PP) de tipo orden parcial estricto, donde PPxy significa que x es propiamente una parte de y (Simons 1987; Varzi 2019). Esta relación cumple tres propiedades formales, a saber, transitividad, irreflexividad y asimetría:

$$\mathrm{PP}xy \wedge \mathrm{PP}yz \rightarrow \mathrm{PP}xz \qquad \text{(trans)}$$

$$\neg\mathrm{PP}xx \qquad \text{(irref)}$$

$$\mathrm{PP}xy \rightarrow \neg\mathrm{PP}yx \qquad \text{(asym)}$$

Antes de proseguir, es necesario notar dos cuestiones. Primero, que la asimetría define la relación PP como un orden (o jerarquía), esto es, como una relación entre desiguales, de igual manera que la simetría define una relación como de igual a igual o sin orden. Segundo, las tres propiedades tienen cierta conjugación en el siguiente sentido: (1) la propiedad de asimetría se deriva de las otras dos propiedades[147] y (2) la propiedad de irreflexividad se concluye de la asimetría.[148]

[147] Por reducción al absurdo:
Supongamos que PP$xy \wedge$ PPyx
Por (trans): PP$xy \wedge$ PP$yx \rightarrow$ PPxx
Por (irref): $\neg$PPxx

[148] Por reducción al absurdo (Simons 1987: 10):
Supongamos PPxx
Por (asym): PP$xx \rightarrow \neg$PPxx

Aun siendo la base de las mereologías clásicas, esas tres propiedades no son suficientes para definir completamente el marco general de estas mereologías. Esto se debe a que permiten relaciones mereológicas problemáticas tales como las que se presentan en las figuras 2.1(a), 2.1(b) y 2.1(c) y sus derivadas (figuras 2.1(d), 2.1(e) y 2.1(f)).[149]

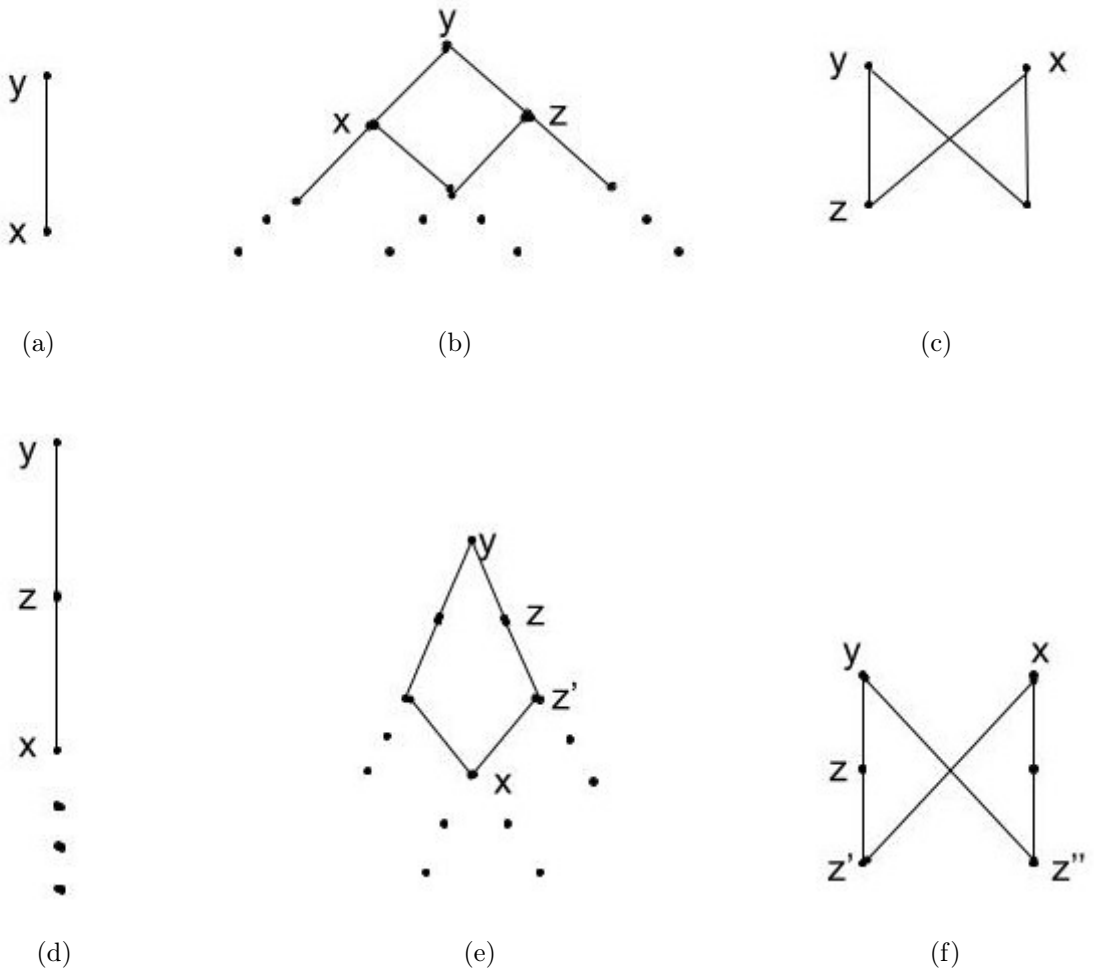


Figura 2.1: Algunas mereologías problemáticas

Estas mereologías son problemáticas, y lo son por diferentes razones. Algunas (fig. 2.1(a), 2.1(b), 2.1(d) y 2.1(e)) reflejan relaciones PPxy que violan la idea intuitiva de que un elemento debe tener más de una parte propia. Otras (fig. 2.1(c) y 2.1(f)) violan la idea intuitiva de que dos elementos distintos (x e y) no pueden tener las mismas partes propias. Si uno quiere borrar del mapa este tipo de mereologías es necesario añadir a las propiedades iniciales (trans), (irref) y (asym) algún requerimiento adicional. Esto nos lleva a la introducción de principios de descomposición (Simons 1987; Varzi 2019) y, en particular, al establecimiento de un principio de descomposición que consigue eliminar de golpe los problemas representados en las figuras anteriores. Se trata del principio de suplementación fuerte.[150] Para ello es necesario introducir antes las nociones de parte y

[149]En todas las figuras presentadas se usa el siguiente criterio: el nodo más alto de la relación representa el todo, mientras que el más bajo representa la parte propia de ese todo. Por ejemplo: en la figura 2.1(a) se representa PPxy.

[150]Entre medias existen distintos principios en los cuales no entraré en profundidad, tales como las

solapamiento, donde Pxy significa que x es parte (no necesariamente propia) de y y Oxy significa que x se solapa (*overlaps*) con y:

$$\mathrm{P}xy \equiv \mathrm{PP}xy \vee (x = y) \tag{part}$$

$$\mathrm{O}xy \equiv \exists z(\mathrm{P}zx \wedge \mathrm{P}zy) \tag{ov}$$

Entonces, el principio de suplementación fuerte se define así:

$$\neg\mathrm{PP}yx \rightarrow \exists z(\mathrm{P}zy \wedge \neg\mathrm{O}xz) \tag{ssp}$$

Las mereologías que aceptan (trans), (irref), (asym) y (ssp) son muy importantes y tienen una gran aplicación en los modelos extensionales. De hecho, se basan en su isomorfismo con las teorías de conjuntos. Por este motivo, Simons (1987) y Varzi (2019)) las denominan mereologías CEM (*classical extensional mereologies*). En la teoría axiomática de conjuntos de Zermelo, el axioma de extensionalidad dice que dado cualquier par de conjuntos X e Y, si todos sus elementos coinciden, entonces ambos conjuntos son idénticos. Formalmente:

$$\forall \mathrm{XY}\,[\forall \mathrm{U}(\mathrm{U} \in \mathrm{X} \leftrightarrow \mathrm{U} \in \mathrm{Y}) \rightarrow \mathrm{X} = \mathrm{Y}] \tag{aext}$$

De igual manera, en mereologías CEM (Simons 1987: 28) a partir de (trans), (irref), (asym) y (ssp) se deriva el principio de partes propias (ppp) que es, por así decirlo, la marca de extensionalidad en mereología (frente a las relaciones PP en otras mereologías, las cuales poseen cierta no-extensionalidad). De acuerdo con (ppp), si las dos entidades referidas por x e y tienen las mismas partes entonces x e y se refieren a la misma entidad y, viceversa, si x e y se refieren a la misma entidad, entonces la entidad a la que se refiere x y la entidad a la que se refiere y tienen las mismas partes. Igual que sucedía con (ssp), (ppp) excluye de la mereología esquemas como los de las figuras 2.1(a) a 2.1(f). Varzi (2019) formula así este principio que relaciona mismidad de partes con igualdad de manera bidireccional:

$$[\exists z\mathrm{PP}zx \vee \exists z\mathrm{PP}zy] \rightarrow [x = y \leftrightarrow \forall z(\mathrm{PP}zx \leftrightarrow \mathrm{PP}zy)] \tag{ppp}$$

2.3.1.2. Problemas con las mereologías clásicas

¿Qué pasa cuando las mereologías clásicas (esto es, las mereologías extensionales) se aplican al análisis de procesos? Lo que sucede es que, de pronto, uno se encuentra en graves aprietos. Aquí voy a formular dos grandes argumentos contra las mereologías CEM. El primero tiene que ver con el principio de partes propias y su inadecuación (tal y como

versiones débiles del principio de suplementación. Estas versiones débiles permiten algunas de las mereología problemáticas mientras que eliminan otras (Varzi 2019). Por ejemplo:

$$\mathrm{PP}xy \rightarrow \exists z\,[\mathrm{PP}zy \wedge \neg(x = z)] \tag{wsp1}$$

$$\mathrm{PP}xy \rightarrow \exists z\,[\mathrm{PP}zy \wedge \neg\mathrm{O}xz] \tag{wsp2}$$

El principio de suplementación débil (wsp1) —también llamado "principio de compañía"— no permite las composiciones de las figuras 2.1(a) y 2.1(d)). Por otro lado, el principio de suplementación débil (wsp2) no permite las composiciones de 2.1(a), 2.1(b), 2.1(d) y 2.1(e).

se define en las mereologías CEM) para ser aplicado en el análisis de procesos. El segundo tiene que ver con algunas composiciones dinámicas especiales que pueden suceder en las relaciones entre procesos y, sin embargo, no son permitidas en las mereologías CEM.

Los problemas del principio de partes propias (ppp) vienen dados en las dos direcciones de la implicación lógica: hay veces que *mismas partes* no implica *identidad*, y hay veces que *identidad* no implica *mismas partes*. El primer problema aparece en la primera dirección: mismas partes no implica siempre identidad. Las palabras "part" y "trap" tienen las mismas partes, pero no son idénticas. Igualmente, un conjunto de piedras amontonadas en el suelo y ese mismo conjunto de piedras formando un puente de piedra no es el mismo proceso pese a que sus partes sean exactamente las mismas. El puente de piedra tiene una serie de propiedades (soporte, durabilidad, funcionalidad) que emanan de la particular forma en que se relacionan sus partes, mientras que las piedras amontonadas en el suelo carecen de estas propiedades.

Este problema parece deberse al hecho de que la extensionalidad de CEM no recoge la noción de tipo de relaciones entre las partes, cosa que parece imprescindible para explicar completamente las partes de los procesos. Uno podría alegar que el problema se puede mitigar mediante el establecimiento de una posición espacio-temporal fija y bien definida para cada una de las partes del proceso. Se podría sostener que no es necesario establecer tipos específicos de relaciones PP entre procesos, sino que basta con determinar suficientemente cada parte y su ubicación espacio-temporal. Por ejemplo, basta con indicar la ubicación de cada parte en la palabra "part" para distinguirla de "trap" y basta con indicar la posición exacta de cada piedra del puente para que el conjunto quede completamente determinado en cuanto a sus propiedades de soporte, durabilidad y funcionalidad.

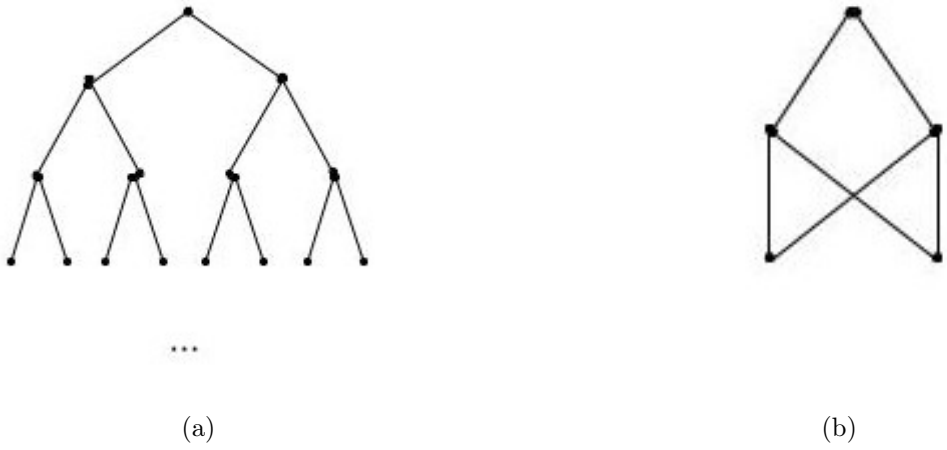
Por desgracia, esto nos conduce a un segundo problema, relacionado con la segunda dirección de la implicación lógica: en CEM la identidad de las cosas implica mismas partes —de acuerdo con el principio de partes propias (ppp)—, lo que contraviene el uso habitual de la noción de identidad de los procesos, en los cuales el proceso persiste pese a que sus partes varían a lo largo del tiempo. Este problema afecta a los procesos de tipo cosa y de tipo desarrollativo. El puente de piedra es el mismo después de reparar el tajamar, un ser humano es el mismo a lo largo del tiempo pese a su transformación (Simons 1987) y el gato Tibbles (Cotnoir y Bacon 2012) sigue siendo el mismo gato una vez que ha perdido su cola por un desafortunado accidente. Pero la cosa se vuelve, si cabe, más problemática cuando se atiende a los procesos de tipo actividad, tales como un torbellino o una tormenta, en los cuales la identidad se basa precisamente en un constante cambio de partes.[151] No es posible identificar los procesos únicamente a través de partes extensionales, por muy bien que se especifiquen las posiciones espacio-temporales de ellas.[152]

El segundo argumento está muy relacionado con el anterior y, aunque es menos evidente, resulta más inquietante. Las mereologías CEM están bien dispuestas para explicar

[151]Se podría argumentar que un torbellino tiene partes (el ojo, la columna) que permanecen idénticas. La pregunta sería entonces, ¿cuáles son las partes de las que están, a su vez, compuestas esas partes?

[152]Por ese motivo Seibt (2009) denomina a su teoría "teoría de procesos generales", lo que no es sino una manera de decir, "teoría de procesos no particulares".

relaciones PP de un tipo concreto (extensión) y especialmente de tipo árbol expansivo, esto es, en las que en la parte inferior el número de partes se expande (ver fig. 2.2(a)). Sin embargo, no parecen estar tan bien preparadas para admitir relaciones PP en forma de bucle (Cotnoir y Bacon 2012). Los procesos, por el contrario, se producen habitualmente en forma de bucle. He aquí un ejemplo paradigmático. Dos horquillas de diferente anchura cosidas por sus extremos producen un proceso total con forma de círculo que mantiene su forma precisamente por el efecto de la tensión que cada parte (es decir, cada horquilla) produce en la otra. Si se entiende cada efecto de tensión como un proceso transformativo (con relaciones PP funcionalmente definidas) se obtiene una relación de procesos con forma de bucle (ver figura 2.2(b) y 2.5). Este tipo de mereologías, de tipo tensionamiento mutuo (*backstretched connection*) son más comunes de lo que parece e incumplen claramente el principio de suplementación fuerte (ssp) y el principio de partes propias (ppp).



(a) (b)

Figura 2.2: Más mereologías problemáticas

2.3.2. Una mereología no clásica para los procesos

Algo pasa con las mereologías clásicas. Describen composiciones expansivas y sin bucles, pero no todas las composiciones son de ese tipo y, además, tampoco están especialmente preparadas para representar las transformaciones y los cambios. Por el contrario, los procesos transformativos suceden a lo largo del tiempo, están en constante cambio, no siempre tienen una composición de árbol expansivo y, en cambio, a menudo adoptan la forma de tensionamiento mutuo. Simons advierte de que las mereologías basadas en procesos conducen a innumerables dificultades (Simons 1987: 2). Sin embargo, aquí estamos obligado a invertir las prioridades. Puesto que los procesos son un "grupo de cambios" y una "familia de ocurrencias", necesitamos una mereología que sirva para analizar los procesos. Y como las partes de los procesos están enlazadas "causalmente" y "funcionalmente", necesitamos una mereología que tenga en cuenta relaciones *es-parte-de* de tipo funcional y no meramente extensional. Y si las mereologías extensionales no valen para

ese propósito, entonces es preciso buscar una alternativa.

2.3.2.1. Propiedades y principios de descomposición

Como las mereologías clásicas están fundadas en la aceptación de (trans), (irref), (asym) y (ssp), las mereologías no clásicas deberán poner en suspenso alguna de estas premisas. Principalmente hay dos opciones, ejemplificadas en Cotnoir y Bacon (2012) y Seibt (2015). Cotnoir y Bacon mantienen la transitividad y revisan la asimetría. Seibt propone una mereología no clásica basada en niveles (LEM, *levelled mereology*) en la que se renuncia a la transitividad mientras que se mantiene la asimetría. A continuación voy a presentar de manera sucinta algunos principios para una nueva mereología enfocada a procesos. Los principios están libremente basados en LEM. Que esta mereología resulte en última instancia (heurísticamente) más adecuada para analizar los procesos transformativos es algo que hay que ver al ponerla en funcionamiento. La sartén se prueba en el fuego. Precisamente eso será lo que haré después. Pero antes es preciso avanzar urgentemente hacia una mereología no clásica. Voy por pasos.

Primero. La primera cuestión atañe a la naturaleza de la relación PP en procesos transformativos. Las mereologías CEM presentan una relación PP de tipo extensional. Los modelos extensionales por excelencia son el espacial y el temporal. La India es una parte propia de Asia. 1995 es una parte propia del siglo XX. Sin embargo, cuando se usan entre procesos transformativos, las relaciones PP entre procesos no son siempre puramente extensionales ni son completamente homogéneas. Por el contrario, se especifican "bien causalmente, bien funcionalmente" en cada caso.

¿Qué significa exactamente aquí "en cada caso"? Quiere decir que las relaciones PP se especifican junto con los *relata* y pueden ser especificadas de diferentes maneras, de acuerdo a los mismos. Por ejemplo, un ladrillo (una viga, un cimiento) es parte de una casa porque cumple cierta función en la misma. Un ladrillo en la mesa de la cocina (el periódico sobre el sofá) no es parte de la casa, mientras que un ladrillo de la pared sí lo es. Seibt distingue claramente la relación *es-parte-propia-de* de la relación *es-parte-propia-extensional-de-la-extensión-de*. La extensión del ladrillo sobre la mesa está en (es parte de) la extensión de la casa. Pero solo el ladrillo de la pared es parte de la casa como proceso. La noción de proceso general que Seibt presenta descansa en esta importante distinción, puesto que los procesos generales no están completamente determinados por su extensión o por su posición espacio-temporal, sino por su funcionalidad. Este es el caso de un torbellino de viento.

Segundo. Las relaciones PP, si bien no son homogéneas, en la medida en que son relaciones del tipo parte propia, están completamente definidas, tanto causal como funcionalmente. Esta idea es importante para distinguir lo que a menudo se llaman "relaciones causales" de las relaciones de *es-parte-propia-de*. Entre dos procesos solo puede existir una única relación PP y no varias. Y esa relación no tiene por qué ser homogénea (con otras relaciones PP), pero debe ser completamente descriptiva, de otro modo no será posible hablar de parte propia.

Las relaciones de parte (no necesariamente propia) Pxy y solapamiento Oxy se pueden

definir de la misma manera que antes:

$$\mathrm{P}xy \equiv \mathrm{PP}xy \vee (x = y) \tag{part2}$$

$$\mathrm{O}xy \equiv \exists z(\mathrm{P}zx \wedge \mathrm{P}zy) \tag{ov2}$$

Nota: las relaciones anteriores se entienden siempre de un nivel, ya que todavía no he aceptado transitividad (ni lo haré de manera general). En particular, Oxy implica cuatro posibilidades: (1) que x e y sean el mismo proceso (esto es, que $z = x = y$), (2) que x sea parte propia de y (esto es, $z = x$), (3) que y sea parte propia de x (esto es, $z = y$) o (4) que exista un z distinto de x e y que sea parte propia de ambos. Dicho de otro modo:

$$\mathrm{O}xy \rightarrow (x = y \vee \mathrm{PP}xy \vee \mathrm{PP}yx \vee \exists z(\mathrm{PP}zx \wedge \mathrm{PP}zy)) \tag{ov3}$$

Tercero. En lo que concierne a la naturaleza funcional/causal de la relación PP, las relaciones PP son asimétricas. Es decir, siendo x e y dos procesos, se cumple que:

$$\mathrm{PP}xy \rightarrow \neg\mathrm{PP}yx \tag{asym2}$$

¿Por qué la relación *es-parte-propia-de* debe ser, antes que nada, asimétrica? En primer lugar, es necesario admitir que la asimetría es un lugar común dentro de la filosofía de procesos. Seibt (2015) la acepta, pero también está en la noción de prehensión (*prehension*) de Whitehead. Hansen (2004: 155) nos recuerda que Whitehead describe las relaciones procesuales como "causales" y con un carácter interno "asimétrico", del cual emana la distinción entre objetividad y subjetividad propia de la prehensión. Como se ha dicho antes, ni la mereología ni el enfoque procesual que se presenta aquí es whiteheadiano. Sin embargo, es significativo esta incidencia en la propiedad (asym) en detrimento de la transitividad.

En segundo lugar, la relación *es-parte-propia-de* entre procesos no puede ser una relación entre iguales porque requiere de cierto orden, de la misma manera que lo funcional y lo causal requieren cierto orden para ser determinados (Seibt 2015: 177). Esta relación asimétrica emana de la sensación de "pertenencia a" y es inseparable de la noción de *es-parte-de*: no es posible una fenomenología de la idea de pertenencia que no pase por cierta asimetría, es decir, que sea completamente simétrica.

En tercer lugar, entre dos procesos no puede haber más de una relación PP. Esto implica que no se puede afirmar que "en un sentido x *es-parte-propia-de* y pero en otro sentido y *es-parte-propia-de* x". Por el contrario, la relación PP, si está suficientemente determinada, excluye explícitamente relaciones en sentido inverso y, en consecuencia, es asimétrica.

Cuarto. Aceptar la asimetría de la relación PP implica automáticamente que se rechaza la reflexividad:

$$\neg\mathrm{PP}xx \tag{irref2}$$

Como se ha dicho antes, para la demostración de (irref2) únicamente se necesita aceptar (asym2). Por reducción al absurdo, supongamos PPxx. Por (asym2) si PP$xx \rightarrow \neg$PPxx.

Rechazar la reflexividad nos conduce a un importante corolario. Cuando se tiene una relación aparentemente de reflexividad en un proceso, tal y como cuando se dice que el sistema endocrino se regula a sí mismo o el sistema vestíbulo-ocular es un mecanismo de autorregulación, lo que se debe entender es lo siguiente: existen procesos intermedios los cuales permiten que se pueda hablar de autorregulación. Dicho de otro modo, la aparente relación "PPxx" implica que es necesario un mayor análisis, esto es, implica que existen otros procesos intermedios u, v, w... que median en esa autorregulación y la labor del buen analista consiste justamente en identificar cuáles son esos procesos. Se puede decir de manera genérica que el mundo se cambia a sí mismo, o el clima se autorregula sin temor a equivocarse. Pero lo que nos dice este corolario es que el problema de estas afirmaciones consiste en que no son lo suficientemente significativas.

Llamaré a este corolario proveniente de la propiedad de (irref2) "corolario de granularidad insuficiente" en la definición de procesos.

Quinto. Si la asimetría era indispensable, la transitividad no lo es. Esto es consecuencia inevitable del primer punto, a saber, que las relaciones *es-parte-de* entre procesos no son completamente homogéneas y, en particular, no son puramente relaciones de extensionalidad. Como las relaciones PP entre procesos no siempre son homogéneas, la transitividad no puede ser una propiedad universal, en el sentido de que no se puede aplicar a todo tipo de relaciones PP.

$$\neg\forall x, y, z(\mathrm{PP}xy \wedge \mathrm{PP}yz \rightarrow \mathrm{PP}xz) \qquad \text{(notrans)}$$

Hay muchos ejemplos de no transitividad en la relación PP. Muchos aparecen de la combinación inveterada de relaciones PP extensionales y procesuales. Mi mano es parte de mí, yo soy parte de la sociedad, pero mi mano no es parte de la sociedad. Otros aparecen dentro de relaciones PP procesuales. Seibt (2015) destaca especialmente dos tipos de ejemplos. El primer tipo está relacionado con la granularidad de la materia. Un átomo de oxígeno es parte propia de una molécula de agua y el agua es parte propia de un río. En ambos casos hacer uso de la relación PP parece legítimo. Sin embargo, no parece tan claro que un átomo de oxígeno sea parte propia de un río, puesto que el uso de esa relación PP no es homogéneo. El segundo tipo de ejemplos se refiere a la granularidad de actividades. Cambiar los pañales a un bebé es parte del proceso consistente en preservar su salud. Abrir la caja de toallitas es parte del proceso de cambiar los pañales. Presionar hacia arriba con el dedo es parte del proceso de abrir la caja de toallitas. Pero no parece aceptable decir que presionar hacia arriba con el dedo en la caja de toallitas es parte del proceso de preservar la salud de un bebé (Seibt 2009: 489). Un tercer ejemplo del mismo tipo: el aire es parte del proceso de respirar, y el proceso de respirar es parte del proceso de cantar, pero no parece aceptable afirmar que el aire es parte del proceso de cantar. Seibt (2015) presenta un listado más amplio de procesos donde se viola la transitividad de la relación *es-parte-propia-de*.

El problema en los casos anteriores es que las relaciones PP entre procesos no siempre se especifican de la misma manera y, desde luego, que la extensionalidad no es suficiente para especificarlas completamente. Cuál es la manera de especificar esta relación PP no es una cuestión cerrada. Si nos fiamos de Seibt hay, por lo menos, cuatro diferentes

maneras de especificación, es decir, hay, por lo menos, cuatro formas para especificar la relación PP entre procesos: espacial, temporal, material y funcionalmente.[153] Lo que nos interesa aquí es retener la idea de que la profundidad de la transitividad depende de la especie de la relación PP (y no del género denominado "relación PP"). Por supuesto, la extensionalidad tiene su espacio de acción. Se puede mantener esa transitividad cuando la relación PP se aplica a los espacios ocupados por cada entidad de manera homogénea, es decir, espacialmente. Si el espacio que ocupa un átomo de oxígeno es parte propia del espacio que ocupa un ladrillo y el espacio que ocupa un ladrillo es parte propia del espacio que ocupa una casa, entonces el espacio que ocupa un átomo de oxígeno es parte propia del espacio que ocupa una casa. El problema es, como se ha dicho, que las relaciones de pertenencia entre espacios y las relaciones de pertenencia entre procesos no tienen la misma especificación.

Las consecuencias de no aceptar transitividad de manera general en la relación PP entre procesos son importantes. La principal es que en casos como los de la figura 2.3, la relación PPxy no es trivial en el sentido de que no está contenida en las relaciones PPxz y PPzy.

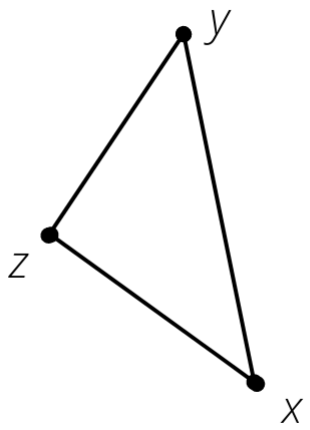


Figura 2.3: Una composición mereológica no clásica

Seibt (2015) establece una mereología de niveles en base a ello y legitima la múltiple realizabilidad en base a la aceptación de un principio de partes propias restringido en niveles. Aquí me centraré en primer lugar en las relaciones PP de primer nivel, es decir, relaciones directas entre procesos donde la transitividad no aplica. Después, identificaré algunos modelos paradigmáticos de relaciones de primer nivel entre procesos.

2.3.2.2. Noción de interferencia y sus composiciones

En esta nueva mereología basada en relaciones heterogéneas es necesario identificar los modelos de relación paradigmáticos entre procesos. Estos modelos, además, deben ser

[153] "La profundidad de cadena *es-parte-de* con la que se puede aplicar el patrón inferencial de transitividad depende de la especificación de la relación *es-parte-de* en cada caso (espacial, temporal, material, funcional, etc.) y del tipo de *relata*." (Seibt 2015: 177)

tales que cumplan que las relaciones PP que se usen definan completamente la relación entre los dos procesos, el proceso total y el proceso parte del proceso total. En su Teoría de Procesos Generales, Seibt (2009) propone el modelo de relación tipo interferencia basado en los siguientes componentes:

- Foco de interferencia: es el proceso que forma el todo, del cual el resto de los procesos que intervienen son parte
- Producto de interferencia: es el proceso producido a causa de la interferencia
- Partes interferentes: son el conjunto de procesos que interfieren para producir el producto de interferencia

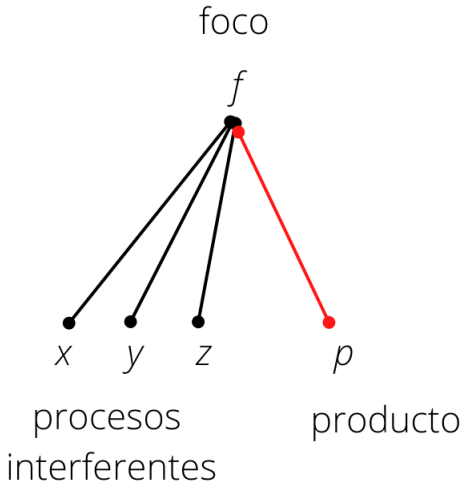


Figura 2.4: Interferencia entre procesos y sus elementos de acuerdo a Seibt (2009). Los procesos interferentes son partes interferentes de un foco que da lugar a un producto de interferencia. El proceso producto también es parte del foco.

El modelo de interferencia es muy útil para el análisis de procesos y bajo su aparente simplicidad se ocultan varias sutilezas. Para empezar, permite una distinción importante con respecto a otras variantes de análisis.[154] En otras variantes únicamente existen dos componentes, es decir, no existe la noción de foco, de tal manera que producto y foco convergen en la misma cosa. Por el contrario, en la analítica de la Teoría de Procesos Generales hay como mínimo tres tipos de componentes: proceso interferente, foco y producto.

Por otro lado, permite una distinción no menos importante con respecto a las mereologías extensionales. En las mereologías extensionales solo existe la operación suma de partes: la suma de las partes es el todo, de acuerdo con el principio de partes propias (ppp). En la mereología no clásica, sin embargo, aparece otra operación no menos importante, el producto de partes. El producto de las partes interferentes x, y, z con foco f es p, esto es, $p = \text{Prod}(x, y, z)$.

Dentro del modelo de interferencia se pueden diferenciar dos subformas de relaciones de parte propia (PP) entre procesos, a saber, la relación de parte propia interferente a foco

[154] Como, por ejemplo, la noción de *prehension* en Whitehead (1978 1929).

($\mathrm{PP}^i xf$) y la relación de parte propia producto a foco ($\mathrm{PP}^p pf$). De igual modo, se puede definir la relación de parte interferente (no necesariamente propia) a foco ($\mathrm{P}^i xy$) y la relación de parte producto (no necesariamente propia) a foco ($\mathrm{P}^p xy$) de la siguiente forma:

$\mathrm{P}^i xy$ es la relación que tiene x ser parte (no necesariamente propia) como interferencia en y:

$$\mathrm{P}^i xy \equiv \mathrm{PP}^i xy \vee (x = y) \qquad \text{(partI)}$$

$\mathrm{P}^p xy$ es la relación que tiene x ser parte (no necesariamente propia) como producto de y:

$$\mathrm{P}^p xy \equiv \mathrm{PP}^p xy \vee (x = y) \qquad \text{(partP)}$$

El modelo de análisis de la interferencia nos permite analizar otros modelos de relación más complejos, como el modelo de tensionamiento mutuo (*palíntonos harmoniē*, *backstretched connection*). En el tensionamiento mutuo se presentan dos procesos diferenciados que se cancelan mutuamente. Quizás el ejemplo más claro es el de las dos horquillas. Las horquillas flexionadas y cosidas entre sí es un ejemplo de proceso transformativo del tipo actividad, con automeridad temporal máxima, que además sigue el modelo *palíntonos harmoniē* o tensionamiento mutuo, en el sentido de que el efecto de la horquilla A en la horquilla B y el efecto de la horquilla B en la horquilla A se cancelan mutuamente. El efecto de estos procesos es fundamental para explicar lo que está sucediendo en el sistema formado por las dos horquillas unidas. Sirve para distinguirlo, entre otras cosas, de un sistema formado por dos horquillas que tuvieran de manera natural esa forma circular sin que hubiese tensión alguna entre ellas. La dinámica del sistema se basa en dos interferencias: la de A en B y la de B en A. En realidad, la mereología del sistema quedaría mejor expresada indicando las diferencias entre la forma en que se producen esas interferencias (ver fig. 2.5). La tensión en la horquilla B es el producto de la tensión en la horquilla A, la conexión 1 y la conexión 2. Al mismo tiempo, la tensión en la horquilla A es el producto de la tensión en la horquilla B, la conexión 1 y la conexión 2.

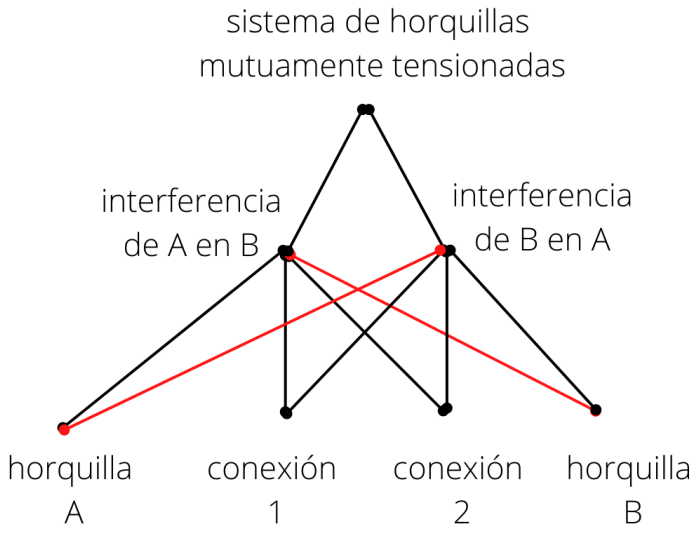


Figura 2.5: Esquema de interferencia entre dos horquillas unidas por los extremos. Este es un ejemplo de tensionamiento mutuo.

La existencia de relaciones *es-parte-propia-de* no homogéneas tales como las que se sugieren en el modelo de interferencia nos obligan a repensar el principio de suplementación fuerte, puesto que es necesario discriminar el foco del proceso de interferencia A en B y el foco del proceso de interferencia de B en A, incluso a pesar del hecho de que los procesos que forman parte de ellos sean los mismos. Dadas dos relaciones de *ser-parte-propia-de* no homogéneas, PP^p y PP^i, y que solo estas dos relaciones están en la composición de un determinado proceso p, el nuevo principio de suplementación fuerte se define así:

$$\neg \mathrm{PP}yx \rightarrow \exists z \left[(\mathrm{P}^i zy \wedge \neg \mathrm{O}^i xz) \vee (\mathrm{P}^p zy \wedge \neg \mathrm{O}^p xz) \right] \tag{ssp2}$$

donde:
$\mathrm{O}^i xy$ es la relación de solapamiento entre x e y como interferencia, que se define exclusivamente a través de relaciones de *es-parte-de* como interferencia:

$$\mathrm{O}^i xy \equiv \exists z \left(\mathrm{P}^i zx \wedge \mathrm{P}^i zy \right) \tag{ovI}$$

$\mathrm{O}^p xy$ es la relación de solapamiento entre x e y como producto, que se define exclusivamente a través de relaciones de *es-parte-de* como producto:

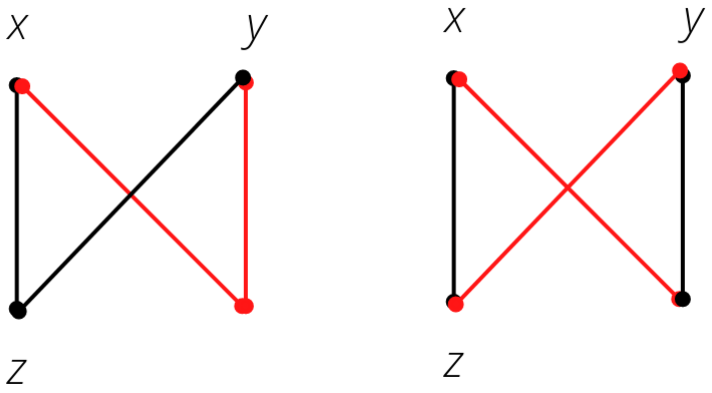
$$\mathrm{O}^p xy \equiv \exists z \left(\mathrm{P}^p zx \wedge \mathrm{P}^p zy \right) \tag{ovP}$$

$\mathrm{P}^i xy$ es la relación que tiene x de *es-parte-de* (no necesariamente propia) como interferencia de y, que se definió en (partI)
$\mathrm{P}^p xy$ es la relación que tiene x de *es-parte-de* (no necesariamente propia) como producto de y, que se definió en (partP)

Es decir, la suplementación fuerte solo aplica a relaciones homogéneas, o sea, aplica a las partes como interferencia (P^i) por separado o a las partes como producto (P^p) por separado. Nótese que todas las relaciones *es-parte-de* de las que he hablado son relaciones de un solo nivel, por (notrans). Determinar qué relaciones *es-parte-de* son heterogéneas y cuáles no lo son no es una tarea sencilla, pero, al menos, se puede aplicar sin problemas en el modelo de interferencia: la relación PP^i entre el foco y las partes interferentes y la relación PP^p entre foco y producto (de las partes interferentes) no son homogéneas. Este nuevo principio establece que la figura 2.6(a) no es válida, mientras que la figura 2.6(b) y 2.5 sí lo son.

De lo anterior se desprende una nueva mereología en la que es posible tener modelos de relación de tipo *tejido*, frente a los modelos de árbol expansivo propios de las mereologías clásicas. Esta nueva mereología está basada en la restricción de la propiedad de transitividad, que dependerá de los *relata* y de la especie de la relación *es-parte-de*. Cuidado, no obstante. Esta mereología no viene a sustituir a CEM. Los procesos siguen sucediendo en el tiempo, y el tiempo en el que los procesos suceden acepta una mereología extensional.[155] Por tanto, si t_1 es parte propia de t_2 y, a su vez, t_2 es parte propia de t_3, entonces t_1 será parte propia de t_3. Del mismo modo, los procesos suceden en un espacio, y el espacio en el que suceden los procesos acepta una mereología extensional. En teoría de procesos, una de las maneras de distinguir un tipo de proceso de otro es a través de la noción de automeridad, la cual se basa precisamente en una mereología clásica aplicada

[155]Por lo menos, si dejamos por un momento de lado toda la problemática derivada de la aceptación de la Teoría Especial de la Relatividad.



(a) (b)

Figura 2.6: Mereologías de interferencia

al tiempo y al espacio en el cual suceden los procesos. Gracias a la mereología CEM se puede diferenciar entre actividades, desarrollos y cosas.

El problema está en aplicar esas mereologías clásicas a las partes de los procesos, que se definen funcional y causalmente, así como en combinar inveteradamente relaciones *es-parte-de* extensionales, que solo se aplican en el espacio y el tiempo, con relaciones *es-parte-de* no extensionales que forman parte del análisis de procesos. Este problema se hace aún mayor en el caso de procesos de tipo actividad o desarrollo, que quedan definidos a través de su cambio durante el tiempo. Las mereologías no clásicas y, en concreto, la mereología no clásica que solo he llegado a esbozar aquí, tienen un ámbito de uso diferente. Trata de establecer algunos principios mínimos para comprender la mereología de los procesos. Este tipo de mereologías son más abiertas y permiten analizar con mayor libertad la complejidad de los procesos transformativos. También están mejor dotadas para analizar un tipo particular de estos procesos del que me ocuparé ahora, a saber, los procesos autotransformativos.

2.4. Autotransformatividad

En palabras de Rescher, los procesos transformativos son aquellos que "únicamente transforman estados de hechos en general y hacen posible que sucedan otros procesos ulteriores sin que resulten en cosas particulares" (Rescher 1996: 41). En general, las filosofías del proceso coinciden en subrayar que este tipo de procesos es fundamental para explicar la realidad. Los procesos autotransformativos son procesos transformativos con cierto cerramiento, en el sentido de que estos procesos transforman (modifican, cambian) otros procesos que a su vez son capaces de transformar (modificar, cambiar) el proceso inicial. Estos procesos son importantes y mucho más abundantes de lo que pudiera parecer a primera vista. Ya hemos visto algunos de ellos: el sistema de dos horquillas, el puente de piedra y, en general, los sistemas con tensionamiento mutuo. A continuación voy a tratar

de definirlos. Para ello, primero estableceré tres características mereológicas que les son propias. Después expondré un ejemplo particular de proceso autotransformativo, el arco tensionado, cuyo análisis pondrá de relieve esas características. Para finalizar expondré una cuarta característica que completa la descripción de este tipo de procesos, a saber, la clausura en la modulación.

2.4.1. Interferencia y clausura

¿Qué composición tienen los procesos autotransformativos? Son un tipo de proceso transformativo con una complejidad especial que se traduce en una composición dinámica específica donde las ocurrencias están enlazadas entre ellas con cierta clausura o ciclicidad.[156] Ahora voy a precisar en qué sentido.

Empecemos por ver qué no es clausura. Clausura no significa aislamiento.[157] Los procesos autotransformativos no son procesos aislados del resto de procesos. Todo lo contrario. Supongamos, por ejemplo, el proceso de autorregulación presa-depredador que se describe por las ecuaciones de Lotka-Volterra:[158]

$$\frac{dx}{dt} = x(\alpha - \beta y)$$
$$\frac{dy}{dt} = -y(\gamma - \delta x)$$

donde x e y representan respectivamente el número de presas y depredadores en un ecosistema dado, α, γ representan la tasa de reproducción de las presas y la tasa de muerte los depredadores respectivamente, y β, δ son parámetros que representan el éxito en la caza, afectando a la presa y al depredador respectivamente. Los parámetros α, β, γ y δ se definen como positivos y pueden depender de otros procesos externos, tales como el clima, el relieve o las enfermedades. Acerca de este proceso de autorregulación descrito por estas ecuaciones se puede decir que tiene cierta clausura, pero no que está aislado.

¿Qué significa entonces clausura? Las ecuaciones de Lotka-Volterra presentan cierta clausura, y no en el sentido de aislamiento, sino en el sentido de dependencia mutua. Las variables x e y dependen mutuamente entre sí, cosa que no sucede con los parámetros α, β, γ y δ. Seibt (2018) llama composición dinámica a la composición que se deriva de aplicar una mereología de niveles (LEM) a los procesos transformativos y distingue dos grandes tipos de composición dinámica: lineal y no lineal. Para Seibt, los procesos autotransformativos tienen cierta clausura en el sentido de que tienen cierta no linealidad en su composición dinámica. Por desgracia, esta denominación es ambigua puesto que se puede confundir con la noción matemática de linealidad, propia de las ecuaciones diferenciales. Por este motivo, aquí la evitaré y, en cambio, retendré el término "clausura".

[156] Se usa el término "composición dinámica" en el sentido de Seibt (2018).

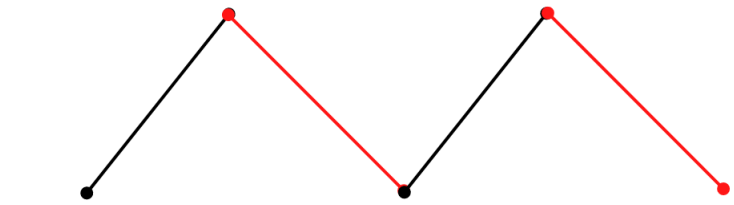
[157] Ni tampoco se usa en el sentido en que McGinn (1991) habla de clausura cognitiva.

[158] Las ecuaciones de Lotka-Volterra forman un sistema no lineal de dos dimensiones en el sentido de Strogatz (1994). Debo este ejemplo a Carlos Sierra-Lechuga.

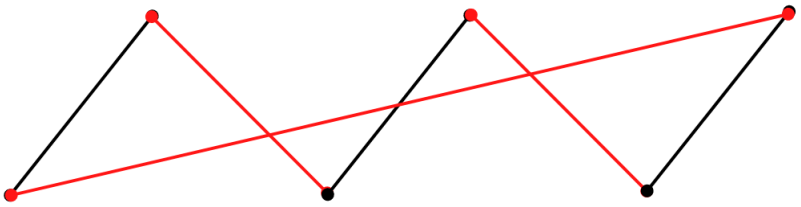
Se han visto dos propiedades de los procesos autotransformativos. La primera es que tienen composición. Es decir, como todo proceso, son complejos. La segunda es que esta composición presenta una característica distintiva, a saber, presenta cierta clausura en el sentido de cierta dependencia mutua entre sus partes componentes. Ahora es necesario decir algo más acerca de esos componentes. Para ello es necesario volver de nuevo a la noción de interferencia en los términos en que se define en la Teoría de Procesos Generales (GPT, *General Process Theory*) de Seibt (2009). De acuerdo con esta teoría se pueden diferenciar dos tipos especiales de interferencias: la interferencia encadenada y la interferencia mutua.

Interferencia encadenada. Se trata de dos o más interferencias, en las cuales el producto de la primera interferencia es, a su vez, parte de la siguiente interferencia con un producto distinto. Este tipo de interferencias encadenadas se pueden denominar, por un símil gráfico, como tipo "diente de sierra" (ver figura 2.7(a)).

Interferencia mutua. En un conjunto de interferencias encadenadas puede suceder que el producto de una de ellas sea una parte de una interferencia anterior. En el caso de la interferencia encadenada, la "mutualidad" puede deberse a un encadenamiento en bucle tipo "diente de sierra" (ver figura 2.7(b)). Nótese que el diente de sierra es compatible con el principio de suplementación fuerte (ssp2), sin importar si se compone de dos o más instancias de interferencia.



(a)



(b)

Figura 2.7: Mereologías complejas de interferencia: interferencia encadenada e interferencia mutua

Un ejemplo clásico de clausura en este sentido, que se remonta a Heráclito, es el arco tensionado. El arco consiste (extensionalmente) en dos partes materiales, cuerpo y cuerda. Estas partes están en tensión, esto es, tienen energía potencial. De acuerdo a la ley

de Hooke, su fuerza es proporcional a su grado de tensionado y a su constante elástica. Estas tensiones, además, interactúan entre sí y se anulan mutuamente. El arco tensionado es un proceso, en el sentido de "familia de ocurrencias sistemáticamente enlazadas entre sí".[159] En concreto, consiste en dos ocurrencias que se oponen entre sí. Además, el arco tensionado es un proceso transformativo, en el sentido de que "transforma estados de hechos en general y hace posible que sucedan otros procesos ulteriores sin que resulten en cosas particulares"[160], a saber, procesos tales como la tensión acumulada que permite lanzar una flecha a largas distancias. Además, el arco está compuesto de otros procesos que se relacionan en forma de bucle, por lo que tiene una cierta clausura propia de los procesos autotransformativos.

La descripción mereológica de un arco en tensión debe contemplar los siguientes puntos:

1. Debe incluir procesos con automeridad temporal alta y automeridad espacial baja, esto es, del tipo "cosa", tales como el cuerpo y la cuerda.
2. Las partes físicas anteriores, sin embargo, no son suficientes. La descripción necesita distinguir entre un arco en tensión y un arco, por así decirlo, congelado. Si se arrojase al suelo un arco congelado, podría romperse en pedazos como la vajilla, y no habría ninguna tensión liberada en ello, al margen de la propia tensión estructural del cuerpo por separado y la cuerda por separado. Por el contrario, el arco tensionado, al romperse, libera otra tensión. Y esa tensión es efecto del cuerpo en la cuerda y de la cuerda en el cuerpo.
3. Por ello, debe contemplar procesos con automeridad temporal alta, del tipo "actividad" como los relacionados con la tensión interna.
4. Si se considera el arco tensionado pero en descanso, debe contemplar procesos de oposición o autorregulación.
5. El arco tensionado, entonces, resultará de la combinación de dos efectos. Primero, el efecto (tensionante) del cuerpo en la cuerda. Segundo, el efecto (curvante) de la cuerda en el cuerpo. Estos son efectos que actúan en la cantidad de tensión (de la cuerda) y en la cantidad de curvatura (del cuerpo).
6. Por consiguiente, en la mereología del arco en tensión se deben considerar dos interferencias de tipo causal. Primero, una interferencia en la cual toma parte la energía del cuerpo y cuyo producto es la cantidad de tensión de la cuerda. Segundo, una interferencia en la cual toma parte la energía de la cuerda y cuyo producto es la cantidad de curvatura del cuerpo. Ambos están descritos por la Ley de Hooke[161]
7. Ambas cantidades se autorregulan. Si el tirador de arco tira de la cuerda con su mano y la suelta, el arco tiende a una posición de descanso. Esta regulación explica la naturaleza "en descanso" del arco tensionado.[162]

[159] Rescher (1996: 38)
[160] Rescher (1996: 41)
[161] *Ut tensio sic vis.*
[162] En términos matemáticos (Strogatz 1994), existe un punto fijo estable. Puede haber otros procesos adicionales, tales como la histéresis elástica (el hecho de que la deformación se retrasa con respecto al esfuerzo a realizar para deformar), la progresiva degradación de la cuerda o la progresiva deformación del cuerpo, que requerirían de mecanismos explicativos adicionales. La mereología propuesta no agota la

Teniendo en cuenta lo anterior, se puede hacer una descripción mereológica del arco como proceso autotransformativo en la cual intervienen las siguientes partes:

Arc: un arco tensionado en descanso

Cpo: cuerpo del arco, que a su vez se compone de dos extremos ($eCpo1$ y $eCpo2$) y de una parte central *centCpo*.

Cda: cuerda del arco, que a su vez se compone de dos extremos ($eCda1$ y $eCda2$) y de una parte central $cCda$.

$Con1$: conexión 1, que une el extremo $eCpo1$ del cuerpo con el extremo $eCda1$ de la cuerda.

$Con2$: conexión 2, que une el extremo $eCpo2$ del cuerpo con el extremo $eCda2$ de la cuerda.

[$tCda$]: cantidad de la tensión de la cuerda. Cuanta más tensión mayor fuerza, de acuerdo con la constante elástica k.

[$pCda$]: cantidad de energía potencial acumulada por la tensión en la cuerda.

[$cCpo$]: cantidad de la curvatura del cuerpo. Cuanta más curvatura, mayor fuerza, de acuerdo con la constante elástica k', donde $k>k'$.

[$pCpo$]: cantidad de energía potencial acumulada por la curvatura del cuerpo.

$Int1$: interferencia en la cual forma parte la energía potencial de la cuerda y cuyo producto es el cambio en la curvatura del cuerpo.

$Int2$: interferencia de la cual forma parte la energía potencial del cuerpo y cuyo producto es la tensión de la cuerda.

Las dos interferencias tienen una composición de tipo diente de sierra. Se trata de una interferencia mutua especial, donde cada interferencia tiene un elemento disjunto, a saber, la energía potencial del cuerpo y la energía potencial de la cuerda.

¿Cuáles son las características de los procesos autotransformativos tipo diente de sierra? Es posible identificar, al menos, tres características de estos procesos autotransformativos.

(1) Complejidad: para que un proceso sea considerado autotransformativo debe estar compuesto, a su vez, por otros procesos. Ningún proceso sin composición, esto es,

descripción del proceso arco. Pero contribuye a su descripción.

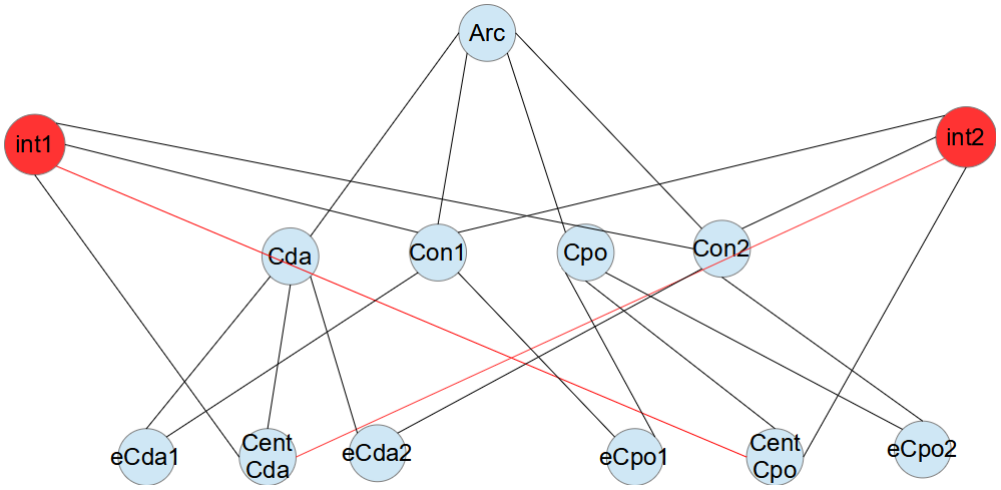


Figura 2.8: Composición de interferencia para el arco tensionado, donde *Arc* es el arco tensionado en descanso, *Cda* es la cuerda del arco, compuesta de dos extremos y un centro y *Cpo* es el cuerpo del arco, compuesto de dos extremos y un centro. Los extremos del cuerpo y la cuerda están conectados mediante dos conexiones *Con*1 y *Con*2. Como resultado, la cuerda interfiere con el arco y a la inversa. Es decir, existe tensionamiento mutuo.

simple, puede ser considerado autotransformativo. Las descripciones de procesos simples con relaciones *es-parte-de* hacia sí mismos son descripciones en las que se está violando (irref2) y, probablemente, (notrans).[163]

(2) Interferencia: para que un proceso sea considerado autotransformativo debe incluir interferencias en las que formen parte los procesos que lo componen.

(3) Clausura: estas interferencias deben tener una composición dinámica con cierto cierre. Como ya se dijo antes, clausura no quiere decir aislamiento. Por el contrario, los procesos autotransformativos están abiertos y tienen relaciones con otros procesos exteriores. Sin embargo, deben poseer una composición dinámica de dependencia cíclica. Se ha identificado una, el tipo diente de sierra, como la básica. La clausura en el tipo diente de sierra es horizontal y se produce entre interferencias. Por ese motivo este bucle solo puede suceder entre dos o más procesos que interfieren mutuamente entre sí. Aquí no se sostiene que la interferencia de tipo diente de sierra sea la única posible para los procesos autotransformativos. De hecho, la mereología no clásica presentada, que se basa en la propiedad (notrans), admite otras formas de combinación con mayor verticalidad tales como las sugeridas en la fig. 2.3. No obstante, el tipo diente de sierra es una de las formas más sencillas y parece también ser una de las más abundantes, a tenor de los ejemplos.[164]

Por simplicidad, voy a denominar temporalmente a este tipo de composición dinámica basada en interferencias con cierta clausura "composición CIC". La composición CIC es necesaria para explicar los procesos autotransformativos, pero quizás no sea suficiente. Pondré un contraejemplo. En una mesa de billar hay dos bolas, A y B. El taco golpea la bola A, la cual golpea la bola B. A su vez, la bola B se mueve de tal manera que acaba golpeando de nuevo a la bola A y, por tanto, modificando su dirección de movimiento.

[163]De acuerdo con el corolario de granularidad insuficiente referido en la sección 2.3.2.1.
[164]Por ejemplo, los casos de retroalimentación, la masa y el muelle o el péndulo, entre otros.

Aparentemente este sistema cumple composición (se basa en diferentes procesos), incluye interferencias (los diferentes golpeos entre elementos) y tiene cierta clausura (acaba con un golpe a la bola inicial A). Es decir, tiene una composición CIC. Sin embargo, no es un proceso autotransformativo en el sentido en el que lo es, por ejemplo, las dos horquillas o el arco tensionado. No hay tensionamiento mutuo. Para distinguir los procesos autotransformativos de otros procesos como el descrito de la mesa de billar es preciso introducir una nueva propiedad. Voy a denominar de manera temporal a esta propiedad (4) "automodulación" y trataré de caracterizarla en los próximos apartados.

2.4.2. Modulación y clausura

Una vez establecidas las nociones de proceso, proceso transformativo y proceso autotransformativo estoy en disposición de introducir un nuevo término. El término "modulación" pretende designar, en primer lugar, una propiedad de los procesos generales de tipo transformativo, en el sentido en que se ha definido proceso transformativo anteriormente, esto es, una "familia organizada de ocurrencias" sistemáticamente relacionadas "bien causalmente, bien funcionalmente" que "transforma un estado de hechos en general y hace posible que sucedan otros hechos, sin resultar en cosas particulares".[165]

La modulación es una propiedad en la intervienen al menos dos procesos transformativos: un proceso A (modulante) y un proceso B (modulado). La conexión entre el proceso A y el proceso B puede ser a través de procesos transformativos intermedios que suceden en el tiempo.

Además de lo anterior, la modulación tiene una particularidad que la caracteriza: durante el tiempo que dura, la variación de cantidad (o cualidad) en el proceso B depende de la cantidad (o cualidad) del proceso A. En otras palabras, A modula a B si el proceso A interviene o tiene una función en cómo varía el proceso B.

La automodulación es a la modulación lo que los procesos autotransformativos son a los transformativos. Un proceso A se automodula cuando tiene la función de variarse a sí mismo.

A continuación voy a presentar dos notas en favor de la utilidad de las ideas de automodulación y modulación, es decir, dos notas que, de cumplirse, subrayan la no trivialidad de estas nociones. Después formularé la noción de modulación y automodulación en su forma mejor definida, esto es, en su forma matemática, que se expresa en la teoría de sistemas dinámicos y se representa a través de ecuaciones diferenciales. Esa presentación es necesaria y se justifica por dos motivos. En primer lugar, porque permite una formulación rigurosa y sistemática para un concepto que, como éste, carece de antecedentes filosóficos. En segundo lugar, porque la modulación matemática será una herramienta útil para caracterizar de manera correcta determinadas teorías de la conciencia de corte matemático (y, más concretamente, la Teoría de la Información Integrada). La formulación de la noción de modulación concluirá con la presentación de una forma canónica

[165] Rescher (1996: 38)

para los sistemas dinámicos que permite visualizar de manera sencilla la modulación y la automodulación matemática del sistema.[166] Para finalizar esta parte, trataré de caracterizar la relación entre la automodulación y los procesos autotransformativos, y daré varias razones en defensa de la idea de que los procesos con una composición dinámica CIC guardan una importante relación con la propiedad de automodulación y modulación clausurada.

Un proceso se automodula cuando tiene la función de variarse a sí mismo. Como quiera que los procesos consisten siempre en una variación, es necesario refinar esta definición, a fin de que no resulte trivial. Se puede decir, por ejemplo, que el todo se automodula, puesto que dentro del todo siempre existen procesos que se modifican a sí mismos. Sin embargo, esa articulación de la noción de automodulación no es útil. Voy a introducir dos notas que, espero, permitan definir de manera no trivial la noción de automodulación.

La primera nota dice que, desde el punto de vista procesual, la función de automodulación debe apoyarse en otros procesos colindantes B_1, $B_2 \ldots$ El proceso A modula al proceso B_1, que, a su vez, modula al proceso $B_2 \ldots$, que, a su vez, modula a un proceso B_N el cual, a su vez, modula de nuevo a A. Sin estos procesos intermedios (B_1, $B_2 \ldots$ B_N), no existiría automodulación. Esta nota proviene directamente del corolario mereológico de granularidad insuficiente, el cual se apoya en la propiedad irreflexiva (irref2). De acuerdo con (irref2), ningún proceso puede ser parte de sí mismo. Por consiguiente, una situación de automodulación debe reflejar cadenas de procesos con cierta clausura, las cuales deberán ser analizadas. En otras palabras, la automodulación indica composición de procesos con cierta clausura.

Esta es, por cierto, una de las diferencias entre la noción de automodulación y la noción de proceso autotransformativo, así como la razón por la cual es necesario introducir esta noción de manera diferenciada de la noción de proceso autotransformativo. Por (irref2), un proceso (autotransformativo o no) nunca puede ser parte de sí mismo, mientras que una variable de un proceso puede automodularse a sí misma.

La segunda nota hace referencia a la idea de variación dentro de la noción de modulación. Variación indica que existe una variable que cambia (varía) de acuerdo a otra variable. Esta variable podría ser, en principio, cuantitativa o cualitativa. No obstante, en términos cuantitativos, la diferencia se observa con claridad. No es lo mismo decir de una variable que $x = f(x)$ o decir que $dx/dt = f(x)$. Lo primero es tautológico, puesto que $x = x$. Lo segundo, en cambio, es una ecuación diferencial que, dadas unas condiciones iniciales, puede describir con precisión cómo varía x con respecto al tiempo t. La variación a la que se refiere la idea de automodulación es del segundo tipo y, en consecuencia, no es trivial.

A continuación, voy a exponer algunos ejemplos de uso de esta noción en diferentes sistemas dinámicos dentro de la mecánica clásica, con la esperanza de que permitan hacernos cargo de su relevancia. Los ejemplos son de modulación en cantidad.

[166] Véase (fc2).

Supongamos, en primer lugar, un sistema inercial como el descrito en la primera ley de Newton, formado por un cuerpo de masa m con movimiento rectilíneo. De acuerdo con la primera ley, si no se le aplica ninguna fuerza, mantendrá su movimiento constante en la dirección y cantidad que tiene, (es decir, $v = cte$). Desde este punto de vista, un sistema tal será no modulado, puesto que no existe ningún proceso que tenga función alguna en su variación. Y los sistemas inerciales de este tipo se pueden considerar sistemas no modulados.

Supongamos ahora que en las cercanías de ese cuerpo de masa m existe otro cuerpo de masa muy superior m'. En este caso la velocidad y dirección del movimiento del primer cuerpo variarán a causa de la fuerza gravitatoria ejercida por el segundo. El cuerpo pequeño será atraído hacia el cuerpo grande, y esa fuerza de atracción será tanto mayor cuanto menor sea la distancia entre ambos (será inversamente proporcional al cuadrado de la distancia). En este caso, se puede decir que existe una modulación y que el cuerpo de masa m' modula el movimiento del cuerpo de masa m. Ambos cuerpos han de entenderse como procesos, esto es, como sistemas dinámicos. Además, si se toma como despreciable el efecto de la masa pequeña en la masa grande ($m << m'$), se tratará de un proceso de modulación sin automodulación.

Por último, supongamos el sistema dinámico de un péndulo simple, formado por un hilo de longitud L y una bola colgante que oscila entre dos posiciones con un movimiento descrito por el ángulo θ del péndulo con respecto a la vertical. Este es un sistema complejo en el cual toman parte diferentes sistemas tales como la bola y la tensión del hilo, la fuerza de gravedad (g) o la posición inicial del péndulo, entre otros. Estos elementos interfieren entre sí de manera tal que el efecto del movimiento tiene la particularidad de que la fuerza producida por la gravedad y opuesta al movimiento pendular ascendente depende del ángulo θ, siendo máxima cuando el ángulo es $\pm 90^{\circ}$. Por consiguiente, el modo en que varía la variación del ángulo θ, esto es, su aceleración $\ddot{\theta}$, depende del valor de ese mismo ángulo θ. Este es un caso de automodulación que se describe a través de la ecuación $\ddot{\theta} + (g/L)\sin\theta = 0$.[167]

2.4.2.1. Un marco teórico: los sistemas dinámicos

La representación formal más elaborada de la automodulación desde el punto de vista *cuantitativo* la proporcionan las ecuaciones diferenciales de las matemáticas. A respecto, Rescher nos recuerda que "ni la lógica de objeto y predicado ni tampoco la gramática de sujeto y verbo imperan en el lenguaje de la naturaleza. Por el contrario, lo que representa mejor el aspecto procesual de la naturaleza es el lenguaje matemático de las ecuaciones diferenciales" (Rescher 1996: 92).

La interpretación de las nociones de modulación y automodulación a través de ecuaciones diferenciales requiere de un marco matemático general. El marco que voy a escoger es la dinámica no lineal de Strogatz (1994). Introduciré mínimamente este marco teórico para

[167] $\ddot{\theta}$ representa la aceleración del ángulo θ y, por tanto, del movimiento pendular. Ver Strogatz (1994: 168).

después definir, dentro del mismo, la noción de automodulación.

Para Strogatz (1994: 6), un sistema dinámico con el tiempo tomado como variable continua se puede expresar mediante ecuaciones diferenciales del tipo siguiente:

$$\begin{aligned} \dot{x_1} &= f_1(x_1, x_2, \ldots, x_n) \\ &\vdots \\ \dot{x_n} &= f_n(x_1, x_2, \ldots, x_n) \end{aligned} \tag{fc1}$$

donde $\dot{x_i} \equiv dx_i/dt$, mientras que $x_1, x_2, \ldots, x_n$ representan aspectos cuantificables relevantes del sistema y f_i representa la dinámica o función que gobierna el cambio de la variable x_i.

Esta forma de representación es muy importante para la caracterización de los sistemas lineales. Temporalmente la llamaré forma canónica (fc1), para distinguirla de otras formas de representación.

En la forma canónica (fc1), se distinguen dos características importantes de los sistemas dinámicos:

(a) dimensión: la dimensión del sistema dinámico viene dada por el número de variables que intervienen en él (entendiendo variables como variables en el tiempo). Por ejemplo, el sistema genérico que se ha representado arriba es n-dimensional. Siguiendo la terminología clásica, Strogatz (1994: 8) llama a esto la dimensión del espacio fásico. Las dimensiones del espacio fásico no tienen por qué coincidir con las dimensiones del espacio físico. Por el contrario, el número de dimensiones del espacio fásico es el número de variables relevantes para describir completamente cómo varía un sistema dinámico.

(b) linealidad: un sistema dinámico es lineal si todas las funciones $f_1, f_2, \ldots, f_n$ son lineales, esto es, si las variables aparecen solo en potencia primera (por ejemplo $a_1x_1 + a_2x_2 + \cdots + a_nx_n$). Un sistema dinámico es no lineal si alguna de las funciones $f_1, f_2, \ldots, f_n$ es no lineal, es decir, contiene productos (x_ix_j), potencias (x_i^3) u otras funciones específicas no lineales (como, por ejemplo, $\sin x_i$). Analíticamente, la linealidad de un sistema es un rasgo de gran importancia: los sistemas lineales son más fáciles de descomponer y analizar. Los sistemas no lineales, por el contrario, son más complejos y difíciles de resolver.

La complejidad del análisis de los sistemas dinámicos se explica precisamente por estas dos características. Los sistemas dinámicos de dimensión alta o no linealidad alta resultan muy complejos de resolver. Brunton y Kutz (2019: 300) nos recuerdan que es posible operar en los sistemas mediante transformaciones de variables para que, sin perder información del sistema, disminuya su no linealidad, a costa de aumentar su dimensión incluso hasta infinito. Estas operaciones ponen de manifiesto la necesidad de un compromiso: si se reduce la no linealidad, aumenta la dimensión y, a la inversa, si reduce la dimensión,

aumenta la no linealidad.[168]

2.4.2.2. Definición de modulación y clausura

A continuación voy a introducir una noción adicional que no está explícitamente indicada en el marco teórico de Strogatz, y que denominaré "modulación matemática" o, simplemente, "modulación". De momento, hablaré de modulación entre variables, y me mantendré al margen de la importante cuestión acerca de si estas variables representan o describen variaciones en procesos.

Supongamos un sistema dinámico de dimensión 1 representado por la siguiente forma canónica:

$$\dot{x_1} = f_1(x_1, \ldots, x_n) \qquad \text{(automodulación)}$$

La modulación (matemática) entre variables sucede cuando la variación de una variable depende de otras. Por ejemplo, en el sistema dado, las variables x_i (para $1 \leqslant i \leqslant n$) modulan a la variable x_1.

La automodulación (matemática) sucede cuando la variación de una variable depende de sí misma. Por ejemplo, en el sistema mencionado, la variable x_1 se automodula.

Supongamos ahora un sistema dinámico de dimensión 2 representado por la siguiente forma canónica:

$$\begin{aligned} \dot{x_i} &= f_i(\ldots x_j \ldots) \\ \dot{x_j} &= f_j(\ldots x_i \ldots) \end{aligned} \qquad \text{(modulación cruzada)}$$

La modulación (matemática) cruzada sucede cuando se produce una modulación matemática cruzada entre variables. Este es el caso del sistema dinámico mencionado, donde las variables x_i y x_j tienen modulación cruzada.[169]

Por último, supongamos un sistema dinámico de dimensión n en el que se cumple lo siguiente para tres variables x_i, x_j y x_k:

[168] Sin entrar en los detalles, baste indicar que la afirmación anterior se basa en las implicaciones del operador Koopman. Dicen Brunton y Kutz: "Bernard O. Koopman demostró que es posible representar un sistema dinámico no lineal en términos de un operador lineal de dimensión infinita que actúa sobre un espacio hilbertiano de funciones de medición de los estados del sistema. Este operador, denominado operador Koopman, es lineal, y su descomposición espectral caracteriza el comportamiento de un sistema dinámico no lineal [...]. Sin embargo, también es de dimensión infinita." (Brunton y Kutz 2019: 300)

[169] En ciencias matemáticas el término habitual para referirse a la modulación cruzada es "acoplamiento" (Strogatz 1994). No obstante, se introduce la noción de modulación para distinguir con más precisión este concepto. El acoplamiento es una forma de modulación con clausura.

$$\begin{aligned} \dot{x_i} &= f_i(\dots x_j \dots) \\ \dot{x_j} &= f_j(\dots x_k \dots) \\ \dot{x_k} &= f_k(\dots x_i \dots) \end{aligned} \qquad \text{(modulación con clausura)}$$

La modulación (matemática) con clausura es aquella en la cual se produce una cadena cerrada de modulaciones entre variables. Por ejemplo, en el sistema anterior existe una modulación con clausura de las tres variables indicadas x_i, x_j y x_k.

En realidad, la clausura es una propiedad tanto de la automodulación como de la modulación cruzada, y éstas son realmente tipos de la modulación con clausura. La automodulación es una modulación con clausura en una única variable. La modulación cruzada es una modulación con clausura en dos variables.

En cualquier caso, es muy importante retener la idea de que los tipos aquí mostrados son solo paradigmas. De igual manera que, por el mero hecho de que un dibujo esté compuesto de rectas y curvas, no es posible clasificar todos los dibujos en dibujos rectos y dibujos curvos (sino que la mayoría serán combinaciones de ambos), tampoco sucede que, por el mero hecho de que se pueda establecer una diferencia entre automodulación y modulación cruzada, es posible clasificar todos los sistemas dinámicos de dos dimensiones en sistemas puramente automodulados y sistemas con modulación cruzada pura. En la realidad lo habitual es encontrar sistemas lineales con modulaciones mixtas o incluso con otras variables, tales como el tiempo.[170]

2.4.2.3. Modulación en sistemas dinámicos sencillos

¿Cuáles son las consecuencias de aplicar esta distinción en el análisis matemático de los sistemas dinámicos? La tarea de identificar estas características es compleja, entra dentro del ámbito de la matemática y excede mis objetivos aquí. A continuación, voy a aplicar los conceptos de automodulación en algunos sistemas dinámicos sencillos. Los sistemas dinámicos más sencillos son aquellos que tienen dimensión baja o no linealidad escasa. Yo voy a escoger únicamente sistemas de dimensión 1 y sistemas lineales de dimensión 2. Esto es solo la punta del iceberg de los sistemas dinámicos, pero sirve como modelo para el análisis de sistemas no lineales a través del método de linealización.

Acerca de los sistemas dinámicos de dimensión 1

Los sistemas dinámicos de primer orden (esto es, de dimensión 1), donde la evolución de la única variable se describe a través de una única ecuación diferencial —$\dot{x} = f(x)$—, son, por definición, siempre sistemas exclusivamente automodulados, esto es, con clausura en una única variable, y siempre tienen un comportamiento monótono, esto es, no existen en ellos oscilaciones (Strogatz 1994: 28-29). Esto quiere decir que, para cada valor de x,

[170] Para una forma más completa de los sistemas dinámicos, ver la distinción entre sistemas autónomos y sistemas no autónomos, así como la forma canónica (fc2), que se presentará más adelante.

o bien la variable tiende a un valor o bien tiende a $\pm\infty$.

Acerca de los sistemas dinámicos lineales de dimensión 2

Los sistemas dinámicos de dimensión 2 de tipo lineal se representan de la siguiente manera:

$$\begin{aligned}\dot{x} &= ax + by \\ \dot{y} &= cx + dy\end{aligned} \qquad \text{(sist. lineal dim2)}$$

En forma de matriz, este tipo de sistemas se pueden representar de forma matricial así:

$$\begin{pmatrix}\dot{x} \\ \dot{y}\end{pmatrix} = \begin{pmatrix}a & b \\ c & d\end{pmatrix}\begin{pmatrix}x \\ y\end{pmatrix} \qquad \text{(forma matric.)}$$

Denominaremos matríz **A** a la matriz que describe la dinámica de este sistema:

$$\mathbf{A} = \begin{pmatrix}a & b \\ c & d\end{pmatrix} \qquad (\mathbf{A})$$

La modulación dentro de sistemas dinámicos lineales de dimensión 2 es relativamente sencilla, por lo menos si se comparan con los sistemas dinámicos no lineales y de dimensión superior a 2. Por ese motivo, es sencillo describirla en términos de modulación. En este caso, la matriz **A** contiene toda la información acerca del tipo de modulación que sucede en el sistema.

Es posible definir un sistema de este tipo como dotado de automodulación pura si b y c son nulos ($b = c = 0$, es decir, no hay modulación cruzada) y alguno de los términos de la diagonal son no nulos ($a \neq 0$ o $d \neq 0$).

$$\begin{pmatrix}\dot{x} \\ \dot{y}\end{pmatrix} = \begin{pmatrix}a & 0 \\ 0 & d\end{pmatrix}\begin{pmatrix}x \\ y\end{pmatrix} \qquad \text{(automod. pura)}$$

A su vez, cada una de las automodulaciones puede ser de dos tipos: automodulación positiva o negativa. En el caso de la variable x, si $a > 0$, la automodulación de la variable x será positiva, mientras que si $a < 0$, la automodulación será negativa. Por otro lado, estos sistemas de dimensión 2 pueden dividirse entre sistemas con automodulaciones opuestas entre sí (si el signo de los términos a y d es distinto, es decir, si $ad < 0$) o no opuesta entre sí (si el signo de los términos a y d es idéntico, es decir, si $ad > 0$).[171] Por último,

[171] Si $ad = 0$, entonces $a = 0$ o $d = 0$, con lo cual no sería posible hablar de oposición entre automodulaciones, puesto que el sistema solo exhibirá una automodulación y no dos.

si $a + d = 0$, tendríamos un caso de automodulación opuesta cancelada.

Por otro lado, un sistema lineal de dimensión 2 se puede definir como de modulación cruzada pura si la matriz **A** tiene la diagonal nula ($a = d = 0$, es decir, no hay automodulación) y los dos términos cruzados son no nulos ($b \neq 0$ y $c \neq 0$).

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{(mod. cruz. pura)}$$

Del mismo modo, los sistemas lineales de dimensión 2 con modulación cruzada pura pueden dividirse entre sistemas opuestos, si el signo de los términos b y c es distinto (esto es, si $bc < 0$) y sistemas no opuestos, si el signo de los términos b y c es el mismo (esto es, si $bc > 0$).

A la hora de describir cómo evolucionará un sistema dinámico es importante determinar cuáles son sus puntos fijos. El punto fijo es el valor de las variables para el cual el sistema se vuelve estacionario. En el caso de sistemas de dimensión dos, el punto fijo es el valor de x e y para el cual $\dot{x} = \dot{y} = 0$. En sistemas lineales de dimensión 2, siempre hay un punto fijo en (0,0).

El marco teórico de sistemas dinámicos dice que es posible conocer qué tipos de punto fijo habrá en el espacio fásico descrito por la matriz **A** con solo analizar su traza ($\tau = a + d$) y su determinante ($\Delta = ad - bc$).[172] El análisis de las soluciones para cada sistema dinámico es complejo y es necesario recurrir a las herramientas adecuadas para ello. Sin embargo, se pueden extraer algunas conclusiones acerca de los sistemas lineales de dimensión 2 únicamente a partir del tipo de modulación que tienen.

(a) En casos de automodulación pura, lo que marca si el punto fijo será estable o inestable es el signo de la oposición. Por ejemplo, la estabilidad se produce solo si la automodulación es no opuesta y negativa.[173] Nota: que el punto fijo sea estable indica que el sistema tiende a ese punto. Esto se interpreta como que, en el caso de automodulación pura, el sistema tiende a estabilizarse solo si ambas variables tienden a disminuir conforme mayor es su valor respectivo.

(b) Que el punto fijo sea un centro quiere decir que el sistema llega a situaciones de bucle donde los mismos estados se repiten una y otra vez de manera cíclica alrededor de ese punto. Esto solo sucede en casos de modulación cruzada opuesta donde la automodulación es nula o está cancelada.[174] Esto se interpreta de la siguiente manera: para que el sistema llegue a situaciones de bucle, es necesario que las variables solo se modulen entre sí y que lo hagan de manera opuesta.

[172] Strogatz (1994: 123).

[173] Puesto que en ese caso la traza de la matriz **A** es negativa ($\tau < 0$), mientras que el determinante es positivo ($\Delta > 0$) (Strogatz 1994: 137).

[174] Los centros se corresponden con sistemas cuya matriz **A** presenta traza nula ($\tau = 0$), es decir, automodulación nula o cancelada y determinante positivo ($\Delta > 0$), que solo se puede conseguir en el caso de modulación cruzada opuesta.

2.4.2.4. Formas canónicas

La modulación con clausura se ha definido siguiendo a Strogatz, quien representa los sistemas dinámicos mediante la forma canónica (fc1). Recordemos una vez más esta forma canónica:

$$\begin{aligned} \dot{x_1} &= f_1(x_1, x_2, \ldots, x_n) \\ &\vdots \\ \dot{x_n} &= f_n(x_1, x_2, \ldots, x_n) \end{aligned}$$

Los sistemas dinámicos, empero, son complejos y esta forma, aparentemente, no parece recoger toda esa diversidad. Ahora mostraré que los sistemas dinámicos se pueden representar en una segunda forma canónica que recoge mejor esa complejidad, y cómo esta segunda forma canónica (fc2) se puede reducir a la forma canónica inicial de Strogatz (fc1).

Cómo reducir a la primera forma canónica

Strogatz (1994) representa los sistemas dinámicos mediante la forma canónica (fc1). Sin embargo, esta forma canónica no recoge toda la complejidad de los sistemas dinámicos. Esta forma canónica únicamente habla de relaciones de modulación con la derivada de primer orden. Sin embargo, es fácil imaginar sistemas en los que aparezcan derivadas de segundo orden tales como $\ddot{x} = x$ o, incluso, con órdenes superiores. ¿Qué sucede con los sistemas en los que aparecen derivadas de segundo, tercer u otro orden superior? Por otro lado, la forma canónica (fc1) no presenta dependencia directa de la variable t. Sin embargo, es perfectamente posible imaginar sistemas no autónomos, esto es, sistemas que dependan del tiempo. ¿Qué sucede con los sistemas que tienen dependencia directa de la variable t, esto es, del tiempo? Ahora voy a mostrar cómo es posible reducir esos dos tipos de sistemas a la forma canónica de manera que, también en ellos, se pueda aplicar la noción de clausura en la modulación.

Con respecto a los sistemas en los que aparezcan derivadas de segundo orden tales como $\ddot{x} = x$ (u otros órdenes superiores), y puesto que la definición de modulación se hizo con respecto a la forma canónica, no es posible garantizar que estos sistemas sean automodulados. De hecho, no lo serán.

Supongamos, por ejemplo, el sistema siguiente, en el cual la aceleración se relaciona con la posición:

$$\ddot{x} = f(x) \tag{ac1}$$

donde $\ddot{x} = d^2x/dt^2$.

En la forma canónica (fc1) estos sistemas no son de una variable, sino de dos variables:

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= f(x) \end{aligned} \tag{ac2}$$

Por consiguiente, este tipo de sistemas tienen dimensión 2 y exhiben modulación cruzada pura. Este es el caso de sistemas lineales que presentan oscilaciones tales como el formado por una masa colgada de un muelle lineal (Strogatz 1994: 124):

$$m\ddot{x} + kx = 0 \qquad \text{(muelle)}$$

donde x representa la posición con respecto al punto de descanso de la masa, m es la masa y k es la constante elástica del muelle.

En forma canónica (fc1), este sistema se representa así:

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -(\frac{k}{m})x \end{aligned} \qquad \text{(muelle fc1)}$$

Por consiguiente, se trata de un sistema con modulación cruzada pura en el que, además, la modulación es siempre opuesta (puesto que, en el muelle lineal, $k > 0$). De hecho, el punto fijo (0,0) es un centro, como es habitual en los sistemas con modulación cruzada pura (y contra lo que sucede en los sistemas con automodulación pura). Por consiguiente, los estados se repiten continuamente y se produce oscilación continua.[175]

La transformación anterior se puede generalizar. Para ello, distinguiré entre modulación de primer orden y modulación de orden superior. La modulación de primer orden es aquella que sucede en la derivada primera. Por ejemplo, el sistema $\dot{x} = f(y)$ representa una modulación de primer orden, y se puede decir que la variable y modula en primer orden a la variable x. La modulación de orden superior es aquella que sucede en la derivada de orden superior (segundo, tercero, etc.). Por ejemplo, el sistema $\ddot{x} = f(y)$ representa una modulación de segundo orden. De igual modo, se podrá decir que el sistema $\ddot{x} = f(x)$ representa una automodulación de segundo orden.

Teniendo esta distinción presente, se puede decir que todo sistema con modulación de segundo orden puede transformarse en un sistema con modulación de primer orden.

Demostración: Para todo $\ddot{x} = f(y)$ se puede definir una variable v que representa la derivada de x con respecto a t tal que:

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= f(y) \end{aligned} \qquad \text{(ec1)}$$

Igualmente, todo sistema con automodulación de segundo orden se puede transformar en un sistema con modulación cruzada de primer orden.

Demostración: Para todo $\ddot{x} = f(x)$ se puede definir una variable v que representa la derivada de x con respecto a t tal que:

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= f(x) \end{aligned} \qquad \text{(ec2)}$$

[175] Un ejemplo similar para un sistema no lineal es el del péndulo, que se representa por la fórmula $\ddot{\theta} + \sin\theta = 0$. Fórmula después de la desdimensionalización. Ver Strogatz (1994: 168).

Igualmente, todo sistema con modulación clausurada de segundo orden se puede transformar en un sistema con modulación de primer orden que mantiene clausura.

Demostración: Para todo $\ddot{x} = f_1(y)$; $\ddot{y} = f_2(x)$ se puede definir una variable v_x que representa la derivada de x con respecto a t y una variable v_y que representa la derivada de y con respecto a t tal que:

$$\begin{aligned} \dot{x} &= v_x \\ \dot{v_x} &= f_1(y) \\ \dot{y} &= v_y \\ \dot{v_y} &= f_2(x) \end{aligned} \qquad \text{(ec3)}$$

Y así sucesivamente.

En cuanto a la segunda cuestión, la forma canónica no presenta dependencia directa de la variable t. Los sistemas dinámicos donde la función f que describe el sistema no depende directamente del tiempo t se denominan sistemas autónomos (Strogatz 1994). Sin embargo, es perfectamente posible imaginar sistemas no autónomos, esto es, sistemas que dependan del tiempo, como, por ejemplo, una perturbación que sucede en determinado momento, sea esta deliberada o no. Strogatz (1994: 8) propone que los sistemas no autónomos, es decir, que dependen del tiempo t, tales como $\dot{x} = f(x,t)$ se describan a través de la forma canónica (fc1), añadiendo una dimensión nueva al sistema (x_{n+1}) tal que su derivada con respecto al tiempo sea igual a 1 ($\dot{x}_{n+1} = 1$).

Supongamos el sistema formado por un muelle con constante elástica k que sostiene una masa m hundida en una sustancia viscosa con constante de amortiguación b que comienza a oscilar con una fuerza impulsora F. Supongamos que el movimiento oscilatorio amortiguado de la masa puede describirse de manera adecuada mediante la siguiente ecuación:

$$m\ddot{x} + b\dot{x} + kx = F\cos t \qquad \text{(m. amort.)}$$

Por tanto, este sistema no es autónomo y, además, tiene automodulación de segundo orden. La ecuación puede transformarse en la forma canónica (fc1) añadiendo dos dimensiones más, una por la automodulación de segundo orden y otra por la dependencia de t:

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= \frac{1}{m}(-kx - by + F\cos z) \\ \dot{z} &= 1 \end{aligned} \qquad \text{(m. amort. fc1)}$$

Segunda forma canónica para representar sistemas dinámicos

Teniendo en cuenta lo anterior, ya estoy en disposición de representar cualquier sistema dinámico en la forma canónica, tenga modulaciones clausuradas de primer orden o superior y sea autónomo o no. En general, se definen cinco elementos para representar completamente los sistemas dinámicos. Siguiendo a Brunton y Kutz (2019) la evolución de un sistema dinámico asociado a un conjunto de variables $\mathbf{x} = (x_1, x_2 \dots x_n) \in \mathbb{R}^n$ se describe de la siguiente manera:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t, u, \beta) \tag{fc2}$$

A esta forma de representar los sistemas dinámicos la denominaré forma canónica (fc2). Los elementos que forman parte de ella son los siguientes:

x: Es el vector de variables, esto es, el mínimo conjunto de estados que representa adecuadamente el sistema cuya dinámica se quiere describir.

f: Es el vector de funciones que describe completamente la dinámica del sistema. Estas funciones pueden ser lineales o no lineales.

t: Representa el tiempo. Por ejemplo, en un determinado momento puede suceder algo que perturbe el sistema. Como se ha indicado más arriba, los sistemas donde no se produce este tipo de dependencia, esto es, donde $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u, \beta)$, se denominan sistemas autónomos.

u: Representa el conjunto de parámetros que determinan el funcionamiento del sistema y son controlables y, por consiguiente, se puede actuar sobre ellos para influir en el sistema. Por ejemplo, en el caso del arco podría ser un regulador que permitiera modificar la extensión del cuerpo del arco o, en el caso del oscilador armónico forzado, podría ser la constante de amortiguación b, si se disminuyera deliberadamente vertiendo una sustancia menos viscosa en la cubeta donde está hundida la masa. Esta acción de control puede estar relacionada con salidas del sistema (bucle cerrado) o no (bucle abierto) y es dependiente del tiempo, es decir, $u(t)$.

β: Representa el conjunto de parámetros que determinan el funcionamiento del sistema y no son controlables, esto es, no se puede actuar sobre ellos para influir en el sistema. Por ejemplo, en el caso del muelle, los parámetros β serían la masa m y la constante elástica k. Generalmente, no son dependientes del tiempo.

¿Qué significa la clausura de modulación matemática en los sistemas dinámicos? La clausura significa que el modo en que las variables cambian en el tiempo depende de las variables mismas con cierto cierre (que se ha descrito antes). Esta característica es importante, como se tratará de demostrar, pero no es suficiente para explicar completamente el sistema. También es necesario conocer **f**, u, β o incluso el tiempo t.

Aquí no se defiende la tesis de que el concepto de modulación clausurada es suficiente para explicar los sistemas dinámicos, ni tampoco la tesis de que es suficiente (por sí solo)

para explicar la conciencia. Eso sería pedirle demasiado a la modulación. La tesis que se defiende aquí es que no se debe ignorar la noción de modulación para explicar los sistemas dinámicos y que, si la conciencia es un sistema dinámico, entonces no se debe ignorar la noción de automodulación y la modulación con clausura para explicar la conciencia. Pero antes de volver a la conciencia, es preciso estudiar la relación entre modulación con clausura y los procesos autotransformativos.

2.4.3. Clausura de interferencia y clausura de modulación

Hasta ahora he caracterizado la noción de proceso autotransformativo en el ámbito de la mereología de procesos a través de tres propiedades (complejidad, interferencia y clausura, esto es, composición CIC) y también he caracterizado la noción de modulación en su forma más definida, a saber, su forma matemática en el marco teórico de los sistemas dinámicos. De entrada, es necesario recalcar que autotransformatividad y automodulación (o clausura en la modulación) no son exactamente lo mismo, por varias razones. Primera, una nos dice cosas acerca de procesos e interferencias, mientras que la otra nos dice cosas acerca de variables y funciones. Segunda, un proceso se autotransforma mediante otros procesos intermedios, mientras que una variable puede modularse a sí misma. La confusión entre ambas ideas puede conducir a aporías o a ver regresiones infinitas donde no las hay.

Sin embargo, hay indicios sólidos de que existe una relación entre la composición dinámica CIC y la modulación con clausura. Por un lado, la idea de clausura está en ambos conceptos. Por otro lado, la interpretación matemática de sistemas físicos complejos donde suceden interferencias cruzadas requiere, a menudo, de sistemas dinámicos representados por modulación matemática con clausura.

Aquí voy a tratar de defender una idea con respecto a esta relación: que la forma natural en que se produce clausura en la modulación es la composición dinámica CIC que se estableció para los procesos autotransformativos.[176] No toda modulación con clausura se corresponde con procesos autotransformativos. Sin embargo, la automodulación y, en general, la modulación con clausura, es una propiedad característica de los procesos autotransformativos, de tal manera que los procesos autotransformativos se caracterizan al menos por cuatro propiedades, tres de tipo mereológico (composición compleja, interferencia y clausura) y una de tipo funcional (modulación con clausura). Eso quiere decir que las tres primeras propiedades no son suficientes, por sí solas, para caracterizar los procesos autotransformativos.

[176] Durante este texto se usarán los términos "composición CIC" y "composición dinámica de interferencia mutua" como sinónimos.

2.4.3.1. Casos de modulación con clausura acompañados de composición dinámica de interferencia mutua

Comenzando con la primera idea, a continuación voy a introducir tres casos concretos en los que la modulación con clausura es un propiedad que acompaña a los procesos con composición dinámica de interferencia mutua (composición CIC). En primer lugar presentaré dos casos descritos matemáticamente: el caso de la ecuación logística y el sistema formado por un muelle y una masa. En ambos casos coincide clausura de modulación y composición CIC. En segundo lugar introduciré un caso basado en la teoría de control, de acuerdo a la cual la composición dinámica CIC es adecuada para producir clausura en la modulación (y alterar, de ese modo, los sistemas dinámicos). En tercer lugar introduciré un caso biológico, que se basa en la idea de que las estructuras de bucle aparecen en la naturaleza con extrema frecuencia.

Ecuación logística y composición CIC

Muchos de los sistemas dinámicos que se presentan en su forma canónica (fc1) o (fc2) con alguna forma de clausura en su modulación admiten una interpretación física a través de una composición de procesos con interferencia mutua (tipo CIC) o, lo que es, lo mismo, admiten una mereología como la característica de los procesos autotransformativos. Probaré esto en algunos de los casos más paradigmáticos de sistemas dinámicos.

El primer ejemplo es la ecuación logística (Verhulst 1838), que representa el crecimiento de poblaciones limitado por los recursos mediante un sistema dinámico no lineal de dimensión 1 de la siguiente manera:[177]

$$\dot{N} = rN(1 - \frac{N}{k}) \qquad \text{(ec. log.)}$$

donde $N(t)$ representa la población dependiente del tiempo, r representa la tasa de crecimiento y k representa la capacidad de soporte (*carring capacity*) del sistema cuyos recursos limitados son consumidos por la población. La solución de este sistema lineal es que, si existe población (si $N > 0$) esta población tenderá a un punto estable marcado por la capacidad de soporte, creciendo si está por debajo de ese punto o decreciendo si está por encima de él. La tasa de crecimiento da una idea de cuán rápido se producirá esa aproximación al punto estable.

La adecuación de la ecuación logística para describir diferentes ecosistemas en biología es una cuestión más o menos controvertida. Strogatz (1994: 23) señala que se adecúa mejor para describir el crecimiento de colonias de bacterias en el laboratorio que para el caso de insectos involucrados en ecosistemas más complejos tales como la mosca de la fruta. En cualquier caso, aquí no se sostiene la adecuación de este sistema dinámico en los casos biológicos, sino el hecho de que la interpretación de la ecuación logística pasa por estipular procesos complejos que interfieren con cierta clausura. En el caso de la ecuación logística hay dos términos que indican automodulación matemática. El primero es positivo (rN), y refleja un crecimiento exponencial de la población. Este crecimiento es atemperado por

[177] Strogatz (1994: 21)

un segundo término negativo ($-\frac{rN^2}{k}$). El efecto combinado de esos dos términos se interpreta a través de procesos con interferencia mutua del tipo diente de sierra. En concreto habrá una interferencia del grupo poblacional en el ecosistema del cual consume recursos complementada con una interferencia del ecosistema consumido en el propio grupo poblacional conforme los recursos escasean. Por consiguiente, la interpretación más habitual de la ecuación logística se basa en procesos compuestos, que tienen relaciones de interferencia y que presentan clausura. En otras palabras, se basa en procesos con composición CIC.

La interpretación como proceso transformativo tiene, además, una ventaja sobre la pura interpretación matemática. Obliga a determinar específicamente cuál es la naturaleza de las dos interferencias. Se puede, por ejemplo, a tenor de la especificación de la segunda interferencia, diferenciar entre una teoría explicativa basada en la interferencia de la falta de recursos en el proceso de reproducción, una teoría basada en la interferencia por inanición o una teoría mixta. En otras palabras, de igual manera que la interpretación matemática del sistema dinámico añade matices a la interpretación mereológico-procesual (tales como la idea de tasa de crecimiento, o la capacidad de soporte), la interpretación mereológico-procesual añade matices que, de ceñirse exclusivamente a la forma canónica del sistema dinámico, quedarían indefinidos.

Muelle y composición CIC

Otro ejemplo paradigmático de sistema dinámico es el formado por el sistema de un muelle y una masa que cuelga de él. Este ejemplo guarda cierta relación con el modelo arco y flecha que se presentó antes. Este sistema presenta diferentes versiones de forma canónica, dependiendo del medio en el que se encuentre la masa (oscilación amortiguada) entre otros factores. La forma más simple es la del oscilador armónico simple. De acuerdo con la fórmula de Hooke, se puede describir así:[178]

$$m\ddot{x} + kx = 0 \qquad \text{(muelle)}$$

donde m es la masa y k es la constante elástica del muelle.

Como se ha visto antes, este sistema (con automodulación de segundo orden) se puede representar también en la forma canónica (fc1) de modulación cruzada pura y opuesta:

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\frac{k}{m}x \end{aligned} \qquad \text{(muelle fc1)}$$

En este caso, la interpretación mereológico-procesual del sistema descrito por la forma canónica de arriba más sencilla es que existe una interferencia mutua de tipo diente de sierra entre el muelle y la masa, de acuerdo a la cual la masa con su gravedad deforma el muelle (interferencia 1), lo que a su vez produce un cambio en la fuerza del muelle que es mayor (en valor absoluto) cuanto mayor es la deformación (interferencia 2), de tal manera que ambas interferencias tienden en este caso a oscilar continuamente y hasta el

[178] Strogatz (1994: 124)

infinito (siempre y cuando no haya otros factores de automodulación).

2.4.3.2. Teoría de control

La teoría de control es una rama de la ingeniería práctica que se ocupa del "estudio y práctica en la manipulación de sistemas dinámicos" (Brunton y Kutz 2019). La teoría de control se basa, de manera generalizada, en mecanismos de reentrada que producen modificaciones en los sistemas dinámicos del sistema. Esto significa que existe una relación entre la composición CIC y la propiedad de la clausura en la modulación, puesto que la forma de manipulación de sistemas dinámicos más eficiente consiste precisamente en introducir procesos autotransformativos adicionales en el sistema, tales como los propios de mecanismos de bucle cerrado por retroalimentación. En forma de argumento, esto se puede esquematizar en cuatro pasos:[179]

(1) De entre los mecanismos de control de la dinámica de sistemas, el más eficiente de todos es el bucle cerrado con retroalimentación.

(2) El bucle cerrado por retroalimentación permite controlar la dinámica de los sistemas fundamentalmente porque permite alterar la forma específica de clausura de la modulación que se da en el sistema.

(3) El bucle cerrado con retroalimentación tiene una composición dinámica CIC tipo bucle diente de sierra, esto es, la propia de los procesos autotransformativos.

(4) Por consiguiente, existe una relación entre la composición dinámica CIC y la propiedad de clausura en la modulación de los procesos. En concreto, la composición dinámica CIC es necesaria para la modificación de la clausura en modulación.

El paso (1) viene dado por la propia experiencia, que se refleja en la teoría de control. En la actualidad existen dos grandes tipos de mecanismos activos de manipulación de sistemas dinámicos (Brunton y Kutz 2019: 326-327). Por un lado, los mecanismos de control basados en bucles abiertos, los cuales introducen un actuador en el sistema basado en un modelado previo del sistema, y en el cual no se tiene en cuenta las perturbaciones externas al sistema, ni el ruido propio de los sensores ni tampoco la incertidumbre o las dinámicas de sistema no contempladas en el modelo previo. Por otro lado, los mecanismos de control basados en bucle cerrado, que introducen actuadores que son capaces de tener en cuenta las perturbaciones externas al sistema, así como el ruido propio de los sensores. La base de este segundo tipo de mecanismo de control es un bucle de retroalimentación que tiene en cuenta la salida del sistema.[180]

[179] Brunton y Kutz (2019)

[180] El ejemplo del control del tráfico urbano sirve para ilustrar la diferencia entre estos dos mecanismos del control, esto es, el mecanismo de bucle abierto y el mecanismo de bucle cerrado. Un mecanismo de bucle abierto consistiría en un sistema de semáforos automatizado y reprogramado de acuerdo a algoritmos que permitan actuar de distintas maneras en determinados tramos horarios, dando prioridad, por ejemplo, en las horas de entrada y salida del trabajo al sentido de tráfico más saturado en detrimento del menos saturado. Este mecanismo estaría preprogramado y no tendría en cuenta perturbaciones especiales tales como un accidente o un evento deportivo. El mecanismo de bucle cerrado, por su parte, tendría sistemas de medición en tiempo real que permitieran a los semáforos modificar su comportamiento de

Las ventajas del mecanismo de bucle cerrado sobre el bucle abierto son evidentes, hasta el punto de que es el mecanismo predilecto para el control de sistemas dinámicos, únicamente limitado por el coste. Sobre un sistema de control sencillo, como el caso del péndulo invertido, dicen Brunton y Kutz lo siguiente:

> El control de bucle abierto [...] usa un modelo del sistema para diseñar una señal de actuación que produzca como salida la referencia deseada. Sin embargo, esta estrategia basada en la preplanificación no es capaz de corregir las perturbaciones externas que sufre el sistema y, fundamentalmente, es incapaz de cambiar la dinámica. De ese modo, es imposible estabilizar un sistema inestable tal como el péndulo invertido con un mecanismo de bucle abierto, ya que sería necesario conocer perfectamente el modelo del sistema y, además, el sistema real debería estar perfectamente aislado de perturbaciones externas. (Brunton y Kutz 2019: 326)

El paso (2) se sigue del paso (1). Con respecto a los mecanismos de control con bucle cerrado Brunton y Kutz dicen que:

- Pueden estabilizar sistemas dinámicos inestables
- Pueden compensar perturbaciones externas
- Pueden corregir dinámicas no modeladas y modelar la incertidumbre (Brunton y Kutz 2019: 327)

Centrémonos en el primer punto. Si los mecanismos de bucle cerrado con retroalimentación son capaces de estabilizar sistemas dinámicos inestables, entonces son capaces de alterar la dinámica de los sistemas, puesto que pueden cambiar puntos inestables en puntos estables. Por ejemplo, en el caso del péndulo invertido, cuya forma canónica se puede expresar en su forma linealizada de la siguiente forma:[181]

$$\begin{aligned} \dot{x_1} &= x_2 \\ \dot{x_2} &= g\frac{x_1}{L} + dx_2 + u \end{aligned} \qquad \text{(pend. inv.)}$$

donde $x_1 = \theta$ (esto es, el ángulo del péndulo con respecto a la vertical), u simboliza una tuerca sujeta en el brazo del péndulo, g es la aceleración debida a la gravedad, L es la longitud del brazo del péndulo y d representa la amortiguación del movimiento en el péndulo debida al rozamiento. Si se desprecia la amortiguación y teniendo en cuenta que $g, L > 0$, el sistema presenta una modulación cruzada pura no opuesta que se traduce en un punto fijo de tipo silla para $\theta = 0$. En otras palabras, el sistema tenderá a alejarse de ese punto a la mínima perturbación, lo que coincide con la idea intuitiva de que el péndulo inverso tiende a caer hacia un lado u otro.

acuerdo al estado actual del tráfico, sea este estado fruto de un modelado anterior o de una perturbación en el sistema. Este segundo mecanismo requiere de retroalimentación.

[181] Brunton y Kutz (2019: 327)

Para transformar esta dinámica inestable en una dinámica estable, Brunton y Kutz proponen transformar la función u en un actuador como el siguiente:

$$u = -4x_1 - 4x_2 \tag{act}$$

Esta función adicional transforma la dinámica del sistema de una modulación cruzada pura a una modulación mixta cuyo punto fijo es estable. La aplicación de ese actuador requiere sensores que midan las variables x_1 y x_2 en salida, y modifique la entrada del sistema. Requiere, en otras palabras, de un mecanismo de bucle cerrado con retroalimentación.

El paso (3) indica que el mecanismo de bucle cerrado por retroalimentación requiere de una composición dinámica con interferencia mutua, como la propia de los procesos autotransformativos. Para demostrar el paso (3) es necesario revisar en qué consiste exactamente el mecanismo de bucle cerrado. De manera general, Brunton y Kutz lo describen por oposición al mecanismo de bucle abierto (ver figura 2.9).

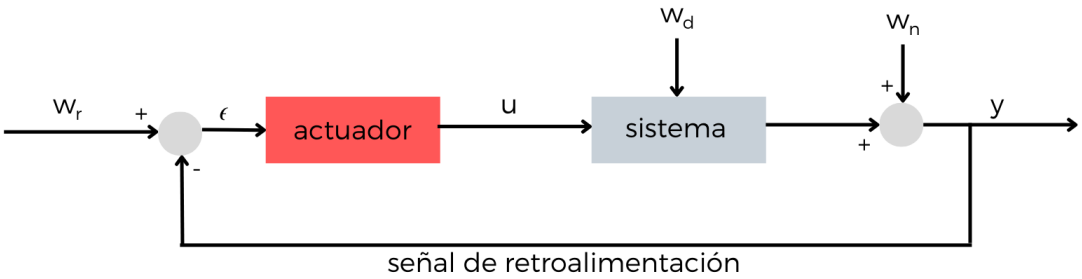


Figura 2.9: Modelo de retroalimentación para control de un sistema dinámico como, por ejemplo, el péndulo invertido. En la figura, y es la señal de los sensores, u es la señal actuadora, w_r es la señal de referencia, cuya diferencia con u da como resultado el error ϵ, y w_d y w_n son respectivamente la señal de perturbación externa y el ruido. Adaptado de Brunton y Kutz (2019: 327).

La base de todo mecanismo de control de bucle cerrado con retroalimentación es la existencia de sensores que recogen información en la salida del sistema y actuadores que modifican la entrada del sistema de acuerdo con esa información recogida. Este bucle es, de hecho, una interferencia en el sentido opuesto a la interferencia original del sistema, en el sentido de que va de salida hacia entrada, precisamente como en las composiciones de bucle tipo diente de sierra de las figuras 2.6(b) y 2.7(b). El flujo de información propio de la retroalimentación representa la clausura propia de la nueva composición dinámica del sistema, que tiene la composición, interferencia y clausura propia de la estructura mereológica de los procesos autotransformativos.

Por consiguiente (4), se puede concluir que la composición dinámica CIC es necesaria para establecer mecanismos de control de la clausura de la modulación que alteren la dinámica del sistema.

2.4.3.3. La biología y la composición con clausura

Si los mecanismos de control de sistemas dinámicos están basados en composiciones dinámicas de interferencia mutua, es natural esperar que ese tipo de composiciones estén presentes en la naturaleza y, en especial, en aquellos sistemas biológicos más complejos y avanzados a través de la evolución, ya que son capaces de compensar y equilibrar sistemas que, de otro modo, serían inestables. En efecto, este es el caso de muchos de los mecanismos de autocontrol dentro de los ecosistemas, que se han de estudiar mediante procesos de interferencia mutua entre especies, tales como la relación depredador-presa descrita por las ecuaciones de Lotka-Volterra o la relación entre consumidor-sistema consumido expresada en la ecuación logística.

Sin embargo, resulta notable la sobreabundancia de mecanismos de retroalimentación dentro de uno de los sistemas que, de manera mayoritaria, se ha tomado siempre como producto paradigmático de la evolución, a saber, el cerebro humano. Por un lado, es un hecho de sobra conocido que la estructura de conexión del cerebro tiene multitud de redes basadas en reentrada, y así se refleja en los estudios al respecto, desde el clásico esquema jerárquico de conexión entre áreas visuales del cerebro del primate (Felleman y van Essen 1991) hasta la complejidad recogida en el Human Connectome Project. Ya Felleman y van Essen señalaban al comienzo de su artículo que "la gran mayoría de los caminos [de conexión entre áreas visuales del cerebro del primate] incluyen conexiones recíprocas entre áreas" (Felleman y van Essen 1991: 1). Estos caminos de reentrada tienen su plasmación en redes funcionales de reentrada. Por mencionar solo dos de ellas, destacaré las conexiones de reentrada talámico-corticales o los caminos de reentrada directo, indirecto e hiperdirecto de los cuales forma parte el córtex motor, los ganglios basales y el cerebelo. En ambos casos parece haber indicios de que una de las funciones de esas conexiones de tipo lazo es estabilizar el sistema, bien sea permitiendo controlar la entrada de los estímulos (la llamada función de activación o *gating function* del tálamo, en el primer caso), bien suavizando y dando precisión a los movimientos (la función de estabilización de la respuesta motora, en el segundo caso).[182]

Por otro lado, los hallazgos de mecanismos de retroalimentación de tipo ya no solo anatómico sino funcional es una tónica dominante en la neurociencia. Los modelos de redes neuronales diferencian entre redes abajo-arriba o de alimentación directa (*bottom-up* o *feedforward*) y redes arriba-abajo o retroalimentadas (*top-down* o *feedback*).[183]

En lo que a los estudios de la conciencia respecta, se puede decir lo mismo. La importancia de los mecanismos de reentrada queda de manifiesto en Edelman (1989), Edelman y

[182]No hay muchas certezas absolutas en neurociencia, y sí, por el contrario, muchas incógnitas. Sin embargo, se resaltan estas dos formas de lazo, una que afecta a la entrada de estímulos en el cerebro y otra a la salida o respuesta motora del cerebro, porque tienen un alto grado de aceptación e incluso forman parte curricular de los estudios en neuroanatomía y fisiología (Cfr. Pineda (2015) y Baars y Gage (2007: 121)). Solo por incluir dos referencias, en la parte del lazo córtico-talámico-cortical, véase Sherman (2016), donde se diferencia la entrada de retransmisión de primer orden (desde la retina al núcleo geniculado lateral) y la de orden superior (desde la capa 5 del córtex). Por lo que al lazo córtex motor-ganglios basales-cerebelo se refiere, véase Middleton y Strick (2000) y recuérdese el papel de la substantia nigra en el control del movimiento puesto de manifiesto por la enfermedad de Parkinson.

[183]Ver Baars y Gage (2007: 73)

Tononi (2000) y Tononi (2008). La propia noción de información integrada Φ se puede tomar (si olvidamos por un momento la tesis fuerte sostenida por Tononi de que representa a la propia conciencia) como una herramienta matemática para medir el nivel de reentrada funcional dentro de un sistema dinámico como el cerebro.

La explicación más plausible para la sobreabundancia de estas estructuras y mecanismos de reentrada en la naturaleza es que esta composición dinámica es más adecuada que otras estructuras (tales como la conexión directa) para el control de la dinámica de los sistemas.

2.4.4. Caracterización completa de los procesos autotransformativos

Hay razones de peso para sostener que los fenómenos de modulación que presentan clausura (tales como automodulación o modulación cruzada) vienen acompañados de manera típica por composiciones dinámicas de interferencia mutua.

La inversa, sin embargo, no se puede afirmar con tanta rotundidad. La composición dinámica basada en interferencia mutua no es suficiente para asegurar que exista procesos autotransformativos en la medida en que no sea posible asegurar transitividad en todas las relaciones de interferencia causal y en todas las relaciones de pertenencia. Un contraejemplo es la rotura de la cadena causal, de acuerdo con la cual la sucesión A causa B y B causa C no garantiza que A cause C, al menos siempre y cuando no hay una cierta homogeneidad. Por ejemplo, que la persona B perciba visualmente a A y la persona C perciba visualmente a B no implica que la persona C perciba visualmente a A. Por otro lado, como ya se indicó antes, el hecho de que una bola de billar A golpee a otra bola B y esta, a su vez, golpee de nuevo a la bola A no asegura un proceso autotransformativo sostenido en el tiempo.

La propiedad de clausura en la modulación, por el contrario, permite identificar de manera funcional la parte autotransformativa de los procesos autotransformativos. Por consiguiente, los procesos autotransformativos quedan caracterizados por dos propiedades, una de tipo composicional, a saber, que tiene una composición de interferencia mutua (complejidad, interferencia y clausura) y otra de tipo causal/funcional, a saber, que sus variables tienen cierta clausura en la modulación.

2.5. Conclusiones

Los estudios acerca de la conciencia están cargados de conceptos tradicionales tales como la idea de estado (estado mental, estado consciente), la idea de inmediatez (percepción inmediata, aprehensión directa) y la noción de sustancia (sujeto, yo sustancial). Estos conceptos, empero, no parecen ser los más adecuados para explicar la conciencia en toda su complejidad y con todas sus propiedades. A lo largo de este capítulo he tratado de

establecer las bases mínimas para un marco conceptual alternativo, desde el cual sea posible explicar en qué consiste la experiencia consciente en nuevos términos. La idea clave en ese marco es el proceso y el marco conceptual es la filosofía de procesos.

De acuerdo con la filosofía de procesos, los rasgos principales de un proceso son dos: (1) que está compuesto y (2) que sus partes se relacionan funcional o causalmente entre sí. Por ese motivo, para entender la idea de proceso es necesario disponer, ante todo, de una analítica de procesos y, más específicamente, de una mereología específica para procesos. Esta mereología no puede ser de corte clásico. Las mereologías clásicas o extensionales se basan, entre otras cosas, en la propiedad de transitividad y el principio de partes propias. Sin embargo, a menudo, la composición dinámica de los procesos viola la transitividad e incumple el principio de partes propias. En consecuencia, es preciso optar por una mereología no clásica.

En filosofía de procesos, la noción clave es el proceso transformativo. A partir de esa noción se puede definir otro tipo de proceso muy importante, a saber, el proceso autotransformativo. Los procesos autotransformativos están por todas partes. Están en la ingeniería: un arco tensionado o un puente de piedra son procesos autotransformativos. Están en la naturaleza: un ecosistema y los mecanismos de reentrada en la activación neuronal son, igualmente, procesos autotransformativos. La cuestión radica en establecer cómo es posible caracterizar estos procesos autotransformativos. Aquí he presentado dos ideas con respecto a esta caracterización.

La primera idea es que se pueden caracterizar a través de su composición, al menos, siempre y cuando se utilice una mereología no clásica. La composición dinámica de los procesos autotransformativos tiene tres rasgos: complejidad, interferencia y clausura. A esta composición la he denominado "composición CIC". Un proceso con composición CIC es un proceso que se compone de otros procesos, los cuales interfieren entre sí con cierta clausura. Es importante destacar que el término "clausura" no se aplica aquí al sistema, sino a las interferencias que se dan entre sus partes. Un sistema con composición CIC es, generalmente, un sistema abierto, que interfiere y es interferido por otros sistemas. Pero es un sistema cuyas partes tienen la peculiaridad de interferir entre ellas en una forma de composición clausurada, esto es, con cierto cierre.

Sin embargo, la composición CIC no es suficiente, por sí sola, para definir completamente los procesos autotransformativos, ya que es posible pensar mecanismos sencillos con composición CIC y sin autotransformatividad. Es necesario añadir una característica adicional. Para ello, he introducido una noción nueva: la modulación. Una variable A de un sistema modula a otra variable B si A tiene un papel en cómo cambia la variable B con respecto al tiempo. Por otro lado, una variable A de un sistema se automodula si su variación con respecto al tiempo depende de su propio valor. Por último, existe modulación con clausura en un sistema si existen variables A, B, C... cuya variación con respecto al tiempo depende del valor de las otras variables de una manera clausurada. Las variables de las que se habla en automodulación y modulación con clausura pueden estar definidas de dos maneras: cuantitativamente o cualitativamente. La teoría matemática de sistemas dinámicos nos ofrece la forma más determinada de modulación con clausura

para las variables cuantitativas, por lo que se tomó como referencia para describir esta noción.

La composición CIC acompaña a menudo a los procesos que presentan automodulación (o modulación con clausura). Esto se puede comprobar en casos particulares, pero también y principalmente en el hecho de que la teoría de control de sistemas dinámicos, que estudia cómo se pueden alterar las dinámicas de los sistemas, establece que el modo más eficiente es mediante mecanismos de retroalimentación, que, al fin y al cabo, no son sino una forma de composición CIC.

En definitiva, un proceso es autotransformativo sí y solo si cumple dos condiciones, (1) tener una composición CIC (complejidad, interferencia y clausura) y (2) que exista modulación con clausura de las variables que describen el proceso. La primera condición es composicional. La segunda es funcional.

La noción de proceso transformativo y proceso autotransformativo nos brindan las bases mínimas para un marco conceptual alternativo. Ahora trataré de aplicar este nuevo marco en la explicación del hecho de la experiencia consciente. Las preguntas a las que toca dar respuesta son ya de otro tipo: ¿Es razonable entender y analizar la conciencia como un proceso? ¿Es razonable analizarla, desde sus propiedades, como un proceso autotransformativo? ¿Ayuda la noción de autotransformación a la hora de comprender las teorías de la conciencia y a la hora de esclarecer algunos de sus problemas internos?

Capítulo 3

Conciencia y proceso

3.1. Introducción: por qué un enfoque de procesos

La tesis que se defiende aquí es que las nociones de conciencia y proceso están íntimamente ligadas. El capítulo 1 comenzó con una pregunta: ¿qué es la conciencia? Esta cuestión derivó en otras dos más: ¿cómo es la conciencia? y ¿cómo se puede explicar? Esto me llevó a describir la conciencia a través de sus propiedades fundamentales y a exponer cuatro de las teorías más relevantes que explican estas propiedades: la Teoría Autorrepresentacional, la Teoría del Espacio de Trabajo Global, la Teoría de la Información Integrada y la Teoría de la Doble Microgénesis.

Después me ocupé de la noción de proceso. El capítulo 2 se ha dedicado enteramente a establecer un marco conceptual analítico basado en la filosofía de procesos y a definir una mereología no clásica para analizar un tipo de procesos que parecen ser muy útiles para explicar la conciencia, a saber, los procesos autotransformativos. Hay dos rasgos que determinan si un proceso es autotransformativo. Primero, debe tener una composición dinámica de interferencia mutua. Esto quiere decir que debe estar formado por otros procesos que interfieren entre sí formando composiciones con cierre. Segundo, las variables cuantitativa o cualitativamente determinadas que describen el proceso como un sistema dinámico deben tener entre sí relaciones de modulación con clausura.

Ahora, en el último capítulo, relacionaré conciencia y proceso. Fundamentalmente trataré de responder a dos preguntas: ¿qué razones hay para optar por un análisis de la conciencia a través de procesos? y ¿qué aporta este tipo de análisis para la resolución de problemas concretos relacionados con la experiencia consciente?

Este capítulo se dedica enteramente a estas dos cuestiones. Con respecto a la primera, fundamentalmente hay dos razones para optar por un análisis procesual de la conciencia. La primera es que las teorías de la conciencia, cuando se les exige un nivel de determinación alto en la explicación de los mecanismos que dan cuenta de la experiencia consciente,

introducen aspectos autotransformativos de manera más o menos explícita. El apartado 3.2 se dedicará a demostrar que esto sucede en las cuatro teorías de la conciencia mencionadas arriba. La segunda razón para optar por un análisis procesual de la conciencia es que sus propiedades fenoménicas, funcionales, estructurales y temporales, tomadas en su conjunto, no son sino propiedades de procesos y, no de procesos cualesquiera, sino de procesos con cierto aspecto autotransformativo. El apartado 3.3 se dedicará a exponer el argumento de intrinsicidad, a través del cual es posible determinar en qué consiste el aspecto autotransformativo de las propiedades de la conciencia.

Una vez justificado el enfoque de procesos, trataré de emplearlo en algunos problemas concretos relacionados con la experiencia consciente. A esto dedicaré el apartado 3.4. He escogido dos problemas con suficiente evidencia empírica para poder ser interpretada de manera procesual: la sensación de certeza en la percepción visual y la sensación de tener algo *en-la-punta-de-la-lengua*. La conclusión es que en ambos problemas el enfoque de análisis de procesos permite revelar aspectos autotransformativos importantes que sirven para explicar con más detalle lo que sucede en estos dos tipos de experiencia consciente.

3.2. Teorías de la conciencia y proceso

A continuación voy a revisar desde el punto de vista procesual las teorías de la conciencia y trataré de destacar los aspectos procesuales que en esas aparecen y, en especial, sus aspectos autotransformativos. La tesis que defiendo es que, cuando estas teorías se someten a estrés, es decir, cuando se les pide un nivel de determinación alto acerca de los mecanismos que subyacen a la experiencia consciente, en algún punto de todas ellas es necesario introducir alguna forma de proceso autotransformativo, y que esa introducción no es casual, sino que constituye una parte fundamental dentro de la propia teoría. He seleccionado cuatro teorías, dispares entre sí pero con dos cosas en común, a saber, que tengan un nivel de determinación alto y que estén bien soportadas por evidencia empírica. Por estos motivos, he escogido la Teoría del Espacio de Trabajo Global, la Teoría de la Información Integrada, la Teoría de la Doble Microgénesis y la Teoría Autorrepresentacional. Ahora voy a identificar los aspectos procesuales y autotransformativos en todas ellas, y lo haré cumpliendo una condición: tratar de no introducir —en la medida de lo posible— elementos ajenos o críticos a la propia teoría. De este modo se quiere mostrar que esas teorías, de acuerdo a cómo fueron elaboradas, incluyen, de por sí, aspectos de autotransformatividad más o menos explícitos.

3.2.1. Autotransformatividad en la Teoría del Espacio de Trabajo Global

La Teoría del Espacio de Trabajo Global[184] y su versión neurológica, la Teoría del Espacio de Trabajo Global Neuronal,[185] asocian la conciencia a la conciencia de acceso y defienden

[184]Baars (1988)
[185]Dehaene y Naccache (2001); Changeux y Dehaene (2008); Dehaene et al. (2011)

que la conciencia de acceso no es sino el proceso de multidifusión de determinada información en lo que se denomina un espacio de trabajo global. Los procesos que dan cuenta del espacio de trabajo global tienen una composición en bucle, a saber, el procesamiento recurrente.

El procesamiento recurrente propio de la difusión en el espacio de trabajo global constituye el aspecto autotransformativo de esta teoría. Para demostrarlo, presentaré un argumento en dos pasos. Primero expondré la paradoja del umbral (Baars 1988), que es una parte fundamental de la Teoría del Espacio de Trabajo Global y cuya resolución obliga a estipular mecanismos de retroalimentación o bucle entre zonas emisoras y zonas receptoras. Segundo, sostendré que estos mecanismos de retroalimentación, que Dehaene et al. (2011) explican como procesamiento recurrente de larga distancia, tienen una composición idéntica a la interferencia mutua propia de la autotransformatividad.

Para finalizar, trataré de mostrar cómo incluso quienes han criticado la asociación entre conciencia y acceso que se asume en la Teoría del Espacio de Trabajo Global Neuronal, como Lamme (2010), también aceptan que el procesamiento recurrente está en la base de los procesos que dan lugar a la conciencia.

3.2.1.1. Paradoja del umbral y procesamiento recurrente

La Teoría del Espacio de Trabajo Global trata de explicar la conciencia a partir de premisas físicas, funcionales y temporales. La premisa física consiste en que la conciencia es propia de los animales con córtex cerebral, y el córtex es un entorno de procesamiento fuertemente jerarquizado, donde existen microprocesadores que, a su vez, se asocian entre sí formando procesadores específicos, los cuales también se asocian entre sí para formar entornos modulares de procesamiento. Las premisas funcionales y temporales son que la conciencia está asociada con la atención, tiene serialidad (al contrario que los procesos no conscientes, no trabaja en paralelo), es ineficiente (las tareas conscientes son más propensas al error) y es lenta (las tareas conscientes cuestan más tiempo) pero, a cambio, tiene gran capacidad de relación (la conciencia permite relacionar tareas antes aisladas, es sensible al contexto y tiene un amplio rango de contenidos). Con estas premisas, esta teoría asocia los estados conscientes con procesos de difusión amplia de determinados contenidos, en lo que se ha denominado el espacio de trabajo global.

El espacio de trabajo global es la clave de esta teoría y la mayoría de los esfuerzos se centran en explicar cómo funciona este espacio de trabajo. La difusión en el espacio de trabajo global tiene recursos limitados, es lenta y más propensa a error que las tareas realizadas por procesadores específicos. Por consiguiente, su uso debe estar regulado de acuerdo a ciertos mecanismos. Uno de los problemas básicos para explicar cómo se produce la difusión es la denominada paradoja del umbral (Baars 1988: 98). El término "paradoja del umbral" se refiere al problema que supone tener una gran cantidad de mensajes candidatos a difusión y un espacio global con recursos limitados, lo que lleva inevitablemente a una competición entre los diferentes mensajes por ese recurso. ¿Cómo se controla esa competición? Por un lado, para que un mensaje involucre a procesadores específicos a través del espacio de trabajo global es necesario que tenga cierto acceso a

ese espacio, es decir, cierto nivel de difusión. Por otro lado, para que exista cierto nivel de difusión, es necesario que el mensaje resulte atractivo (esto es, sea informativo) a los ojos de los diferentes procesadores específicos. ¿Cómo definir el umbral de acceso al espacio de trabajo global? La cuestión no está cerrada, pero Baars propone dos hipótesis que pueden ser complementarias. En la primera, el acceso al espacio de trabajo global se realiza a través de pequeños espacios parciales (no globales) de difusión, de tal modo que, con la reentrada adecuada de los procesadores específicos, estos espacios parciales se van haciendo más grandes hasta que, finalmente, el mensaje llega a ser difundido de manera global. En la segunda, cada mensaje disfruta de una pequeña ranura de tiempo para usar el espacio de trabajo global y probar de ese modo su efecto en los procesadores específicos, los cuales, a su vez, retroalimentan positiva o negativamente la difusión del mensaje. Baars las llama la "Hipótesis de la Sala de Espera" y la "Hipótesis del Acceso Momentáneo", respectivamente.

Estas dos hipótesis tienen en común que es necesario estipular mecanismos regulativos para estimular la difusión cuando el mensaje es informativamente relevante. Estos mecanismos se explican a través de bucles de procesamiento recurrente en los cuales intervienen los procesadores emisores y los procesadores receptores. Baars (1988) ha propuesto formas de procesamiento recurrente en las que intervienen partes del córtex cerebral, el tálamo y la formación reticular mesencefálica.[186] Sin embargo, quizás es Dehaene[187] quien ha detallado mejor en qué consiste ese procesamiento recurrente propio del espacio de trabajo global, como se mostrará a continuación.

3.2.1.2. Autotransformatividad y procesamiento recurrente

Dehaene et al. (2001) desarrollan una versión neuronal de la Teoría del Espacio de Trabajo Global que está especialmente centrada en la conciencia visual, en la cual proponen substratos neuronales concretos para el espacio de trabajo global. Estos substratos son procesos de activación basados en bucles autosostenidos de larga distancia, donde autosostenidos significa que pueden perdurar en el tiempo durante decenas de segundos (en la memoria de trabajo) y larga distancia significa que intervienen en ellos áreas del cerebro tan diversas y lejanas como el córtex prefrontal, el cingulado anterior y las regiones visuales. De acuerdo con esta teoría, la conciencia es un proceso compuesto al menos de tres etapas que están diferenciadas por dos umbrales. El primer umbral distingue la activación que no llega a ser procesada del procesamiento subliminal, esto es, el procesamiento que sucede sin conciencia y cuyo contenido todavía no es accesible por parte del resto de sistemas de procesamiento. El segundo umbral, llamado "umbral de conciencia", distingue el procesamiento subliminal del procesamiento consciente, que se caracteriza por la accesibilidad del contenido por parte del resto de sistemas de procesamiento y se basa en los bucles de procesamiento mencionados. Este procesamiento consciente es un "sistema neuronal distribuido o espacio de trabajo con conectividad de larga distancia que po-

[186] Es el sistema ERTAS (*Extended Reticular-Thalamic Activating System*) (Baars 1988: 122-123). La función del tálamo como puerta de control de las señales de entrada al córtex es bien conocida, así como la importancia de los mecanismos de reentrada que comparte con diferentes zonas del córtex (Baars y Gage 2007: 42).

[187] Dehaene y Naccache (2001); Changeux y Dehaene (2008); Dehaene et al. (2011)

tencialmente puede interconectar diversas áreas cerebrales especializadas de una manera coordinada, pero variable" (Dehaene et al. 2001: 13). La accesibilidad de la información en este procesamiento atiende principalmente a dos rasgos dinámico-estructurales: "Primero, la información debe presentarse de una manera activa, a través del disparo de una célula o un conjunto de células. Segundo, deben existir conexiones bidireccionales entre ese conjunto de células y el conjunto de neuronas del espacio de trabajo, de tal modo que se pueda establecer un bucle de amplificación sostenido en el tiempo". (Dehaene et al. 2001: 15)

La solución de la paradoja del umbral propuesta por Dehaene y Naccache tiene exactamente la estructura de un proceso autotransformativo que puede ser explicado mediante sistemas dinámicos. Las partes que forman parte del espacio de trabajo global, a saber, el conjunto de células emisoras y el conjunto de células del espacio de trabajo, interfieren entre sí. Además, esa interferencia tiene una composición con clausura, como se refleja en las conexiones bidireccionales entre las diversas áreas cerebrales que forman parte del espacio de trabajo global. El resultado es el establecimiento del bucle de amplificación, es decir, una modulación positiva entre variables del sistema. Este bucle de amplificación requiere de cierto tiempo de ignición. Los procesos conscientes, por tanto, suceden en el tiempo.

En consecuencia, la Teoría del Espacio de Trabajo Global requiere para su explicación de procesos autotransformativos.

3.2.1.3. El procesamiento recurrente es necesario

Los mecanismos neuronales de la Teoría del Espacio de Trabajo Global Neuronal no están exentos de crítica. Lamme (2010) rechaza la asociación entre conciencia y acceso y defiende que existe una etapa en la que se da cierta conciencia que es previa a los bucles de larga distancia propios de la memoria de trabajo. Esta etapa se denomina la "memoria visual frágil de corto plazo" (*fragile visual short-term memory*, VSTM), y se caracteriza por una forma de conciencia no accesible, no directamente reportable, de duración corta (pocos segundos) y frágil, en el sentido de ser fácilmente borrada por la entrada de nuevos estímulos visuales.

No obstante, incluso también desde estas posiciones críticas con la Teoría del Espacio de Trabajo Global, se ha aceptado que, para que exista conciencia, es necesario que haya procesos autotransformativos. Como en el caso de la memoria de trabajo, los substratos neuronales de esta memoria visual frágil de corto plazo son también procesos recurrentes. Pero, en esta ocasión, los bucles son de corta distancia, participando únicamente áreas visuales occipitales y temporales, esto es, sin intervención del córtex prefrontal o el cingulado anterior. De hecho, Lamme sostiene "que el procesamiento recurrente es necesario para la conciencia visual es algo que está ya razonablemente bien establecido y contrastado por diferentes experimentos" (Lamme 2010: 219).

En conclusión, los procesos recurrentes, que son, a su vez, procesos autotransformativos,

forman parte de la explicación de la conciencia en dos de las teorías más determinadas a nivel neurofisiológico. Queda por ver si hay rasgos de procesos autotransformativos en la otra teoría informativa que tiene gran determinación neurofisiológica, a saber, la Teoría de la Información Integrada.

3.2.2. Autotransformatividad en la Teoría de la Información Integrada

La Teoría de la Información Integrada (*Integrated Information Theory*, IIT)[188] sostiene que la conciencia es, fundamentalmente, una estructura conceptual (a saber, *in-formación* con cierto nivel de integración) dentro de un sistema complejo de elementos y que el carácter cualitativo de la conciencia viene precisamente dado por las relaciones estructurales (que, a su vez, son también *in-formacionales*) que existen entre los elementos que forman parte de ese sistema complejo. Por sí sola, la integración de la información no requiere que exista clausura y, por tanto, no requiere de aspectos autotransformativos en el sistema complejo. Algunos sistemas sencillos basados en unas pocas neuronas, por ejemplo, tienen cierto nivel de integración de la información sin presentar por ello ningún rasgo de clausura.[189] Sin embargo, otros sistemas como el cerebro humano no son simples, sino complejos y están formados por redes de procesamiento hacia delante (*feedforward*) para las cuales, la manera más efectiva de integrar la información consiste en crear, a su vez, relaciones de retroalimentación (*feedback*). Por consiguiente, los sistemas complejos propuestos por esta teoría para dar cuenta de la conciencia son, en última instancia, sistemas con cierto aspecto autotransformativo.

A continuación, voy a exponer las razones para sostener un argumento como el anterior. En primer lugar mostraré que los sistemas complejos de elementos propios de la Teoría de la Información Integrada son realmente mapas iterativos, es decir, sistemas dinámicos discretos, donde los elementos representan variables que pueden tomar diferentes valores en cada iteración y donde las relaciones entre esos elementos representan modulaciones entre esas variables. Para ello, irremediablemente tendré que entrar en los aspectos matemáticos de la teoría. Por ser más sencilla y accesible, me centraré sobre todo en la versión 2.0 de esta teoría.[190] En segundo lugar defenderé la idea de que, aunque la integración de la información y clausura en la modulación no son conceptos idénticos, los sistemas complejos que son necesarios para dar cuenta del procesamiento propio de los sistemas cognitivos requieren, en efecto, de clausura en la modulación. La consecuencia inevitable de todo esto es que, cuando se les exige un nivel de determinación alto, los sistemas complejos con información integrada que propone esta teoría como la base de la conciencia resultan ser sistemas dinámicos con cierto aspecto autotransformativo.

[188] Tononi (2004); Balduzzi y Tononi (2008); Oizumi et al. (2014)
[189] Un perceptrón, por ejemplo.
[190] Balduzzi y Tononi (2008).

3.2.2.1. Sistemas complejos como información

La Teoría de la Información Integrada trata de explicar la conciencia partiendo de varias premisas, denominadas axiomas. En este momento voy a destacar solo dos, una de tipo informacional y la otra de tipo dinámico-estructural. La primera es que la experiencia consciente surge dentro de un amplio repertorio de experiencias conscientes posibles. Este es el axioma de información o diferenciación. La segunda premisa es que, cuando la experiencia consciente sucede, lo hace con un aspecto de unidad o, lo que es lo mismo, lo hace de manera integrada. Este es el axioma de integración. ¿Cómo pueden compatibilizarse ambos axiomas? La respuesta está en la noción de información integrada, Φ, la cual mide al mismo tiempo dos aspectos: la cantidad de información de un sistema y el grado de integración de esa información dentro del sistema. La información integrada Φ es el reflejo de la divergencia entre la cantidad de información generada por un sistema complejo de elementos tomado en su conjunto y la cantidad de información generada por las particiones de sistema tomadas por separado. Para entender esta noción es necesario comprender antes en qué sentido se habla de sistema de información en esta teoría.

Para empezar, un sistema complejo de elementos se entiende como un sistema formado por varios elementos $1, 2, 3 \cdots n$, los cuales se describen mediante estados $\langle 1 \rangle, \langle 2 \rangle, \langle 3 \rangle \cdots \langle n \rangle$ y se encuentran interrelacionados entre sí mediante un mecanismo *mech* bien definido. Por ejemplo, la figura 3.1 presenta un sistema sencillo formado solo por dos elementos, cada uno de los cuales solo puede estar en dos estados, activo "1" o no activo "0". Supongamos un estado final del sistema particular x_1=(1,1). El repertorio *a priori* de estados posibles del sistema es (0,0), (0,1), (1,0) y (1,1), todos ellos con idéntica probabilidad (es decir, 1/4). El repertorio de estados del sistema *a posteriori*, es decir, conocido cuál es el estado final x_1 y dada la mecánica del sistema es, por el contrario, menor. En este caso, los estados previos que pueden conducir el sistema al estado final x_1 son solo dos, a saber, (1,0) y (1,1). Ambos con idéntica probabilidad (1/2).

La información eficiente (*ei*) del sistema con respecto a un estado final x_1 es la entropía relativa, es decir, la divergencia entre dos probabilidades de distribución, a saber, la distribución de probabilidad del sistema para llegar a ese estado x_1, $p(X_0(mech, x_1))$, y la distribución de probabilidad en el caso de entropía máxima del sistema, $p(X_0(maxH))$. En otras palabras:

$$ei(X(mech, x_1)) = H\left[p(X_0(mech, x_1))|p(X_0(maxH))\right] \tag{ei}$$

donde $H[p|q]$ es la entropía relativa (o divergencia de Kullback-Leibler) de una distribución de probabilidad p con respecto a otra distribución de probabilidad q, y se calcula de la siguiente manera:

$$H[p|q] \equiv \sum_{i=1}^{n} p_i \log_2(\frac{p_i}{q_i}) \tag{D_{KL}}$$

Para un sistema definido por un mecanismo y conformado por un número conocido de elementos, la información efectiva indica el grado de información que el sistema, tal y como es, aporta para llegar a un estado final concreto. Cuando el sistema no aporta un mecanismo que reduzca la incertidumbre del sistema (si, por ejemplo, el sistema resulta

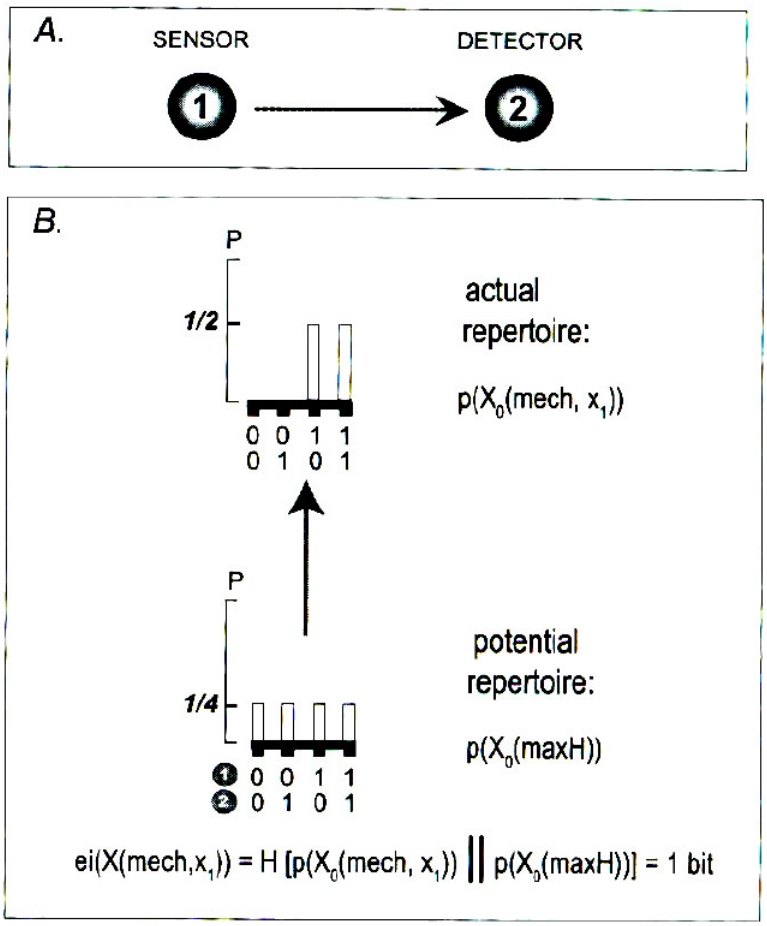


Figura 3.1: Sistema de fotodiodo con un mecanismo sencillo de copia de valor entre los elementos 1 y 2 y con estado final de activación de ambos elementos $x_1 = (1, 1)$. La información eficiente (ei) se calcula como la entropía del sistema para llegar al estado x_1 en relación con la entropía máxima del sistema, esto es, sin mecanismo. El sistema puede llegar al estado x_1 a través de dos estados iniciales $x_0 = (1, 0)$ o $x_0 = (1, 1)$, ambos con idéntica probabilidad de 0'5. Por tanto, la información eficiente para ese sistema y ese estado final es 1 bit. Extraído de Tononi (2008: 219).

ser aleatorio), entonces la información eficiente del sistema será nula. Si, por el contrario, el mecanismo reduce el repertorio y, por tanto, la incertidumbre (como es el caso de la figura 3.1), entonces la información eficiente será distinta de cero.

Sin embargo, la información efectiva no dice nada acerca del grado de integración de la información que tiene el sistema. Para ello es necesario introducir una nueva noción, la información integrada Φ, que se define así:

$$\Phi(X(mech, x_1)) = H\left[p(X_0(mech, x_1))|\prod p({}^kM_0(mech, \mu_1))\right] \qquad (\Phi)$$

donde: ${}^kM_0 \in MIP$; MIP es la partición de información mínima, es decir, la partición para la cual las partes son más independientes (o, en otras palabras, la partición para la cual las partes están menos integradas entre sí)[191] y $\prod p({}^kM_0(mech, \mu_1))$ es el producto del repertorio actual de cada una de las partes ${}^kM_0$ tomadas por separado.[192]

La información integrada permite medir de manera cuantitativa el grado de integración de los elementos, como se puede ver en la figura 3.2.

3.2.2.2. Los sistemas complejos como mapas iterativos

¿Es posible aplicar el análisis de procesos y la noción de autotransformatividad a la Teoría de la Información Integrada? A simple vista se puede intuir que, puesto que el sistema A' de la figura 3.2 tiene un mayor número de interconexiones mutuas que el sistema (A), deberá tener, también, una Φ más alta. Sin embargo, no se puede concluir por ello que Φ sea una medida del grado de autotransformatividad de un sistema. Recordemos, Φ mide únicamente el grado de información integrada del sistema, no su grado de autotransformatividad. Para mostrar que los procesos autotransformativos y, más concretamente, la automodulación y la modulación cruzada, tienen un papel importante en esta teoría, se presentará a continuación un argumento basado en dos pasos. En primer lugar, se razonará que los sistemas complejos de elementos se pueden entender como una subespecie de los sistemas dinámicos, a saber, sistemas dinámicos discretos o mapas iterativos y que, por consiguiente, están sujetos a las mismas características de modulación que aplicaban a los sistemas dinámicos de ecuaciones diferenciales. En segundo lugar, se razonará que, desde la propia teoría, se llega a la conclusión de que existe una relación íntima entre la información integrada Φ y la modulación con clausura para determinados tipos de sistemas, que son precisamente los sistemas típicos del procesamiento cognitivo.

[191] Intuitivamente, si debo realizar la partición del sistema formado por la rama de un naranjo y su naranja en dos partes, lo más sencillo es partirlo por el punto más débil (y de menor integración), esto es, por el tallo que las une, en lugar de hacerlo partiendo la naranja en dos o partiendo la rama en dos. Esta idea intuitiva aplica para la definición de la partición de información mínima. El cálculo exacto de la MIP o partición de información mínima está matemáticamente establecido en la versión 2.0 de esta teoría y requiere de una normalización. Véase Balduzzi y Tononi (2008).

[192] La manera de solapar dos o más distribuciones de probabilidad independientes es el producto $\prod$, igual que para dos sucesos independientes A y B, la probabilidad de $P(A \cap B) = P(A)P(B)$. Este es el caso de las distribuciones de probabilidad resultantes de tomar cada una de las partes de la partición por separado.

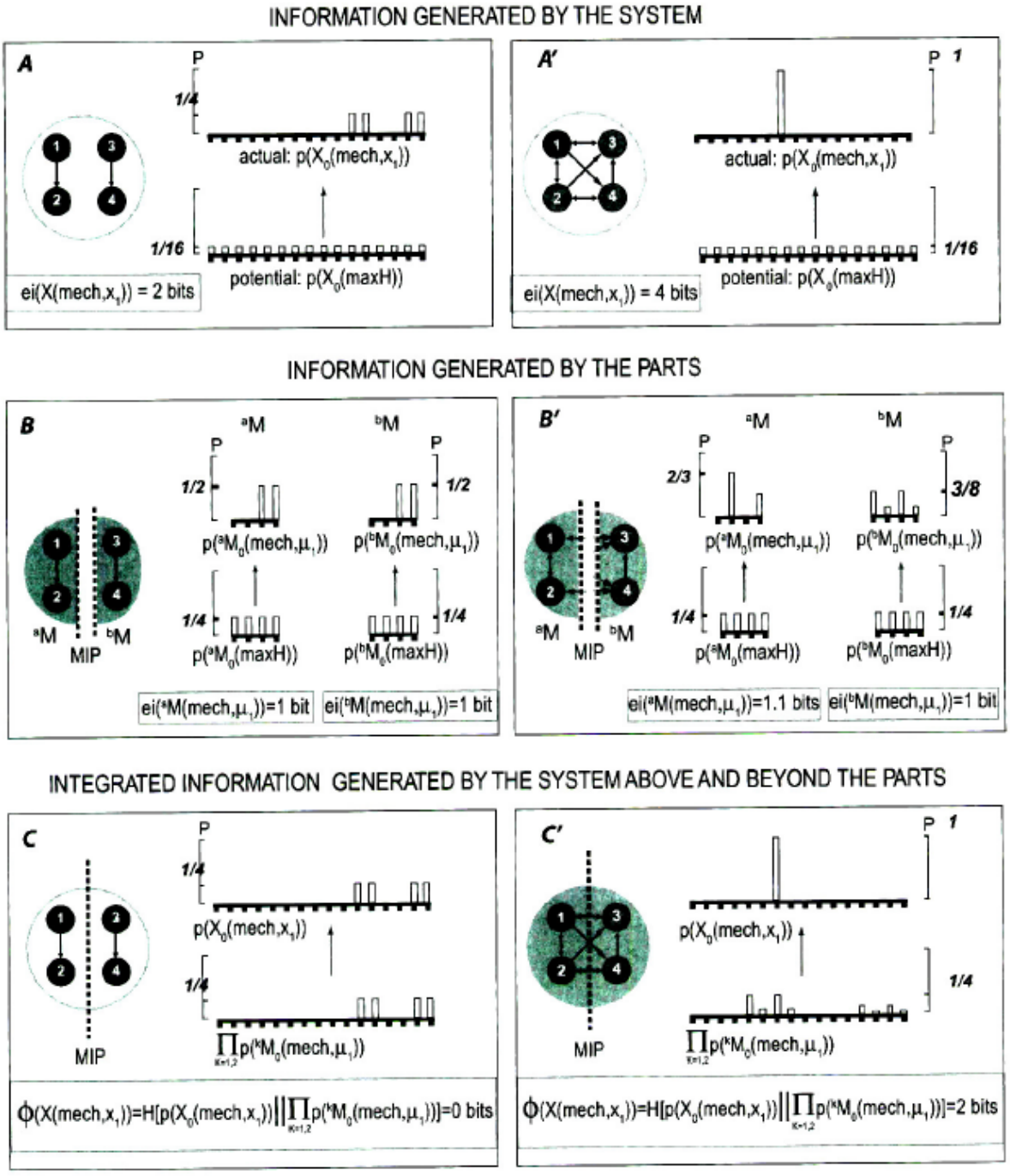


Figura 3.2: Cálculo de información eficiente *ei* e información integrada Φ para dos sistemas, ambos de cuatro elementos. El primer sistema (A) tiene un mecanismo sencillo de copia de estado entre los elementos, y su estado final es $x_1 = (1, 1, 1, 1)$, esto es, todos los elementos activos. El segundo sistema (A') tiene un mecanismo de activación si se reciben dos más inputs positivos (*two or more spikes*) y su estado final es $x_1 = (1, 0, 0, 0)$, esto es, únicamente queda activo el elemento "1". El sistema (A') tiene una Φ más alta que el sistema (A), puesto que tiene un grado de integración mayor, como se puede deducir visualmente del mecanismo de conexión entre sus elementos. Extraído de Tononi (2008: 222).

De acuerdo con Strogatz (1994), hay dos tipos de sistemas dinámicos: los continuos, descritos por las ecuaciones diferenciales (y expresados en la forma canónica (fc1)) y los discretos, descritos mediante mapas iterativos y que, de una forma canónica, se podrían representar así:

$$\begin{aligned} x_{n+1} &= f_x(x_n, y_n, z_n \dots) \\ y_{n+1} &= f_y(x_n, y_n, z_n \dots) \\ z_{n+1} &= f_z(x_n, y_n, z_n \dots) \\ &\vdots \end{aligned} \qquad \text{(mapa it.)}$$

donde $x, y, z \dots$ son los elementos cuyos estados describen el sistema dinámico discreto y n indica el número de iteración. Nótese que el sistema se denomina mapa iterativo porque, dado un estado inicial $(x_0, y_0, z_0 \dots)$, es posible calcular el valor de $x_i, y_i, z_i \dots$ para todo $i > 0$ simplemente mediante iteración.

En sistemas dinámicos discretos es posible definir la clausura de modulación de la misma manera que se hizo en el sistema dinámico continuo.

Automodulación pura:

$$x_{n+1} = f(x_n) \qquad \text{(automod. pura)}$$

Modulación cruzada pura:

$$\begin{aligned} x_{n+1} &= f_x(y_n) \\ y_{n+1} &= f_y(x_n) \end{aligned} \qquad \text{(mod. cruz. pura)}$$

Modulación con clausura para tres elementos:

$$\begin{aligned} x_{n+1} &= f_x(y_n) \\ y_{n+1} &= f_y(z_n) \\ z_{n+1} &= f_z(x_n) \end{aligned} \qquad \text{(mod. claus.)}$$

Y así sucesivamente.

Si volvemos un momento a la definición de sistemas complejos de elementos de la Teoría de la Información integrada, se puede comprobar que es posible entender estos sistemas de manera funcional como si estuvieran descritos por mapas iterativos (o sistemas dinámicos discretos) cuyos elementos son los elementos del sistema complejo y cuya mecánica *mech* queda completamente descrita por el conjunto de funciones que describen el sistema complejo f_x, f_y ,f_z, etc.[193]

[193]De hecho, en la versión 2.0 de la Teoría de Información Integrada, los sistemas complejos de elementos se describen como procesos de Markov de primer orden, donde "la salida de un elemento en el tiempo t depende únicamente de la entrada en el tiempo $t-1$" (Balduzzi y Tononi 2008: 3).

Por ejemplo, el sistema simple de fotodiodo de la figura 3.1 es un sistema dinámico discreto formado por dos variables 1 y 2 sin clausura, que se describe del siguiente modo:

$$\langle 2\rangle_{n+1} = COPY(\langle 1\rangle_n) \qquad \text{(fotodiodo)}$$

Igualmente, el sistema (A) de la figura 3.2 es un sistema dinámico discreto formado por cuatro variables sin clausura que se describe de esta forma:

$$\begin{aligned}\langle 2\rangle_{n+1} &= COPY(\langle 1\rangle_n)\\ \langle 4\rangle_{n+1} &= COPY(\langle 3\rangle_n)\end{aligned} \qquad \text{(Sist. A)}$$

Mientras que el sistema (A') de la figura 3.2 es un sistema dinámico discreto formado por cuatro variables que tiene la siguiente clausura:

$$\begin{aligned}\langle 1\rangle_{n+1} &= TWOORMORESPIKES(\langle 2\rangle_n, \langle 3\rangle_n)\\ \langle 2\rangle_{n+1} &= TWOORMORESPIKES(\langle 1\rangle_n, \langle 4\rangle_n)\\ \langle 3\rangle_{n+1} &= TWOORMORESPIKES(\langle 1\rangle_n, \langle 2\rangle_n, \langle 4\rangle_n)\\ \langle 4\rangle_{n+1} &= TWOORMORESPIKES(\langle 1\rangle_n, \langle 2\rangle_n)\end{aligned} \qquad \text{(Sist. A')}$$

Las analogías entre los sistemas complejos de elementos de esta teoría y los sistemas dinámicos discretos son múltiples. Ambos describen un sistema real de manera funcional, esto es, al margen de la especificación física de la interferencia (al margen de si, por ejemplo, las relaciones entre elementos están realizadas mediante puertas AND o mediante sinapsis neuronales). Y en ambos es posible visualizar de manera sencilla aspectos de la relación tales como la clausura de la modulación o la integración. La modulación con clausura de los sistemas dinámicos discretos es un reflejo de los mecanismos de retroalimentación que ocurren en los sistemas complejos de elementos en IIT.

3.2.2.3. Autotransformatividad en los sistemas complejos

La cuestión ahora es si la información integrada y la clausura tienen alguna relación. Una vez determinada la relación entre sistemas dinámicos discretos y sistemas complejos de elementos propios de la Teoría de la Información Integrada, es posible responder a la siguiente cuestión: ¿qué es lo que mide la información integrada Φ? La información integrada no mide el nivel de autotransformatividad de un sistema dinámico, sino el grado de información y de integración de esa información. La pregunta es, entonces, la siguiente. Supongamos un sistema de 9 elementos (ver figura 3.3), donde el noveno elemento tiene una integración como la descrita a continuación:

$$\langle 9\rangle_{n+1} = f(\langle 1\rangle_n, \langle 2\rangle_n, \ldots \langle 8\rangle_n) \qquad \text{(Sist. estrella)}$$

¿Podría un sistema simple con 9 elementos como el siguiente tener un Φ no nulo a pesar de carecer de clausura de modulación?

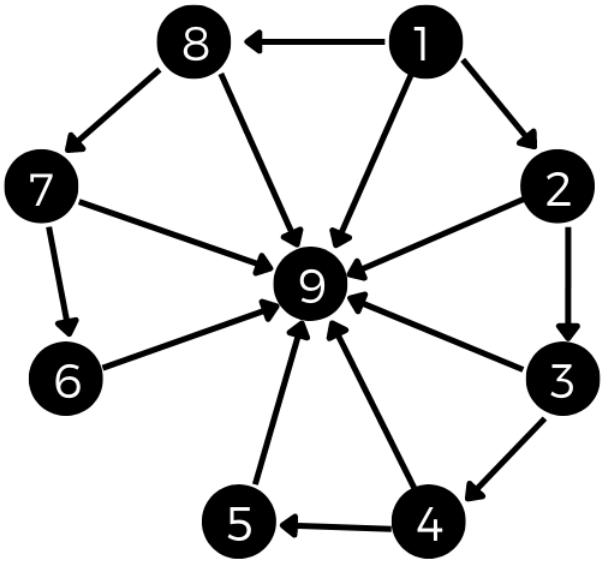


Figura 3.3: Sistema de nueve elementos en estrella. Intuitivamente tiene cierta integración, pero carece de clausura de modulación entre sus componentes.

Este tipo de sistemas simples, si bien presentan cierta integración de la información, no son aceptables desde el punto de vista de la IIT, y mucho menos si lo que se quiere es formalizar cómo funcionan los sistemas cognitivos propios del cerebro. Los sistemas cognitivos deben ser complejos en el procesamiento y deben ser robustos frente a posibles problemas. Sin embargo, los sistemas simples como el anterior no son los más adecuados. En primer lugar, presentan una plasticidad escasa. En este caso, por ejemplo, si fallara el elemento 9, la dinámica del sistema colapsaría. En segundo lugar, tienen una capacidad de procesamiento limitada porque carecen de complejidad.[194] Por otro lado, si se intenta hacer más complejos estos sistemas mediante nuevas capas de procesamiento puramente *feedforward* se pierde la integración de la información. ¿Cómo es posible conseguir compatibilizar valores de información integrada altos con la plasticidad, resiliencia y eficiencia de procesamiento requerida para los sistemas cognitivos?

Balduzzi y Tononi nos dan la solución. En estos casos, "añadir conexiones de retroalimentación (*feedback*) y conexiones laterales a una red de alimentación directa (*feedforward*) puede incrementar potencialmente la información integrada generada por el sistema, al aumentar el número de interacciones causales no redundantes" (Balduzzi y Tononi 2008: 14). Las conexiones de retroalimentación son precisamente las propias de la clausura de modulación de los sistemas dinámicos discretos. Por ejemplo, las conexiones de retroalimentación entre los elementos 1, 2 y 3 del sistema (A') de la figura 3.2 se traducen en relaciones de clausura en la modulación de las variables que representan los valores de esos elementos $\langle 1 \rangle$, $\langle 2 \rangle$ y $\langle 3 \rangle$, tal y como se muestra en la forma canónica presentada. Esta dependencia entre clausura de la modulación e información integrada es, si cabe, más acusada en la versión 3.0 de la teoría, donde no es posible tener un sistema complejo

[194]Minsky y Papert (1969) ya mostraron las ventajas de esa arquitectura de red con una sola neurona, y también sus limitaciones en cuanto a lo que se puede procesar con ella.

con información integrada en ausencia de clausura de la modulación.[195]

En suma, ¿cuál es la relación entre la clausura de modulación y Φ y, por consiguiente, entre los procesos autotransformativos e IIT? La relación consiste en que la modulación con clausura aumenta Φ en el caso de redes con procesamiento cognitivo (es decir, de tipo *feedforward*) tales como el cerebro.

3.2.3. Autotransformatividad en la Teoría de la Doble Microgénesis

La Teoría de la Doble Microgénesis[196] establece que las experiencias conscientes son el producto de un proceso compuesto de dos etapas: la microgénesis representacional y la microgénesis fenoménica, y que solo en la segunda etapa se produce propiamente aquello que llamamos experiencia consciente fenoménica. Esta segunda etapa está relacionada con procesos en los que interactúan sistemas específicos (en el caso de la conciencia visual, por ejemplo, aquellos que reciben y procesan los estímulos visuales) y sistemas no específicos (independientes de la modalidad del estímulo, esto es, independientes de si el estímulo es visual auditivo o táctil). La forma concreta en que se producen estos procesos de interacción es mediante una modulación que se lleva a cabo a través de bucles de procesamiento con reentrada. Esto es, tiene aspecto autotransformativo. Sin embargo, la fórmula que propone para describir matemáticamente esos bucles,[197] $C = L \times D$, que denota modulación, no parece denotar que exista una clausura en la modulación. Para probar si hay aspectos autotransformativos en la teoría de Bachmann es necesario reparar en la composición y presumir, al mismo tiempo, que la fórmula matemática propuesta no refleja la complejidad del sistema dinámico que conforma esa composición.

Para argumentar que existen aspectos autotransformativos en esta teoría, procederé en dos pasos. Expondré, primero, en qué consiste la modulación propia de los sistemas no específicos en los sistemas específicos, base de la microgénesis fenoménica de la conciencia. Después, definiré la forma de clausura que Bachmann establece como la composición dinámica de esos sistemas. Por último, introduciré el problema que se deriva de no incluir en esa composición cierta clausura de la modulación.

[195] Tononi y Koch (2015: fig. 5) presentan un ejemplo de complejo solo con alimentación directa (sin retroalimentación) que carece completamente de integración de la información. Cfr. también Mayner et al. (2018), donde se pueden comprobar diferentes combinaciones de sistemas complejos.

[196] Bachmann (2000)

[197] Bachmann y Hudetz (2014). Por recordar, C significa la conciencia, L es el sistema de nivel, esto es, el nivel crítico de actividad neuromodulatoria, de carácter no específico, y D es el sistema de datos, que refleja los contenidos de la conciencia provenientes de los sistemas de representación, de carácter específico.

3.2.3.1. Modulación entre sistemas no específicos y sistemas específicos

La Teoría de la Doble Microgénesis —así como sus versiones, Teoría de Recontacto Perceptual, teoría del doble proceso, teoría de la integración dendrítica, etc.— surge de la necesidad de integrar en la explicación de la conciencia las propiedades temporales y, en especial, todas aquellas evidencias puestas de manifiesto en la abundante literatura basada en experimentos de enmascaramiento de la percepción visual mediante metacontraste. Por consiguiente, es una teoría especialmente enfocada en la conciencia visual. Trata de dar respuesta a cuestiones tales como ¿cómo es posible que la génesis de la conciencia de la representación cueste más tiempo que la génesis de la representación (y no sea el mismo tiempo, por ejemplo, como dirían las teorías representacionales de primer orden)?, ¿por qué existe facilitación (o aceleración) del proceso de la conciencia en casos de estímulos visuales distintos e incluso divergentes? y, en general, ¿por qué la ignición de la conciencia tiene unas características dinámicas específicas?.[198]

Para dar cumplida respuesta a estas cuestiones Bachmann propone diferentes mecanismos fisiológicos que siempre tienen una característica en común, a saber, que la percepción visual consciente (y, en general, cualquier experiencia consciente con contenido) está formada por dos procesos bien diferenciados. Por un lado, la microgénesis representacional, de la cual forman parte los sistemas específicos de activación talámico-cortical. Por otro lado, la microgénesis fenoménica, en la cual participan sistemas no específicos de activación talámico-cortical. Estos dos procesos tienen propiedades diferentes porque su base neurofisiológica es también diferente.

¿Existe alguna relación entre esta teoría y los procesos transformativos? De entre todas las teorías vistas hasta ahora, la Teoría de la Doble Microgénesis es quizás aquella en la que más explícitamente se puede contemplar esa relación. En primer lugar, es una teoría puramente procesual, basada en las propiedades temporales y los mecanismos fisiológicos particulares de la conciencia. En segundo lugar, es la única teoría que usa explícita y continuamente el término "modulación". Cuando formula su principal postulado acerca de los mecanismos de interferencia que suceden entre los sistemas específicos de activación y los sistemas no específicos de activación, Bachmann afirma literalmente que en la Teoría de Recontacto Perceptual lo que sostiene la sensación fenoménica de percibir es "la modulación que recibe la actividad neuronal específica por parte de las entradas talámicas no específicas" (Bachmann 2000: 195). Esta importancia que se concede a los procesos de modulación no es circunstancial, sino que acompaña a todas las versiones de la teoría presentadas a lo largo del tiempo. Bachmann y Hudetz sostienen que, para que haya conciencia, es necesaria la activación orquestada de dos sistemas, "(1) el sistema de datos D, de tipo abajo-arriba" en el cual se recibe la aferencia de datos en las neuronas piramidales de la capa 5 del córtex y "(2) el sistema modulación, que envía señales asociativas intracorticales de tipo arriba-abajo así como señales talámico-corticales no específicas, y conduce a un nivel de activación necesario para hacer posible el nivel de conciencia" (Bachmann y Hudetz 2014: 3). Igualmente, Bachmann, Suzuki y Aru afirman que "el tálamo está en disposición de llevar a cabo la función principal en la modulación del estado y el nivel de conciencia" (Bachmann et al. 2020: 8).

[198] Bachmann (2000: 201)

Está claro que la modulación juega un papel trascendental para explicar la interacción entre los dos procesos que dan lugar a la percepción consciente. Sin embargo, sigue siendo necesario aclarar dos aspectos. Primero, si esa noción de modulación tiene que ver con la noción de modulación que se estipuló en sistemas dinámicos. Segundo, si la noción de modulación de los sistemas no específicos en los sistemas específicos incluye algún tipo de clausura como la propia de los procesos autotransformativos.

Con respecto a la primera cuestión, la respuesta no es sencilla, puesto que la naturaleza exacta de esa modulación es el aspecto más complejo y menos definido de toda la teoría de Bachmann. No obstante, en Bachmann (2000: 158) se hace una descripción bastante aproximada del efecto a nivel de célula individual de la modulación que se produce desde sistemas no específicos a sistemas específicos (ver figura 3.4). La modulación de los sistemas no específicos modifica los niveles de depolarización de las neuronas correspondientes a los sistemas específicos, y su efecto en estas neuronas se puede resumir en tres aspectos:

(a) Puede provocar que la neurona se dispare.

(b) Puede controlar o incluso modular la frecuencia de disparo de esa neurona.

(c) Puede reducir la latencia con la que esa neurona comienza a dispararse.

Esto es debido a que la activación proveniente de los sistemas no específicos tiene mucha menor resolución temporal y espacial que la activación proveniente de los sistemas específicos.

Los tres efectos implican modificaciones en la forma en que los sistemas específicos varían. Por tanto, el efecto de esa modulación se puede expresar así en forma canónica:

$$\dot{\mathbf{x}} = \mathbf{f_m}(\dots \mathbf{y} \dots) \qquad \text{(mod. sist. esp.)}$$

donde $\mathbf{x}$ describe el vector los estados de activación de las neuronas pertenecientes al sistema específico, $\mathbf{y}$ describe el vector los estados de activación de las neuronas pertenecientes al sistema no específico y f_m la función que describe los tres efectos modulatorios listados arriba.

3.2.3.2. Autotransformatividad entre sistemas no específicos y sistemas específicos

Esta forma canónica que se acaba de presentar manifiesta una modulación sin clausura, por lo que cabe preguntarse si la Teoría de la Doble Microgénesis requiere o no de formas de modulación que presenten clausura, tales como las definidas para los sistemas dinámicos. De nuevo, la respuesta parece ser afirmativa, a tenor de la importancia que se da en la teoría a los mecanismos de reentrada desde los sistemas específicos hacia los sistemas no específicos. Bachmann describe la arquitectura completa de interferencia entre los sistemas específicos y los sistemas no específicos, como una "reentrada" basada en tres

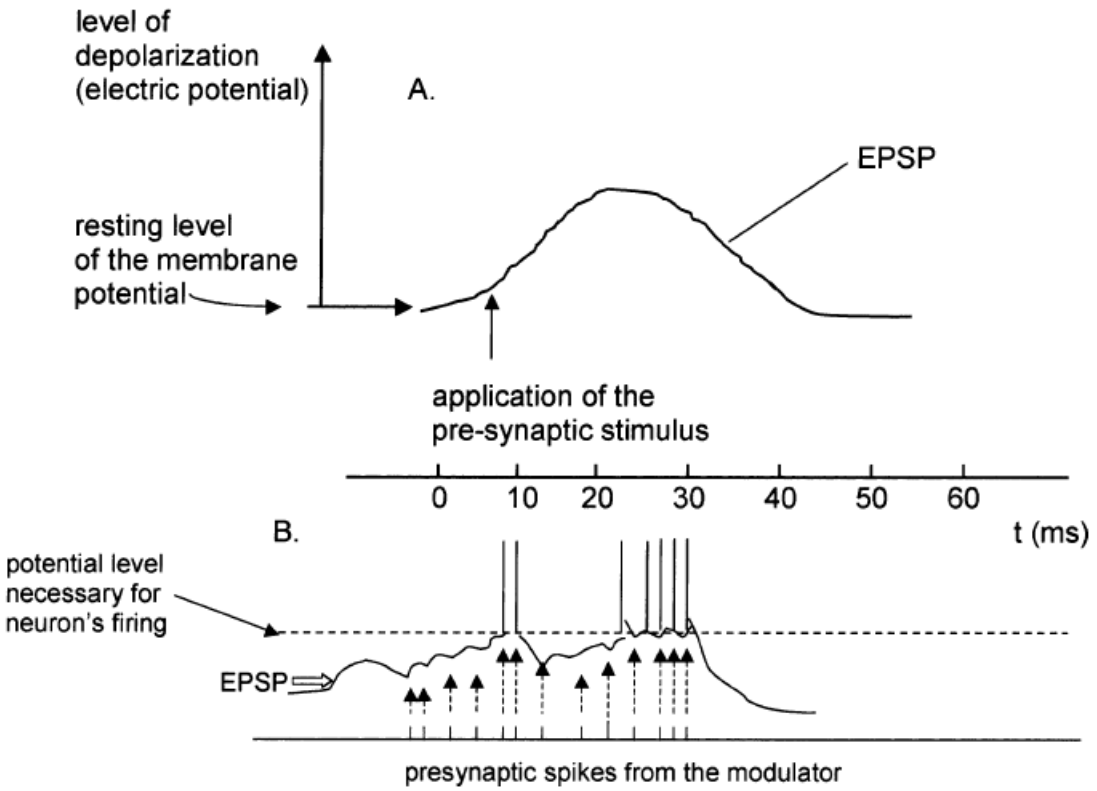


Figura 3.4: Ejemplo de potencial excitatorio postsináptico (EPSP) de una neurona. En A se ilustra el aumento de su depolarización ante un estímulo presináptico para una neurona cualquiera. En B se ilustra el efecto modulatorio proveniente de los sistemas talámicos no específicos en una neurona cortical perteneciente al sistema representacional. Esta modulación puede provocar que la neurona se dispare, pero también puede cambiar su frecuencia de disparo y su latencia de disparo. Extraído de Bachmann (2000: 157).

acciones. La primera es "la acción dirigida desde el núcleo talámico hasta los nodos específicos corticales". La segunda es "la acción proveniente de la entrada específica y dirigida hacia neuronas talámicas no específicas". La tercera es la "actividad neuronal reentrante y cíclica entre las partes no específicas del tálamo y las partes específicas del córtex" (Bachmann 2000: 195). Como se ha dicho, la naturaleza exacta de esa arquitectura es algo por determinar, aunque queda acreditada la existencia fundamental de mecanismos de reentrada y retroalimentación sin los cuales no es posible explicar la modulación de los sistemas no específicos en los sistemas específicos.

3.2.3.3. ¿Existe clausura en la modulación?

Bachmann y Hudetz (2014) proponen la siguiente formulación simbólica para explicar la conciencia:

$$C = L \times D$$

Aquí, C simboliza la conciencia, la cual es cualitativamente determinable (a través de su contenido, en la experiencia consciente) y cuantitativamente determinable (como nivel de conciencia, que va desde "lo inconsciente hasta la representación totalmente consciente y fenoménicamente estable").[199] L es el sistema de nivel, esto es, el nivel crítico de actividad neuromodulatoria que es de carácter no específico, de tipo arriba-abajo (top-down) y en el que toman parte "señales de activación intracorticales y/o señales no específicas talámico-corticales".[200] Por último, D es el sistema de datos, esto es, los contenidos de la conciencia provenientes de los sistemas de representación, de carácter específico y tipo abajo-arriba (bottom-up), en el que toman parte las señales aferentes recibidas en el córtex.[201] El signo de producto ($\times$) expresa el hecho de que sin el nivel de actividad neuromodulatoria (sin L) no habría conciencia, igual que sin contenido (sin D) tampoco la habría.

La formulación anterior admite una doble interpretación, cuantitativa o matemática y cualitativa o simbólica. La interpretación matemática de la formulación propuesta resulta insuficiente para expresar la naturaleza autotransformativa de los sistemas dinámicos que forman parte de la experiencia consciente, puesto que no garantiza, por sí sola, la existencia de una clausura en la modulación. Esto se tratará de argumentar a continuación.

Tomada en sentido matemático, esta formulación permite identificar cierto componente de modulación de las variables L y D, que miden niveles de activación en la variable C, que mide el nivel de conciencia:
Por la regla de producto:

$$\dot{C} = L \cdot \dot{D} + \dot{L} \cdot D$$

[199] Bachmann y Hudetz (2014: 10).
[200] Bachmann y Hudetz (2014: 3).
[201] Especialmente en la "capa 5 de las neuronas piramidales" (Bachmann y Hudetz 2014: 3).

luego:

$$\dot{C} = f(L, D, ...)$$

donde $\dot{C}$, $\dot{L}$ y $\dot{D}$, simbolizan la variación en el tiempo de C, L y D respectivamente. En consecuencia, el modo en que varía el nivel de conciencia depende del nivel de activación de los dos sistemas de nivel y de datos.

Esta formulación no garantiza, por sí misma, que exista una clausura en la modulación, lo que contrasta con el aspecto autotransformativo de la composición dinámica de los procesos que forman parte de este sistema, que está compuesto por interferencias *top-down* combinadas con interferencias *bottom-up*. La explicación puede ser que la formulación propuesta por Bachmann y Hudetz no recoge toda la complejidad de las interacciones entre los sistemas propuestos y que sea necesario incluir otras formulaciones que podrían incluir esa clausura tales como: $\dot{L} = f(L, \dots)$, $\dot{L} = f(C, D, \dots)$, etc. Este tipo de formulación permitiría explicar mejor los tiempos de ignición propios de la experiencia consciente, así como sus tiempos de estabilidad, que no dependen de los valores específicos de los sistemas de nivel y de datos.

En conclusión, los sistemas de activación formados por los sistemas específicos y los sistemas no específicos, que son la base de la microgénesis fenoménica (y, por extensión, de la experiencia consciente) tiene una composición de interferencia de tipo clausurado y, en consecuencia, tienen aspectos autotransformativos. Estos aspectos parecen sembrar cierta sombra de dudas sobre la formulación simbólica de Bachmann y Hudetz. Aquí se sugiere que esa formulación no refleja la complejidad de relaciones de modulación que existen entre las variables que describen a este sistema dinámico.

3.2.4. Autotransformatividad en la Teoría Autorrepresentacional

La Teoría Autorrepresentacional[202] explica el carácter subjetivo de la conciencia como un estado mental representacional apropiado, en el sentido de que es capaz de representar un contenido al mismo tiempo que se representa a sí mismo. Para entender esta teoría es necesario encuadrarla en el contexto del resto de teorías representacionales. Las teorías representacionales de primer orden asocian los estados mentales a estados representacionales y se centran en definir correctamente la noción de representación, si bien no acaban de introducir criterios concluyentes para diferenciar los estados mentales no conscientes de los estados mentales conscientes[203] ni tampoco parecen explicar de manera convincente cómo sucede el carácter subjetivo propio de la experiencia consciente. También tienen dificultades para explicar fenómenos como la visión ciega,[204] donde las tareas de discri-

[202]Kriegel (2009)

[203]Cfr. Tye (1995)

[204]La visión ciega (*blindsight*) es un síndrome poco frecuente que sucede en pacientes con lesión en las áreas visuales primarias (V1 y V2). Estos pacientes no son capaces de ver en determinadas zonas del campo visual. Sin embargo, son capaces de hacer discriminaciones acertadas basadas en determinados estímulos básicos en esas zonas, como por ejemplo intuir la dirección de un movimiento. A pesar de

minación o detección de estímulos se disocian de la propia conciencia del estímulo. Las teorías representacionales de orden superior, por su parte, introducen un criterio para diferenciar los estados conscientes, a saber, la representación de orden superior y, de este modo, ofrecen una explicación del carácter subjetivo de la experiencia consciente. Sin embargo, no son del todo satisfactorias. Por un lado, no acaban de presentar ningún rasgo adicional que diferencie el representar correspondiente al primer orden del representar correspondiente al orden superior y explique por qué uno es consciente y el otro no. Esto puede constituir un problema a la hora de identificar cuáles son los mecanismos neurofisiológicos específicos para la conciencia. Por otro lado, pueden conducir a regresiones infinitas, derivadas del hecho de que los estados de orden superior son, a su vez, tomados como estados conscientes.[205] En este panorama, las teorías representacionales del mismo orden tratan de retener los aspectos más positivos de las teorías de primer orden y de orden superior, al tiempo que intentan resolver sus problemas. Este es el caso de la Teoría Autorrepresentacional.

La Teoría Autorrepresentacional parte de una distinción fenoménica relativamente nueva (Nagel 1974; Crane 2000), a saber, que la fenomenología de una experiencia consciente se compone de dos caracteres: el carácter cualitativo, es decir, cómo es estar en una experiencia *en particular*, y el carácter subjetivo, es decir, cómo es estar en esa experiencia *para mí*. Esta teoría sostiene que en todo estado consciente coexisten dos representaciones que explican esos dos caracteres: una acerca del objeto y la otra acerca de sí misma.[206] La conciencia que surge de esta autorrepresentación es la conciencia subjetiva y tiene rasgos bien definidos: es un *ser-consciente-de* interno y periférico. Interno, porque se refiere a la experiencia misma, y no al objeto de la experiencia, sea este último externo o interno.[207] Periférico, porque no está en el foco de la atención, al contrario de lo que sucede con la conciencia cualitativa propia de la representación de primer nivel.

¿Existen elementos procesuales o autotransformativos dentro de la Teoría Autorrepresentacional? La respuesta no es sencilla. Por un lado, la teoría tiene claramente un aspecto de clausura representacional, puesto que establece que una representación consciente es una representación que se representa a sí misma. Por otro lado, es necesario aclarar cuál es la idea de representacionalidad que Kriegel maneja, y Kriegel no siempre demuestra tener una postura fija y definida al respecto, si bien se inclina por una noción reductivista dependiente de la respuesta del sujeto. A continuación trataré de interpretar con la mayor fidelidad posible la noción de representación y autorrepresentación de Kriegel. Después presentaré un argumento en favor de la idea de que, para definir la teoría autorrepresentacional, es preciso adoptar un cierto enfoque de proceso y no cualquiera, sino uno con aspecto autotransformativo.

discriminar correctamente, los pacientes con visión ciega afirman no haber sido conscientes de ver esos estímulos en ningún momento. (Weiskrantz 1986, 1996)

[205] Cfr. Rosenthal (2005), donde se trata de dar solución a algunos de estos problemas.

[206] Por tanto, esta teoría se debe en gran parte a Brentano y su distinción entre objeto primario y objeto secundario de representación (Brentano 1874: 167).

[207] En términos de Brentano, la autorrepresentación tiene un aspecto interno porque se refiere al objeto secundario, y no al objeto primario de la representación. Y esto sucede al margen de si el objeto primario es externo o interno.

La caracterización de la noción de representación y autorrepresentación en Kriegel es un largo y sinuoso camino que se puede dividir en dos etapas. La primera etapa consiste en la caracterización de la noción de representación (que tiene aspecto reductivista y teleosemántico) y de la noción de representación apropiada (a saber, autorrepresentación). Este es el denominado "enfoque estricto". La segunda etapa consiste en la caracterización de la ontología de la conciencia y el papel que la representación apropiada juega en ella. Este es el denominado "enfoque sinuoso".[208]

3.2.4.1. Representación apropiada o enfoque estricto

Comenzaré por la caracterización de la noción de representación. Kriegel presenta sus argumentos de manera progresiva, a modo de escalera. Esta "escalera argumentativa" se puede esquematizar en cinco grandes pasos.

El primero consiste en caracterizar la representación. Kriegel parte de una noción contemporánea de representación. La noción clásica de representación como imagen (Locke, Hume) es sustituida por la noción contemporánea de representación como aquello que tiene valor informativo. Como ya se mencionó más arriba, hay dos ideas que allanan el camino hacia esta nueva idea de representación. La primera se debe a Brentano (1874: 104), quien sostiene que la representación no es lo representado, sino el acto (o el hecho) de representar. Por tanto, la representación es, ante todo, un acto, un proceso. La segunda se debe a Dretske (1995: xii) quien, al definir su tesis representacional, sostiene que todos los hechos representacionales son hechos acerca de funciones informacionales. Por lo tanto, el representar se define de manera funcional. Representar es tener la función de dar información.

Kriegel es perfectamente conocedor de este acervo. Sin embargo, no pone al proceso en el centro de la idea de representación, sino que se centra en la idea de información: la representación es una propiedad dependiente de respuesta. Una propiedad dependiente de respuesta es una disposición hacia una cierta respuesta R que posee cierto sujeto que responde bajo ciertas circunstancias C. Por ejemplo, la representación del cielo con un determinado matiz de azul viene dada por la disposición del sujeto para responder de una determinada manera bajo determinadas circunstancias (por ejemplo, para decir que "el cielo es menos azul cuando se compara con el mar que está bajo él" o para protegerse los ojos al mirarlo directamente). Por tanto:

> Paso 1: La propiedad dependiente de respuesta (entendida como disposición de un cierto sujeto que responde a dar una cierta respuesta R bajo unas ciertas circunstancias) da cuenta de la representación

[208] La caracterización de la representación propia del enfoque estricto se corresponde con Kriegel (2009: 87–93), mientras que la caracterización de la ontología de la conciencia propia del enfoque sinuoso se encuentra en Kriegel (2009: 215–230). La presentación de Kriegel tiene forma de argumento en desarrollo, con abundantes rectificaciones y justificaciones.

Esta concepción de la representación, no obstante, no puede ser irreductible, de acuerdo con Kriegel (2009: 89). La razón reside en el viejo argumento del espectro invertido y el espectro desplazado. No es compatible que la representación quede completamente caracterizada de manera funcional y que la representación sea suficiente para explicar las propiedades cualitativas de la conciencia. Si la representación quedase completamente caracterizada funcionalmente, dos sujetos que fueran funcionalmente idénticos disfrutarían de la misma experiencia consciente. Sin embargo, de acuerdo con el argumento del espectro desplazado, eso no tiene por qué suceder en todos los casos.

Igualmente, tampoco es aceptable que las disposiciones queden completamente definidas por su soporte físico o neurofisiológico (por ejemplo, por determinados sistemas de activación talámico-corticales), ya que, de ser así, la conciencia carecería de múltiple realizabilidad. Por este motivo, Kriegel apuesta por una noción de conjunción de disposiciones. Para determinado tipo de individuos, la disposición se determina por un soporte neurofisiológico, sin óbice de que, para otro tipo de respondientes, los soportes puedan también ser distintos. Sea cual sea el caso, sucede que:

> Paso 2: Determinados soportes físicos dan cuenta de la propiedad dependiente de respuesta

Tomemos un poco de aire. Kriegel acaba de presentar una noción de representación disposicional (en el sentido de que es dependiente de respuesta), reductivista (en el sentido de que queda determinada por el nivel físico) y compatible con el principio de múltiple realizabilidad. Pero esta representación disposicional, reducible y múltiplemente realizable no es suficiente, por sí misma, para explicar las propiedades cualitativas que forman parte de la experiencia consciente. Su principal problema es que no introduce ningún criterio para explicar el carácter subjetivo propio de esa experiencia. Por tanto, únicamente da cuenta de propiedades no cualitativas, esto es, propiedades que no pertenecen estrictamente al ámbito de la experiencia consciente. Esto es:

> Paso 3: La representación dependiente de respuesta da cuenta únicamente de las propiedades no manifiestas (que Kriegel denomina *schmalitativas*)

Para dar cuenta de las propiedades cualitativas es necesario un tipo de representación "apropiada" (Kriegel 2009: 110). De acuerdo con la tradición representacional, la representación "apropiada" solo puede ser una representación de la representación. Por otro lado, por el argumento maestro de Kriegel se concluye que esa representación de una representación solo puede ser una autorrepresentación.[209] Esta es la base de la experiencia consciente. Es decir:

[209]De acuerdo con el argumento maestro, todo estado consciente debe ser representado conscientemente. Si todo estado consciente es representado conscientemente, entonces existen tres alternativas: o bien es representado (conscientemente) por ese mismo estado mental o bien es representado (conscientemente) por un estado mental diferente del estado mental propio de la representación o bien sucede una mezcla de ambos. Sin embargo, las dos últimas opciones conducen a regresiones infinitas.

> Paso 4: La representación dependiente de respuesta y apropiada, es decir, con autorrepresentación, es condición de la experiencia consciente (porque le confiere su carácter subjetivo)
>
> Paso 5: La representación dependiente de respuesta y apropiada, es decir, con autorrepresentación, da cuenta de las propiedades cualitativas de la experiencia

Como admite el mismo Kriegel, en este punto nos acechan más dudas que certezas.[210] Algunas de esas dudas son: ¿Cuál es la naturaleza de la autorrepresentación? ¿Es también "dependiente de respuesta"? ¿Y cuáles son los mecanismos fisiológicos que la apuntalan? La certeza es que, sean cuales sean los mecanismos fisiológicos, esos mecanismos que soportan la autorrepresentación soportarán también la representación apropiada propia de la experiencia consciente. Y como quiera que esos mecanismos no pueden ser otra cosa que procesos concretos de algún tipo, la representación apropiada propia de la experiencia consciente quedará definida por esos procesos.[211]

3.2.4.2. Ontología de la conciencia o enfoque sinuoso

La segunda etapa de la argumentación de Kriegel trata de despejar las dudas anteriores y se centra en determinar cuál es la ontología de la conciencia a través de las nociones de representación y representación apropiada.

Aquí los renglones en los que se escribe el argumento son torcidos. Kriegel lo denomina "el enfoque representacional sinuoso" (*crooked representational account*). Esta vez no voy a acompañar a Kriegel en su largo y proceloso camino, sino que lo voy a esperar casi al final de la meta, en la versión (CR7) del enfoque sinuoso. Para ello, primero voy a describir los conceptos básicos en los que se encuadra este enfoque, a saber, noción de parte lógica, la distinción complejo/suma y la noción de representación indirecta.

Noción de parte lógica. Kriegel necesita una mereología con la cual analizar la conciencia por medio de sus partes (es decir, por descomposición). Lo hace de la siguiente manera: (1) acepta una mereología clásica, donde la relación PP (parte propia) es transitiva, irreflexiva y asimétrica; (2) las partes, no obstante, no pueden ser partes temporales ni partes espaciales, ya que esas partes no parecen funcionalmente adecuadas para analizar qué es la conciencia (esto es, para su ontología). Por ese motivo, se define una nueva noción de parte de tipo funcional: la parte lógica. Para Kriegel, los estados mentales se componen

[210] "El problema del enfoque representacional estricto es que no está claro cómo puede un estado mental tener una genuina relación causal consigo mismo, en virtud de la cual se represente a sí mismo." (Kriegel 2009: 220)

[211] Las cuestiones anteriores admiten una problematización en términos de procesos. Brentano (1874: 104) sostiene que representación es acto de representar, es decir, proceso. Por consiguiente, la autorrepresentación solo puede ser un proceso reflexivo, esto es, un proceso que es parte de sí mismo. De nuevo, asoma aquí el problema del corolario de granularidad insuficiente referido en la sección 2.3.2.1. Como, por (irref2), un proceso no puede ser parte de sí mismo, es necesario estipular procesos intermedios que medien en la explicación de esa autorrepresentacionalidad. El problema de Kriegel es definir con una granularidad suficiente en qué consiste, realmente, la autorrepresentacionalidad propia de los estados conscientes. La dificultad añadida es que lo hace en términos no procesuales.

de partes lógicas y los estados conscientes, en tanto que mentales, también.

Distinción complejo/suma. Inspirado en Simons (1987), Kriegel distingue entre dos tipos de todo: sumas y complejos. El complejo es un todo cuyas partes están relacionadas entre sí de manera esencial. La suma es un todo cuyas partes no están relacionadas entre sí o, de estarlo, es de una manera contingente. Las moléculas son complejos de átomos y no su mera suma. Una sociedad es un complejo de personas y no su mera suma. Hay una característica adicional que diferencia los complejos de las sumas: para que una suma deje de existir es necesario reducir su número de partes, por ejemplo, quitar una de ellas. Sin embargo, un complejo puede dejar de existir incluso en el caso de que se mantengan cada una de sus partes. La idea de complejo se introduce para distinguir la Teoría Autorrepresentacional de las teorías representacionales de orden superior.

Noción de representación indirecta. El hecho de que un estado se represente a sí mismo encierra innumerables complicaciones y conduce a regresiones infinitas. Kriegel resuelve este problema postulando dos partes de ese estado, una de las cuales representa a la otra. Para que esa representación sea autorrepresentación, es necesario refinar la noción de representación. Kriegel introduce una distinción entre representación directa y representación indirecta. Una representación directa tiene dos elementos (Rte y Rdo) y está determinada teleosemánticamente (por ejemplo, mediante propiedades dependientes de respuesta). Una representación indirecta, en cambio, tiene tres elementos (Rte, RdoI y RdoD, donde Rte representa indirectamente a RdoI a través de su representación directa de RdoD) y requiere de dos condiciones: (i) que exista una relación naturalística relevante entre la representación (Rte) y parte de lo representado (RdoD) y (ii) que exista una relación de *ser-parte-propia-de* relevante entre la parte de lo representado (RdoD) y el representado (RdoI).[212] Según Kriegel, en la experiencia consciente existe una representación directa de una de las partes del estado consciente dirigida hacia la otra parte y una representación indirecta del estado total, esto es, de la experiencia consciente como un todo.

El modelo (cuasi) final del enfoque sinuoso es el siguiente:[213]

> (CR7) Para cualquier estado mental M de un sujeto S, M es consciente ssi existen dos $M*$ y $M\blacklozenge$, tales que (i) $M*$ es una parte propia de M, (ii) $M\blacklozenge$ es una parte propia de M, (iii) M es un complejo de $M*$ y $M\blacklozenge$, y (iv) $M*$ representa a M mediante su representación de $M\blacklozenge$.

donde el complejo entre $M*$ y $M\blacklozenge$ se basa en la unidad sincrónica de la conciencia, y las bases de la representación indirecta son la relación naturalística entre $M*$ y $M\blacklozenge$ así como la relación *ser-parte-propia-de* entre $M\blacklozenge$ y M.

[212] Kriegel (2009: 226)

[213] Lo he denominado modelo "(cuasi) final" porque aún existe una refinación mayor del modelo, que se deriva de la necesidad de evitar la regresión infinita que proviene del hecho de que $M*$ es, a su vez, consciente (Kriegel 2009: 226). Esta circunstancia no altera el argumento posterior.

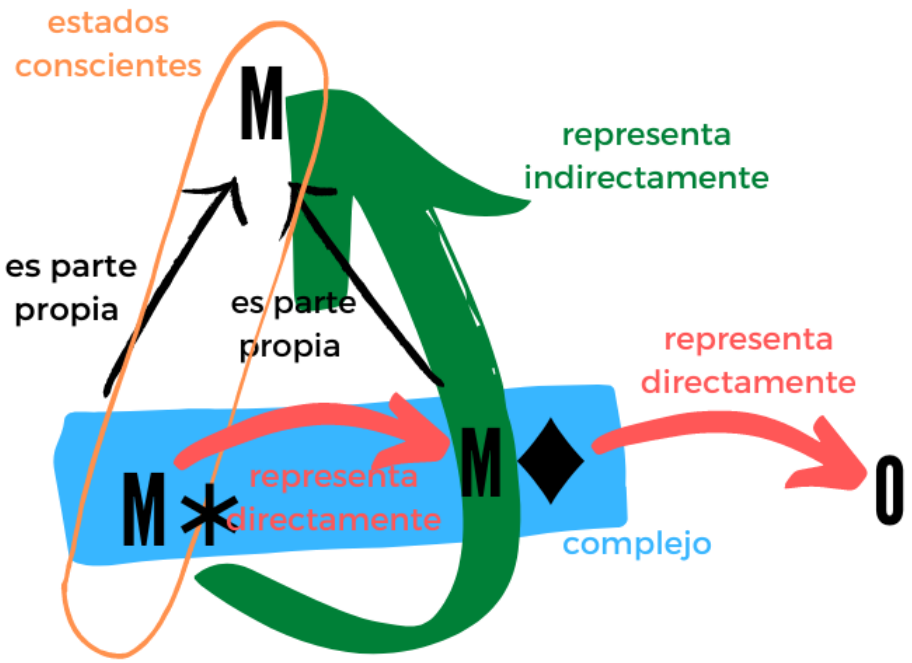


Figura 3.5: Esquema visual del enfoque representacional sinuoso de acuerdo a su versión (CR7). M, $M\ast$ y $M\blacklozenge$ están definidas en esta versión. Por su parte, O representa el objeto (o contenido) de la representación teleosemántica $M\blacklozenge$.

3.2.4.3. Clausura en la autorrepresentación

El análisis de la noción de autorrepresentación que acabo de presentar y que concluye en la versión (CR7) del enfoque representacional sinuoso es la complicada historia de cómo tratar de explicar una clausura (la propia del hecho autorrepresentacional) a través de relaciones no clausuradas, basándose en una mereología clásica y prescindiendo de la noción de proceso. Lo que trataré de hacer ahora es dar una serie de argumentos que apoyen la tesis de que esta explicación, este enfoque, no evita la clausura y está abocado a una revisión desde el punto de vista de proceso y, más en concreto, de proceso autotransformativo.

La arquitectura relacional formalizada en (CR7) y esquematizada en la figura 3.5 no presenta, al menos en apariencia, bucles sospechosos. Se podría argumentar que, puesto que el estado mental $M\ast$ representa a M y, al mismo tiempo, es una parte de M, esto podría generar un bucle. Si $M\ast$ representa a M, entonces su soporte físico debe estar causalmente determinado por M, mientras que por otro lado, si $M\ast$ es una parte lógica de M, entonces el soporte físico de M estaría determinado (en parte) por $M\ast$. Se diría que Kriegel es muy consciente de esta dificultad y que el problema queda resuelto al introducir la noción de representación indirecta: $M\ast$ representa a M solo en virtud de representar a $M\blacklozenge$. $M\ast$ representa a M solo de manera indirecta. No hay clausura por este lado.

Los detalles, sin embargo, parecen delatar otro tipo de clausura. Proviene de la conjunción de dos ideas irrenunciables para el enfoque sinuoso, a saber, que $M\ast$ representa

directamente a $M\blacklozenge$ (lo que viene dado por iv) y que M es un complejo de $M*$ y $M\blacklozenge$ (lo que viene dado por iii). Primero definiré por qué se introducen estas dos ideas en el enfoque sinuoso y después analizaré las consecuencias de su introducción conjunta.

La primera idea, que $M*$ representa directamente a $M\blacklozenge$, es la base de la autorrepresentacionalidad que busca explicarse. Si esto no se diese, si $M*$ no representase directamente a $M\blacklozenge$, no podría haber autorrepresentación en absoluto. De hecho, ni siquiera habría representación de orden superior.

En cuanto a la segunda idea, esto es, que M es un complejo de $M*$ y $M\blacklozenge$, ¿por qué es necesaria? Kriegel la justifica porque desea diferenciar nítidamente el enfoque autorrepresentacional de las teorías de orden superior. Si $M*$ y $M\blacklozenge$ no formasen un complejo (a saber, M), entonces el enfoque sinuoso se diferenciaría de un enfoque de orden superior únicamente de manera "superficial".[214] La idea de complejo, por tanto, es necesaria y está justificada. De ella depende que la teoría sea autorrepresentacional y no de orden superior. Pero ¿en qué sentido sostiene exactamente Kriegel que M es un complejo de $M*$ y $M\blacklozenge$?

Para Kriegel, el complejo consiste en la unidad sincrónica de la conciencia que, probablemente, queda apuntalada por activaciones neuronales de tipo sincrónico. ¿Constituye esta unidad un complejo? Por definición, un complejo es un todo cuyas partes están conectadas entre sí de manera esencial. Hay aquí dos opciones: La opción (a) es que la unidad sincrónica de la conciencia no sea sino una unificación que no altera esencialmente los elementos que unifica, es decir, $M*$ y $M\blacklozenge$. Si esto es así, entonces no habría complejo y, en consecuencia, el enfoque sinuoso colapsaría en un enfoque de orden superior, perdiendo su carácter autorrepresentacional. La opción (b) consiste en que la unidad sincrónica sea realmente un complejo, de tal manera que produce una interconexión esencial entre lo que unifica, a saber, $M*$ y $M\blacklozenge$.

La opción (b), sin embargo, conduce irremediablemente a la existencia de un bucle o clausura dentro del enfoque sinuoso. El bucle sucede del siguiente modo. Por (iv) más la noción reductivista de representación, $M*$ representa a $M\blacklozenge$ y, por tanto, existe una relación causal RC1 entre los soportes físicos de $M\blacklozenge$ y los soportes físicos de $M*$. Por otro lado, por (iii), si $M*$ y $M\blacklozenge$ son un complejo, y eso tiene un significado no trivial, entonces por fuerza debe existir una relación causal RC2 entre $M*$ y $M\blacklozenge$ distinta de la relación causal RC1 en virtud de la cual ambos estarán interconectados esencialmente.

En definitiva, el enfoque sinuoso conlleva relaciones causales internas con clausura.

Las relaciones internas con clausura no son algo problemático, ni invalidan el enfoque sinuoso. Sin embargo, ponen de manifiesto que es necesario introducir la noción de clausura al sistema. Teniendo en cuenta que las relaciones causales de las que se habla aquí solo pueden ser procesos, la inevitable conclusión es que los procesos que soportan el enfoque sinuoso deben ser procesos compuestos, interferentes y con clausura. Es decir, deben tener

[214] Kriegel (2009: 221).

justamente el aspecto de un proceso autotransformativo.

3.2.5. Conclusiones

La idea de que la conciencia es un proceso autotransformativo se encuentra de forma más o menos explícita en muchas teorías de la conciencia, y no de manera meramente anecdótica, sino de manera fundamental. Esto se puede observar en cuatro de las teorías más importantes. La Teoría del Espacio de Trabajo Global establece que los procesos que dan cuenta de la conciencia son recurrentes y, por tanto, tienen la clausura de composición propia de los procesos autotransformativos. La Teoría de la Información Integrada, por su parte, explica la integración de la información a través de sistemas dinámicos discretos y establece que esos sistemas deben tener una composición dinámica y una modulación con la clausura propia de la autotransformatividad. La Teoría de la Doble Microgénesis plantea también que la conciencia se explica a través de dinámicas basadas en procesos de reentrada en los que toman parte los sistemas específicos y los sistemas no específicos del cerebro, y estos procesos de reentrada tienen, de nuevo, un aspecto autotransformativo. Por último, la Teoría Autorrepresentacional, que no hace uso de la noción de proceso sino de la de estado, se ve abocada, en última instancia, a introducir aspectos autotransformativos cuando necesita explicar cuál es la naturaleza concreta de la autorrepresentación propia de los estados conscientes.

La idea de aplicar un enfoque de procesos para explicar la experiencia consciente no es un conejo que sale de la chistera, ni tampoco es algo nuevo en los estudios acerca de la conciencia. Más bien ha estado en el ambiente desde hace mucho tiempo. Ahora voy a dar más razones para emplear este enfoque procesual, solo que, esta vez, estas razones no estarán basadas en las teorías de la conciencia, sino en las mismas propiedades de la experiencia consciente.

3.3. Propiedades de la conciencia y proceso

Una vez más, un proceso es "una familia organizada de ocurrencias que están sistemáticamente enlazadas entre sí, bien causalmente, bien funcionalmente" (Rescher 1996: 38). Si se acepta como válido el marco de la filosofía de procesos, esto es, si todo en la realidad son procesos y la experiencia consciente constituye una parte de la realidad, entonces, de manera trivial, la conciencia debe tratarse como un proceso. No obstante, y dejando a un lado argumentos que suponen aquello que quieren demostrar, existen razones poderosas para poner bajo sospecha todo enfoque no procesual acerca de la conciencia. De nuevo voy a defender las mismas dos ideas: (1) que la conciencia debe ser comprendida como un proceso y (2) que este proceso tiene cierto aspecto autotransformativo. Sin embargo, esta vez defenderé estas ideas exponiendo dos argumentos basados en cómo es la conciencia, es decir, en sus propiedades. El primero es el argumento de las propiedades no fenoménicas, que se basará en propiedades dinámico-estructurales (la composicionalidad), propiedades funcionales y propiedades temporales para defender que es preciso analizar la conciencia como un proceso. El segundo argumento es el argumento de la intimidad, de acuerdo con

el cual la explicación del carácter subjetivo de la conciencia requiere introducir cierto aspecto autotransformativo a los procesos que dan cuenta de la experiencia consciente.

3.3.1. Argumento de las propiedades no fenoménicas

Si la conciencia es un proceso, entonces aquellas que se definen como sus propiedades deberán tener el aspecto de propiedades de procesos.[215] La definición de proceso es muy amplia, pero, al menos, señala dos requisitos fundamentales que todo proceso, y la conciencia entendida como tal, debe cumplir. Por un lado, debe ser compuesto. Por otro lado, sus componentes deben estar relacionados causal o funcionalmente. Ahora analizaré ambos requisitos, poniendo énfasis en una propiedad dinámico-estructural, a saber, la composicionalidad de la conciencia (que defiende su carácter compuesto) y en las denominadas propiedades funcionales (de acuerdo con las cuales la conciencia no es un epifenómeno). Además, todo proceso se entiende como algo que sucede en el tiempo. Por consiguiente, acabaré esta argumentación exponiendo las propiedades temporales, sin las cuales no es posible explicar completamente la conciencia y que nos dan, quizás, la pista definitiva para comprender la conciencia como un proceso.

El argumento de las propiedades no fenoménicas se puede esquematizar de la siguiente manera:

> Premisa 1: un proceso es todo aquello que tiene cierta composición, esto es, tiene partes que se relacionan entre sí de manera funcional o causal, y sucede en el tiempo (y no de manera instantánea)
>
> Premisa 2: la conciencia cumple la propiedad de composición, esto es, tiene partes, pero se vive como con cierta unidad.
>
> Premisa 3: las partes de la conciencia tienen relaciones funcionales o causales relevantes entre sí.
>
> Premisa 4: la conciencia tiene propiedades de tipo temporal, es decir, que son propias de aquello que sucede en el tiempo. Igualmente, estas propiedades son relevantes.
>
> Por consiguiente, la conciencia es un proceso y debe ser analizada como tal.

Premisa 1: sobre lo que define a un proceso

El primer paso viene dado por la definición de proceso como "familia organizada" de ocurrencias enlazadas "causal o funcionalmente". La definición de Rescher es, por supuesto,

[215] De nuevo, no entraré aquí a juzgar la adecuación o no del término "propiedad" referido a procesos, incluso admitiendo su gran carga filosófica dentro de la metafísica de la sustancia (sustancia como soporte de propiedades). Simplemente emplearé la noción con la esperanza de que se entienda de un modo neutro y tratando de mantenerme en todo momento en los límites de esa neutralidad.

amplia, como corresponde a un enfoque que filosofía de procesos, en el cual todo debe ser explicado mediante relaciones, y nada debe ser explicado mediante la noción de sustancia (entendida como soporte de propiedades).

Premisa 2: la conciencia tiene composición y unidad

Si la conciencia fuese simple, es decir, si fuese monolítica y estuviera formada de un solo componente, entonces se escaparía del análisis de procesos o, dicho en otras palabras, el análisis de la conciencia en términos de proceso carecería de interés. Si, por ejemplo, se entendiera la conciencia como un puro fenómeno que se puede separar de todas las propiedades dinámico-estructurales (manteniendo de ella tan solo las propiedades fenoménicas y la propiedad de la unidad, que vendría directamente dada a partir de su simplicidad) y si se entendiera, además, que ese fenómeno carece de partes experienciales y carece de relación, en cuanto a su *complejidad* de composición, con la *complejidad* de composición del mundo que nos rodea, entonces no tendría sentido invertir nuestros esfuerzos en un análisis procesual de la conciencia. Este modelo de la conciencia no es mayoritario, pero existe, y podría ser representado (con matizaciones) por el enfoque PANIC de Tye (1995, 2003), de acuerdo con el cual la experiencia es abstracta, intencional y carece de contenido conceptual. No obstante, negar a la experiencia consciente la propiedad de composición tiene un coste demasiado alto que no todo el mundo está dispuesto a pagar. De hecho, equivale a invalidar toda fenomenología, por modesta que sea. Los mismos argumentos que hablan a favor de la propiedad de la composición son también argumentos en contra de este tipo de modelos de la conciencia. Aquí recordaré estos argumentos. Para empezar, hay una tendencia relativamente clásica[216] que distingue, al menos, dos caracteres en la experiencia fenoménica: el carácter subjetivo y el carácter cualitativo. Esta distinción ya introduce una cierta composicionalidad que puede ser analizada en términos de proceso. Además, hay otra tendencia, si cabe tan clásica como la anterior,[217] que entiende la conciencia como una corriente o flujo continuo del cual no solo forman parte los estados más estables y tradicionales, sino también otros estados más escurridizos que actúan, por así decirlo, en el margen de la conciencia. De nuevo, aceptar la distinción entre parte nuclear y parte marginal de la conciencia equivale a aceptar que existe cierta composicionalidad. Por último, existe una tendencia —que proviene de las teorías informacionales, tales como la Teoría de la Información Integrada, pero también compartida por otros autores tan divergentes como Searle (1992) o Kriegel (2009)— que establece una relación entre la *complejidad* del mundo y la *complejidad* de la experiencia consciente. Este tipo de teorías, que se podrían llamar de "composicionalidad fuerte", asocian percibir a *percibir-como* y ser consciente a *ser-consciente-de-algo-como* y, en consecuencia, implican que la experiencia consciente debe poseer cierta composicionalidad.[218] Lo importante, sin embargo, no es aquí tanto el hecho de que sea preferible optar por teorías de composicionalidad débil o teorías de composicionalidad fuerte, sino el hecho de que la composicionalidad es una propiedad irrenunciable para la mayoría de las teorías de la conciencia y que ello, junto con el resto de propiedades, hace necesario el análisis mediante procesos.

Al mismo tiempo, la conciencia se vive como dotada de unidad. En otras palabras, sus

[216] James (1931); Nagel (1974); Crane (2000); Kriegel (2009); Sebastián (2020).
[217] James (1931); Mangan (1993).
[218] Searle (1992: 133).

partes tienen una cierta integración. El argumento principal para esto es que la sensación de unidad (de no interrupción, de conciencia única) es parte de la fenomenología de la experiencia consciente.

Premisa 3: las partes de la conciencia tienen una función relevante

La funcionalidad en la conciencia se puede entender de dos formas. Primero, se puede entender como funcionalidad por debajo de la conciencia, es decir, que existen elementos funcionales que determinan la experiencia consciente. Segundo, se puede entender como funcionalidad en las partes de la experiencia consciente entre sí. La distinción es importante, puesto que hay autores que aceptan la primera forma de causalidad mientras que niegan la segunda.

De hecho, dos de las posiciones epifenoménicas más conocidas son explícita o implícitamente funcionalistas. Jackendoff, por ejemplo, acepta al mismo tiempo la hipótesis de suficiencia computacional —de acuerdo con la cual "toda distinción fenoménica está causada por/está soportada por/se proyecta desde una distinción computacional correspondiente" (Jackendoff 1994: 24)— y la hipótesis de no eficacia de la conciencia —de acuerdo con la cual "la conciencia de una entidad no puede tener ningún efecto en la mente computacional" (Jackendoff 1994: 25)—. Chalmers (1996), por su parte, es funcionalista en el sentido de que propone una explicación funcional, con finura de grano indeterminada (probablemente, la neurona) para la conciencia y, al mismo tiempo, rechaza que la experiencia consciente tenga ninguna función relevante.[219]

Con objeto de diferenciar correctamente en qué sentido se habla de funcionalidad, definiré el principio de funcionalidad como la asunción de que la experiencia consciente o sus partes experienciales (a) tienen una función y (b) esa función es relevante para los procesos cognitivos. Aceptar este principio abre la posibilidad de un análisis basado en procesos desde la noción de partes de la experiencia consciente. Negar (b) pero aceptar (a) es una posición de epifenomenalismo débil.[220] Negar (a) y (b) es una posición de epifenomenalismo fuerte.

¿Cuáles son las razones para aceptar el principio de funcionalidad? Existe una razón inicial para aceptar cierta funcionalidad de la conciencia y es que el epifenomenalismo fuerte es contradictorio *in terminis*.[221] Es decir, no es posible mantener al mismo tiempo que la conciencia existe y que no tiene ninguna función en absoluto puesto que, si no tuviese

[219]La posición de Chalmers (1996) resulta, como es bien sabido, una *rara avis* en el ecosistema de las teorías de la conciencia. No hay tiempo es este momento para entrar en la aparente dificultad de compatibilizar el argumento zombi (que justifica su posición epifenoménica) y el argumento de los *qualia* danzantes y desvanecientes (que justifica su posición funcionalista). Simplemente quiero resaltar que, si bien la posición de Chalmers (en la medida en que se sostiene que la experiencia es afuncional), no admite un análisis procesual de las partes experiencial, si admite y, de hecho, exige un análisis procesual del que formen parte los elementos funcionales que producen la conciencia como tal, sea este análisis de la finura de grano que sea.

[220]Quizás el exponente más célebre del epifenomenalismo débil es Huxley (1874), quien sostiene que la conciencia es a los sistemas cognitivos lo que el silbido al motor del tren, es decir, algo sin una función relevante.

[221]La idea proviene de Searle, quien la expone en Searle et al. (1997).

ninguna función en absoluto, sería imposible siquiera hablar sobre ella y, si solo fuese una ilusión, sería necesario estipular que no existe. En cualquier caso, la disputa acerca del epifenomenalismo fuerte no puede resolver el problema que se dirime aquí, a saber, si el enfoque de procesos es adecuado para comprender la conciencia. Para esto es necesario discutir el epifenomenalismo débil.

La primera razón para discutir el epifenomenalismo débil viene por la vía de la conciencia de acceso. La experiencia consciente, en su forma completa, posee propiedades funcionales, y estas propiedades no siempre son beneficiosas.[222] Para empezar, la conciencia ralentiza el procesamiento cognitivo y reduce su eficiencia, ya que lo hace más propenso a errores. Por ejemplo, pensar en cada palabra reduce la fluidez del discurso, igual que ser conscientes de cada paso nos vuelve más torpes al caminar. En contraposición, la conciencia aumenta la capacidad para relacionar contenidos dispares entre sí y aumenta también la capacidad de relacionar contenidos conscientes con contextos no conscientes. Por ejemplo, hace que se pueda reaccionar mejor ante situaciones novedosas, tales como frenar ante la presencia de un peatón mientras conducimos, y permite trazar conexiones de ideas en el curso del estudio de una materia. Por ese motivo, Baars (1988) defiende que existen dinámicas amplias (o macrodinámicas de la conciencia) en las cuales se pueden alternar fases donde el contenido de la cognición forma parte de la conciencia con fases donde el contenido de la cognición no forma parte de la conciencia. Este es el caso del proceso de aprender a tocar el piano, durante el cual quien aprende comienza siendo consciente de la posición de las manos para, paulatinamente, comenzar a perder conciencia de esa posición conforme adquiere habilidad. Eventualmente, a causa de un vicio postural, puede ser necesario volver a ser consciente de la postura de la mano para poder corregirla.

Contra lo anterior se podría objetar que las propiedades funcionales son propiedades únicamente de una forma de conciencia, a saber, la conciencia de acceso, mientras que otras formas de conciencia exclusivamente fenoménicas suceden en etapas previas al acceso sin presentar ninguna propiedad funcional (Block 2011; Lamme 2010). Esta objeción parece estar reforzada por las evidencias empíricas de que existe una memoria icónica o una memoria visual de corto plazo frágil que podría ser (fenoménicamente) consciente pero que todavía no tiene las características propias de la conciencia de acceso.[223]

[222] Baars (1988).

[223] Véase, entre otros, Sperling (1960) y Vandenbroucke et al. (2014). Sperling ideó un sencillo experimento. Presentó a varios sujetos un estímulo breve basado en matrices de letras (o letras y números) y midió la capacidad de reporte posterior al estímulo de los sujetos en dos casos: reporte completo y reporte parcial. La tarea de reporte completo consistía en recordar el máximo número de ítems y su posición. En esta tarea los sujetos eran capaces de reportar una media de 4,3 letras de un total de 12 letras. La tarea de reporte parcial era idéntica, pero incluyendo retroseñalización, esto es, introduciendo inmediatamente después del estímulo (y no de manera simultánea) una señal acústica que señalaba la línea de la matriz que se había de reportar, a saber, la de arriba, la de en medio o la de abajo. Al ser una señal acústica y no visual se trataba de garantizar la menor interferencia posible. En el caso de reporte parcial los sujetos eran capaces de reportar unas 3 letras por fila. En otras palabras, la media de letras reportadas por los sujetos subió drásticamente hasta 9,1 de un total de 12. Sperling concluyó que había una considerable capacidad de información visual que el sujeto no era capaz de reportar y aceptó la dificultad de explicar los mecanismos que subyacen a ese exceso de información no reportable. Vandenbroucke et al. (2014), por su parte, realizaron experimentos similares basados en tareas de detección de cambios en estímulos visuales en condiciones de presencia y ausencia de retroseñalización. De nuevo, los resultados sugerían un mayor rendimiento en los casos de retroseñalización. Block (2011) y Lamme (2010) sugieren que existe

En contra de esta objeción caben dos réplicas.

La primera réplica es que los casos de conciencia fenoménica sin conciencia de acceso no representan la mayoría de las experiencias conscientes, sino que, en todo caso, representan una forma de experiencia consciente muy específica que se produce en condiciones de presentación experimentales (brevedad de presentación del estímulo, enmascaramiento), esto es, mediante una intervención artificiosa que condiciona el proceso consciente o lo detiene en una de sus etapas iniciales. Por consiguiente, como mínimo se podrá decir que el enfoque de procesos es pertinente de manera típica, es decir, en la mayoría de los casos.

La segunda réplica consiste en que los experimentos en los que se han provocado estas situaciones de conciencia fenoménica sin aparente conciencia de acceso (tales como la memoria icónica o la memoria visual de corto plazo frágil) se sirven de mecanismos con cierta funcionalidad reducida para inferir que existe experiencia consciente en ellas.[224] Por ejemplo, se basan en la mejora del rendimiento de la tarea de identificación de estímulos mediante retroseñalización. Por consiguiente, estas situaciones no son exactamente casos de conciencia sin acceso, sino, a lo sumo, casos de conciencia sin acceso para reporte, pero con cierta conciencia de acceso indirecto, esto es, acceso que puede ser suscitado mediante procedimientos indirectos. Por consiguiente, este tipo de experiencia consciente todavía posee cierto tipo de propiedades funcionales que avalan el enfoque de procesos.

La consecuencia de la existencia de macrodinámicas de la conciencia es que la conciencia no puede ser algo funcionalmente irrelevante, sino una herramienta que tiene funciones concretas en los sistemas cognitivos.

Premisa 4: las propiedades temporales de la conciencia son relevantes

La idea de que la conciencia puede ser descrita de manera temporal es relativamente nueva y proviene de las evidencias empíricas recogidas sobre todo a finales del siglo XX y en el siglo XXI. Los experimentos de enmascaramiento hacia atrás mediante metacontraste indican que la experiencia consciente requiere de cierto tiempo de ignición y que ese tiempo de ignición puede ser variable.[225] Otros experimentos sencillos como el efecto de retardo del destello demuestran que ese tiempo de ignición puede acelerarse bajo determinadas circunstancias.[226] Estas evidencias reflejan las propiedades de lentitud y aceleración de la conciencia. Los experimentos de rivalidad binocular[227], así como otros experimentos

cierta memoria visual de corto plazo frágil (o memoria icónica) y que esta memoria sería un caso de conciencia fenoménica sin conciencia de acceso. En otras palabras, los sujetos son conscientes de más información de la que pueden reportar.

[224]Cfr. Sperling (1960); Vandenbroucke et al. (2014), respectivamente.

[225]Bachmann y Francis (2014).

[226]Flash lag effect: es el efecto de retardo con el que se percibe un estímulo presentado de manera fugaz con respecto a otro idéntico en forma pero en movimiento. Este retardo es percibido claramente pese a que, en realidad, ambos estímulos se presentan simultáneamente. Véase Watanabe (2022); Watanabe et al. (2010). Ver también Bachmann (2000: 112) y Bachmann (2000: 136)

[227]Logothetis y Schall (1989); Leopold y Logothetis (1996); Blake y Logothetis (2002).

relacionados con fenómenos más sencillos y reproducibles, como el cubo de Necker,[228] demuestran que determinados contenidos de la experiencia consciente no son compatibles entre sí y deben producirse en serie (es decir, uno detrás del otro), incluso en los casos de persistencia del estímulo. Esta es la propiedad de la serialidad de la experiencia consciente. La experimentación psicofísica y la basada en técnicas de medición de la actividad en el cerebro tales como fRMI, ERP y electrodos en la corteza cerebral confirman que la conciencia se asocia con determinados procesos que suceden en el tiempo y que estos procesos son específicos de la experiencia consciente y no de otras tareas cognitivas.[229]

Lentitud, serialidad y especificidad son tres de las propiedades de la conciencia que tienen que ver con cómo es la experiencia consciente a lo largo del tiempo. ¿Son estas propiedades relevantes o son simplemente propiedades irrelevantes para el análisis de la experiencia consciente? Si fuesen irrelevantes, probablemente el análisis de la conciencia como proceso que sucede en el tiempo no tendría sentido o tendría poco sentido. Voy a dar tres razones de que esto no es así: (1) la temporalidad afecta a las funciones de la conciencia, (2) la serialidad muestra que la experiencia consciente tiene un carácter agónico o competitivo y (3) ambas propiedades reflejan una cierta especificidad de la experiencia consciente que la diferencian del estímulo y la tarea de representación.

En cuanto a lo primero, la temporalidad de la conciencia es importante porque condiciona las funciones de la misma. Los experimentos basados en metacontraste demuestran que la detención de los procesos propios de la conciencia pueden prevenir la accesibilidad propia de la conciencia de acceso únicamente mediante manipulación del tiempo transcurrido entre la presentación de los estímulos (Dehaene et al. 2006). Igualmente, el efecto Broca y Sulzer (1902) demuestra que la intensidad con que se percibe un destello de luz puede ser mayor cuando el tiempo de exposición del estímulo es menor, al menos durante una franja de tiempo. El tiempo propio de la conciencia es relevante también para explicar la conciencia de movimiento y el efecto cinematográfico, de acuerdo con el cual las imágenes sucesivas pasan a verse como una animación. Y la conciencia de presente solo puede entenderse bajo el prisma de la temporalidad en la que suceden los procesos conscientes.

La serialidad hace referencia a que existen ciertas experiencias conscientes que son incompatibles entre sí y, por consiguiente, solo se pueden experimentar de una en una y nunca al mismo tiempo. Esta propiedad se da especialmente en la conciencia visual y se puede entender en dos sentidos. El primero es que, cuando los estímulos cambian, la representación varía y, con ella, la experiencia visual también varía. Por ejemplo, cuando se presentan al sujeto dos imágenes distintas de manera sucesiva bajo unas condiciones específicas (tales como fijación de la mirada y control de los tiempos de duración de los estímulos así como del tiempo transcurrido entre la presentación de ambos estímulos[230]), la segunda imagen interfiere en la percepción consciente de la anterior, en un fenómeno que se denomina "enmascaramiento hacia atrás" (*backward masking*). Este fenómeno se puede explicar en términos representacionales.

[228] Ver figura 1.1.
[229] Dehaene (2014).
[230] Esto es, dependen del SOA (*stimulus onset asynchrony*).

El segundo sentido de serialidad, sin embargo, da información relevante acerca de cómo funciona específicamente la conciencia. La experiencia consciente puede cambiar incluso sin alteración en los estímulos o en las representaciones. Exsiten numerosas evidencias empíricas de existencia de serialidad en la conciencia, es decir, de que existen contenidos de la experiencia consciente que no son compatibles entre sí. A continuación recordaré algunas de ellas.

El cubo de Necker es un caso de estímulo ambiguo, en el que la experiencia consciente oscila entre dos manifestaciones, de las cuales solo una puede ser (conscientemente) manifiesta en cada instante. En el cubo de Necker el sujeto solo puede ver una perspectiva en cada momento. Sin embargo, el efecto de interferencia entre procesos no es atribuible a una alteración de los estímulos, ya que el estímulo es siempre el mismo, y el efecto se produce igualmente en casos de fijación de mirada. A menudo, este fenómeno va acompañado de cierta oscilación entre las diferentes opciones.

Por su parte, los experimentos de rivalidad binocular también demuestran que existe cierta competición por el espacio de la conciencia en la percepción visual.[231] La rivalidad binocular se produce cuando se presenta a cada ojo un estímulo diferente de manera simultánea. En estos casos, el sujeto solo puede ser consciente de uno de los dos contenidos en cada momento, quedando la percepción del otro contenido completamente extinguida. De nuevo, se produce un efecto de interferencia entre procesos que no es atribuible a una alteración de los estímulos, sino a una competición por el espacio de la conciencia. La explicación de estos fenómenos recuerda a los fenómenos de biestabilidad propios de los sistemas dinámicos (por ejemplo, la reacción Beloúsov-Zhabotinsky) y parecen llevarnos de nuevo al ámbito del análisis de procesos.

Igualmente, hay evidencias empíricas que nos permiten determinar el tiempo de ignición de la nueva conciencia cuando hay competición por la misma. El enmascaramiento mediante metacontraste es un caso especial de enmascaramiento en el cual se presentan dos estímulos que no ocupan el mismo espacio visual.[232] Sucede bajo ciertas condiciones temporales de presentación de los estímulos (tales como un tiempo entre estímulos inferior a 100ms, y que el segundo estímulo, esto es, la máscara, sea corto), las cuales hacen posible que el segundo estímulo enmascare la percepción del primero, incluso a pesar de que no exista solapamiento del área visual entre ambos. En estos casos no es posible atribuir el efecto de enmascaramiento únicamente a una interferencia representacional. Por consiguiente, debe tratarse de una interferencia que atiene específicamente a la percepción consciente de los estímulos. Los efectos de enmascaramiento son variables, dependen de diversos factores y, en especial, dependen del tiempo transcurrido entre la presentación de los estímulos. Todo esto sugiere que debe existir un tiempo específico de ignición de la conciencia y permite hacer estimaciones sobre el mismo.

La lentitud y la serialidad ayudan también a poner de manifiesto un aspecto muy relevante de la conciencia, a saber, que la conciencia es un hecho específico y no una propiedad secundaria pegada a procesos representacionales y completamente determinada por ellos.

[231] Logothetis y Schall (1989); Leopold y Logothetis (1996); Blake y Logothetis (2002)
[232] Bachmann y Francis (2014)

Con "específico" quiero decir que la conciencia se atiende a procesos concretos que pueden ser distinguidos de otros procesos que son puramente representacionales o puramente cognitivos. Por ese motivo, es necesario sostener que las propiedades temporales son relevantes para la experiencia consciente.

La conclusión final es que las propiedades de la conciencia no solo son compatibles con el análisis de procesos, sino que ponen de manifiesto la necesidad de comprender y analizar la experiencia consciente como un proceso, es decir, como algo que tiene partes que se relacionan entre sí de manera funcional, que sucede en el tiempo y que está compuesto de etapas.

3.3.2. Argumento de la intimidad

Hasta ahora he echado mano de las propiedades dinámico-estructurales (composición), funcionales y temporales de la conciencia para sostener la idea de que es necesario aplicar el análisis de procesos. Sin embargo, he dejado de lado un tipo de propiedades de la conciencia que, paradójicamente, es quizás el más significativo y destacado, a saber, las propiedades fenoménicas y, en concreto, su carácter subjetivo. Ahora trataré de argumentar que el carácter subjetivo de la conciencia, asociado con la intimidad, requiere de un análisis específico de procesos que es el propio de los procesos autotransformativos.

3.3.2.1. Estructura argumental

El argumento de la intimidad tiene una forma algo más compleja que el argumento acerca de las propiedades no fenoménicas. Se basa en propiedades de tipo fenoménico para establecer que los procesos que dan cuenta de la experiencia consciente, y de la intimidad como parte de esa experiencia, deben tener un aspecto autotransformativo. Se entiende intimidad como el carácter subjetivo de la experiencia consciente, es decir, aquello que hace que esa experiencia sea *para-mí* o *mía*, en oposición a otros procesos que suceden sin esa característica, tales como el proceso digestivo, el metabólico o el respiratorio.

El argumento de la intimidad consta de dos partes: el argumento de intrinsicidad y su instanciación en la intimidad como parte de la experiencia consciente.

El argumento de intrinsicidad muestra que existe una relación entre determinadas partes de la experiencia consciente (que, de hecho, pueden llegar a ser muy características de los procesos en los que hay conciencia como, por ejemplo, la intimidad) y la autotransformatividad en los procesos que dan cuenta de ellas. Esta relación se basa en aquello que se viene a denominar procesos intrínsecos, esto es, procesos que surgen por la propia experiencia consciente, que forman parte de ella, pero que no pueden explicarse únicamente de manera relacional-representacional de tipo convencional, a saber, a través de la representación o información que dan acerca del contenido primario de la experiencia

consciente.

La instanciación en la intimidad, por su parte, consiste en demostrar que el argumento de intrinsicidad se puede aplicar a la intimidad, entendida como una parte de la experiencia consciente.

La consecuencia de aceptar ambos argumentos es que la intimidad, entendida como una parte de la experiencia consciente, está relacionada con la existencia de un aspecto autotransformativo en los procesos que dan cuenta de esa experiencia consciente.

3.3.2.2. Argumento de intrinsicidad y argumento de intimidad

Este argumento de intrinsicidad es un argumento que se puede usar para determinadas experiencias conscientes, a saber, aquellas que tienen alguna parte experiencial que cumpla dos criterios fenomenológicos concretos que ahora definiré.[233] El resultado de este argumento es que los procesos que dan cuenta de ese tipo de experiencias conscientes deben tener las características de un proceso autotransformativo. En la medida en que esas partes experienciales son comunes a la experiencia consciente, el grado de generalización de este argumento será mayor.

De acuerdo con el argumento de intrinsicidad, si una parte de una experiencia consciente que tiene un contenido primario cumple el criterio de particularidad (a saber, que es de ese contenido primario particular y no otro) y el criterio de independencia (a saber, pero que no depende del contenido primario de esa experiencia), entonces (paso 1) existen procesos específicos a la experiencia consciente que son intrínsecos, no preexistentes y forman parte de ella y (paso 2) estos procesos específicos que dan cuenta de dicha experiencia consciente deben tener características de autotransformatividad.

Este argumento hace algunas asunciones. En primer lugar, asume que la conciencia tiene ciertas propiedades (fenoménicas, dinámico-estructurales, funcionales y temporales). Asume, en concreto, que la experiencia consciente tiene partes experienciales. En segundo lugar, requiere de la noción de proceso y, por tanto, asume que existen elementos procesuales en el análisis de la experiencia consciente.[234] El objetivo del argumento es defender que esos elementos procesuales no son cualesquiera, sino que deben estar dotados de cierto aspecto autotransformativo.

Este argumento de intrinsicidad tiene la siguiente forma:

[233] En principio, no tiene por qué aplicar a todas las experiencias conscientes.
[234] Es decir, asume el argumento de las propiedades no fenoménicas.

Premisa: Supongamos una experiencia consciente con un contenido primario. Supongamos que esa experiencia consciente tiene una parte experiencial que cumple dos criterios:

- criterio de independencia (CI): esa parte no depende del contenido primario de la experiencia consciente (esto es, la parte acompaña a la experiencia *al margen de* si el contenido primario de esa experiencia es introspectivo o no, conceptual o no, visual, auditivo, acerca de una cosa u otra).

- criterio de particularidad (CP): esa parte es *particular* de ese contenido primario y no otro.

Paso 1: para que una parte de la experiencia consciente cumpla ambos criterios es necesario definir procesos específicos, que deben cumplir estas tres condiciones: (1) ser procesos intrínsecos de esta experiencia, en el sentido de no ser relacionales con el contenido primario de la experiencia, o dicho de otro modo, no consistir en representar nada que estuviese antes de que comenzase la experiencia consciente (por el criterio de independencia); (2) no tener preexistencia (por el criterio de particularidad) y (3) formar parte de la experiencia consciente que tiene ese contenido primario (por premisa establecida).

Paso 2: los procesos intrínsecos antes referidos, es decir, no relacionales con el contenido primario, sin preexistencia y con participación en la experiencia consciente tienen la composición con clausura de interferencia propia de los procesos autotransformativos. Los procesos intrínsecos se disparan por los mismos procesos que producen la experiencia consciente y, al mismo tiempo, interfieren en esa experiencia consciente por ser parte de ella.

Por consiguiente, los procesos que dan cuenta de cualquier experiencia consciente que tenga alguna parte experiencial que cumpla el criterio de independencia y el criterio de particularidad tienen características autotransformativas.

A esta instanciación de este argumento en la propiedad de la intimidad la denominaré argumento de intimidad y tiene el siguiente aspecto:

La intimidad es una parte de la experiencia consciente que cumple los criterios de independencia y particularidad.

Por consiguiente, por el argumento de intrinsicidad, los procesos que dan cuenta de toda experiencia consciente con intimidad tienen características autotransformativas.

Ahora voy a analizar uno por uno los pasos del argumento de intrinsicidad.

Premisa: parte experiencial particular e independiente

La experiencia consciente está formada por partes experienciales. Estas partes se pueden caracterizar fenomenológicamente de acuerdo a dos criterios: independencia y particularidad.

Primero introduciré el criterio de independencia. Para ello usaré la distinción entre objeto primario y objeto secundario de Brentano (1874):

> Podemos llamar objeto primario del oír al sonido y objeto secundario al oír mismo. Pues temporalmente se hacen presentes a la vez, aunque sea el sonido lo primero realmente. (Brentano 1874: 167/libro II, cap 2, §8) (Benítez 2005)

Si se acepta esta distinción, es posible distinguir entre dos tipos de partes experienciales en la experiencia consciente. Por un lado, las partes experienciales que dependen del contenido primario de la experiencia consciente.[235] Por otro lado, las partes experienciales que no dependen del contenido primario de la experiencia consciente. La rojidad (o redondidad) de un tomate es una parte experiencial de la experiencia consciente de ver un tomate que depende del contenido primario de esa experiencia. Sin embargo, la certeza de percibir (o la intimidad con que se percibe) no son partes experienciales que dependan del contenido primario de esa experiencia, sino del contenido secundario, es decir, del "percibir" o "tener" esa experiencia consciente.

Una parte experiencial cumplirá el criterio de independencia si esa parte experiencial no depende del contenido primario de la experiencia consciente. Esto es, si una parte experiencial cumple el criterio de independencia, entonces no depende de si la experiencia está representando un contenido introspectivo o no, un contenido visual, auditivo, sensorial, e igualmente no depende de si el contenido es conceptual o no, ni tampoco de su determinación particular. En otras palabras, esta parte no representa nada que fuese anterior al momento en que comenzó la experiencia consciente.

Ahora introduciré el criterio de particularidad. Este criterio se refiere al hecho de que una parte experiencial puede ser particular de un contenido primario en concreto o no. En general, se puede afirmar que toda parte experiencial que depende del contenido primario es particular de ese contenido. Por ejemplo, la rojidad del tomate es particular del contenido primario de la experiencia consciente consistente en percibir ese tomate. Sin embargo, en el caso de las partes experienciales que no dependen del contenido primario (esto es, que cumplen el criterio de independencia) hay más opciones. Algunas partes experienciales pueden ser particulares de ese contenido primario. Por ejemplo, la certeza con que se percibe el tomate es una certeza de esa percepción de ese tomate en concreto, y no otra. En cambio, existen otras partes experienciales que no dependen del contenido primario y, además, no son particulares de ese contenido, incluso a pesar de que formen parte de la experiencia consciente en su conjunto. Por ejemplo, la sensación de ansiedad o tristeza puede impregnar todo un conjunto de percepciones (entre ellas, la percepción

[235] Por ser más asépticos, denominamos contenido primario de la experiencia consciente a lo que Brentano llama objeto primario de la conciencia.

	¿Cumple CI?	¿Cumple CP?
Rojidad en la experiencia de percibir un tomate	No (depende del contenido, a saber, "tomate rojo")	Sí (es particular de esa experiencia —de tomate—, y no de otra)
Sensación de depresión, que impregna toda la experiencia de percepción visual del sujeto	Sí (no depende de lo que veo en cada momento)	No (no es particular de la percepción, impregna igualmente todas las percepciones)
Intimidad en la percepción visual (vivo el tomate como para mí)	Sí (no depende del contenido de la experiencia en cada momento)	Sí (es particular a la experiencia con ese contenido)
Certeza en la percepción visual (siento que estoy percibiendo el tomate)	Sí (no depende del contenido de la experiencia en cada momento)	Sí (es particular a la experiencia con ese contenido)

Cuadro 3.1: Algunos tipos de partes experienciales de la experiencia consciente de acuerdo al criterio de independencia (CI) y al criterio de particularidad (CP). Una parte experiencial cumple el criterio de independencia si no depende del contenido primario de esa experiencia. Una parte experiencia cumple el contenido de particularidad si es particular de ese contenido primario de la experiencia y no otro.

del tomate), sin que sea particular de cada una de esas percepciones (en concreto, sin ser particular de la percepción de ese tomate).

Una parte experiencial cumplirá el criterio de particularidad si esa parte experiencial es particular del contenido primario de esa experiencia consciente y no de ningún otro contenido.

Paso 1: para dar cuenta de la parte experiencial particular e independiente hacen falta procesos específicos

¿Cómo es una experiencia consciente con partes particulares e independientes? Denominaré X a los procesos que dan cuenta del contenido primario de la experiencia.

Por el criterio de particularidad es necesario establecer que los procesos que dan cuenta de esta parte de la conciencia no son preexistentes a X, puesto que deben ser particulares a ellos. Si fueran preexistentes, entonces podrían asociarse tanto a los procesos X como a otros procesos Y o Z posteriores y con un contenido primario completamente distinto. Por ejemplo, podrían ser parte de la experiencia consciente de ver un tomate rojo y, después, ser parte de la experiencia consciente de ver una manzana verde. De ese modo, perderían la particularidad, a saber, que son parte de esa experiencia concreta de ver el tomate rojo y no de otra experiencia.

Por consiguiente, por el criterio de particularidad, en condiciones normales,[236] los pro-

[236]Ver notas anteriores sobre los casos de facilitación de la experiencia consciente (Bachmann 2000: 122).

cesos de esta parte experiencial serán disparados por los procesos X de los cuales serán particulares.

Por otro lado, debido al criterio de independencia, los procesos de esta parte experiencial deben ser diferentes de los procesos X. Por consiguiente, deben ser no relacionales o intrínsecos.[237] Para definir independencia es preciso recurrir a la mereología no clásica de procesos. De acuerdo con el principio de suplementación fuerte, los procesos que dan cuenta de esa parte experiencial deben estar compuestos por partes de interferencia de primer nivel que, a su vez, no forman parte de los procesos X.

A esos procesos intrínsecos, no preexistentes y que forman parte de la experiencia consciente los denominaré temporalmente procesos "específicos" de la experiencia consciente.

Paso 2: estos procesos específicos tiene aspectos autotransformativos

Este paso consiste en establecer si estos los procesos específicos que dan lugar a la experiencia consciente, tal y como se han definido en el proceso anterior, tienen características autotransformativas. Dicho de otro modo, se trata de establecer que no existen procesos que, sin tener características autotransformativas, son capaces de dar cuenta por sí mismos de la experiencia consciente con partes experienciales que cumplen los criterios de particularidad e independencia.

Las tres propiedades composicionales de la autotransformatividad son composicionalidad, interferencia y clausura. A continuación voy a tratar de demostrar que estos procesos cumplen estas tres propiedades.

- Por intrinsicidad, los procesos que dan cuenta de la experiencia consciente constan, al menos, de los procesos X y los procesos intrínsecos, que difieren de X. Por tanto, existe cierta composicionalidad.
- Por no preexistencia, los procesos intrínsecos se disparan *en los procesos* que dan cuenta de la experiencia consciente. Eso quiere decir que no existirían *si no existiera* esa experiencia consciente. Por tanto, existe interferencia.
- Por parte experiencial, esos procesos intrínsecos interfieren, a su vez, en la experiencia consciente de manera particular. Son procesos distintos que influyen entre sí. Por tanto, existe clausura en la interferencia.

Por consiguiente, de acuerdo con este argumento, si hubiera una parte de la experiencia consciente que cumpliese los criterios de particularidad e independencia, los procesos de

[237] Es necesario recordar una vez más en qué sentido se usa aquí la noción de intrinsicidad. Lo intrínseco es lo no relacional. Carruthers (2005) usa la noción en sentido absoluto, como aquello que es absolutamente no relacional. Este no es el caso. En el caso de esta parte de la experiencia metacognitiva y periférica, intrínseco a la experiencia consciente significa no relacional con la intencionalidad del estado consciente al que acompaña, es decir, no relacional con respecto a la característica propia del estado consciente de ser acerca del contenido primario

esta experiencia tendrían las características de autotransformatividad.

El último paso es comprobar si la intimidad, en cuanto parte de la experiencia consciente, cumple estos criterios.

3.3.2.3. Instanciación en la intimidad

Que el carácter subjetivo forma parte de la experiencia consciente es una idea recurrente de la fenomenología. Independientemente de la fuerza o debilidad con la que se defienda, así como independientemente de su forma específica (*mí-idad*, *para-mi-idad*, perspectividad, inmunidad a error de identificación, etc.), el carácter subjetivo es un lugar común en la literatura.[238] Por otro lado, concuerda con la propia experiencia fenoménica de conciencia, la cual se asocia con cierta perspectiva, a saber, la perspectiva de primera persona. Este carácter subjetivo es lo que denominaré aquí "intimidad" de la experiencia consciente.

Las teorías de la conciencia se ven obligadas a explicar por qué o cómo sucede que los contenidos de la conciencia se presenten o se manifiesten con una perspectiva subjetiva, es decir, con una cierta intimidad. Esto implica dar cuenta de dos rasgos del carácter subjetivo con el que se presenta la experiencia consciente, en relación con los contenidos de esa experiencia consciente. El primer rasgo (1) consiste en que el carácter subjetivo sucede para todo tipo de contenidos conscientes y sin importar si son introspectivos o no, conceptuales o no, visuales, auditivos o intermodales, es decir, independientemente de su modalidad. La intimidad no depende de ese contenido. Denominaré a esto el rasgo de independencia (independencia del carácter subjetivo con respecto a los contenidos).

El segundo rasgo (2) es que, pese a (1), el carácter subjetivo de la conciencia no se puede separar de su contenido, en el sentido de que es un carácter subjetivo asociado a esos contenidos concretos que quien tiene la experiencia consciente vive como *para-mí* o como *míos* y no otros. Denominaré a esto el rasgo de particularidad (particularidad del carácter subjetivo con respecto a los contenidos).

El rasgo (1) asegura que la experiencia de ver un paisaje tiene un grado de intimidad de la misma manera que lo tiene la experiencia de sentir dolor y, por lo tanto, la explicación de ese grado de intimidad no puede residir en el hecho de que el contenido de esa experiencia sea más o menos externo, o sea de un tipo u otro. Por otro lado, el rasgo (2) asegura que la experiencia subjetiva, por ejemplo, la experiencia de ver como sujeto el paisaje o de sentir como sujeto el dolor, son experiencias de ese paisaje en concreto y no otro (para mí, como mía, desde cierta perspectiva) y de ese dolor en concreto y no de otro (para mí, como mío, desde cierta perspectiva).[239]

[238] Shoemaker (1968); Nagel (1974); Searle (1992); Crane (2000); Kriegel (2009); Zahavi (2005); Sebastián (2020).

[239] Que el dolor se presenta como con cierta perspectiva o cierta intimidad puede resultar paradójico. Sin embargo, existen casos clínicos (o bajo efecto de estupefacientes) donde el sujeto habla de experiencia de dolor sin intimidad, como si no le importase, o como si estuviese sucediendo a otro.

Cuando los rasgos (1) y (2) se aplican a las teorías de la conciencia, son dos criterios que debe cumplir todo proceso que sea definido como soporte de la propiedad de la intimidad.

Una de las grandes dificultades de la explicación de la intimidad reside, precisamente, en que es necesario unificar y compatibilizar estos dos criterios, a saber, independencia y particularidad para cada experiencia consciente y única. ¿Cómo puede haber algo que es independiente del contenido de los procesos conscientes pero que los acompaña de una manera particular a ellos?

Las soluciones son diversas, pero tienen ciertos puntos en común. Desde las teorías representacionales, el carácter subjetivo de la experiencia consciente se explica mediante procesos de representación específicos, esto es, mediante un "representar" con unas características especiales. Para cumplir el criterio de independencia, esta nueva representación debe ser diferente de la representación de los objetos previos, esto es, distinta de una representación de primer orden. Por otro lado, para cumplir con el criterio de particularidad debe ser un representar particular del representar de primer orden, esto es, que se produce precisamente en el momento de la conciencia y no como algo preexistente y, además, debe participar de la experiencia consciente. Rosenthal (2005) propone que debe tratarse de un nuevo estado representacional de nivel superior, lo cual conduce a una regresión infinita de estados, al menos si se quiere aceptar que el estado representacional superior también es consciente. Por este motivo, Kriegel (2009) propone que se trata del mismo estado original, el cual, además de representar lo exterior, está en un "representarse a sí mismo". Para explicar esa autorrepresentación Kriegel necesita defender la existencia de partes del estado mental que, a su vez, representan a otras partes. Este es el motivo por el cual el enfoque sinuoso hace la distinción entre $M*$ y $M\blacklozenge$.

Desde las teorías biológicas tales como la Teoría de la Doble Microgénesis se define un proceso posterior y específico de la conciencia, la microgénesis fenoménica, que cumple el criterio de independencia (como se deduce del hecho de que exista facilitación de la microgénesis fenoménica de contenidos representacionales a partir de otros contenidos distintos, tal y como se observa en los experimentos de metacontraste)[240] y al mismo tiempo cumple con el criterio de particularidad (ya que sucede de manera particular para cada contenido representacional).

Aunque las soluciones son diferentes, hay ciertos rasgos comunes. En general, para cumplir el criterio de independencia es necesario definir procesos específicos, esto es, que son intrínsecos (no relacionales con el contenido primario de la experiencia). Por otro lado, para cumplir el criterio de particularidad es necesario que estos procesos se produzcan en la experiencia consciente, es decir, que no sean preexistentes a ella y que se extingan con ella. Por último, estos procesos deben formar parte del conjunto de la experiencia consciente, en el sentido de que configuran de manera relevante cómo será esa experiencia consciente.

[240]La facilitación de la microgénesis fenoménica de unos contenidos por otros distintos y anteriores es un fenómeno muy reseñable que se observa en los experimentos basados en enmascaramiento por metacontraste. Cfr. Bachmann (2000: 122)

A los procesos que cumplen las condiciones anteriores, es decir, intrinsicidad, no preexistencia y participación, los denominaré procesos específicos de la experiencia consciente. La microgénesis fenoménica de la conciencia o la parte experiencial $M*$ dentro del enfoque sinuoso de la teoría autorrepresentacional se podría considerar como dos formas de instanciación de estos procesos específicos en dos teorías de la conciencia concretas.

¿Cómo se puede garantizar que existan procesos específicos a la experiencia consciente de este tipo? Este tipo de procesos, puesto que no son preexistentes, en condiciones normales, deben ser causados en la propia experiencia consciente, quizás en una etapa inicial, pero nunca antes que ella.[241] En otras palabras, su ignición es un producto de interferencia del proceso que determina el contenido primario de la experiencia, a pesar de que este nuevo proceso no interviene en la determinación de ese contenido primario de la experiencia consciente. Por otro lado, este proceso es parte de la experiencia consciente, y para que forme parte de la experiencia consciente es necesario que tenga cierta causalidad en el proceso global de la experiencia consciente. Funcionalmente debe ser relevante en el proceso consciente y, al mismo tiempo, originariamente deben provenir de ese mismo proceso consciente. Esta propiedad de provenir de un proceso y actuar al mismo tiempo en ese proceso tiene el aspecto de clausura, es decir, de un proceso que es producto (bajo unas determinadas circunstancias) de otros procesos iniciales al mismo tiempo que es capaz de interferir de manera relevante en estos procesos iniciales. Los procesos que dan cuenta de la intimidad tienen, pues, el aspecto de una parte de un proceso autotransformativo:

- Son parte del proceso global de la experiencia (existe composicionalidad).
- Son producto de una interferencia concreta de procesos (existe interferencia).
- La forma de interferencia es de ida y vuelta, esto es, son producto del proceso originario al tiempo que lo modifican (existe clausura de interferencia).

En consecuencia, la explicación de la sensación de intimidad que acompaña a la experiencia consciente requiere de procesos específicos que tienen características autotransformativas.

3.3.2.4. Comentarios sobre el argumento de intrinsicidad

El argumento de intrinsicidad relaciona determinadas partes de la experiencia consciente (a saber, aquellas que cumplen el criterio de particularidad e independencia) con un aspecto autotransformativo de los procesos que dan cuenta de ellas. El argumento de la intimidad, por su parte, instancia el argumento de intrinsicidad para la intimidad como parte de la experiencia consciente. Para finalizar esta parte, voy a hacer dos comentarios acerca de estos argumentos.

[241] De nuevo. Se añade esta coletilla de "en condiciones normales" para tener en cuenta que podrían suceder casos de facilitación como los indicados en Bachmann (2000: 122).

En primer lugar, el argumento de intrinsicidad introduce la noción de procesos intrínsecos, esto es, de procesos independientes del contenido primario de la experiencia consciente y, por consiguiente, no relacionales en términos de esa representacionalidad. Supongamos que se adopta para la explicación de la conciencia una postura representacionalista, del tipo "los estados conscientes son estados representacionales". Una consecuencia del argumento de intrinsicidad es que la experiencia consciente no puede ser explicada *única y exclusivamente* a través de su representacionalidad del mundo exterior. En otras palabras, el argumento no es compatible con una posición puramente externalista acerca de la experiencia consciente.[242]

En segundo lugar, el argumento de la intimidad suscita una cuestión adicional: ¿cómo se lleva a cabo la participación de esa parte experiencial de intimidad en la experiencia consciente? Esta cuestión lleva a una cierta reflexión acerca de la noción de modulación. Hasta el momento he establecido que la modulación matemática es una de las propiedades de los procesos autotransformativos y he ofrecido varios argumentos por los cuales es preciso pensar que todos o la mayoría de los procesos conscientes son procesos autotransformativos. Cabe la posibilidad, sin embargo, de que el enfoque autotransformativo se redujese a los sistemas dinámicos tales como los que conforman las conexiones cerebrales que se estudian en neurociencia, donde las propiedades fenoménicas no tienen relevancia, y donde la clausura de la modulación es exclusivamente matemática.[243]

Ahora, sin embargo, estoy en disposición de dar un paso más. El argumento de intrinsicidad nos enseña que los procesos específicos de la experiencia consciente son capaces de formar parte de la misma experiencia y, de algún modo, modificarla. Para explicar este hecho es necesario establecer que existe una modulación que no es estrictamente cuantitativa (que no es estrictamente matemática), sino de tipo cualitativo: la experiencia consciente participa en su propia variación a través de una parte de esa experiencia. La cuestión ahora es ¿es posible caracterizar fenoménicamente este tipo especial de modulación?

Comenzaré por definir una modulación cualitativa sencilla. Una experiencia consciente e_1 modula a otra experiencia consciente e_2 si e_1 tiene una función en el modo en que e_2 cambia. Esta modulación puede ser muy diversa. Para explicarla, por el momento, es preciso recurrir a términos algo metafóricos. Estas son algunas posibilidades:

- Puede atenuarlo. Por ejemplo: escuchar una música tranquila puede aliviar la sensación de ansiedad o dolor.[244]
- Puede hacerlo más controlable. Por ejemplo: algunos tipos de meditación pueden ayudar a controlar otros estados conscientes.[245]

[242]No es compatible, por ejemplo, con Dretske (1995) o Tye (1995).

[243]Por ejemplo, Izhikevich (2007).

[244]Cfr. Krishnaswamy y Nair (2016).

[245]Cfr. Eppley et al. (1989), donde se presenta un metaestudio comparativo acerca de la influencia de las distintas formas de relajación y meditación en la reducción de la ansiedad de rasgo (esto es, la ansiedad entendida como la tendencia a tener estados de ansiedad). Más en concreto, acerca de meditación trascendental la literatura es muy diversa, de distinta naturaleza y no siempre fiable. Véase

- Puede teñirlo de una tonalidad especial. Por ejemplo: la sensación de melancolía o de ansiedad parece teñir la experiencia consciente.
- Puede interferir en él hasta la extinción. Por ejemplo: el cubo de Necker o, en general, los estímulos visuales contrapuestos.

Los tres primeros casos de modulación están relacionados con partes de la experiencia que son compatibles entre ellas. Para que esto sea posible, una o las dos deben formar parte de la conciencia periférica. El segundo se corresponde con la naturaleza serial o agónica de la conciencia, y afecta principalmente a contenidos focales no compatibles entre sí.

Esta distinción permite diferenciar entre dos tipos de modulación por parte de una experiencia consciente. La primera es una modulación que extingue a otras partes experienciales. La segunda es una modulación que simplemente varía a otras partes experienciales sin extinguirlas. Este segundo tipo de modulación es relevante, puesto que permite una diversidad informacional que no es posible en el primer tipo. Para diferenciarlas, denominaré "modulación agónica" al primer tipo y "modulación no agónica" al segundo tipo.

El tipo de modulación que tiene lugar en el caso de la intimidad como parte experiencial es de tipo no agónico. Esto es, la parte experiencial que modula la experiencia consciente la tiñe sin modificar el contenido primario de la misma. El símil que se podría usar es, quizás, el de un filtro de color, que mantiene los rasgos representacionales de una imagen si bien baña todo su contenido con un aspecto experiencial. Esa experiencia de "tinción", se entiende ahora, cumple el criterio de independencia y particularidad. Es independiente del contenido primario, pero es particular a ese contenido y no a otro.

3.3.2.5. Conclusiones

En esta parte he dado argumentos en favor de la idea de que la noción de proceso y, en concreto, la noción de proceso autotransformativo, es necesaria para explicar la conciencia. Estos argumentos están basados en las propiedades fenoménicas de la conciencia. Los argumentos son dos: el argumento de intrinsicidad y su instanciación en la intimidad. ¿Por qué se hizo así? Para poder emplear este argumento en otros casos de partes experienciales de la conciencia en las cuales existan procesos específicos en el sentido descrito aquí, esto es, procesos no preexistentes, intrínsecos y que forman parte de la experiencia consciente.

A continuación, voy a usar el argumento de intrinsicidad para resolver problemas relativos a otros tipos de experiencia consciente. Dos de estos problemas se refieren a tipos de experiencia consciente que tienen que ver con ciertos aspectos metacognitivos, a saber, la sensación de certeza en la percepción visual y el efecto *punta-de-la-lengua*. A menudo, estos dos tipos de experiencia se han tratado de explicar exclusivamente por medio de estados representacionales, lo cual conduce a ciertos problemas, ya que ambos parecen

Travis y Pearson (2000) para un estudio basado en EEG (electroencefalograma) que carece de elementos de control.

suceder de manera independiente del contenido representacional, a pesar de que son particulares de ese contenido. El enfoque de procesos permitirá interpretar estas experiencias conscientes a través de procesos específicos de la experiencia consciente que tendrán las características de un proceso autotransformativo.

3.4. El enfoque de procesos en dos problemas concretos

El objeto de este capítulo es explicar la conciencia como un proceso. En la primera parte he tratado de justificar que el marco conceptual de los procesos es útil, válido y necesario para el estudio y análisis de la experiencia consciente. Ahora voy a usar este marco para resolver algunos problemas concretos. Me centraré en dos tipos de experiencia consciente, la experiencia de certeza en la percepción y la experiencia de inminencia de recuperación de ítem (o efecto *punta-de-la-lengua*). He escogido estas dos formas de la experiencia consciente, principalmente, por dos motivos. Primero, porque son muy representativas del tipo de problemas que se encuentran cuando se opta por un análisis basado en la noción de estado en lugar de usar un análisis basado en la noción de proceso. Segundo, porque existe la suficiente evidencia empírica acerca de ambas experiencias como para poder llegar a un nivel de determinación alto de los procesos que dan cuenta de ellas.

El método que usaré es el análisis de procesos. Dado que los procesos son familias de ocurrencias, el análisis de procesos pasa obligatoriamente por identificar cuáles son las partes de esos procesos y qué relaciones existen entre ellas. Para ello tenemos a nuestra disposición dos mereologías, una mereología clásica para distinguir las partes temporales y las partes espaciales y una mereología no clásica para distinguir las partes funcionalmente relevantes. La primera mereología me servirá para hacer un análisis de procesos en partes temporales y espaciales. La segunda mereología me permitirá hacer el análisis funcional de procesos. A modo introductorio, antes de comenzar con el análisis funcional, voy a definir una serie de criterios generales acerca de qué tipo de proceso puede constituir la experiencia consciente con respecto a sus partes temporales y espaciales.

3.4.1. Introducción: la conciencia como proceso en el espacio y en el tiempo

Como se ha visto, la filosofía de procesos nos ofrece un criterio para tipificar procesos de acuerdo con la relación que mantiene el todo con sus partes temporales y espaciales.[246] Esto es posible gracias a las nociones de automeridad y la homomeridad, tanto en su versión espacial como temporal. Ahora voy a tratar de tipificar la conciencia entendida como proceso de acuerdo con este criterio.

[246]Este criterio nos lo ofrece la Teoría de Procesos Generales (Seibt 2009)

Antes de tipificar la conciencia de modo general, es necesario volver al debate terminológico y precisar, una vez más, a qué me refiero con el término "conciencia". De entre los diferentes análisis, Kriegel (2004) ofrece una distinción bastante completa. Kriegel propone tres ejes de distinción de la conciencia. El primero es el eje criatura/estado y sirve para distinguir entre la conciencia de criatura, esto es, la conciencia predicada de una criatura o individuo, y la conciencia de estado, esto es, la conciencia predicada de un estado mental, esto es, un estado consciente. El segundo es el eje foco/periferia, que sirve para distinguir la conciencia focal o conciencia donde el contenido es el núcleo del estado consciente y conciencia periférica donde el contenido se encuentra en el margen de la conciencia. El tercero es el eje transitividad/intransitividad, que sirve para diferenciar entre la conciencia transitiva (cuando posee intencionalidad, es decir, es "sobre algo") de la conciencia intransitiva (esto es, sin contenido específico). La sensación de rojez al ver un tomate es una experiencia consciente de estado, focal y transitiva. La sensación de ansiedad puede ser una experiencia consciente de estado, periférica e intransitiva. La conciencia de criatura es aquello que se intenta medir a través de la escala de Glasgow. [247]

3.4.1.1. Partes temporales de la conciencia

Gracias a la distinción eje criatura/estado es posible diferenciar el tipo de automeridad y homomeridad temporal de la conciencia.

Si se entiende la conciencia como conciencia de criatura, esto es, como un estar consciente que se predica del individuo, entonces esa propiedad tiene homomeridad temporal alta, puesto que las partes temporales del individuo mantienen en todo momento la propiedad de ser consciente, por lo menos desde que el individuo se levanta hasta que se acuesta y duerme sin sueños, de acuerdo con la célebre definición de Searle (1999: 40) o, quizás, incluso cuando duerme sin sueños, de acuerdo con otros autores (Thompson 2015). La conciencia como estar consciente es una actividad.[248]

En cambio, si la conciencia se entiende como experiencia consciente, es decir, como un proceso particular el cual tiene (al menos, si es transitiva) un contenido concreto, entonces los procesos conscientes tienen una automeridad temporal más bien baja: no sucede que todas las partes temporales del proceso consciente son, por sí mismas, procesos conscientes. Por el contrario, el proceso tiene más bien el aspecto de un desarrollo, con un tiempo de ignición y un tiempo de desvanecimiento, en los cuales el proceso no es todavía consciente.[249] El tiempo de ignición de la experiencia consciente no es compatible con una noción del proceso consciente como actividad, donde cada parte temporal es también un proceso consciente. Por el contrario, durante el periodo de ignición no hay todavía experiencia consciente propiamente dicha. Igualmente, la serialidad de la conciencia no es compatible con una noción de proceso consciente como actividad, donde cada parte

[247] En medicina, la escala de coma de Glasgow ofrece una forma normalizada y procedimentada de medir el grado de conciencia de un paciente de acuerdo a tres criterios: respuesta ocular, respuesta verbal y respuesta motora.

[248] Es el tipo 1 de la clasificación de Seibt (2009). Véase cuadro 2.1.

[249] Esto es el tipo 4 de de la clasificación de Seibt (2009). De nuevo, véase cuadro 2.1.

temporal de la experiencia consciente es, a su vez, la experiencia consciente. Por el contrario, la serialidad de las experiencias conscientes implica que las experiencias conscientes compiten entre sí por el espacio temporal de la conciencia.

La idea de entender la conciencia como un desarrollo es recurrente en las teorías de la conciencia. Muchas de las teorías de la conciencia defienden que la conciencia se desarrolla a lo largo del tiempo, y se diferencian en la manera en que se define este proceso. Esto es especialmente claro en el ámbito de la conciencia visual. Pondré tres ejemplos. La Teoría de la Doble Microgénesis establece que hay dos grandes etapas en la producción de la conciencia visual: la microgénesis representacional, durante la cual no hay conciencia, y la microgénesis fenoménica, en la que se produce la experiencia consciente.[250] La Teoría del Espacio de Trabajo Global Neuronal, por su parte, establece que la conciencia visual se produce en tres fases, siendo las dos primeras no conscientes, mientras que la tercera es consciente y se asocia con procesamientos de reentrada de larga distancia.[251] Por último, la teoría basada en la memoria visual frágil de corto plazo defiende igualmente que la conciencia visual requiere de fases, si bien no son tres sino cuatro las etapas que componen este proceso. En este caso, solo las dos últimas fases se producen con conciencia y se asocian con procesamiento recurrente que, o bien se reduce a las zonas visuales, o bien sea extiende hasta áreas diversas del córtex cerebral a través de la acción del córtex prefrontal.[252]

3.4.1.2. Partes espaciales de la conciencia

La mereología espacial de los procesos conscientes tiene una fuerte dependencia de la teoría de la conciencia escogida, hasta el punto de que no es posible pronunciarse acerca de ella sin conocer cuáles son los mecanismos fisiológicos propios de los procesos conscientes. Por este motivo, no tiene sentido hablar de mereología espacial de la conciencia en teorías que no tienen un soporte fisiológico o neuronal bien definido.

Entre las teorías que sí lo tienen se encuentran las teorías biológicas. En este aspecto, las teorías biológicas de Koch o Bachmann, son ejemplares. Al margen de las diferencias que hay entre ellas acerca de cuáles son los mecanismos biológicos concretos que están a la base de los procesos conscientes, ambas aceptan de una manera u otra la idea de que hay correlatos neuronales de la conciencia, los cuales pueden ser de dos tipos: correlatos habilitadores de la conciencia (NCCe, *enabling*) y correlatos específicos de la conciencia (NCCs, *specific*). Los correlatos habilitadores son necesarios para que haya conciencia, y se suelen asociar con sistemas de activación en los que intervienen, entre otras áreas, la formación reticular mesencefálica y los sistemas no específicos del tálamo.[253] Los corre-

[250] Bachmann (2000)
[251] Dehaene et al. (2001)
[252] Lamme (2010)
[253] Pese a la coincidencia, es importante no confundir correlatos neuronales específicos de la conciencia, término acuñado por Koch (2004) que designa funcionalmente a aquellas partes y procesos que hacen falta para que la conciencia tenga un contenido concreto (sean cuales sean) con los sistemas específicos, término que, en neurociencia, designa a determinadas partes del tálamo (y sus procesos de activación asociados) que reciben la señal proveniente de los órganos de percepción (visual, auditivo, etc.).

latos específicos, por su parte, son los que hacen que la conciencia sea de determinada manera y tenga determinado contenido, y se suelen asociar con activación cortical y/o con los sistemas específicos del tálamo. Es posible distinguir entre ambos tipos de correlatos neuronales a la hora de evaluar la mereología espacial.

Los correlatos neuronales de la conciencia de tipo habilitador (NCCe) parecen poseer una automeridad espacial media/baja. Koch (2004) destaca dos zonas especialmente asociadas con los NCCe: la formación reticular mesencefálica, que es parte del tronco del encéfalo, y los núcleos talámicos no específicos, concretamente los núcleos intralaminares. Existe abundante literatura acerca de la criticidad de estas estructuras para la existencia de experiencia consciente.[254] Por el contrario, los correlatos neuronales de la conciencia de tipo específico (NCCs) presentan una mayor plasticidad, de tal modo que las lesiones en el córtex no evitan completamente la conciencia salvo que sean de gran extensión. Si los NCCe parecen poseer una automeridad espacial media/baja, los NCCs parecen poseer una automeridad espacial más bien media/alta. El mismo Koch lo confirma al indicar que "las personas pueden perder trozos grandes de su corteza cerebral sin una pérdida generalizada de consciencia. De hecho, unos cuantos centenares de pacientes viven razonablemente bien con un solo hemisferio cerebral. No obstante, ciertas lesiones bilaterales, comparativamente pequeñas, en los NIL [núcleos intralaminares] talámicos o en ciertas partes del tronco del encéfalo pueden eliminar por completo toda consciencia" (Koch 2004: 92). La cuestión de cuáles son los correlatos neuronales específicos de la experiencia consciente, no obstante, no está completamente cerrada. Únicamente centrándonos en la conciencia visual, la literatura ya diverge. Pollen (1995) defiende que el área V1 está involucrada en la conciencia visual al tiempo que minimiza el papel del córtex prefrontal. Otros, en cambio, defienden la relevancia del córtex prefrontal, y especialmente de la zona dorsolateral.[255] Dehaene (2014), por su parte, destaca las evidencias empíricas del papel del córtex parietal, que se añaden a las propias del córtex frontal. La cuestión acerca de la ubicación es la cuestión de la automeridad, pero es cierto que las evidencias no son todavía concluyentes y que estas evidencias se refieren, con frecuencia, a lesiones de muy amplio alcance (como en el caso del paciente DB de Weiskrantz (1986)), o bien a disminuciones del nivel de conciencia visual de tipo parcial, esto es, sin extinción completa de la experiencia consciente (como en Lau y Passingham (2006) y Rounis et al. (2010)). Por tanto, la cuestión acerca del nivel de automeridad espacial en la conciencia visual y, en general, de la conciencia, no está completamente cerrada.

El enfoque de procesos permite analizar la experiencia consciente a través de sus partes espaciales y temporales mediante una mereología clásica. Esta forma de análisis es importante, puesto que permite identificar cómo es la conciencia como proceso general. No obstante, no permite identificar los aspectos autotransformativos que puedan dar cuenta de las partes funcionales de la experiencia consciente ya que, en mereología clásica, el todo nunca es nada más que la mera adición de sus partes.[256] A continuación voy a tratar dos partes de la experiencia consciente que no son ni meramente espaciales ni meramente temporales, sino que son partes funcionalmente relevantes de esa experiencia consciente. La primera de ellas es la experiencia de certeza que forma parte de la experiencia de

[254] Cfr. Moruzzi y Magoun (1949) y Koch (2004).
[255] Lau y Passingham (2006); Rounis et al. (2010); Lau y Rosenthal (2011).
[256] Al menos, no de la manera en que se ha definido aquí la autotransformatividad.

percepción visual. La segunda es la experiencia de tener algo *en-la-punta-de-la-lengua*, que forma parte de la experiencia de tratar de recuperar un ítem de la memoria. La explicación de estas dos partes experienciales requiere de aspectos autotransformativos.

3.4.2. La conciencia de la certeza en la percepción

Sobre el papel, la metacognición y la conciencia no son la misma cosa. Por un lado, se puede ser plenamente consciente solo mirando el paisaje que nos rodea y ajeno a todo lo demás. Por otro lado, el hecho de que mi estado consciente sea sobre una percepción (por ejemplo, la sensación haber visto algo antes) o sobre un pensamiento (por ejemplo la sensación de que "algo no cuadra" aquí) no me hace, *a priori*, ser más consciente de ese contenido, igual que, *a priori*, no es más consciente un psicólogo que un físico por el mero hecho de que uno se ocupa de los estados internos y el otro de lo que sucede en el mundo exterior. Sin embargo, existe una idea que atraviesa las teorías representacionales y que cobra importancia en las teorías de orden superior y de mismo orden, de acuerdo con la cual la conciencia de un estado consiste precisamente en que el sujeto sea consciente de ese estado. ¿Cómo se hacen compatibles ambas ideas? ¿A qué especie de metacognición nos referimos cuando se dice que la conciencia se basa en un estado de orden superior (o de mismo orden)? Esta metacognición tiene algunos rasgos particulares: no pertenece, por lo general, al contenido focal del estado consciente. Es algo que suele estar, más bien, en la periferia de la experiencia consciente y, por decirlo en términos metafóricos, la envuelve. Aquello que envuelve a una experiencia puede ser muy variopinto: puede ser una sensación de certidumbre o incertidumbre, puede ser una intuición de ir tras la pista correcta o de estar dando círculos sin avanzar, puede ser la sensación de que el pensamiento está inacabado o la sensación de que todo está completo y en orden, puede ser la sensación de estar a punto de recordar algo, etc. Todas estas experiencias pertenecen a la periferia de la conciencia, y se podría pensar que no forman el núcleo duro de la explicación de la experiencia consciente, es decir, aquello que es necesario explicar en primer lugar si se quiere saber en qué consiste una experiencia consciente. No son la sensación de rojo al ver un tomate, ni la sensación de tener un orgasmo, ni tampoco la sensación de dolor al pincharse el dedo con una aguja. Sin embargo, forman parte de la diversidad de la conciencia y, *de facto*, a veces se han asociado con su parte más emblemática, esto es, con el carácter subjetivo de la conciencia.[257]

A continuación voy a analizar un ejemplo paradigmático de este tipo de conciencia, la conciencia periférica y metacognitiva acerca de la certeza de estar percibiendo visualmente un estímulo. Para ello, tomaré como referencia dos estudios empíricos en neurociencia. Se trata de los trabajos de Lau y Passingham (2006) y Rounis et al. (2010). Ambos tratan de reproducir estados liminales de conciencia visual en tareas de discriminación visual realizadas en situación de enmascaramiento por metacontraste. El objetivo de ambos trabajos es disociar la tarea de discriminación visual —por ejemplo, distinguir una forma de cuadrado de una forma de diamante— de la experiencia metacognitiva de certidumbre en la respuesta —es decir, cómo de seguro está el sujeto de que su respuesta es la correcta—. A través de esta disociación se tratará de buscar correlatos neuronales específicos para

[257] Cfr. Lau y Rosenthal (2011).

esa experiencia consciente de tipo metacognitivo. La conclusión a la que llegan ambos estudios es que estos correlatos están relacionados con determinados procesos de activación en el córtex prefrontal.

Para el análisis de esta forma de experiencia consciente procederé de la siguiente manera. Primero expondré brevemente los dos experimentos. Para ello explicaré previamente una distinción general de relevancia para la labor experimental, a saber, la distinción entre criterio metacognitivo y criterio puro de tarea. Después compararé los dos estudios. Una vez suficientemente descrita la experiencia consciente que es objeto de estos dos estudios, a saber, estados conscientes periféricos y metacognitivos, trataré de aplicar los conceptos de análisis de proceso a ese tipo de estados, para lo cual usaré una instanciación del argumento de intrinsicidad. Por último, aplicaré la noción de modulación cualitativa no agónica para explicar la fenomenología concreta de la experiencia de percepción visual con certeza. Con todo esto obtendré un marco de referencia procesual para los estados objeto de estudio de esos dos trabajos experimentales con el cual se puedan establecer unas pautas adecuadas de interpretación de los mismos.

3.4.2.1. Distinción entre criterio metacognitivo y criterio puro de tarea

Existe un problema metodológico para determinar si hay conciencia o no. Este problema tiene diferentes vertientes, una de las cuales viene de la denominada "hipótesis de acceso". Recordemos el enunciado de esta hipótesis, que habla acerca de un estado S que pertenece a un sujeto y tiene un contenido p:

> Hipótesis de acceso: S es un estado consciente solo si su contenido p es 'directamente' accesible por el sujeto (es decir, si es utilizable sin necesidad de procesamiento adicional) para llevar a cabo un amplio repertorio de tareas cognitivas tales como reportar p, razonar de acuerdo con p o actuar en base a p. (Phillips 2018: 1)

Como criterio para establecer si existe o no conciencia, la hipótesis de acceso adolece fundamentalmente de dos problemas. Primero, es una hipótesis que se basa en la necesidad, es decir, asegura que, si no hay acceso, no hay conciencia. Pero no es una hipótesis de suficiencia, es decir, no asegura que, si hay acceso, haya conciencia. Podría haber estados mentales no conscientes que fueran accesibles directamente (por ejemplo, mi representación no consciente de la farola de esta calle me permite esquivar la farola mientras voy caminando). Por tanto, no sirve para asegurar que hay estado consciente. Segundo, la hipótesis contempla tres posibilidades de acceso: acceso para reportar, acceso para razonar y acceso para actuar. Por tanto, no define exactamente cuál de esos criterios se ha de tomar, o si debe ser más de uno el que se adopten.

El primer problema es complejo de resolver. Si la accesibilidad por parte del sujeto no es suficiente para asegurar que haya conciencia, no se pueden determinar cuáles son los correlatos neuronales específicos de la conciencia. Por otra parte, si no se conocen cuáles son los correlatos neuronales específicos de la conciencia, no es posible determinar si hay conciencia al margen de su accesibilidad por parte del sujeto. Sin embargo, la urgencia

empírica obliga. A menudo, a falta de otros criterios neurológicos más fiables y medibles de manera sencilla, la urgencia en un trabajo de campo de distinguir si existe o no experiencia consciente hace que, en determinadas condiciones, se tome la accesibilidad como un criterio suficiente para establecer la existencia de experiencia consciente. En otras palabras, en la práctica, a menudo la hipótesis de acceso se usa como una hipótesis de necesidad y suficiencia.[258]

El segundo problema es también complejo y expresa una disyuntiva metodológica: ¿qué tipo de accesibilidad se debe escoger como criterio de existencia de experiencia consciente en cada caso?

Tradicionalmente se han distinguido dos formas de verificar si hay experiencia consciente.[259] La primera es el criterio (de acceso) metacognitivo. De acuerdo con este criterio, la experiencia consciente se verifica si existe un reporte del sujeto de ser consciente de una o otra manera. La segunda es el criterio puro de tarea, de acuerdo con el cual la experiencia consciente se verifica si el sujeto es capaz de llevar a cabo ciertas tareas concretas de discriminación por encima de lo aleatorio. Ambos criterios pueden ser ajustados y mejorados mediante procedimientos estadísticos propios de Teoría de Detección de Señal para reducir o eliminar el sesgo subjetivo.[260] Estos dos criterios también se pueden usar de forma combinada.

El primer criterio es más subjetivo y se relaciona con la accesibilidad para el reporte. El segundo es más objetivo y se relaciona con la accesibilidad para la acción. La adopción de uno u otro criterio lleva a resultados distintos y es todavía objeto de controversia. De manera muy amplia y siempre teniendo en cuenta el riesgo de generalización que la afirmación que sigue conlleva, quienes han adoptado posiciones cercanas a las teorías de orden superior o mismo orden se han inclinado por el criterio metacognitivo que, combinado con el criterio de tarea en determinadas circunstancias, permite disociar ambos resultados y, de ese modo, permite también identificar correlatos neuronales de la conciencia específicos de los reportes metacognitivos.[261] Por el contrario, quienes no han adoptado posiciones tan marcadas con respecto al orden de la conciencia han optado a menudo por evidencias empíricas apoyadas en el criterio puro de tarea.[262] En la actualidad no existe consenso acerca de qué opción tomar, lo que conduce a *divergencias empíricas* en los resultados y a *divergencias teóricas* en las interpretaciones.[263] Por ejemplo, el primer criterio parece llevar a correlatos neuronales de la conciencia visual basados en la actividad en el córtex prefrontal, mientras que el segundo criterio conduce a correlatos neuronales más amplios que incluyen también el córtex parietal.

[258]Este es un viejo problema metodológico en el trabajo empírico. Existen, sin embargo, intentos de discriminar la conciencia por métodos alternativos al de la accesibilidad, a menudo inspirados en la Teoría de la Información Integrada. Ese es el caso del índice de conciencia que se propone en Casali et al. (2013).

[259]Boly et al. (2013: 5) y Phillips (2018: 5).

[260]La Teoría de Detección de Señal (*Signal Detection Theory*, SDT) define una serie de procedimientos matemáticos para analizar los datos provenientes de tareas de categorización de estímulos ambiguos en psicología. En particular, en SDT no solo se tienen en cuenta los resultados positivos, sino también los falsos positivos.

[261]Ver Lau y Rosenthal (2011).

[262]Ver Dehaene (2014).

[263]Boly et al. (2013).

Al margen de cuestiones que no voy a poder dirimir aquí, se puede llegar a una conclusión: los dos criterios parten de asunciones distintas para llegar a resultados (en parte, al menos) distintos. De ahí que, antes de analizar e interpretar cualquier evidencia empírica, resulte imprescindible identificar meridianamente cuál de los dos criterios se está adoptando.

3.4.2.2. Dos estudios empíricos

Una vez establecida la distinción anterior, me voy a centrar ahora en un tipo concreto de conciencia, a saber, la conciencia periférica y metacognitiva que presenta el sujeto en relación con la certeza acerca de su percepción. Para algunos autores este tipo de conciencia resulta especialmente interesante. Kriegel (2009), Rosenthal (2005) y Lau y Rosenthal (2011), por ejemplo, la asocian con el carácter subjetivo propio de la experiencia consciente, es decir, precisamente con aquello que caracteriza todos los estados conscientes. Sin dar este paso, si algo parece fuera de toda duda es que esta conciencia constituye, por sí misma, una parte relevante de la experiencia consciente visual.

Desde el punto de vista empírico, se pueden destacar dos estudios que se han centrado específicamente en esta conciencia periférica y metacognitiva en el ámbito de la percepción visual. Me refiero a los trabajos de Lau y Passingham (2006) y Rounis et al. (2010). Lau y Passingham (2006) identifican cierta actividad en la zona dorsolateral del córtex prefrontal como un correlato neuronal de esta conciencia a través de resonancia magnética funcional (fRMI). Por su parte, Rounis et al. (2010) confirman el papel de esa actividad a través de estimulación transmagnética craneal (TMS).

Los dos trabajos coinciden en adoptar el criterio metacognitivo, basado en el reporte del sujeto. En una situación de presentación de un estímulo visual objetivo enmascarado mediante metacontraste, se solicita al sujeto dos cosas. En primer lugar, que realice una tarea discriminativa acerca de la forma de ese estímulo visual objetivo, que puede ser un cuadrado o un diamante. En segundo lugar, que ofrezca una evaluación subjetiva de la certeza que tiene acerca de la tarea discriminativa que acaba de realizar. Esta evaluación subjetiva y metacognitiva puede tener dos valores. Si el sujeto cree que ha visto el estímulo marcará "*seen*" o "*clear*". Si el sujeto cree que estaba probando al azar, marcará "*guess*" o "*unclear*". El análisis de los correlatos neuronales se basa en la disociación de ambos resultados (discriminativo y metacognitivo), para distinguir la experiencia consciente metacognitiva de la mera tarea de discriminación. Por tanto, la conciencia que se aísla empíricamente es de carácter metacognitivo. Por otro lado, a la pregunta de si esa conciencia es focal o periférica, la respuesta más plausible es que se trata de una conciencia periférica, ya que se realiza en el transcurso de una tarea de discriminación de un estímulo visual que es el contenido focal de la experiencia y se suscita de manera indirecta mediante una interrogación explícita que fuerza la respuesta del sujeto.

En el cuadro 3.2 se presenta una comparación entre ambos experimentos.

	Lau y Passingham (2006)	Rounis et al. (2010)
Objeto	identificar correlatos neuronales específicos de conciencia metacognitiva en percepción visual mediante fRMI	identificar correlatos neuronales específicos de conciencia metacognitiva en percepción visual mediante alteración basada en estimulación TMS tipo "ráfaga theta"
Condiciones	fijación, respuesta háptica (teclado), sujetos sanos despiertos	fijación, respuesta háptica (teclado), sujetos sanos despiertos
Tarea de discriminación (estímulo objetivo S_1)	formas S_1: diamante vs. cuadrado (33 ms)	formas S_1: cuadrado-diamante vs. diamante-cuadrado (33 ms)
Máscara (estímulo S_2)	metacontraste S_2: máscara en la zona exterior al contorno diamante/cuadrado (50 ms)	metacontraste S_2: máscara en la zona exterior al contorno cuadrado-diamante/diamante-cuadrado (50 ms)
SOA[264]	variable (de -50 a 133 ms)	Fijo: 100 ms
Criterio metacognitivo	"Los voluntarios tuvieron que realizar juicios de discriminación con elección forzada acerca de sí el objetivo era un cuadrado o un diamante. El nivel de conciencia se estableció inmediatamente después del juicio discriminativo, solicitándoles que presionaran teclas para indicar si realmente habían visto el objetivo o simplemente estaban probando suerte" (Lau y Passingham 2006: 18763).	"Solicitamos a los voluntarios que realizarán una tarea de discriminación visual con elección forzada entre dos alternativas, para identificar la disposición espacial de dos estímulos (un diamante y un cuadrado) [...] Al mismo tiempo, los voluntarios tenían que categorizar la visibilidad subjetiva del estímulo (claro/no claro)" (Rounis et al. 2010: 166).
Análisis de datos	8 participantes en la primera etapa, 13 participantes en la segunda etapa. Resultados promediados. (*two-tailed Wilcoxon signed-rank test*). Se incluye cálculo de d' (Teoría de Detección de Señal) únicamente para el juicio de discriminación.	20 participantes. Resultados promediados con ANOVA. Se incluye Teoría de Detección de Señal para la tarea de discriminación (d') y para la respuesta metacognitiva (meta-d')

[264] *Stimulus-onset asynchrony*: es el tiempo transcurrido entre el inicio del primer estímulo S1 y el inicio del segundo estímulo S2. SOA $= t(S_2) - t(S_1)$.

Metodología y resultados	Primera etapa (estudio conductual): variación de SOA. Se identifican dos valores de SOA para los cuales el rendimiento en la tarea de discriminación era prácticamente idéntico, mientras que el juicio metacognitivo era significativamente distinto: SOA corto: 33 ms (68,27 % correcto y 49,17 % visto) y SOA largo (70,22 % correcto y 60,13 % visto). Segunda etapa (estudio fRMI): se comparan los resultados fRMI para las situaciones de SOA corto y SOA largo. Se detecta correlación entre el juicio metacognitivo y la actividad en el córtex prefrontal dorsolateral.	Doble condición: TMS falsa y TMS real. Se usa estimulación TMS mediante ráfagas, que permite efectos a largo plazo (20 m). La estimulación es bilateral y dirigida al córtex prefrontal dorsolateral. Se mide el coeficiente de correlación phi, que es la relación entre la discriminación de estímulos y la visibilidad de los mismos. TMS reduce significativamente phi. Se mide también la relación meta-d'-d', esto es, la divergencia de la sensibilidad metacognitiva óptica. TMS afecta negativamente a esa sensibilidad de manera significativa.
Conclusión principal	"Nuestros resultados sugieren que el córtex prefrontal es importante para los aspectos subjetivos de la percepción consciente" (Lau y Passingham 2006: 18763). "En condiciones de SOA largo existe un nivel de percatación subjetiva significativamente superior al que hay en condiciones de SOA corto, lo que está asociado con un incremento significativo de la actividad en el córtex prefrontal dorsolateral medio." (Lau y Passingham 2006: 18765).	"Nuestros resultados muestran que la estimulación TMS de ráfaga theta aplicada al córtex prefrontal dorsolateral bilateral puede reducir la sensibilidad metacognitiva, es decir, la eficacia con la que el sujeto distingue entre juicios correctos e incorrectos mediante categorizaciones subjetivas de visibilidad [...] El efecto de TMS es específico de la sensibilidad metacognitiva; TMS no afecta al rendimiento de clasificación de estímulos" (Rounis et al. 2010: 175).

Cuadro 3.2: Comparación de dos estudios empíricos acerca de la certeza de la percepción visual: Lau y Passingham (2006) y Rounis et al. (2010).

El criterio escogido por ambos estudios es idéntico (criterio metacognitivo), por lo que es de esperar que la experiencia consciente objeto de estudio sea la misma, a saber, una experiencia relacionada con la metacognición, periférica y asociada a la percepción visual. Sin embargo, las conclusiones del primer estudio son más fuertes que las del segundo, ya que asimilan los correlatos neuronales de esa experiencia a "los aspectos subjetivos de la percepción consciente". Esto se debe a que la asunción teórica de Lau es que la conciencia subjetiva es precisamente metacognitiva.[265] Rounis y otros, en cambio, se muestran más cautos con respecto a ese aspecto subjetivo y hablan simplemente de sensibilidad metacognitiva para distinguir los juicios correctos de los incorrectos.

Por el momento, mi posición aquí se acerca más a la del segundo estudio. Los estudios anteriores tratan de replicar un carácter específico de la conciencia, que no agota la diversidad de la experiencia consciente. Podría haber, por ejemplo, estados subjetivos que

[265] Véase Lau y Rosenthal (2011).

acompañaran justamente a escenarios de fracaso cognitivo, tales como el efecto *punta-de-la-lengua*, el efecto Zeigarnik o la sensación de no estar tras la pista buena. Este tipo de estados conscientes tienen también un aspecto subjetivo y son diferentes del estado del sujeto acerca de si ha visto o no determinado estímulo.

3.4.2.3. Autotransformatividad en la conciencia de certeza de percepción

Ahora voy a hacer es un esfuerzo adicional de interpretación de la experiencia consciente metacognitiva y periférica que es objeto de estudio en los dos trabajos anteriores. Para ello, me apoyaré en el marco conceptual de procesos que he presentado durante este capítulo. El primer paso consistirá en hacer una nueva instanciación del argumento de intrinsicidad, demostrando que se puede aplicar igualmente a experiencias conscientes como las anteriores, a saber, experiencias metacognitivas y periféricas. El siguiente paso consistirá en emplear la noción de modulación cualitativa para distinguir entre procesos conscientes metacognitivos periféricos y estados mentales no conscientes. Presumiblemente, los primeros formarán parte de procesos con modulación cualitativa, mientras que los segundos, no.

Si recordamos, el argumento de intrinsicidad define dos criterios para establecer si los procesos que dan cuenta de una parte de la experiencia consciente son específicos y, por tanto, si tienen un aspecto autotransformativo. Estos son los criterios de independencia y particularidad:

> Supongamos una experiencia consciente con un contenido primario, la cual tiene una parte experiencial. Esta parte experiencial puede cumplir dos criterios:
>
> - criterio de independencia (CI): la parte no depende del contenido primario de la experiencia consciente (esto es, la parte acompaña a la experiencia *al margen de* si el contenido primario de esa experiencia es introspectivo o no, conceptual o no, visual, auditivo, acerca de una cosa u otra).
>
> - criterio de particularidad (CP): la parte es *particular* de ese contenido primario y no otro.

Voy a verificar el cumplimiento del criterio de independencia y del criterio de particularidad en el caso de la experiencia consciente de certeza en la percepción visual.

Criterio de independencia. La experiencia de certeza en la percepción visual del cuadrado y el diamante es una parte de la experiencia de percepción visual que tiene una característica específica. En el caso de Lau y Passingham (2006), se produce independientemente de que la percepción visual sea de cuadrado (en los casos en que el estímulo es un cuadrado) o de diamante (en los casos en que el estímulo es un diamante). En el caso de Rounis et al. (2010), la experiencia de certeza se produce independientemente de que la percepción sea de una disposición cuadrado-diamante (en el caso de estímulo cuadrado-diamante) o de una disposición diamante-cuadrado (en el caso de estímulo diamante-cuadrado). Ninguno

de los dos estudios hace una distinción de los correlatos específicos de acuerdo con esa distinción representacional. Por consiguiente, los correlatos específicos de la experiencia consciente de certeza en la percepción visual son independientes del contenido primario de la experiencia consciente.

Criterio de particularidad. Sin embargo, la experiencia de certeza en la percepción visual del cuadrado o del diamante son experiencias de cada estímulo concreto, en el sentido de que se refieren a la exposición concreta de dicho estímulo y se mide justo al mismo tiempo que la tarea de discriminación de manera háptica. La experiencia de mayor o menor certeza sobre la percepción es concreta y particular para cada una de las experiencias de percepción.

En consecuencia, los procesos que dan lugar a esa parte de la experiencia concreta, a saber, la parte experiencial de la certeza de percibir deben ser intrínsecos, es decir, no son relacionales con respecto al contenido primario de esa experiencia y, al mismo tiempo, no son preexistentes, sino que se disparan a partir de esa experiencia concreta y se disuelven con ella. Igualmente, estos procesos constituyen parte de la experiencia consciente puesto que son reportados y vividos por el sujeto como tales. Esto implica un aspecto autotransformativo en los procesos que dan cuenta de la experiencia de percepción visual que se suscita en ambos casos.

3.4.2.4. Modulación no agónica en la fenomenología de la conciencia de certeza en la percepción

Uno de los problemas con respecto a la fenomenología de la experiencia consciente de certeza en la percepción consiste en que debe ser posible establecer distinciones fenomenológicas entre (1) la experiencia de percibir visualmente sin tener la completa certeza de que se ha percibido, por ejemplo, en los casos de estímulos débiles o difusos, (2) la experiencia de tener confianza o certeza en uno mismo y sus percepciones en todo momento y (3) la experiencia de tener confianza o certeza sobre una experiencia en concreto, que es distinta de una experiencia general de certeza o de una experiencia de percepción sin completa certeza.

La mera suma de (1) y (2) no parece suficiente para explicar (3). Por ejemplo: puede suceder que quien tiene mayor confianza o certeza en sí mismo es porque sabe diferenciar correctamente cuándo tiene certeza sobre lo que percibe y cuándo no tiene esa certeza. En otras palabras, es capaz de distinguir con claridad cuándo debe fiarse de su percepción y cuándo no. La seguridad en uno mismo y la certeza acerca de una percepción concreta son cosas distintas.

En definitiva, es necesario estipular una relación entre las diferentes partes experienciales que componen (3) más allá de la suma mereológica. Denominaré a las dos partes de (3) PE_1 y PE_2. Por ende, los procesos que dan cuenta de PE_1 y PE_2 no pueden ser procesos completamente separados, sino que deben interferir entre sí, y esa interferencia no debe

ser destructiva, en el sentido de que debe ser compatible y estable, al menos durante el tiempo que se prolonga la experiencia de percepción visual. La aceptación de procesos autotransformativos permite estipular procesos de modificación de la experiencia consciente por sí misma en los cuales mantener el carácter cualitativo del contenido primario (PE_1) y añadir una parte experiencial correspondiente a la certeza de la percepción (PE_2) no sea incompatible. Por tanto, de los dos tipos de modulación posible entre experiencias, el agónico y el no agónico, es necesario optar por una modulación no agónica, en la cual el carácter cualitativo que proviene del contenido primario de la experiencia consciente se vea modificado sin eliminar las distinciones propias de ese carácter. La modulación de la parte experiencial correspondiente a la certeza de la percepción es una modulación de grado, por usar términos metafóricos, a falta de mejores términos, que tiñe la experiencia consciente al mismo tiempo que es compatible con ella y es un producto de ella.

3.4.3. La experiencia *punta-de-la-lengua*

De entre las experiencias conscientes de carácter metacognitivo, el estado *punta-de-la-lengua* (TOT, *tip-of-the-tongue*) es una de las que más atención han recibido en la literatura científica. El estado TOT se define como la sensación de cercanía o inminencia en la recuperación de un ítem (una palabra, por lo general) que, sin embargo, en ese instante no puede ser recordado (Schwartz y Metcalfe 2014). Hoy en día, estos estados distan mucho de ser considerados algo anecdótico o carente de importancia. Por el contrario, la historia del estudio de los TOTs se remonta como mínimo hasta James. James los describe como el "espectro que rodea a la palabra" y produce un "hormigueo" por la cercanía del término que casi se puede tocar, pero no se alcanza a recordar (James 1931: 251). Pero lo que le otorga la categoría de fenómeno analizable y descriptible empíricamente es el meticuloso trabajo de Brown y Mcneill (1966), en el que, a través del método de *prospección*, se recopilan y describen las múltiples aristas del fenómeno, tales como la recuperabilidad parcial de rasgos fonológicos del término (número de sílabas, la letra por la que empieza, si es un nombre propio, el número de palabras, etc.). En la actualidad el número de estudios es considerable, especialmente conductuales, aunque también desde el ámbito de la neurociencia.[266]

¿Qué hace de los TOTs un fenómeno tan interesante? Hay más de un factor. Voy a destacar tres.[267] Por un lado, es un fenómeno universal, es decir, que se observa en sujetos sanos y enfermos y al margen de su edad, desde niños a ancianos, y en diferentes culturas e idiomas distintos, desde los hablados hasta las lenguas de signos. Por otro lado, es un fenómeno muy frecuente, que forma parte de la cotidianidad y que, además, se puede reproducir en el laboratorio de manera sencilla y controlada, a saber, simplemente a través de preguntas.[268] Por último, es un fenómeno fácilmente referenciable, a través de la recuperación de una palabra o término en concreto, por lo que es posible establecer criterios puramente discriminativos para su evaluación y análisis (como, por ejemplo,

[266]Ver Schwartz y Metcalfe (2014) para una recopilación de estudios acerca del estado TOT.

[267]Schwartz y Metcalfe (2014).

[268]Para darnos cuenta de la extraordinaria frecuencia de este fenómeno, basta ver cualquier concurso de preguntas de la televisión y comprobar cuántas veces se suscita el fenómeno en nosotros como espectadores o en el concursante mismo.

también sucedía en la conciencia de la certeza de haber percibido, y en oposición a otras experiencias conscientes más escurridizas tales como el *déjà vu* o el efecto Zeigarnik).

Desde el punto de vista fenoménico, los TOTs suponen el contrapunto de la conciencia metacognitiva y periférica de la certeza de la percepción. Como aquella, es también una experiencia consciente metacognitiva. Como aquella, es también una experiencia periférica, que acompaña a la experiencia focal de imposibilidad de recordar. Pero, al contrario de aquella, su contenido no es directamente, sino indirectamente accesible. Lo que el TOT indica es, precisamente, la posibilidad cercana, casi inminente, de acceder a un contenido que, por lo demás, permanece directamente inaccesible.

Ahora voy a presentar dos estudios acerca de los correlatos neuronales de los TOTs. Los estudios de este tipo no son muy abundantes y, desgraciadamente, carecen de la meticulosidad de los estudios que presenté acerca de la certeza de la percepción. A pesar de ello, permiten definir un marco genérico para el análisis de los estados TOT a través de la noción de proceso, que será aquello que nos ocupará después. Para este análisis, primero argumentaré que los TOTs tienen cierto carácter intrínseco propio del argumento de intrinsicidad que se presentó más arriba y, en consecuencia, tienen un aspecto autotransformativo. Para ello me valdré de las evidencias provenientes del estudio conductual de Schwartz (2008). Por último, analizaré el papel de la modulación cualitativa en el estado TOT.

3.4.3.1. Dos estudios acerca de los estados TOT

La mayoría de los estudios acerca de los estados TOT son de tipo conductual y los estudios desde la neurociencia son escasos. Maril et al. (2001) y Maril et al. (2005) son dos ejemplos en los que, mediante el criterio metacognitivo y basado en el reporte del sujeto, se trata de identificar algunos correlatos o, al menos, algunos marcadores neuronales de los estados TOT mediante medición basada en fRMI. Los estados TOTs son suscitados en los sujetos por medio de pistas que señalan, generalmente, un nombre propio, por ejemplo "Chinatown + director". Se emplea esta técnica por ser especialmente eficiente a la hora de suscitar el efecto *punta-de-la-lengua*. La principal diferencia entre ambos estudios es que el segundo incluye también la distinción entre estados TOT y estados FOK (*feeling-of-knowing* o juicio de reconocimiento). Un sujeto tiene un estado FOK si tiene la convicción de que, en caso de ver la respuesta en una lista o en un mapa, sería capaz de reconocerla. Los estados FOK pueden suceder acompañados o no de estados TOT. Es decir, es posible tener la convicción de que se podría reconocer la respuesta sin tener la sensación de que esta respuesta se vaya a recuperar de manera inminente.

	Maril et al. (2001)	Maril et al. (2005)
Objeto	identificar los correlatos neuronales específicos de TOT mediante fRMI	identificar correlatos neuronales específicos de TOT y FOK mediante fRMI
Tarea de recuperación	Basada en pistas mostradas sobre una pantalla, 4 s. de tiempo	Basada en pistas mostradas sobre una pantalla, 4 s. de tiempo

Diseño del experimento	"Hemos desarrollado un paradigma conductual que suscita estados TOT con una frecuencia comparable a la literatura anterior. Se presenta al sujeto dos pistas semánticas (e.g. Chinatown + director; Iraq + capital) que convergen en un ítem objetivo (e.g. Roman Polanski, Bagdag). La mayoría de los ítems objetivo son nombres propios [...]. Durante el escaneado, se les dio a los sujetos 4 segundos para indicar si conocían el objetivo, si no lo conocían y si experimentaban un TOT. Entonces realizamos análisis fRMI relacionado a evento para examinar la actividad neuronal durante el estado TOT en comparación con las situaciones de "conocer" y "no conocer" [...] En un test [...] posterior [...] se verificó el 93 % de las respuestas." (Maril et al. 2001: 654)	"Durante el experimento los sujetos respondieron en cada intento presionando una de cuatro posibles teclas de respuesta para indicar el resultado—recuperación correcta ("conocido"), recuperación no conseguida no acompañada de TOT o FOK ("no conocido"), recuperación no conseguida acompañada de TOT ("TOT") o recuperación no conseguida acompañada de FOK ("FOK"). Antes del experimento se les dieron instrucciones precisas para clarificar qué constituye un estado TOT, qué constituye un estado FOK y cuál es la diferencia entre ambos. Para verificar [...] las preguntas se les presentaron 24 horas después" (Maril et al. 2005: 1131)
Criterio metacognitivo	de reporte háptico (teclado)	de reporte háptico (teclado)
Análisis de datos	Media SEM. Sin Teoría de Detección de Señal	Media SEM. Sin Teoría de Detección de Señal
Resultados	"los intentos donde se suscitó TOT vinieron acompañados de una actividad significativamente mayor en el córtex frontal medio derecho [...] y en el córtex cingulado anterior [...]" (Maril et al. 2001: 655)	"La activación en el córtex cingulado anterior, el córtex prefrontal dorsolateral derecho, el córtex prefrontal inferior derecho y el córtex prefrontal anterior era específica de TOT. Las regiones medial y lateral del córtex parietal se activaron tanto en TOT como en FOK" (Maril et al. 2005: 1137)
Conclusión principal	"El aumento de activación en ACC y PFC asociado con TOT [...] sugiere que los sujetos necesitan monitorizar activamente los diferentes tipos de información que acuden a la mente durante el TOT, incluido el conocimiento contradictorio" (Maril et al. 2001: 658)	"El resultado de este estudio sugiere que diversas activaciones en las áreas frontales reflejan procesos que se asocian específicamente con TOT, más que con otros procesos metacognitivos más generales" (Maril et al. 2005: 1137)

Cuadro 3.3: Comparación de dos estudios empíricos acerca del efecto *punta-de-la-lengua*: Maril et al. (2001) y Maril et al. (2005).

Los estudios anteriores no son perfectos. Por ejemplo, carecen de control del sesgo subjetivo, no distinguen entre TOTs verdaderos (aquellos que realmente conducen al ítem

objetivo) de TOTs falsos (aquellos que conducen a otro ítem distinto del ítem objetivo) y obligan al sujeto a escoger entre TOTs y FOKs (sin poder escoger ambos). A pesar de ello, y con todas las reservas provenientes de estas limitaciones, los resultados parecen apuntar a algunas conclusiones generales:

- Los correlatos neuronales de los estados TOT se corresponden con zonas previamente identificadas con señales metacognitivas (Schwartz y Metcalfe 2014: 26).
- Hay correlatos específicos de los estados reportados como TOT que no se correlan con los estados reportados como FOK.
- Los estudios de neuroimagen no permiten identificar que haya mecanismos cognitivos específicamente relacionados con los estados TOT, esto es, no permiten indicar si estos estados se activan por recuperación parcial semántica, fonológica, por familiaridad de pistas, etc. (Schwartz y Metcalfe 2014: 25). Algunas de las teorías acerca de los TOTs sugieren que podría ser una suma de cosas.

3.4.3.2. Los estados TOT y FOK como proceso

Los estudios anteriores apuntan con cierta probabilidad a que procesos específicos de carácter metacognitivo dan cuenta de los estados TOTs. Ahora voy a analizar si estos procesos pueden caracterizarse un poco mejor a través de las evidencias conductuales que de ellos se tienen. Primero revisaré las dos teorías más importantes que describen la naturaleza de los estados TOTs, a saber, la Teoría del Acceso Directo y la Teoría Heurístico-Metacognitiva. Después, aprovecharé las evidencias que hay en favor de la segunda (Schwartz 2008) para interpretar los procesos propios de los estados TOTs como procesos intrínsecos que difieren de otros estados tales como los estados FOK.

Si se recopila la investigación acerca del fenómeno TOT desde Brown y Mcneill (1966) hasta la actualidad, se pueden distinguir dos teorías bien diferenciadas para explicar este fenómeno: la Teoría del Acceso Directo y la Teoría Heurístico-Metacognitiva.[269]

La Teoría del Acceso Directo (Gollan y Brown 2006) defiende que los estados TOT son causados por el acceso no consciente al ítem objetivo que debe ser recordado en una situación especial en la que algo impide que se produzca esa recuperación. En otras palabras, los estados TOT se caracterizan porque el ítem objetivo tiene un nivel de activación suficiente para disparar el estado TOT pero insuficiente para disparar el recuerdo de dicho ítem. En concreto, lo que sucede es que el nivel de activación es suficiente para recuperar la información semántica acerca del ítem y, quizás, alguna parte aislada de la información fonémica del ítem (tal como con qué letra empieza o si es una palabra compuesta o simple), pero insuficiente para recuperar completamente la información fonológica.

La Teoría Heurístico-Metacognitiva, por su parte, no se opone completamente a la Teoría del Acceso Directo, puesto que, como ella, acepta que ciertas partes de la información fonológica o semántica están accesibles durante el TOT, al menos de manera fragmentaria.

[269] Schwartz y Metcalfe (2014: 18)

Sin embargo, permanece agnóstica con respecto a las causas que impiden que se produzca la recuperación completa del ítem, y precisa que el estado TOT debe ser explicado únicamente a través de esas partes de información u otras semejantes, como resonancias o pistas asociadas, y no por un nivel de activación liminal. De acuerdo con esta teoría, cuando, por ejemplo, una persona trata infructuosamente de recordar una palabra y se inicia el proceso cognitivo para intentar recuperar ese ítem, al mismo tiempo se ponen en marcha también una serie de procesos metacognitivos cuya labor es, principalmente, recopilar información y notificar al sujeto si esa palabra será potencialmente recuperable o no. Los mecanismos que subyacen a estos procesos no están precisados, si bien Schwartz y Metcalfe (2014) proponen el diagrama funcional de la figura 3.6.

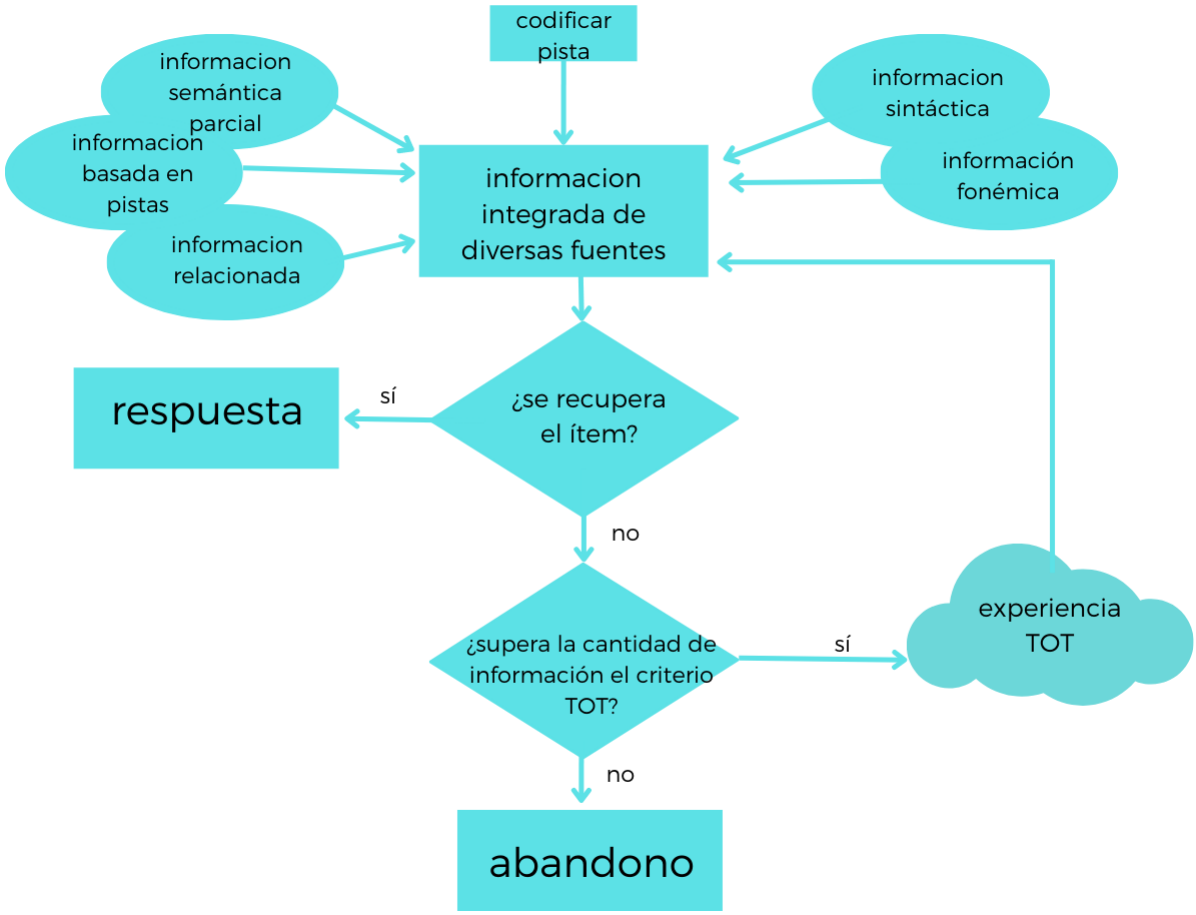


Figura 3.6: Esquema de cómo la monitorización heurístico-metacognitiva sopesa la información asociada al proceso fallido de recuperación y determina si se dispara o no un TOT. La experiencia TOT es, a su vez, una fuente más de información. Adaptado de Schwartz y Metcalfe (2014).

Entre estas dos teorías, la que mejor encaja con el análisis procesual es la segunda, puesto que explica el estado TOT como un proceso que se sostiene en el tiempo y se autorregula en función de la información recopilada acerca del ítem objetivo. Una de las evidencias empíricas más sólidas en su favor proviene de la disociación entre los estados TOTs y FOKs a través del efecto diferenciado que produce en ambos la carga en la memoria de trabajo (Schwartz 2008). Esta evidencia opera, además, en favor de la interpretación procesual que estoy defendiendo aquí, basada en procesos intrínsecos con un aspecto au-

totransformativo. Voy a presentarla para después ofrecer una interpretación procesual plausible.

Como se ha visto antes, TOT y FOK designan estados que giran alrededor del estado consciente de no poder recordar un ítem determinado, generalmente, una palabra. El término TOT (*tip-of-the-tongue*) se refiere, como se ha señalado ya, a la sensación que tiene el sujeto de que la palabra será recuperada de manera inminente (la cual, a su vez, puede producir un juicio de inminencia como "pronto me vendrá la palabra a la cabeza"). El término FOK (*feeling-of-knowing*), por su parte, se refiere a la sensación de que, si el sujeto viera la palabra entre otras en una lista, la reconocería sin mayores problemas (la cual, a su vez, produce un juicio de reconocimiento como "si la viera, la reconocería").

Schwartz diseñó cuatro experimentos, si bien la base de todos ellos era la misma. En la condición de carga de la memoria de trabajo, los sujetos recibían primero una serie de dígitos. Después se les mostraba una pregunta genérica (generalmente solicitando un nombre propio, para suscitar los TOTs; por ejemplo: "¿quién compuso Don Giovanni?"). Si la tarea de recuperación era fallida, es decir, si el sujeto no podía recordar la respuesta a la pregunta, entonces se le preguntaba alternativamente si había experimentado un estado TOT o un estado FOK. Finalmente, se le solicitaba que reportara los dígitos inicialmente mostrados. En la condición sin carga de la memoria de trabajo, no existía la tarea de recordar los dígitos.

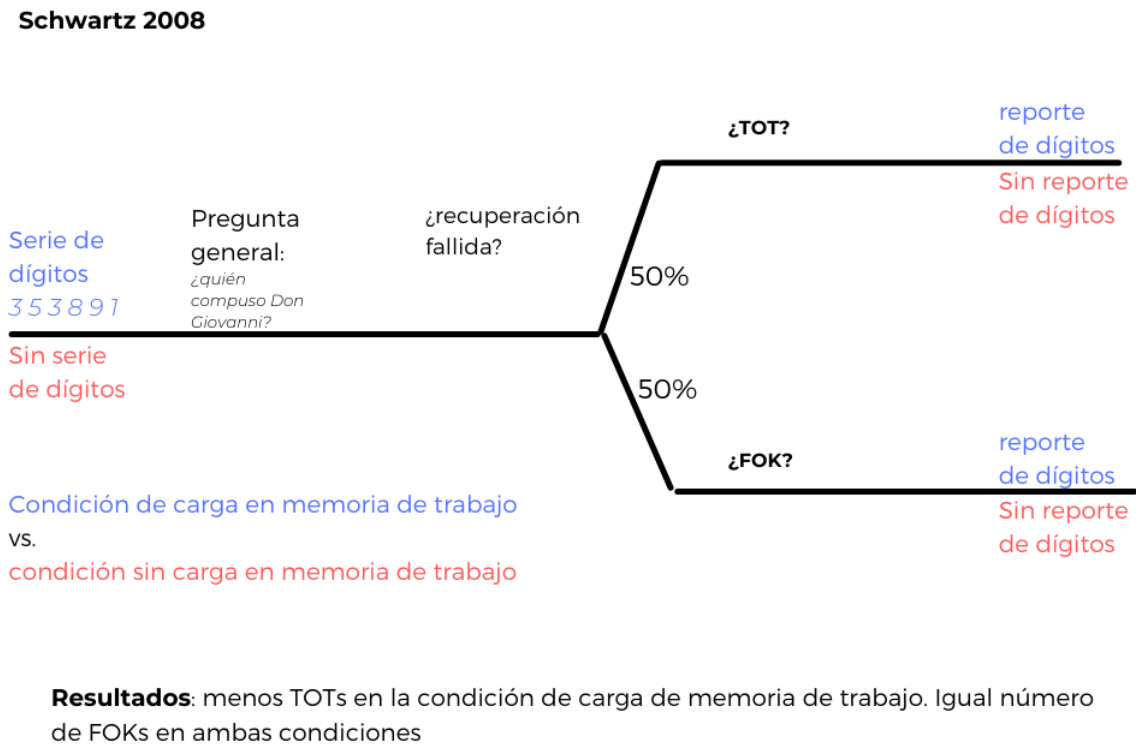


Figura 3.7: Diseño del experimento de Schwartz (2008), donde se evalúa el efecto que tiene la carga en la memoria de trabajo en la frecuencia de estados TOT (*tip-of-the-tongue*) y FOK (*feeling-of-knowing*).

El resultado de este experimento fue que el número de TOTs disminuyó levemente (5 %) pero de manera efectiva en la condición de carga en memoria de trabajo, mientras que el número de FOKs era similar en ambas condiciones. Schwartz extrae dos conclusiones principales: (1) que los procesos que subyacen a los juicios TOT y FOK son distintos (tal y como sugerían Maril et al. (2005)) y (2) que existe un efecto bidireccional entre el rendimiento de la memoria de trabajo y los juicios TOT. Además, Schwartz y Metcalfe (2014) destacan que estos resultados apoyan la Teoría Heurístico-Metacognitiva frente a la Teoría del Acceso Directo, ya que la primera establece procesos de autorregulación en los que la experiencia consciente propia de la sensación TOT tiene un papel importante y que, por consiguiente, se pueden ver interferidos en caso de carga de la memoria de trabajo, mientras que la segunda no otorga un papel tan importante a la experiencia consciente.

3.4.3.3. Autotransformatividad en los estados TOT

Los resultados del experimento de Schwartz (2008) son relevantes y, aunque no es cierto que solo admitan una única interpretación, apuntan fuertemente hacia la posibilidad de que los procesos que dan cuenta de los estados TOT sean de naturaleza heurístico-metacognitiva. Pero, ¿qué significa exactamente esto?

La Teoría Heurístico-Metacognitiva no se distingue, como ya se ha dicho, de la Teoría de Acceso Directo porque niegue que la información parcial (semántica, fonológica) acerca del ítem que hay que recuperar sea parte del estado TOT. Se distingue principalmente porque establece que el estado TOT se basa en procesos específicos en los cuales toma parte tanto esa información parcial como otras informaciones tales como las basadas en pistas, sintácticas, etc., que se recopilan *precisamente* durante el proceso de intentar recuperar el ítem en cuestión y no antes. Además, la Teoría Heurístico-Metacognitiva sugiere que la conciencia (fenoménica y de acceso) que se asocia con los estados TOT no es algo afuncional, sino que tiene un papel, a saber, regula el intento de recuperación, lo refuerza y advierte al sujeto de que esa recuperación será inminente.

Un argumento en favor de esta teoría es que, si hubiera pistas acerca de los propios procesos de recuperación —y no solo acerca del ítem en cuestión— que pueden ayudar a dicha recuperación, no tendría sentido que no formarán parte de los procesos que dan cuenta del estado TOT.

La interpretación entonces es casi directa. Los estados TOT y FOK son estados sobre un contenido primario que no está accesible. Pero son distintos. Los estados TOT monitorizan el mismo proceso de recuperación de ese ítem, el cual incluye información parcial asociada al ítem (semántica, sintáctica, fonológica) y otras pistas relacionadas con el propio proceso, mientras que los estados FOK no monitorizan, sino que simplemente reflejan un metaconocimiento acerca de la accesibilidad de dicho ítem. Los estados TOT requieren de procesos específicos en los cuales la conciencia es una parte activa y juega un papel decisivo, y esos procesos específicos autorregulan el proceso global de recuperación y lo prolongan, si es necesario. Los estados FOK, por el contrario, se basan en procesos que no están monitorizando la propia tarea de recuperación en ese momento dado. Han de

tomarse, más bien, como juicios metacognitivos independientes del estado en el que se encuentra el propio proceso de recuperación.

La parte experiencial correspondiente al estado TOT, al menos de acuerdo con la interpretación heurístico-metacognitiva, admite la instanciación en el argumento de intrinsicidad. Cumple el criterio de independencia, ya que esa parte experiencial se refiere a la cercanía de recuperación del ítem, al margen de qué ítem sea. De hecho, Schwartz y Metcalfe (2014: 25) descartan que se deban solo a una recuperación parcial de información, bien sea semántica, fonológica o de otro tipo, sino a algo así como una evaluación conjunta del proceso de recuperación en curso que no tiene por qué privilegiar algunas informaciones parciales sobre otras. Podría ser, por ejemplo, una sensación de familiaridad con el ítem, o una pista acerca de cómo recuperar esa información (por ejemplo, haber visto el mismo ítem recientemente).

Por otro lado, de acuerdo con la interpretación heurístico-metacognitiva, el estado TOT cumple también el criterio de particularidad, puesto que indica la cercanía de recuperación de ese ítem, y no otro. Incluso en el caso de TOTs falsos, esto es, aquellos TOTs que refieren a ítems que realmente no son la respuesta correcta, existe esa particularidad. El estado TOT cumple, además, que es parte de la experiencia consciente, puesto que viene dado por el estado consciente de fracaso en la recuperación de ese ítem, no sucede en ausencia de conciencia y es capaz de modificar esa experiencia consciente concreta. Por consiguiente, es necesario estipular cierta intrinsicidad en los procesos que dan lugar al estado TOT. Estos procesos intrínsecos son los propios de la monitorización del proceso de recuperación y están reflejados por los bucles de la figura 3.6. Estos procesos no son preexistentes, puesto que monitorizan específicamente el proceso de recuperación en su misma actividad de recuperar, en tiempo real. Por otra parte, estos procesos forman parte de la experiencia consciente, puesto que el estado TOT se caracteriza por una fenomenología (que James describió como un "hormigueo"). Por último, los estados TOT tienen cierta independencia del contenido primario de la experiencia consciente, en la medida en que son conciencia periférica, puesto que este ítem no puede ser recuperado (si bien es cierto que, gracias a ese estado se puede llegar a determinada información acerca del ítem una vez que el estado TOT pasa a ser focal).

Por tanto, para comprender los procesos que dan cuenta del estado TOT y, de ese modo, poder distinguirlos adecuadamente de otros procesos tales como los FOK, es necesario recurrir a su aspecto autotransformativo. Esto quiere decir que estos procesos se producen justamente a causa del proceso de recuperación del ítem y, al mismo tiempo, interfieren en ese proceso de manera positiva, mediante la monitorización e intensificación de ese proceso de acuerdo con los niveles de información recopilados. Según esta interpretación, este es precisamente el mecanismo que se estipula en la Teoría Heurístico-Metacognitiva.

3.4.3.4. Modulación no agónica en la fenomenología de los estados TOT

Entender los estados TOT como parte experiencial de procesos con cierto aspecto autotransformativo abre la puerta a analizar la fenomenología de los estados TOT como una

fenomenología en la que la modulación tiene un papel relevante.

Como parte de la experiencia consciente que acompaña al intento de recuperar un ítem de la memoria, los estados TOT son conciencia periférica, en el sentido de que acompañan a otra experiencia consciente focal, a saber, la de no poder recuperar dicho ítem. Por otro lado, puesto que la información que aportan es acerca de esa recuperación más que sobre el ítem en sí, son conciencia metacognitiva. Esta conciencia metacognitiva y periférica, además, cumple la nueva versión de la hipótesis de acceso, ya que acompaña a un contenido que solo es accesible indirectamente, a saber, el ítem.[270] La forma en la que el estado TOT modula la experiencia consciente a la que está asociado es, probablemente, cualitativa y no agónica. ¿Por qué cualitativa? Porque el estado TOT se basa precisamente en la conciencia y no en un estado representacional acerca del ítem ausente. ¿Por qué no agónica? Porque su función es precisamente mantener el proceso de recuperación de ese ítem. Por consiguiente, no tiene sentido que destruya o modifique significativamente el contenido primario de la experiencia consciente con la cual va asociado.

De algún modo, el estado TOT tiñe la experiencia consciente con la cual está ligado (la experiencia de no poder recuperar el ítem) de un aspecto específico que James (1931: 251) definió como "hormigueo". Este estado nos da una cierta información. Pero esa información es intrínseca, en el sentido de que no es sobre aquello que es el contenido de la representación, a saber, el ítem que se intenta recuperar. Por el contrario, es una información en tiempo real acerca del estado en el que se encuentra el proceso de recuperación.

3.5. Conclusiones

En este capítulo he tratado de defender la idea de que existe una relación relevante entre conciencia y proceso. Hay dos grandes razones para defender un enfoque procesual de la conciencia. La primera razón es que las teorías de la conciencia, al menos cuando se les exige un grado alto de determinación acerca de los mecanismos que subyacen a la experiencia consciente, recurren a nociones procesuales con ciertos aspectos autotransformativos. Esto sucede de manera implícita en la Teoría del Espacio de Trabajo Global, donde los procesos neurofisiológicos son recurrentes y en la Teoría de la Información Integrada, la cual usa sistemas dinámicos discretos con clausura en la modulación para explicar la noción de información integrada que, a su vez, explica la conciencia. Esto también sucede de manera explícita en la Teoría de la Doble Microgénesis, la cual utiliza la noción de modulación clausurada de dos sistemas para explicar la creación de la experiencia consciente. La Teoría Autorrepresentacional es, en los términos en que está presentada, refractaria a la noción de proceso, puesto que habla de estados. Sin embargo, he tratado de demostrar que, cuando se le exige a esa teoría un nivel de determinación alto en la explicación de cómo es posible que un estado representacional se represente a sí mismo, es necesario introducir aspectos autotransformativos que resuelvan los conflictos internos de este enfoque.

[270] Ver 1.2.2.2

La segunda gran razón que justifica el empleo de un enfoque de procesos en el estudio de la conciencia proviene de las propiedades de la conciencia. La conciencia sucede a través del tiempo, como los procesos. Cuando se estudian con detenimiento los mecanismos físicos concretos que dan cuenta de la experiencia consciente, uno se encuentra con que constan de diversas etapas, como sucede en los procesos desarrollativos. La conciencia, además, tiene partes experienciales que, si se analizan en detalle, resultan ser funcionalmente relevantes, como es propio de los procesos transformativos.

Más aun, las relaciones que se dan entre esas partes de la experiencia consciente necesitan de procesos intrínsecos, es decir, procesos que no están relacionados con los procesos propios del contenido primario de la experiencia consciente. Esos procesos intrínsecos dotan a la conciencia de cierto aspecto autotransformativo en dos sentidos. Primero, en el sentido de que tienen una composición basada en partes que interfieren entre sí con clausura. Segundo, en el sentido de que existe una modulación no agónica de unas partes en las otras, que permite mantener la significatividad de cada parte dentro de un todo que resulta ser mayor que sus partes.

La intimidad, como parte experiencial que refleja el carácter subjetivo de la experiencia consciente, se explica a través de procesos intrínsecos como los indicados. La intimidad es un contraste con respecto a un contenido. Surge junto con el contenido primario en la experiencia consciente, pero es independiente de él.

El análisis de procesos se muestra eficaz también cuando se tratan problemas concretos de la conciencia, tales como la sensación de certeza en la percepción visual o la sensación de tener algo *en-la-punta-de-la-lengua*.

Los análisis de las partes experienciales de intimidad, sensación de certeza y sensación de tener *en-la-punta-de-la-lengua* son solo la punta del iceberg dentro de la inmensa riqueza fenoménica de la conciencia. Resulta muy difícil imaginar cómo se podría explicar la conciencia de movimiento o la conciencia temporal de otra manera que no sea a través de procesos que se sostienen en el tiempo. Igualmente, resulta difícil imaginar experiencias como la sensación de haber olvidado algo, el efecto Zeigarnik o la sensación de que algo cobra sentido sin que estas experiencias estén basadas en procesos con cierto aspecto de autotransformatividad. Estos son solamente algunos ejemplos de la potencialidad de este enfoque orientado a procesos.

Conclusiones

Dice Rescher que la complejidad es uno de los rasgos "más profundamente característicos de lo real", y que esta complejidad opera a diferentes niveles, entre ellos, a nivel estructural y a nivel funcional.[271] La autotransformatividad no es sino una nueva herramienta para explicar esa complejidad. Lo que aquí se ha expuesto se resume brevemente en una idea fundamental: es necesario entender la conciencia como un proceso, pero no como un proceso cualquiera, sino como un proceso autotransformativo, esto es, como un proceso que (1) tiene partes que interfieren mutuamente y (2) se describe a través de variables que tienen relaciones de modulación con clausura. Las conclusiones generales son las siguientes.

1. La conciencia es algo que despierta un interés creciente. A mediados del siglo pasado la situación de lo que luego se denominó los "Estudios de la Conciencia" era francamente mejorable. La conciencia apenas formaba parte de los programas de estudio en la psicología conductual y tampoco lo hacía de manera sustancial en el ámbito de una neurociencia que estaba todavía en ciernes. Desde 1974 hasta el día de hoy, una de las mayores contribuciones de la filosofía a este debate ha consistido en subrayar incansablemente la importancia del carácter fenoménico de la experiencia consciente y defenderlo constantemente de aquellos enfoques que, de uno u otro modo, tratan de reducir la conciencia a otras cosas. La experiencia consciente es, fundamentalmente, el "cómo es" estar en esa experiencia para el sujeto que la tiene. "Cómo es" estar en la experiencia de percibir el rojo del tomate, "cómo es" estar en la experiencia de dolor al pincharse con una aguja o "cómo es" estar en la experiencia de tener un orgasmo.

Pese a ello, y por mucho que sea una tarea imprescindible, la labor principal en el estudio de la conciencia no puede consistir *solamente* en preservar estas formas únicas e irrepetibles de la experiencia consciente como un tesoro. La tarea principal en el estudio de la conciencia debe consistir en comprender y desentrañar este hecho complejo en todas sus dimensiones. Esto significa comprender la conciencia como una experiencia con carácter fenoménico, pero también en su interacción con los sistemas cognitivos, también en cuanto a cómo sucede en el tiempo, también a través de los procesos que dan cuenta de ella y también en sus manifestaciones menos paradigmáticas, como pueden ser las experiencias de conciencia periférica tales como la intuición, la sensación de certeza y la sensación de error.

[271] Rescher (1998: xiii).

2. La conciencia es un hecho específico y complejo de explicar. Por este motivo, este hecho debe ser tomado en todas sus dimensiones. Aplicando esta idea a las propiedades de la conciencia, esto significa que no es posible aceptar algunas propiedades, las que tienen un sabor más filosófico como el carácter cualitativo o la subjetividad, al mismo tiempo que se rechazan o ignoran otras propiedades relacionadas con su función, su estructura o su temporalidad. Y menos aún debe hacerse esto *a priori*. La filosofía basada en la experiencia de percibir un tomate, de sufrir un dolor o sentir un orgasmo nos recuerda que la experiencia consciente es algo irreducible, pero esas tres experiencias no agotan la riqueza fenoménica de la conciencia.

Igualmente, no es posible aceptar *solamente* las propiedades funcionales, estructuras y temporales e ignorar las propiedades fenoménicas. La conciencia es un hecho que viene con todas las propiedades y no solo con algunas de ellas. Para entender la conciencia de movimiento hace falta entender también cuál es el tiempo de procesamiento y en qué consisten efectos como la facilitación. Para entender la sensación de certeza acerca de nuestra percepción de algo hace falta comprender los procesos que permiten que el contenido de esa experiencia sea accesible para hablar, actuar o razonar de acuerdo a él.

3. La consecuencia irremediable de tomar la conciencia con *todas* sus propiedades es que hay que entenderla como un proceso. La experiencia consciente sucede en el tiempo, tiene partes experienciales y se manifiesta como una única experiencia que es algo más que la mera adición de esas partes. Las estrategias que se han adoptado para entender la experiencia consciente quizás no siempre hayan sido las más adecuadas. A menudo se han usado términos como "estado consciente", "sustancia", "sujeto" o "inmediatez". Estos términos no reflejan el dinamismo propio de la experiencia consciente. Si se han de usar, debe ser con mucha cautela, evitando caer en una ontología de bodegón, donde los elementos permanecen inmóviles unos junto a otros. Si esto no es posible, lo mejor es evitarlos.

4. A primera vista, podría parecer que adoptar el análisis de procesos obliga a derribar todo lo anterior y empezar a construir la casa desde cero. Pero esto no es exactamente así.

Adoptar el análisis de procesos para el estudio de la conciencia no es remar a contracorriente. Por el contrario, es remar a favor de la corriente de la gran mayoría de las teorías de la conciencia. Las teorías representacionales de la conciencia, por ejemplo, enfatizan la idea de que la representación es un representar, esto es, un dinamismo. Las teorías informacionales, por su parte, introducen aspectos procesuales que forman parte fundamental de su *corpus* teórico. Otras teorías de tipo exclusivamente biológico también incluyen explícitamente aspectos procesuales, relacionados especialmente con la necesidad de procesos recurrentes, donde las reentradas adquieren gran importancia.

En definitiva, el análisis de procesos permite atender a aspectos genuinamente imbricados en muchas de las teorías de la conciencia que, quizás, habían quedado olvidados y desatendidos bajo el peso de la denominada metafísica de la sustancia.

5. La autotransformatividad es una parte importante de la explicación de la conciencia. Permite, por ejemplo, comprender mejor su carácter subjetivo. La subjetividad propia de la experiencia consciente acompaña a su contenido siendo independiente de ese contenido y, al mismo tiempo, siendo particular de ese contenido y no de otro. Esto lleva a la conclusión de que no hay una subjetividad previa y, menos aún, una subjetividad como sustancia. La subjetividad es proceso. Es un efecto de contraste entre un contenido, el contenido primario de la conciencia, y el propio proceso consciente.

6. Sin embargo, aunque la autotransformatividad es una parte importante de la explicación de la conciencia, no es la solución definitiva a todos los problemas de la conciencia. Aquí se ha presentado bajo la forma de una composición específica, a saber, la interferencia mutua entre procesos. Esta forma es, quizás, la más sencilla de detectar. Pero no tengo razones para pensar que sea la única. Si nos detenemos, por poner solo un ejemplo, en la noción de procesamiento recurrente en redes neuronales, se pueden distinguir múltiples formas de recurrencia. Podría ser que en la conciencia sucediera lo mismo.

En suma, la autotransformatividad es compleja, y aquí solo se han esbozado una de sus formas más sencillas. Se trata, al fin y al cabo, de tener nuevas herramientas para poder atacar al mismo problema desde distintos flancos.

7. El análisis presentado aquí se ha basado, fundamentalmente, en la experiencia consciente con contenido. Sin embargo, eso no significa que no pueda existir conciencia sin contenido, esto es, experiencia pura de conciencia. Simplemente refleja el hecho innegable de que la mayoría de los trabajos que se han hecho hasta la fecha (así como la mayoría de las evidencias empíricas que se han recabado hasta el momento) giran en torno de este tipo de experiencia consciente. No obstante, es posible que haya procesos conscientes sin contenido o, por así decirlo, procesos conscientes con un grado de intrinsicidad incluso mayor del aquí presentado. De nuevo, será necesario refinar nuestras herramientas para poder analizarlos.

Índice de figuras

1.1. Tridente del diablo, cubo de Necker y *subitización* 21
1.2. Funciones de enmascaramiento por metacontraste 34
1.3. Teorías de la conciencia por grupos 39

2.1. Algunas mereologías problemáticas 76
2.2. Más mereologías problemáticas 79
2.3. Una composición mereológica no clásica 83
2.4. Noción de interferencia entre procesos 84
2.5. Esquema de interferencia entre dos horquillas unidas 85
2.6. Mereologías de interferencia 87
2.7. Interferencia encadenada e interferencia mutua 89
2.8. Composición de tensionamiento mutuo en arco tensado 92
2.9. Modelo de control para el péndulo inverso 110

3.1. Cálculo de *ei* para un sistema de fotodiodo 122
3.2. Cálculo de *ei* y Φ para sistemas de cuatro elementos 124
3.3. Sistemas de nueve elementos en estrella 127
3.4. Ejemplo de potencial excitatorio postsináptico (EPSP) de una neurona. . 131
3.5. Esquema del enfoque representacional sinuoso de Kriegel (2009) 139
3.6. Experiencia TOT de acuerdo con Schwartz y Metcalfe (2014) 176
3.7. Diseño del experimento de Schwartz (2008) 177

Índice de cuadros

1.1. Propiedades funcionales en procesos conscientes y no conscientes 20
1.2. Formas del problema de la unidad 29

2.1. Procesos por automeridad y homomeridad 74

3.1. Intrinsicidad para diferentes tipos de experiencia consciente 153
3.2. Comparación de Lau y Passingham (2006) y Rounis et al. (2010) 169
3.3. Comparación de Maril et al. (2001) y Maril et al. (2005) 174

Bibliografía

Aaronson, S. (2014). Why I am not an integrated information theorist (or the unconscious expander) https://www.scottaaronson.com/blog/?p=1799.

Adam, C. & Tannery, P., Eds. (1903). *Oeuvres de Descartes: Correspondance V (Mai 1647-Février 1650)*, volume V. Paris: Léopold Cerf.

Adam, Charles y Tannery, P., Ed. (1904). *Oeuvres de Descartes: Meditationes de Prima Philosophia VII*, volume VII. Paris: Léopold Cerf.

Albert, D. & Loewer, B. (1988). Interpreting the many worlds interpretation. *Synthese*, *77*(2), 195–213.

Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America*, *43*(8), 648–657.

Anderson, N. D. (2015). Teaching signal detection theory with pseudoscience. *Frontiers in psychology*, *6*, 762, https://doi.org/10.3389/fpsyg.2015.00762.

Angell, J. R. (1907). The province of functional psychology. *Psychological review*, *14*(2), 61.

Aquila, R. E. (1992). On Plotinus and the Togetherness of Consciousness. *Journal of the History of Philosophy*, *30*(1), 7–32.

Aramendia, E. (2013). La visión en Marr y Berkeley. El problema de perderse el principio de la película. *Daimon*, (59), 125–144.

Arias, A. (2011). ¿Ha logrado Dennett quinear los qualia?: una revisión naturalista. *A Parte Rei: revista de filosofía*, (73), 10.

Arias, A. (2021). *Introducción a la ciencia de la conciencia: El estudio de la experiencia subjetiva en filosofía, psicología y neurociencias*. Madrid: Catarata.

Aristóteles (1988). *Acerca del alma*. Biblioteca Clásica Gredos 14. Gredos.

Aristóteles (1996). *Física*. Alma Mater Colección de autores griegos y latinos. Madrid: Consejo Superior de Investigaciones Científicas.

Armstrong, D. M. (1981). *The nature of mind, and other essays*. Ithaca N.Y.: Cornell University Press.

Arnim, H. F. A. v. & Adler, M. (1964). *Stoicorum veterum fragmenta*, volume Volumen III. Stutgardiae: Teubneri, editio stereotypa editionis primae (mcmiii) edition.

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge University Press.

Baars, B. J. & Gage, N. M., Eds. (2007). *Cognition, brain and consciousness. Introduction to cognitive neuroscience*. Elsevier.

Bach-y Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969). Vision Substitution by Tactile Image Projection. *Nature*, *221*(5184), 963–964, https://doi.org/10.1038/221963a0.

Bach-y Rita, P., Danilov, Y., Tyler, M. E., & Grimm, R. J. (2005). Late human brain plasticity: vestibular substitution with a tongue BrainPort human-machine interface. *Intellectica*, *1*(40), 115–122.

Bachmann, T. (2000). *Microgenetic approach to the conscious mind*, volume 25. Amsterdam/Philadephia: John Benjamins Publishing Company.

Bachmann, T. & Francis, G. (2014). *Visual Masking: Studing Perception, Attention and Consciousness*. Oxford: Academic Press. Elvesier.

Bachmann, T. & Hudetz, A. G. (2014). It is time to combine the two main traditions in the research on the neural correlates of consciousness: C = L D. *Frontiers in psychology*, *5*(Article 940), 1–13, https://doi.org/10.3389/fpsyg.2014.00940.

Bachmann, T., Suzuki, M., & Aru, J. (2020). Dendritic integration theory: A thalamocortical theory of state and content of consciousness. *Philosophy and the Mind Sciences*, *1*(II), https://doi.org/10.33735/phimisci.2020.II.52.

Balduzzi, D. & Tononi, G. (2008). Integrated Information in Discrete Dynamical Systems: Motivation and Theoretical Framework. *PLoS Computational Biology*, *4(6): e1000091*, https://doi.org/10.1371/journal.pcbi.1000091.

Basmajian & John V., Eds. (1979). *Biofeedback: Principles and practice for clinicians*. Oxford, England: Williams & Wilkins.

Bayne, T. (2018). On the axiomatic foundations of the integrated information theory of consciousness. *Neuroscience of Consciousness*, *2018*, https://doi.org/10.1093/nc/niy007.

Bayne, T. & Chalmers, D. J. (2003). What is the unity of consciousness? En A. Cleeremans (Ed.), *The unity of consciousness* (pp. 23–58). Oxford: Oxford University Press.

Benítez, A. (2005). El concepto de representación en Brentano. *Contrastes. Revista Internacional de Filosofía* http://www.revistas.uma.es/index.php/contrastes/article/download/1370/1320.

Bergson, H. (1888). *Essai sur les données immédiates de la conscience*. Québec: Bibliothèque Paul-Émile-Boulet de l'Université du Québec à Chicoutimi.

Berkeley, G. (1732). *An essay towards a new theory of vision. En The Works of George Berkeley, Bishop of Cloyne. Vol. one Editores Luce, A.A.;Jessop, T.E.; 1948-1957*, volume 1. London: Thomas Nelson and Sons Ltd.

Beth Warriner, A. & Humphreys, K. R. (2008). Learning to fail: Reoccurring tip-of-the-tongue states. *The Quarterly Journal of Experimental Psychology*, *61*(4), 535–542, https://doi.org/10.1080/17470210701728867.

Bhalla, M. & Proffitt, D. R. (1999). Visual–motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1076.

Bickhard, M. H. (2003). Process and Emergence: Normative Function and Representation. En J. Seibt (Ed.), *Process Theories* (pp. 121–155). Dordrecht: Springer Netherlands.

Blake, R. & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*(1), 13–21.

Bloch, A. M. (1885). Experience sur la vision. *Comptes Rendus de Seances de la Societe de Biologie, Paris*, *37*, 493–495.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in cognitive sciences*, *15*(12), 567–575, https://doi.org/10.1016/j.tics.2011.11.001.

Block, N. J. (1980). Troubles with functionalism. *Readings in philosophy of psychology*, *1*, 268–305.

Block, N. J. (1995). On a Confusion About a Function of Consciousness. *Brain and Behavioral Sciences*, *18*(2), 227–247.

Block, N. J. (2009). Comparing the major theories of consciousness. En M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1111–1123). Cambridge Mass.: MIT press.

Block, N. J. & Fodor, J. A. (1972). What psychological states are not. *The Philosophical Review*, (pp. 159–181).

Bogen, J. (2009). Theory and observation in science https://plato.stanford.edu/entries/science-theory-observation/.

Boly, M. & Seth, A. K. (2012). Modes and models in disorders of consciousness science. *Arch.Ital.Biol*, *150*(2-3), 172–184.

Boly, M., et al. (2013). Consciousness in humans and non-human animals: recent advances and future directions. *Frontiers in psychology*, *4*, 1–20, https://doi.org/10.3389/fpsyg.2013.00625.

Bostrom, N. (2003). Are we living in a computer simulation? *The Philosophical Quarterly*, *53*(211), 243–255.

Bower, J. M. (2002). The Organization of Cerebellar Cortical Circuitry Revisited. Implications for Function. *Annals of the New York Academy of Sciences*, *978*(1), 135–155, https://doi.org/10.1111/j.1749-6632.2002.tb07562.x.

Boyle, R. (1666). *The Origine of Formes and Qualities: according to the Corpuscular Philosophy, Illustrated by Considerations and Experiments*. Oxford: H. Hall.

Bremer, F. (1966). Neurophysiological correlates of mental unity. En J. C. Eccles (Ed.), *Brain and conscious experience* (pp. 283–297). Berlin, Heidelberg, New York: Springer-Verlag.

Brentano, F. (1874). *Psychologie vom empirischen Standpunkte*. Leipzig: Duncker & Humblot.

Brewer, W. F. & Lambert, B. L. (2001). The theory-ladenness of observation and the theory-ladenness of the rest of the scientific process. *Philosophy of Science*, (pp. S176–S186).

Broad, C. D. (1923). *Scientific thought: A philosophical analysis of some of its fundamental concepts*, volume I of *International library of philosophy*. Oxfordshire [England] and New York, New York: Routledge, repr edition.

Broad, C. D. (1938). *An Investigation of McTaggart's Philosophy, Vol. II, Part I*. Cambridge: Cambridge University Press.

Broadbent, D. E. (1958). *Perception and communication*. Oxford: Pergamon Press.

Broca, A. & Sulzer, D. (1902). La sensation lumineuse en fonction du temps. *Journal de Physiologie et Pathologie Generale, Paris*, *4*, 632–640.

Brook, A. & Raymont, P. (2017). The Unity of Consciousness. En Zalta, Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Brown, R. & Mcneill, D. (1966). The "Tip of the Tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325–337, `https://doi.org/10.1016/S0022-5371(66)80040-3`.

Bruner, J. S. & Goodman, C. C. (1947). Value and need as organizing factors in perception. *The journal of abnormal and social psychology*, *42*(1), 33.

Bruner, J. S. & Postman, L. (1949). On the perception of incongruity: A paradigm. *Journal of personality*, *18*(2), 206–223.

Brunton, S. L. & Kutz, J. N. (2019). *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press.

Bueno, G. (1990). *Materia*. Pentalfa.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature neuroscience*, *7*(3), 308–313.

Carruthers, P. (2004). HOP over FOR, HOT theory. En R. J. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology*, volume 56 (pp. 115). John Benjamins Publishing.

Carruthers, P. (2005). *Consciousness: Essays from a higher-order perspective*. Oxford and New York: Clarendon Press and Oxford University Press.

Casali, A. G., et al. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science translational medicine*, *5*(198), 198ra105, `https://doi.org/10.1126/scitranslmed.3006294`.

Ceurstemont, S. (2022). Motion blindness `https://www.youtube.com/watch?v=Dnw40nrI8sE`.

Chacón, P. (2000). Consciencia. En J. Muñoz & V. J (Eds.), *Compendio de Epistemología*. Madrid: Trotta.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford paperbacks. New York: Oxford Univ. Press, 1. issued as an oxford university press paperback edition.

Chalmers, D. J. (2000). What Is a Neural Correlate of Consciousness? En T. Metzinger (Ed.), *Neural Correlates of Consciousness : Neural correlates of consciousness: Empirical and conceptual questions* (pp.1̃7). MIT press.

Chalmers, D. J. (2002). Does Conceivability entail Possibility? En T. S. Gendler & J. Hawthorne (Eds.), *Conceivability and possibility* (pp. 145–200). Oxford University Press.

Chalmers, D. J. (2010). *The character of consciousness*. Philosophy of mind. Oxford: Oxford Univ. Press.

Chalmers, D. J. & McQueen, K. J. (2021). Consciousness and the Collapse of the Wave Function `https://arxiv.org/pdf/2105.02314`.

Changeux, J. P. & Dehaene, S. (2008). The Neuronal Workspace Model: Conscious Processing and Learning. *Learning and Memory: A Comprehensive Reference*, *1*, 729–758, `https://doi.org/10.1016/B978-012370509-9.00078-4`.

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. MIT press.

Cleeremans, A., Ed. (2003). *The unity of consciousness: Binding, integration, and dissociation*. Oxford: Oxford University Press, reprinted. edition.

Cohen, M. A. & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in cognitive sciences*, *15*(8), 358–364, `https://doi.org/10.1016/j.tics.2011.06.008`.

Corballis, M. C. (2011). *The recursive mind: The origins of human language, thought, and civilization*. Princeton University Press.

Corkin, S. (2002). What's new with the amnesic patient HM? *Nature Reviews Neuroscience*, *3*(2), 153–160.

Cotnoir, A. & Bacon, A. (2012). Non-Well founded Mereology. *The Review of Symbolic Logic*, *5*, 187–204, `https://doi.org/10.1017/S1755020311000293`.

Couturat, L., Ed. (1903). *Opuscules et fragments inédits de Leibniz : extraits des manuscrits de la bibliothèque royale de Hanovre*. Paris: Félix Alcan.

Crane, T. (2000). The Origins of Qualia. En T. Crane & S. Patterson (Eds.), *History of the mind-body problem*, London studies in the history of philosophy (pp. 169–194). London and New York: Routledge.

Crane, T. (2003). *The mechanical mind: A philosophical introduction to minds, machines, and mental representation*. London and New York: Routledge, 2nd ed. edition.

Crane, T. & Patterson, S., Eds. (2000). *History of the mind-body problem*, volume 3 of *London studies in the history of philosophy*. London and New York: Routledge.

Crick, F. (1994). *The astonishing hypothesis: the scientific search for the soul*. Simon & Schuster.

Crick, F. & Koch, C., Eds. (1990). *Towards a neurobiological theory of consciousness*, volume 2.

Crick, F. & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, *375*(6527), 121–123.

Dainton, B. (2000). *Stream of consciousness: Unity and continuity in conscious experience*. International library of philosophy. London and New York: Routledge, rev. pbk. ed. edition.

Dainton, B. (2018). Temporal Consciousness. En Zalta, Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Damasio, A. (2010). *Self comes to mind: Constructing the conscious brain*. Pantheon/Random House.

Debruyne, H., Portzky, M., van den Eynde, F., & Audenaert, K. (2009). Cotard's syndrome: a review. *Current psychiatry reports*, *11*(3), 197–202.

Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts*. New York New York: Viking.

Dehaene, S., Changeux, J.-P., & Naccache, L. (2011). The global neuronal workspace model of conscious access: from neuronal architectures to clinical applications. En S. Dehaene & Y. Christen (Eds.), *Characterizing Consciousness: From Cognition to the Clinic?* (pp. 55–84). Springer.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in cognitive sciences*, *10*(5), 204–211.

Dehaene, S. & Christen, Y., Eds. (2011). *Characterizing Consciousness: From Cognition to the Clinic?* Springer.

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(24), 14529–14534.

Dehaene, S. & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, *79*(1), 1–37, https://doi.org/10.1016/S0010-0277(00)00123-2 https://www.sciencedirect.com/science/article/pii/S0010027700001232.

Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J. F., Poline, J. B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature neuroscience*, *4*(7), 752–758.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness. *PLoS Biology*, *5(10): e260*, https://doi.org/10.1371/journal.pbio.0050260.

Del Pin, S. H., Skóra, Z., Sandberg, K., Overgaard, M., & Wierzchoń, M. (2021). Comparing theories of consciousness: why it matters and how to do it. *Neuroscience of Consciousness*, *2021*(2), niab019, https://doi.org/10.1093/nc/niab019.

Dennett, D. C. (1971). Intentional systems. *The Journal of Philosophy*, (pp. 87–106).

Dennett, D. C. (1988). Quining qualia. En A. J. Marcel & E. Bisiach (Eds.), *Consciousness in Contemporary Science* (pp. 42–77). Oxford University Press.

Dennett, D. C. (1991). *Consciousness explained*. Penguin Science, philosophy. London: Penguin Books.

Dennett, D. C. (2017). *From bacteria to Bach and back: The evolution of minds*. WW Norton & Company.

Dewey, J. (1967). *The early works of John Dewey, 1882-1898: Volume 2: 1887 Psychology*. USA: Southern Illinois University Press.

DeWitt, B. S. (1970). Quantum mechanics and reality. *Physics Today*, *23*(9), 30–35, `https://doi.org/10.1063/1.3022331`.

Diderot, D. & Escobar, J. (2002). *Carta sobre los ciegos seguido de carta sobre los sordomudos*. Pre-textos.

Diels, H., Ed. (1903). *Die fragmente der Vorsokratiker griechisch und deutsch*. Berlin: Weidmannsche Buchhandlung.

Donald, M. J. (1997). On many-minds interpretations of quantum theory. *arXiv preprint quant-ph/9703008*.

Dretske, F. (2006). Representation and Self-Knowledge. En G. Macdonald & D. Papineau (Eds.), *Teleosemantics* (pp. 69–84). Oxford: Clarendon.

Dretske, F. I. (1995). *Naturalizing the mind*, volume 1994 of *The Jean Nicod lectures*. Cambridge Mass.: MIT press.

Duffau, H. (2011a). Brain hodotopy: new insights provided by intrasurgical mapping. En H. Duffau (Ed.), *Brain mapping. From neural basis of cognition to surgical applications* (pp. 335). Austria: SpringerWienNewYork.

Duffau, H., Ed. (2011b). *Brain mapping. From neural basis of cognition to surgical applications*. Austria: SpringerWienNewYork.

Eastman, T. & Keeton, H., Eds. (2004). *Physics & Whitehead: Quantum, Process and Experience*. Albany: State University of New York Press.

Eccles, J. C., Ed. (1966). *Brain and conscious experience*. Berlin, Heidelberg, New York: Springer-Verlag.

Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.

Edelman, G. M. & Mountcastle, V., Eds. (1978). *The mindful brain*. The MIT Press.

Edelman, G. M. & Tononi, G. (2000). Reentry and the dynamic core: neural correlates of conscious experience. *Neural correlates of consciousness, empirical and conceptual questions.MIT Press, Cambridge, Mass*, (pp. 139–151).

Efron, R. (1973). Conservation of temporal information by perceptual systems. *Perception & Psychophysics*, *14*, 518–530.

Eppley, K. R., Abrams, A. I., & Shear, J. (1989). Differential effects of relaxation techniques on trait anxiety: A meta-analysis. *Journal of Clinical Psychology*, *45*(6), 957–974, https://doi.org/10.1002/1097-4679(198911)45:6<957::AID-JCLP2270450622>3.0.CO;2-Q.

Epstein, R. (2000). The Neural-Cognitive Basis of the Jamesian Stream of Thought. *Consciousness and cognition*, *9*(4), 550–575.

Erismann, T. & Kohler, I. (1950). Die Umkehrbrille und das aufrechte Sehen. *educational film*, *16*.

Farrell, J. & McClelland, T. (2017). Editorial: Consciousness and Inner Awareness. *Review of philosophy and psychology*, *8*(1), 1–22, https://doi.org/10.1007/s13164-017-0331-x.

Felleman, D. J. & van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, N.Y. : 1991)*, *1*(1), 1–47, https://doi.org/10.1093/cercor/1.1.1-a.

Feyerabend, P. K., Feigl, H., & Maxwell, G. (1962). Explanation, reduction, and empiricism. *Minnesota Studies in the Philosophy of Science*, *3*, 28–97.

Flanagan, O. J. (1992). *Consciousness reconsidered*. London, England: MIT press Cambridge, MA.

Fodor, J. A. (1975). *The language of thought*, volume 5. Harvard University Press.

Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. MIT press.

Forman, R. K. C. (1990a). Introduction: Mysticism, Constructivism, and Forgetting. En R. K. C. Forman (Ed.), *The Problem of Pure Consciousness: Mysticism and Philosophy* (pp. 3–49). New York: Oxford University Press.

Forman, R. K. C., Ed. (1990b). *The Problem of Pure Consciousness: Mysticism and Philosophy*. New York: Oxford University Press.

Foster, J. (1979). In Self-Defence. En G. F. Macdonald (Ed.), *Perception and Identity: Essays Presented to A. J. Ayer with his Replies to them* (pp. 161–185). London: Macmillan Education UK.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie Und Philosophische Kritik*, *100*(1), 25–50.

Freud, S. (1915). Das Unbewusste. *Internationale Zeitschrift für Psychoanalyse*, *3*(4), 189–203.

Gabbiani, F., Krapp, H. G., & Laurent, G. (1999). Computation of object approach by a wide-field, motion-sensitive neuron. *The Journal of neuroscience*, *19*(3), 1122–1141.

Galilei, G. (1623). Il Saggiatore, nel quale con bilancia esquisita e giusta si ponderano le cose contenute nella libra astronomica e filosofica di Lotario Sarsi Sisgensario. *Rome: Mascardi*.

Gall, E. P. & Gooden, M. A. (1982). *Examination of the musculoskeletal system*. Tucson Ariz.: University of Arizona Health Sciences Center.

Galton, F. (1888). Co-relations and their measurement, chiefly from anthropometric data. *Proceedings of the Royal Society of London*, *45*(273-279), 135–145.

Gazzaniga, M. S., Ed. (2009). *The cognitive neurosciences*. Cambridge Mass.: MIT press, 4th ed. edition.

Gendler, T. S. & Hawthorne, J., Eds. (2002). *Conceivability and possibility*. Oxford University Press.

Gennaro, R. J. (2004a). *Higher-order theories of consciousness: An anthology*, volume 56. John Benjamins Publishing.

Gennaro, R. J., Ed. (2004b). *Higher-order theories of consciousness: An anthology*. John Benjamins Publishing.

Gollan, T. H. & Brown, A. S. (2006). From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology: General*, *135*(3), 462–483, https://doi.org/10.1037/0096-3445.135.3.462 https://login.bucm.idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/tip-tongue-tot-data-theoretical-implications-two/docview/614503264/se-2?accountid=14514.

Gomes, A. O. (1966). The brain-consciousness problem in contemporary scientific research. En J. C. Eccles (Ed.), *Brain and conscious experience* (pp. 446–469). Berlin, Heidelberg, New York: Springer-Verlag.

Gonzalo, L. M., Irujo, A. M., & Insausti, R. (1993). *Anatomía funcional*. Ediciones Universidad de Navarra. EUNSA.

González-Castán, O. (1999). *La conciencia errante: Introducción crítica a la filosofía de la psicología*. Tecnos.

González-Castán, O. (2008). Fenomenología de la conciencia y naturalismo. *Teorema*, *27*(3), 197–210.

Good, I. J., Ed. (1961). *The Scientist Speculates*. Heineman.

Goodale, M. A. (1988). Modularity in visuomotor control: From input to output. En Z. W. Pylyshyn (Ed.), *Computational processes in human vision: An interdisciplinary perspective* (pp. 262–285). Ablex Norwood, NJ.

Goodale, M. A. & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, *15*(1), 20–25.

Graziano, M. S. A. (2014). *Consciousness and the Social Brain*. Cary: Oxford University Press.

Gregory, R. L. (1988). Consciousness in science and philosophy: Conscience and conscience. En A. J. Marcel & E. Bisiach (Eds.), *Consciousness in Contemporary Science*. Oxford University Press.

Gross, O. (1902). *Die Cerebrale Sekündarfunction*. Leipzig: Vogel.

Gross, O. (1907). *Das Freudsche Ideogenitätsmoment und seine Nedeutung im manisch-depressivem Irresein Kraepelins*. Leipzig: F.C.W. Vogel.

Guillot, M. (2017). I Me Mine: on a Confusion Concerning the Subjective Character of Experience. *Review of philosophy and psychology*, *8*(1), 23–53, https://doi.org/10.1007/s13164-016-0313-4.

Hameroff, S. & Penrose, R. (1996). Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. *Mathematics and Computers in Simulation*, *40*(3-4), 453–480, https://doi.org/10.1016/0378-4754(96)80476-9.

Hampson, M., Ruiz, S., & Ushiba, J. (2020). Neurofeedback. *NeuroImage*, *218*, 116473, https://doi.org/10.1016/j.neuroimage.2019.116473.

Hansen, N. V. (2004). Spacetime and becoming: Overcoming the contradiction between special relativity and the passage of time. En T. Eastman & H. Keeton (Eds.), *Physics & Whitehead: Quantum, Process and Experience* (pp. 136–163). State University of New York Press.

Hanson, N. R. (1958). *Patterns of Discovery An Inquiry into the Conceptual Foundations of Science*. Cambridge: Cambridge Unversity Press.

Harman, G. (1990). The Intrinsic Quality of Experience. *Philosophical Perspectives*, *4*, 31, https://doi.org/10.2307/2214186.

Hayles, N. K. (2014). Cognition Everywhere: The Rise of the Cognitive Nonconscious and the Costs of Consciousness. *New Literary History*, *45*(2), 199–220, https://doi.org/10.1353/nlh.2014.0011.

Hebb, D. O. (1949). *The Organization of Behavior. A Neuropsychological Theory.* New York: John Wiley and Sons, Inc.

Held, R., Ostrovsky, Y., de Gelder, B., Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011). The newly sighted fail to match seen with felt. *Nature neuroscience*, *14*(5), 551–553.

Hennig, B. (2006). *"Conscientia" bei Descartes: Zugl.: Leipzig, Univ., Diss., 2004 u.d.T.: Hennig, Boris: Was bedeutet "conscientia" bei Descartes?*, volume 127 of *Symposion*. Freiburg: Alber, orig.-ausg edition.

Henry, P. & Schwyzer, H.-R. (1964). Plotini opera, 3 vols. *Oxford, Clarendon Press*, *1976*, 1982.

Hertz, H., Ed. (1894). *Die Prinzipien der Mechanik: in neuem zusammenhange Dargestellt*. Leipzig: Johann Ambrosius Barth (Arthur Meiner).

Heymans, C. & de Schaepdryver, A. (1966). Dopamine and central neurotransmission. En J. C. Eccles (Ed.), *Brain and conscious experience* (pp. 506–521). Berlin, Heidelberg, New York: Springer-Verlag.

Hierocles (2014). Ἠθικὴ στοιχείωσις. *Helmantica: Revista de filología clásica y hebrea*, *65*(193), 39.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research. *ACM Transactions on Computer-Human Interaction*, *7*(2), 174–196.

Hume, D. (1888 [1739]). *A treatise of human nature*. Oxford: Clarendon Press.

Humphrey, N. (1992). *A history of the mind*. Simon & Schuster.

Husserl, E. (1913). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*, volume Erster buch of *Jahrbuch für Philosophie und phänomenologische Forschung*. Halle: Max Niemeyer.

Husserl, E. (1964 [1928]). *Phenomenology of Internal Time-Consciousness*. Bloominston, IN: Indiana University Press, ed. m. heidegger trans. j. s. churchill edition.

Husserl, E. & Serrano de Haro, A. (2002). *Lecciones de fenomenología de la conciencia interna del tiempo*. Madrid: Trotta.

Huxley, T. H. (1874). On the hypothesis that animals are automata, and its history. *Collected essays*, *1*.

Izhikevich, E. M. (2007). *Dynamical systems in neuroscience: The geometry of excitability and bursting*. Computational neuroscience. Cambridge Mass.: MIT press.

Jackendoff, R. (1994). *Consciousness and the computational mind*, volume 3 of *A Bradford book*. Cambridge, Mass.: MIT press, 4. print edition.

Jackson, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly (1950-)*, *32*(127), 127–136, https://doi.org/10.2307/2960077 http://www.jstor.org/stable/2960077.

James, W. (1931). *The principles of psychology, Vol I*. New York: Henry Holt and Co.

Jasper, H. H. (1966). Pathophysiological studies of brain mechanisms in different states of consciousness. En J. C. Eccles (Ed.), *Brain and conscious experience* (pp. 256–282). Berlin, Heidelberg, New York: Springer-Verlag.

Johnson-Laird, P. N. (1983). A computational analysis of consciousness. *Cognition & Brain Theory*.

Jones, E. G., Ed. (2007). *The Thalamus*. Cambridge University Press, 2nd edn edition.

Jung, G. (1933). Syneídesis, conscientia, Bewußtsein. *Archiv für die gesamte Psychologie*, *89*, 525–540.

Katz, S. T. (1978a). Language, Epistemology, and Mysticism. En S. T. Katz (Ed.), *Mysticism and philosophical analysis* (pp. 22–74). New York: Oxford University Press.

Katz, S. T., Ed. (1978b). *Mysticism and philosophical analysis*. New York: Oxford University Press.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The Discrimination of Visual Number. *American Journal of Psychology*, *62*(4), 498 https://login.bucm.idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/discrimination-visual-number/docview/1289819292/se-2?accountid=14514.

Kitcher, P. (1988). Marr's computational Theory of Vision. *Philosophy of Science*, *55*, 1–24.

Klein, D. B. (1989). *El concepto de la conciencia*. México DF: Fondo de Cultura Económica.

Koch, C. (2004). *The Quest for Consciousness*. Roberts and Company.

Koch, C. (2012). *Consciousness: Confessions of a romantic reductionist.* MIT press.

Koch, C. & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in cognitive sciences*, *11*(1), 16–22.

Kriegel, U. (2004). Consciousness and Self-Consciousness. *Monist*, *87*(2), 182–205, https://doi.org/10.5840/monist20048725.

Kriegel, U. (2009). *Subjective consciousness: A self-representational theory.* Oxford and New York: Oxford University Press.

Kriegel, U. (2011). Self-Representationalism and the Explanatory Gap. En J. Liu & J. Perry (Eds.), *Consciousness and the Self: New Essays.* Cambridge University Press.

Krishnaswamy, P. & Nair, S. (2016). Effect of Music Therapy on Pain and Anxiety Levels of Cancer Patients: A Pilot Study. *Indian journal of palliative care*, *22*(3), 307–311, https://doi.org/10.4103/0973-1075.185042 https://pubmed.ncbi.nlm.nih.gov/27559260.

Krpan, D. & Schnall, S. (2014). When perception says "no" to action: Approach cues make steep hills appear even steeper. *Journal of experimental social psychology*, *55*, 89–98.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Kutas, M. & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, *4*(12), 463–470.

La Mettrie, J. O. d. (1748). *L'homme machine.* Leyden: Elie Luzac.

Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive neuroscience*, *1*(3), 204–220, https://doi.org/10.1080/17588921003731586.

Lando, G. (2017). *Mereology: A philosophical introduction.* London UK and New York: Bloomsbury Academic.

Lashley, K. S. (1960a). Cerebral organization and behaviour. En F. A. Beach, D. O. Hebb, C. T. Morgan, & H. W. Nissen (Eds.), *The Neuropsychology of Lashley.* (pp. 529–543). New York: McGraw-Hill Book Company.

Lashley, K. S. (1960b). *The Neuropsychology of Lashley. Selected papers of K. S. Lashley.* New York: McGraw-Hill Book Company.

Lau, H. & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in cognitive sciences*, *15*(8), 365–373, https://doi.org/10.1016/j.tics.2011.05.009.

Lau, H. C. & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(49), 18763–18768, https://doi.org/10.1073/pnas.0607716103.

Leibniz, G. W. (1886). *Nouveaux essais sur l'entendement humain.* Librairie Hachette.

Lennie, P. (2003). The cost of cortical computation. *Current biology*, *13*(6), 493–497.

Leopold, D. A. & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, *379(6565)*(Feb 8), 549–553.

Levine, J. (2001). *Purple haze: The puzzle of consciousness*. Oxford University Press.

Lewis, C. I. & Broad, C. D. (1925). Scientific Thought. *The Philosophical Review*, *34*(4), 406, https://doi.org/10.2307/2179384.

Lewis, D. (1983). Extrinsic properties. *Philosophical Studies*, *44*(2), 197–200, https://doi.org/10.1007/BF00354100.

Libet, B. (1966). Brain stimulation and the threshold of conscious experience. En J. C. Eccles (Ed.), *Brain and conscious experience* (pp. 165–181). Berlin, Heidelberg, New York: Springer-Verlag.

Libet, B., Wright, JR, E. W., Feinstein, B., & Pearl, D. K. (1979). Subjective referral of the timing for a conscious sensory experience: a functional role for the somatosensory specific projection system in man. *Brain*, *102*(1), 193–224, https://doi.org/10.1093/brain/102.1.193.

Liu, J. & Perry, J., Eds. (2011). *Consciousness and the Self: New Essays*. Cambridge University Press.

Locke, J. (1825). *An essay concerning human understanding*. London: Thomas Davidson, Whitefriars, twenty-fifth edition edition.

Lockwood, M. (1989). *Mind, brain and the quantum: The compound 'I'*. Basil Blackwell.

Lockwood, M. (1996). 'Many Minds'. Interpretations of Quantum Mechanics. *The British Journal for the Philosophy of Science*, *47*(2), 159–188.

Logothetis, N. K. & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science*, *245*(4919), 761–763.

London, F. & Bauer, E. (1983). The theory of observation in quantum mechanics. *Quantum theory and measurement*, (pp. 217–259).

Lucas, J. R. (1961). Minds, Machines and Gödel. *Philosophy*, *36*(137), 112–127, https://doi.org/10.1017/S0031819100057983.

Lycan, W. G. (1987). *Consciousness*. Cambridge Mass.: MIT press.

Macdonald, G. & Papineau, D. (2006a). Introduction: Prospects and Problems for Teleosemantics. En G. Macdonald & D. Papineau (Eds.), *Teleosemantics* (pp. 1–22). Oxford: Clarendon.

Macdonald, G. & Papineau, D., Eds. (2006b). *Teleosemantics: New philosophical essays*. Oxford: Clarendon.

Macdonald, G. F., Ed. (1979). *Perception and Identity: Essays Presented to A. J. Ayer with his Replies to them*. London: Macmillan Education UK.

Machery, E. (2015). Cognitive Penetrability: a No-Progress Report. En J. Zeimbekis & A. Raftopoulos (Eds.), *The Cognitive Penetrability of Perception: New Philosophical Perspectives : .(2015).T* (pp. 59–71). Oxford Scholarship Online.

Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly*, *2*(4), 245–264 http://www.jstor.org/stable/20009173.

Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Oxford: Clarendon Press.

Mahon, B. Z. & Wu, W. (2015). Cognitive Penetration of the Dorsal Visual Stream? En J. Zeimbekis & A. Raftopoulos (Eds.), *The Cognitive Penetrability of Perception: New Philosophical Perspectives* (pp. 201–217). Oxford Scholarship Online.

Mangan, B. (1993). Taking Phenomenology Seriously: The "Fringe" and Its Implications for Cognitive Research. *Consciousness and Cognition*, *2*(2), 89–108, https://doi.org/10.1006/ccog.1993.1008.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive psychology*, *15*(2), 197–237.

Marcel, A. J. (1988). Phenomenal experience and functionalism. En A. J. Marcel & E. Bisiach (Eds.), *Consciousness in Contemporary Science* (pp. 121–158). Oxford University Press.

Marcel, A. J. & Bisiach, E., Eds. (1988). *Consciousness in Contemporary Science*. Oxford University Press.

Marcel, A. J. & Bisiach, E., Eds. (1992). *Consciousness in contemporary science*. Oxford: Clarendon Press, repr.with corrections edition.

Marietta, D. E. (1970). Conscience in Greek Stoicism. *Numen*, (pp. 176–187).

Maril, A., Simons, J. S., Weaver, J. J., & Schacter, D. L. (2005). Graded recall success: an event-related fMRI comparison of tip of the tongue and feeling of knowing. *NeuroImage*, *24*(4), 1130–1138, https://doi.org/10.1016/j.neuroimage.2004.10.024 https://www.sciencedirect.com/science/article/pii/S1053811904006317.

Maril, A., Wagner, A. D., & Schacter, D. L. (2001). On the Tip of the Tongue. *Neuron*, *31*(4), 653–660, https://doi.org/10.1016/S0896-6273(01)00396-8 https://www.sciencedirect.com/science/article/pii/s0896627301003968.

Marr, D. (1969). A theory of cerebellar cortex. *The Journal of physiology*, *202*(2), 437–470.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge. Massachusets.: The MIT Press, 2010 edition.

Martínez-Freire, P. (2001). Aproximaciones científicas al problema de la conciencia. En J. A. Mora (Ed.), *Neuropsicología cognitiva: algunos problemas actuales*. Málaga: Aljibe.

Mayner, W. G. P., Marshall, W., Albantakis, L., Findlay, G., Marchman, R., & Tononi, G. (2018). PyPhi: A toolbox for integrated information theory. *PLoS Computational Biology*, *14*(7), e1006343, https://doi.org/10.1371/journal.pcbi.1006343.

McGinn, C. (1991). *The problem of consciousness: Essays towards a resolution*. Oxford: B. Blackwell.

McTaggart, J. E. (1908). The unreality of time. *Mind*, *XVII*(4), 457–474, https://doi.org/10.1093/mind/XVII.4.457.

Merrit, S. (1999). Experimental music love.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 851–861, https://doi.org/10.1037/0278-7393.19.4.851 https://login.bucm.idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/cue-familiarity-heuristic-metacognition/docview/614341956/se-2?accountid=14514.

Metzinger, T., Ed. (2000). *Neural Correlates of Consciousness : Neural correlates of consciousness: Empirical and conceptual questions*. MIT press.

Metzinger, T. (2009). *The ego tunnel: The science of the mind and the myth of the self*. Basic Books.

Metzinger, T., Ed. (2015). *Open MIND*. Mainz: Frankfurt am Main : MIND Group.

Middleton, F. & Strick, P. L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Research Reviews*, *31*(2-3), 236–250, https://doi.org/10.1016/S0165-0173(99)00040-5.

Millikan, R. (2006). Useless Content. En G. Macdonald & D. Papineau (Eds.), *Teleosemantics* (pp. 100–114). Oxford: Clarendon.

Minsky, M. & Papert, S. (1969). *Perceptrons*. Perceptrons. Oxford, England: M.I.T. Press.

Mishkin, M. & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural brain research*, *6*(1), 57–77.

Montemayor, C. (2013). *Minding time : a philosophical and theoretical approach to the psychology of time*, volume v. 5 of *Supplements to the study of time*. Leiden: Brill http://booksandjournals.brillonline.com/content/9789004236172.

Montemayor, C. (2017). Conscious awareness and time perception. *PsyCh journal*, *6*(3), 228–238, https://doi.org/10.1002/pchj.173.

Mora, J. A., Ed. (2001). *Neuropsicología cognitiva: algunos problemas actuales*. Málaga: Aljibe.

Moruzzi, G. & Magoun, H. W. (1949). Brain stem reticular formation and activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, *1*(1-4), 455–473, https://doi.org/10.1016/0013-4694(49)90219-9.

Mossbridge, J. A., Tressoldi, P., Utts, J., Ives, J. A., Radin, D., & Jonas, W. B. (2014). Predicting the unpredictable: critical analysis and practical implications of predictive anticipatory activity. *Frontiers in human neuroscience*, *8*(146), 1–10, https://doi.org/10.3389/fnhum.2014.00146.

Mountcastle, V. (1978). An organizing principle for cerebral function: The unit model

and the distributed system. En G. M. Edelman & V. Mountcastle (Eds.), *The mindful brain* (pp. 7–50). The MIT Press.

Murakami, I. (2001). A flash-lag effect in random motion. *Vision Research*, *41*(24), 3101–3119, https://doi.org/10.1016/S0042-6989(01)00193-6 https://www.sciencedirect.com/science/article/pii/S0042698901001936.

Muñoz, J. & J, V., Eds. (2000). *Compendio de Epistemología*. Madrid: Trotta.

Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, *83*(4), 435–450, https://doi.org/10.2307/2183914.

Neander, K. (2006). Content for Cognitive Science. En G. Macdonald & D. Papineau (Eds.), *Teleosemantics* (pp. 167–194). Oxford: Clarendon.

Neisser, U. (1967). *Cognitive psychology*. Century psychology series. New York: Appleton-Century-Crofts.

Neumann, J. V. (1955). *Mathematical foundations of quantum mechanics*. Princeton University Press.

Nicholson, D. J. & Dupré, J., Eds. (2018). *Everything flows: Towards a processual philosophy of biology*, volume 1. Oxford United Kingdom: Oxford University Press, first edition edition.

Niikawa, T. (2020). A Map of Consciousness Studies: Questions and Approaches. *Frontiers in psychology*, *11*, https://doi.org/10.3389/fpsyg.2020.530152 https://www.frontiersin.org/article/10.3389/fpsyg.2020.530152.

Noë, A. (2004). *Action in perception*. MIT press.

Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology*, (pp. May 8, 2014)., https://doi.org/10.1371/journal.pcbi.1003588.

Okon, E. & Sebastián, M. A. (2018). A consciousness-based quantum objective collapse model. *Synthese*, *40*(12), 2935, https://doi.org/10.1007/s11229-018-1887-4.

Papineau, D. (1984). Representation and Explanation. *Philosophy of Science*, *51*(4), 550–572, https://doi.org/10.1086/289205.

Penfield, W. (1966). Speech, perception and the uncommitted cortex. En J. C. Eccles (Ed.), *Brain and conscious experience* (pp. 217–237). Berlin, Heidelberg, New York: Springer-Verlag.

Penfield, W. & Jasper, H. (1954). *Epilepsy and the functional anatomy of the human brain*. Epilepsy and the functional anatomy of the human brain. Oxford, England: Little, Brown & Co.

Penrose, R. (1994). Shadows of the mind: A search for the missing science of consciousness.

Penrose, R. (2002). Consciousness, computation, and the chinese room. En J. Preston & M. Bishop (Eds.), *Views into the Chinese room: New Essays on Searle and artificial intelligence*. Oxford, New York: Oxford University Press.

Phillips, I. (2010). Perceiving Temporal Properties. *European Journal of Philosophy*, *18*(2), 176–202, https://doi.org/10.1111/j.1468-0378.2008.00299.x.

Phillips, I. (2016). No watershed for overflow: Recent work on the richness of consciousness. *Philosophical Psychology*, *29*(2), 236–249, https://doi.org/10.1080/09515089.2015.1079604.

Phillips, I. (2018). The methodological puzzle of phenomenal consciousness. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *373*(1755), https://doi.org/10.1098/rstb.2017.0347.

Pierce, C. A. (1955). *Conscience in the New Testament: A Study of Syneidesis in the New Testament*. Alec R. Allenson.

Pineda, J. A. (2015). COGS 107A - Neuroanatomy and Physiology - https://podcast.ucsd.edu/watch/fa15/cogs107afa15/18.

Plotino, E. (1989). *Enneads. Plotinus in seven volumes*, volume 1. Loeb Classical Library.

Plotino, E. (1992). Enéadas I-VI, vers. de Jesús Igal, 3 vols. *Madrid, Editorial Gredos*.

Poggio, T. (1990). A theory of how the brain might work. *Cold Spring Harbour Symposia on Quantitative Biology*, *Cold Spring Harbour Laboratory Press*, 899–910.

Pollen, D. A. (1995). Cortical areas in visual awareness. *Nature*, *377*(6547), 293–295, https://doi.org/10.1038/377293b0.

Preston, J. & Bishop, M., Eds. (2002). *Views into the Chinese room: New Essays on Searle and artificial intelligence*. Oxford, New York: Oxford University Press.

Prinz, J. J. (2012). *The conscious brain: How attention engenders experience*. Philosophy of mind. Oxford and New York: Oxford University Press.

Prston, J. & Bishop, M., Eds. (2002). *Views into the Chinese Room : Views into the chinese room*. Clarendon Press Oxford.

Pylyshyn, Z. W. (1984). *Computation and cognition: Towards a foundation for cognitive science*. Cambridge, MA: MIT Press.

Pylyshyn, Z. W., Ed. (1988). *Computational processes in human vision: An interdisciplinary perspective*. Ablex Norwood, NJ.

Pylyshyn, Z. W. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(03), 341–365.

Raftopoulos, A. (2015). Cognitive Penetrability and Consciousness. En J. Zeimbekis & A. Raftopoulos (Eds.), *The Cognitive Penetrability of Perception: New Philosophical Perspectives* (pp. 268–295). Oxford Scholarship Online.

Ramachandran, V. D. & Blakeslee, S. (1998). *Phantoms in the brain: Probing the mysteries of the mind*. Harper Perennial.

Rashbrook, O. (2013). Diachronic and synchronic unity. *Philosophical Studies*, *164*(2), 465–484, https://doi.org/10.1007/s11098-012-9865-z.

Reimers, J. R., McKemmish, L. K., McKenzie, R. H., Mark, A. E., & Hush, N. S. (2009). Weak, strong, and coherent regimes of Fröhlich condensation and their applications to

terahertz medicine and quantum consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(11), 4219–4224, https://doi.org/10.1073/pnas.0806273106.

Rescher, N. (1996). *Process metaphysics: An introduction to process philosophy*. Suny Press.

Rescher, N. (1998). *Complexity: A philosophical overview*. Transaction Publishers.

Rescher, N. (2000). *Process philosophy: A survey of basic issues*. University of Pittsburgh Pre.

Revonsuo, A. (1999). Binding and the phenomenal unity of consciousness. *Consciousness and cognition*, *8*(2), 173–185, https://doi.org/10.1006/ccog.1999.0384.

Rieiro, H., Martinez-Conde, S., Danielson, A. P., Pardo-Vazquez, J. L., Srivastava, N., & Macknik, S. L. (2012). Optimizing the temporal dynamics of light to human perception. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(48), 19828–19833, https://doi.org/10.1073/pnas.1213170109.

Riggs, G. K. & Gall, E. P. (1984). *Rheumatic diseases: Rehabilitation and management*. Boston: Butterworth Publishers.

Robinson, H. (2021). Substance. En Zalta, Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Rodríguez González, M. (2018). *Más allá del rebaño: Nietzsche, filósofo de la mente*. Avarigani Editores.

Rosenthal, D. M. (2005). *Consciousness and mind*. Oxford and New York: Oxford University Press.

Rosenthal, D. M. (2008). Consciousness and its function. *Neuropsychologia*, *46*(3), 829–840.

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive neuroscience*, *1*(3), 165–175, https://doi.org/10.1080/17588921003632529.

Ryder, D. (2006). On Thinking of Kinds: A Neuroscientific Perspective. En G. Macdonald & D. Papineau (Eds.), *Teleosemantics* (pp. 115–145). Oxford: Clarendon.

Sasanuma, S. (1975). Kana and Kanji processing in Japanese aphasics. *Brain and language*, *2*, 369–383.

Schaffer, J. (2016). The Metaphysics of Causation. En Zalta, Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Schrödinger, E. (1958). *Mind and matter: the Tarner lectures delivered at Trinity College, Cambridge, in Oct. 1956*. University Press.

Schwartz, B. L. (2008). Working memory load differentially affects tip-of-the-tongue states and feeling-of-knowing judgments. *Memory & cognition*, *36*(1), 9–19, https://doi.org/10.3758/mc.36.1.9.

Schwartz, B. L. & Brown, A. S., Eds. (2014). *Tip-of-the-tongue states and related phenomena*. New York: Cambridge University Press.

Schwartz, B. L. & Metcalfe, J. (2011). Tip-of-the-tongue (TOT) states: retrieval, behavior, and experience. *Memory & cognition*, *39*(5), 737–749, https://doi.org/10.3758/s13421-010-0066-8.

Schwartz, B. L. & Metcalfe, J. (2014). Tip-of-the-Tongue (TOT) States: Mechanisms and Metacognitive Control. En B. L. Schwartz & A. S. Brown (Eds.), *Tip-of-the-tongue states and related phenomena* (pp. 15–31). New York: Cambridge University Press.

Schwartz, B. L., Travis, D. M., Castro, A. M., & Smith, S. M. (2000). The phenomenology of real and illusory tip-of-the-tongue states. *Memory & cognition*, *28*(1), 18–27, https://doi.org/10.3758/BF03211571.

Scoville, W. B. & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry*, *20*(1), 11–21.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.

Searle, J. R. (1984). *Minds, brains, and science*. The 1984 Reith lectures. Cambridge Mass.: Harvard University Press.

Searle, J. R. (1992). *The rediscovery of the mind*. Representation and mind. Cambridge and London: MIT press.

Searle, J. R. (1994). The connection principle and the ontology of the unconscious: a reply to Fodor and Lepore. *Philosophy and Phenomenological Research*, *54*(4), 847–855.

Searle, J. R. (1995). Consciousness, the Brain and the Connection Principle: A Reply. *Philosophy and Phenomenological Research*, *55*(1), 217–232, https://doi.org/10.2307/2108322.

Searle, J. R. (1998). How to study consciousness scientifically. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *353*, 1935–1942 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1692422/pdf/9854266.pdf.

Searle, J. R. (1999). *Mind, language, and society: Philosophy in the real world*. MasterMinds series. New York, NY: Basic Books, 1. paperb. ed. edition.

Searle, J. R. (2002). Twenty-one years in the Chinese Room. En J. Prston & M. Bishop (Eds.), *Views into the Chinese Room : Views into the chinese room* (pp.5̃1). Clarendon Press Oxford.

Searle, J. R., Dennett, D. C., & Chalmers, D. J. (1997). *The mystery of consciousness*. New York: New York Review of Books, 1st ed. edition.

Sebastián, M. A. (2012). Experiential Awareness: Do You Prefer "It" to "Me"? *Philosophical Topics*, *40*(2), 155–177.

Sebastián, M. A. (2018). Consciencia. *Enciclopedia de la Sociedad Española de Filosofía Analítica*.

Sebastián, M. A. (2020). Perspectival self-consciousness and ego-dissolution. *Philosophy*

and the Mind Sciences, *1*(I), 1–27, https://doi.org/10.33735/phimisci.2020.I.44.

Sebastián, M. A. & Artiga, M. (2017). Can Informational Theories Account for Metarepresentation? *Topoi*, *20*(14), 97, https://doi.org/10.1007/s11245-017-9514-4.

Sebastián, M. (2022). *El problema de la conciencia: Una introducción crítica a la discusión filosófica actual*. Madrid: Cátedra.

Seibt, J., Ed. (2003). *Process Theories: Crossdisciplinary Studies in Dynamic Categories*. Dordrecht: Springer Netherlands.

Seibt, J. (2009). Forms of emergent interaction in General Process Theory. *Synthese*, *166*(3), 479, https://doi.org/10.1007/s11229-008-9373-z.

Seibt, J. (2015). Non-Transitive Parthood, Leveled Mereology, and the Representation of Emergent Parts of Processes. *Grazer Philosophische Studien*, *91*(1), 165–190, https://doi.org/10.1163/9789004302273{_}008.

Seibt, J. (2018). Ontological Tools for the Process Turn in Biology: Some Basic Notions of General Process Theory. En D. J. Nicholson & J. Dupré (Eds.), *Everything flows*, volume 1 (pp. 113–136). Oxford United Kingdom: Oxford University Press.

Seibt, J. (2020). Process Philosophy. En Zalta, Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Seibt, J., Imaguire, G., & Gerogiorgakis, S. (2017). *Handbook of Mereology*. Philosophia Verlag München.

Seth, A. K. (2022). What are the four main theories of consciousness? https://www.newscientist.com/question/four-main-theories-consciousness/#ixzz7PNmb4mGN.

Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge University Press.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *American Telephone and Telegraph Co*, 379–423–623–656.

Sheinberg, D. L. & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences*, *94*(7), 3408–3413.

Sherman, S. M. (2016). Thalamus plays a central role in ongoing cortical functioning. *Nature neuroscience*, *19*(4), 533–541, https://doi.org/10.1038/nn.4269.

Shoemaker, S. S. (1968). Self-Reference and Self-Awareness. *The Journal of Philosophy*, *65*(19), 555, https://doi.org/10.2307/2024121.

Siegel, S. (2016). The Contents of Perception. En Zalta, Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Simons, P. M. (1987). *Parts: A study in ontology*. New York: Oxford University Press.

Singer, W. & Gray, C. M. (1995). Visual feature integration and the temporal correlation

hypothesis. *Annual review of neuroscience*, *18*, 555–586, https://doi.org/10.1146/annurev.ne.18.030195.003011.

Sokolov, E. N. (1963). *Perception and the Conditioned Reflex*. Oxford: Pergamon Press.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29, https://doi.org/10.1037/h0093759.

Sperry, R. W. (1966). Brain bisection and mechanisms of consciousness. En J. C. Eccles (Ed.), *Brain and conscious experience* (pp. 298–313). Berlin, Heidelberg, New York: Springer-Verlag.

Stokes, D. (2012). Perceiving and Desiring: A New Look at the Cognitive Penetrability of Experience. *Philosophical Studies*, *158*(3), 477–492.

Stokes, D. (2015). Towards a Consequentialist Understanding of Cognitive Penetration. En J. Zeimbekis & A. Raftopoulos (Eds.), *The Cognitive Penetrability of Perception: New Philosophical Perspectives* (pp. 75–98). Oxford Scholarship Online.

Strawson, P. (1959). *Individuals: An Essay in Descriptive Metaphysics*. Routledge.

Strogatz, S. H. (1994). *Nonlinear dynamics and chaos: With applications to Physics, Biology, Chemics and Engineering*. Readings, Massachusetts: Perseus Books https://scholar.google.es/citations?user=fxyrwlcaaaaj&hl=es&oi=sra.

Sumiko, S. (1974). Impairment of written language in Japanese aphasics: Kana versus kanji processing. *Journal of Chinese linguistics*, *2*(2).

Sumner, P. & Husain, M. (2008). At the edge of consciousness: automatic motor activation and voluntary control. *The Neuroscientist*, *14*(5), 474–486.

Thiel, U. (2011). *The early modern subject: self-consciousness and personal identity from Descartes to Hume*. Oxford University Press.

Thompson, E. (2015). Dreamless Sleep, the Embodied Mind, and Consciousness: The Relevance of a Classical Indian Debate to Cognitive Science. En T. Metzinger (Ed.), *Open MIND*, volume 37(T) (pp. 1–19). Mainz: Frankfurt am Main: MIND Group.

Tirapu-Ustárroz, J., Muñoz-Céspedes, J. M., & Pelegrín-Valero, C. (2003). Hacia una taxonomía de la conciencia. *Revista de Neurología*, *36*(11), 1083–1093.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, *5*(1), 42, https://doi.org/10.1186/1471-2202-5-42 https://link.springer.com/article/10.1186/1471-2202-5-42.

Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *The Biological bulletin*, *215*(3), 216–242.

Tononi, G. (2012). The Integrated Information Theory of Consciousness: An Updated Account. *Archives italiennes de biologie*, *150*(2/3), 290–326, https://doi.org/10.4449/aib.v149i5.1388 http://architalbiol.org/index.php/aib/article/download/15056/23165867.

Tononi, G. (2014). Why Scott should stare at a blank wall and reconsider (or, the conscious grid).

Tononi, G. & Koch, C. (2015). Consciousness: Here, There and Everywhere? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *370*, https://doi.org/10.1098/rstb.2014.0167.

Travis, F. & Pearson, C. (2000). Pure consciousness: distinct phenomenological and physiological correlates of "consciousness itself". *International Journal of Neuroscience*, *100*(1-4), 77–89.

Treisman, A. M. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97–136, https://doi.org/10.1016/0010-0285(80)90005-5 http://www.sciencedirect.com/science/article/pii/0010028580900055.

Tye, M. (1995). *Ten problems of consciousness: A representational theory of the phenomenal mind*. Representation and mind. Cambridge, Mass.: MIT press.

Tye, M. (2003). *Consciousness and persons: Unity and identity*. A Bradford book. Cambridge, Mass.: MIT press.

Tye, M. (2009). *Consciousness revisited: Materialism without phenomenal concepts*. The MIT Press.

Uzgalis, W. L., Ed. (2011). *The Correspondence of Samuel Clarke and Anthony Collins, 1707-08*. Broadview Editions.

Vandenbroucke, A. R. E., Sligte, I. G., Barrett, A. B., Seth, A. K., Fahrenfort, J. J., & Lamme, V. A. F. (2014). Accurate Metacognition for Visual Sensory Memory Representations. *Psychological Science*, *25*(4), 861–873 http://www.jstor.org.bucm.idm.oclc.org/stable/24543526.

Varzi, A. (2019). Mereology. En Zalta, Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Verhulst, P. F. (1838). Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathématique et Physique*, *10*, 113–126 https://ci.nii.ac.jp/naid/10015246307/.

Voltaire (1996). *Elementos de la Filosofía de Newton*. Universidad del Valle.

von der Malsburg, C. (1994). The Correlation Theory of Brain Function. *Models Neural Netw*, *2*, https://doi.org/10.1007/978-1-4612-4320-5{_}2.

von Helmholtz, H. (1853). Ueber einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern, mit Anwendung auf die thierisch-elektrischen Versuche. *Annalen der Physik*, *165*(7), 353–377.

von Helmholtz, H. (1894). Vorwort. En H. Hertz (Ed.), *Die Prinzipien der Mechanik* (pp. vii–xxii). Leipzig: Johann Ambrosius Barth (Arthur Meiner).

Warren, E. W. (1964). Consciousness in Plotinus. *Phronesis*, *9*(2), 83–97.

Watanabe, E. (2022). Flash-lag Effect: Visual Illusion 3D https://www.youtube.com/watch?v=DUBM-GG0gAk.

Watanabe, E., Matsunaga, W., & Kitaoka, A. (2010). Motion signals deflect relative positions of moving objects. *Vision Research*, *50*(23), 2381–2390, https://doi.org/10.1016/j.visres.2010.09.021.

Weiskrantz, L. (1986). *Blindsight: A case study and implications*. Oxford: Clarendon.

Weiskrantz, L. (1996). Blindsight revisited. *Current Opinion in Neurobiology*, *6*(2), 215–220, https://doi.org/10.1016/S0959-4388(96)80075-4.

Wertheimer, M. (1912). Experimentelle Studien uber das Sehen von Bewegung. *Zeitschrift fur Psychologie*, *61* https://ci.nii.ac.jp/naid/10024048816/.

Whitehead, A. N. (1978 [1929]). *Process and reality : an essay in cosmology*. New York: Free Press.

Wiesel, T. N. & Hubel, D. H. (1963a). Effects of visual deprivation on morphology and physiology of cells in the cat's lateral geniculate body. *Journal of Neurophysiology*, *26*(978), 6.

Wiesel, T. N. & Hubel, D. H. (1963b). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of Neurophysiology*, *26*(1003), 17.

Wigner, E. P. (1961). Remarks on the Mind-Body Question. En I. J. Good (Ed.), *The Scientist Speculates* (pp. 284–302). Heineman.

Williford, K. (2015). Representationalisms, Subjective Character, and Self-Acquaintance. En T. Metzinger (Ed.), *Open MIND*, volume 39(T) (pp. 1–27). Mainz: Frankfurt am Main : MIND Group.

Windt, J. M. (2015). Just in Time—Dreamless Sleep Experience as Pure Subjective Temporality: A Commentary on Evan Thompson. En T. Metzinger (Ed.), *Open MIND*, volume 37(C) (pp. 1–34). Mainz: Frankfurt am Main : MIND Group.

Wittgenstein, L. (1953). *Philosophische Untersuchungen*. Oxford: Blackwell.

Woodman, G. F. (2010). A Brief Introduction to the Use of Event-Related Potentials (ERPs) in Studies of Perception and Attention. *Attention, Perception & Psychophysics*, *72*(8), https://doi.org/10.3758/APP.72.8.2031.

Woodward, J. & Ross, L. (2021). Scientific Explanation. En Zalta, Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Yu, F., Jiang, Q.-j., Sun, X.-y., & Zhang, R.-w. (2014). A new case of complete primary cerebellar agenesis: clinical and imaging findings in a living patient. *Brain*, https://doi.org/10.1093/brain/awu239.

Yunesian, M., Aslani, A., Vash, J. H., & Yazdi, A. B. (2008). Effects of Transcendental Meditation on mental health: a before-after study. *Clinical Practice and Epidemiology in Mental Health*, *4*(1), 25, https://doi.org/10.1186/1745-0179-4-25.

Zabaleta, P. (1994). *Ukoreka*. Tafalla: Txalaparta.

Zahavi, D. (2005). *Subjectivity and selfhood: Investigating the first-person perspective*. Cambridge Mass.: MIT press.

Zalta, Edward N., Ed. (2016a). *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2016 edition.

Zalta, Edward N., Ed. (2016b). *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition.

Zalta, Edward N., Ed. (2017). *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, summer 2017 edition.

Zalta, Edward N., Ed. (2018). *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, winter 2018 edition.

Zalta, Edward N., Ed. (2019). *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, spring 2019 edition.

Zalta, Edward N., Ed. (2020). *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, summer 2020 edition.

Zalta, Edward N., Ed. (2021a). *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, fall 2021 edition.

Zalta, Edward N., Ed. (2021b). *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, summer 2021 edition.

Zeimbekis, J. & Raftopoulos, A., Eds. (2015a). *The Cognitive Penetrability of Perception: New Philosophical Perspectives.* Oxford Scholarship Online, 2015 edition.

Zeimbekis, J. & Raftopoulos, A., Eds. (2015b). *The Cognitive Penetrability of Perception: New Philosophical Perspectives : .(2015).T.* Oxford Scholarship Online, 2015 edition.

Zeki, S. & Bartels, A. (1999). Toward a theory of visual consciousness. *Consciousness and cognition*, *8*(2), 225–259.

Zubiri, X. (1989). *Estructura dinámica de la realidad.* Madrid: Alianza Editorial and Fundación Xavier Zubiri.

Zucker, F. (1928). Syneidesis-Conscientia. *Ein Versuch zur Geschichte des sittlichen Bewußtseins im griechischen und griechisch-römischen Altertum.*

Madrid
2022
©Enrique Aramendia Muneta