

**UNIVERSIDAD NACIONAL DE CÓRDOBA
FACULTAD DE FILOSOFÍA Y HUMANIDADES
ESCUELA DE FILOSOFÍA**



Tesis de grado

**LÍMITES Y POSIBILIDADES DE UNA
CIENCIA DE LA CONCIENCIA**

Autor: Horacio A. Chiarella

Director: José Ahumada

Codirector: Aarón Saal

2014

Índice de contenidos

Introducción.....	1
1. La conciencia fenoménica como explanandum de la nueva ciencia de la conciencia.....	5
1.1 Conciencia fenoménica vs. conciencia de acceso.....	6
1.2. Problema difícil.....	7
1.3. Antecedentes del problema difícil: Dos argumentos antimaterialistas.....	9
1.3.1 El argumento de la concebibilidad.....	11
1.3.2. El argumento del conocimiento.....	13
1.4 Evidencia empírica: el caso del overflow fenoménico.....	14
2 Teorías neurocientíficas de la conciencia.....	19
2.1. Global Work-Space Theory (GWST).....	22
2.1.1. Evidencia a favor de la GWST.....	25
2.1.2. Lo que la GWST explica. Y lo que no.....	26
2.2. Teoría de la conciencia como procesamiento recurrente.....	30
2.2.1. Explicatividad del procesamiento recurrente.....	34
3. La estrategia eliminativista.....	37
3.1 ¿Qué son los qualia?.....	38
3.2. Qualia y conciencia fenoménica.....	41
3.3. Quining Qualia.....	43
3.3. ¿Eliminando la conciencia?.....	47
3.4. Eliminando el problema difícil.....	49
4. Ventajas de la estrategia eliminativista.....	51
4.1. Estrategias conceptuales para los argumentos epistémicos.....	51
4.1.1. Matando zombies.....	52
4.1.2. Lo que Mary sabe.....	55
4.2. Teorías rivales y evidencia empírica.....	56
4.2.1. El overflow fenoménico revisado.....	58
4.2.2. El experimento perfecto.....	62
4.2.3. Overflow y subjetividad.....	65

5. Desafíos para el eliminativismo.....	68
5.1. El nuevo explanandum de la ciencia de la conciencia.....	68
5.2. Evidencia de procesamiento cognitivo inconsciente.....	68
5.2.1. Evidencia de procesamiento semántico inconsciente de números.....	69
5.2.2. Integración inconsciente de la información visual.....	70
5.2.3. Condicionamiento afectivo e inconsciente de la conducta.....	71
5.3. La distinción entre los procesos conscientes e inconscientes.....	72
CONCLUSIONES.....	78
REFERENCIAS.....	81

Introducción

Cada día nos despertamos y el mundo se abre a nuestros sentidos: la luz del sol entrando por la ventana, la textura de las sábanas sobre nuestro cuerpo, el olor a café viniendo de la cocina y el despertador, irritante, escondido tras algún libro que ayuda a conciliar el sueño cada noche. Si es viernes nos sentimos afortunados, porque es día de ir al “Sargento” a bailar con La Mona hasta que el calor agobiante producto de la muchedumbre y del precario techo de chapa nos exija descanso, momento en el cual procedemos a calmar el agobio de nuestra garganta con un “pritiado”. Si es domingo juega Talleres, los colores de la camiseta que ya son los colores de tu sangre, el aliento que te pone la piel de pollo, y el destino de una pelota que condiciona las emociones de tantos y tantas...

No puedo evitar preguntarme por qué es que transitamos de este modo la vida, entre goces estéticos, prácticas religiosas, placeres hedónicos, y sufrimientos diversos. En palabras más severas ¿qué es lo que nos hace humanos? El género humano ha reflexionado desde siempre acerca de su condición en el mundo. Dicha condición viene dada por una particularidad que es a la vez objeto y condición de posibilidad de toda reflexión: estamos constituidos por nuestra experiencia consciente y a causa de ella aprehendemos lo que nos rodea desde nuestra propia subjetividad.

En el presente trabajo quiero poner en consideración la posibilidad de abordar la experiencia consciente desde el discurso científico, en particular el de las neurociencias cognitivas.

La historia de las neurociencias cognitivas tiene su origen en el seno de las ciencias cognitivas, por la década del 50. Durante la mayor parte de su corta -y al mismo tiempo rica- historia armó su programa de investigación en torno a la descripción y comprensión del funcionamiento de los distintos sistemas cognitivos, poniendo en relación las funciones y capacidades del nivel psicológico con las estructuras y procesos cerebrales, principalmente a través de explicaciones funcionales y computacionales. Así, el avance de las neurociencias nos ha permitido comprender cada vez con mayor precisión el modo en que las distintas capacidades psicológicas se implementan a partir de procesos neuronales. Los estudios sobre el

modo en el que procesamos la información visual, en el que la memoria almacena la información o en el que procesamos el lenguaje, son algunos de los ejemplos más representativos del aumento significativo en la comprensión que hoy tenemos de las bases físicas de nuestras funciones cognitivas. Pero hubo que esperar hasta los '90 para que las neurociencias intenten dar una explicación de los procesos cerebrales que subyacen a la conciencia.

Bajo el supuesto generalmente aceptado de que la conciencia se origina en procesos cerebrales, y explotando la aparición de nuevos métodos de neuroimagen -cuyo caso paradigmático es la resonancia magnética funcional (fMRI)- y de técnicas para la experimentación como los procesos de enmascaramiento -por ejemplo, las técnicas de *masking* por *continuous flash* suppression- numerosos científicos intentaron fundar una nueva ciencia de la conciencia. Pensadores y científicos de la talla de Crick, Edelman, Dennett, Kandel, Baars, y tantos otros, se lanzaron a la aventura de descifrar el misterio de la conciencia en clave neuronal. Como podremos imaginar todos han dado respuestas muy disímiles: entienden cosas distintas por conciencia y no hay acuerdo en relación al lugar que pueda tener en la explicación de la conciencia el *correlato neural*.

En el año 1995 se publican dos importantes trabajos para el desarrollo de las neurociencias de la conciencia: *On a confusion about the function of consciousness*, de Block, y *Facing up to the problem of consciousness*, de Chalmers. Estos trabajos configuran una parte importante del desarrollo posterior de las neurociencias, fijando como explanandum la conciencia *fenoménica*, pero al mismo tiempo considerando que dicha tarea constituye un *problema difícil*, siendo difícil un eufemismo para imposible. Es decir que estos trabajos suponen un límite para la tarea *explicativa* de las neurociencias. Así las cosas hubo quienes enfocaron desarrollos teóricos aceptando estas limitaciones y orientaron su tarea a la *identificación* de los mecanismos necesarios -aunque no por ellos suficientes- para la producción de la conciencia, de los *correlatos neuronales de la conciencia* [CNC]; mientras que, por el otro lado, también los hubo rebeldes que han pretendido encontrar mecanismos *explicativos* para la conciencia, y no meramente correlacionales.

En la presente tesis me interesa discutir la posición sentada por Chalmers al sostener que **las propiedades interesantes de la conciencia no son caracterizables en**

términos estructurales o funcionales, y de allí que el problema se vuelve inasible al discurso científico. Para ello desarrollaré una estrategia eliminativista a partir de la cual se puede mostrar que el problema difícil no es *tal*, y que aún cuando el problema de la experiencia sea un desafío verdaderamente acuciante no es una tarea inabordable para las ciencias, e.e. se puede comprender en términos de función y estructura. Las neurociencias no sólo están capacitadas para determinar los correlatos neurales de la experiencia consciente (las meras co-ocurrencias neuronales de los estados fenoménicos), sino que además están en condiciones de dar una explicación acerca de cómo y por qué dichos correlatos dan lugar o constituyen la experiencia consciente.

Para defender esta posición es necesario identificar el concepto de conciencia tal como ha pasado a constituir el *explanandum* de las neurociencias de la conciencia: como conciencia fenoménica en tanto distinta de la conciencia de acceso. Esta dicotomía, a su vez, es consistente con aquella que distingue entre problema difícil y problemas fáciles, respectivamente. Esto es desarrollado en el primer capítulo, donde también se presentan una serie de argumentos filosóficos y evidencia empírica que dan apoyo a la idea de que la conciencia es un problema *difícil*.

El segundo capítulo está destinado a mostrar cómo algunas de las teorías actuales de la conciencia más importantes, como lo son la teoría del espacio de trabajo global y la teoría de la conciencia como recurrencia neuronal, fracasan en el intento de explicar las particularidades de la experiencia consciente. Las dos teorías que se consideran son casos válidos y representativos en tanto son teorías materialistas, cuyo propósito es explicativo, y en tanto representan las dos posiciones posibles en relación al vínculo entre la conciencia y el sistema cognitivo (o la conciencia es cognitiva o no lo es).

En el tercer capítulo se introduce un argumento eliminativista respecto de la noción de *qualia* y a partir de allí de la noción de conciencia fenoménica, en tanto contrapuesta a una conciencia de acceso. Me interesa mostrar cómo esta eliminación nos lleva a la disolución del problema difícil.

En el cuarto capítulo presento las ventajas de la estrategia eliminativista en relación a dos elementos de análisis: por un lado se pretende salvar la *brecha explicativa* (Levine, 1983) considerada en el problema difícil -y en los argumentos filosóficos que lo sustentan-; y por el otro evaluar y decidir entre aquellas teorías o

familias de teorías rivales que forman parte del panorama actual en neurociencias de la conciencia.

Por último, el quinto capítulo presenta una serie de desafíos a los que debe responder el eliminativismo si es que ha de erigirse en una estrategia válida para sortear el problema difícil.

La conciencia fenoménica como explanandum de la nueva¹ ciencia de la conciencia

Parece que nada de lo que nos importa nos importaría del mismo modo si no tuviéramos ciertos estados de *vigilia* durante los cuáles podemos reconocernos a nosotros mismos como sujetos de las experiencias más disímiles. Es manifiesto que somos conscientes mientras estamos despiertos y hasta que uno se duerme (Searle, 1990) o, si se es menos afortunado, hasta que uno se desmaya, cae en coma o muere. En estos últimos casos decimos que estamos inconscientes o perdimos la conciencia. A veces también decimos que alguien es consciente de algo cuando tiene conocimiento de ese algo, cuando puede decírnoslo, o porque atiende y responde a los estímulos y tareas que le proponemos. Hay un uso variopinto, flexible, y por ello algo confuso, del concepto de conciencia. ¿Cuánto de explicar el estar despierto es explicar la conciencia? ¿Cuánto de explicar la atención se requiere para explicar la conciencia? Si explicamos cómo es que alguien puede responder nuestras preguntas o cumplir una serie de tareas ¿habremos explicado por qué esa persona es consciente?

El objetivo del presente capítulo es identificar el *explanandum* de la nueva de la ciencia de la conciencia, es decir, determinar a qué nos referimos cuando hablamos de conciencia, cuál es el fenómeno que nos interesa explicar. Una vez caracterizado nuestro objeto de estudio presentaré, a través de Chalmers, la particular dificultad a la que dicho estudio se enfrenta. Por último mostraré de qué modo se haya inserto este problema en la tradición filosófica reciente, a partir de argumentos que lo anteceden y lo complementan.

1 Al referirme a una “nueva ciencia de la conciencia” estoy haciendo extensivas las consideraciones que hace Gardner (1985) al hablar de la “nueva ciencia de la mente”. El estudio de la conciencia se ha mantenido hasta hace unos pocos años al margen del estudio de las ciencias cognitivas, y sólo con la aparición de nuevos métodos de neuroimagen y con la aparición de teorías neurobiológicas destinadas a identificar los patrones neuronales de la conciencia, ésta pasó a conformar parte de las ciencias cognitivas en un sentido pleno.

1.1 Conciencia fenoménica vs. conciencia de acceso

A través de nuestros sentidos el mundo se nos presenta como siendo de una determinada manera. Somos conscientes del particular azul de la camiseta de Talleres, o percibimos las (des)afinaciones de La Mona como constitutivas de su estilo particularmente popular. No sólo atendemos y respondemos a estímulos, sino que experimentamos los gustos, las texturas, sonidos, colores y olores como poseyendo cierta cualidad y no otra. Nuestro sentido del oído, por ejemplo, no es sólo un radar que capta ciertas ondas que vibran a cierta frecuencia y las procesa de modo tal que puede establecer cierta respuesta conductual; percibimos esas ondas como poseyendo rasgos cualitativos distintos del proceso físico, a saber, una altura distintiva (la nota La a 880 Hz suena una octava más aguda que un La a 440 Hz), intensidad (o en un lenguaje más coloquial “volumen”), duración y timbre (dos notas de misma altura, intensidad y duración aún pueden tener un *color* distintivo: sonar más metálica, por ejemplo, si es ejecutada cerca del puente de una guitarra, o más dulce y redonda si la cuerda se pulsa sobre la boca del instrumento). Del mismo modo ocurre con nuestros estados emocionales. Cuando digo que siento miedo no me refiero a un *mero* estado corporal de alerta ejecutado por mi organismo, un mecanismo de autodefensa que me previene del peligro. Cuando siento miedo dicho estado se manifiesta como cualitativamente significativo para mi, o para quien sea que lo sufra. Percibimos el mundo que nos rodea como siendo de tal y cual modo y dicha percepción es un estado subjetivo. En esta dirección se dirige Nagel (1974) cuando dice que “un organismo tiene estados mentales conscientes si y sólo si hay algo que es *ser como* ese organismo – algo que es como ser ese organismo *para* ese organismo” (p. 436).

Hablar de rasgos cualitativos en este sentido nos remite al carácter fenoménico de la experiencia. Haciendo uso de la definición de Nagel antes citada acerca de la conciencia decimos que un estado fenoménico comporta un conjunto de propiedades que determinan cómo se siente [*what it is like*] tener esa experiencia. A su vez, y por esto mismo, la descripción de la experiencia en términos fenoménicos está ligada a su carácter subjetivo: toda experiencia se manifiesta en primera persona, desde una perspectiva subjetiva, la experiencia y sus cualidades fenoménicas se presentan como imprescriptiblemente nuestras. Es este carácter

experiencial de la conciencia lo que constituye su rasgo más distintivo y, por ello mismo, el que nos ocupará en adelante. Explicar la conciencia, entonces, no es otra cosa que explicar su naturaleza fenoménica.

El primer intento de sistematizar la noción de conciencia como conciencia fenoménica, dentro del debate en neurociencias, lo debemos a Block (1995), quien distingue la idea de conciencia fenoménica, que hemos retratado hasta aquí, de otras originadas en usos frecuentes, pero que no tienen éxito en capturar el fenómeno que nos interesa estudiar sino que lo confunden con algo distinto, en particular con lo que Block llama conciencia de acceso.

Un estado consciente es de acceso si su contenido está disponible para el control del razonamiento y de la conducta. La conciencia de acceso está conformada por las distintas habilidades y funciones cognitivas por las que accedemos a los estados fenoménicos de nuestra experiencia, quedando claro que aquello a lo que se accede es distinto del mecanismo cognitivo que provee el acceso. La conciencia de acceso a diferencia de la conciencia fenoménica, es eminentemente funcional y en tal sentido es relativa a un sistema o subsistema cognitivo particular (Block, 1995). Hay conciencia de acceso cuando el contenido de un estado mental está disponible para el funcionamiento de algún o algunos sistemas cognitivos: memoria, habilidades lingüísticas y muy particularmente atención (tan particularmente que algunos [Lamme, 2003n; Koch y Tsuchiya, 2007] consideran la atención como el caso único de acceso, pero este es tan craso error como el que Block procura evitar al distinguir conciencia fenoménica de conciencia de acceso).

1.2. Problema difícil

La distinción entre conciencia fenoménica y conciencia de acceso no sólo es interesante porque nos permite identificar las propiedades de la conciencia que interesan a nuestra investigación, a saber, el carácter cualitativo, subjetivo, experiencial de los estados conscientes. Esta distinción además condiciona el modo en que podemos explicar tales estados: cuando decimos que lo que nos ocupa es la conciencia fenoménica, decimos que son las propiedades fenoménicas, y no otras, las que requieren explicación. Al preguntarnos si es posible una ciencia de la conciencia, nos estamos preguntando si las neurociencias pueden explicar la conciencia

fenoménica, y no alguno de los tipos de conciencia que Block llama *de acceso*.

Chalmers acepta la distinción entre conciencia de acceso y conciencia fenoménica, y establece en torno a ella la posibilidad de éxito por parte de las ciencias para explicar la conciencia: Si por conciencia nos referimos a conciencia de acceso, entonces las neurociencias cognitivas pueden, y de hecho es lo que hacen, explicar tales estados. Si lo que nos interesa, en cambio -y éste es nuestro caso- es la conciencia *qua* conciencia fenoménica, entonces las ciencias no tienen ninguna chance de alcanzar una explicación genuina del problema de la conciencia. Para decir esto mismo Chalmers habla de problemas fáciles de la conciencia, en el caso de aquellos originados por la conciencia de acceso, y del problema difícil, que no es otro que el problema de la experiencia consciente (siendo difícil un modo de decir imposible).

Son problemas fáciles aquellos que pueden ser explicados en términos de función² y estructura del modo en que lo han hecho usualmente las ciencias cognitivas y particularmente las neurociencias cognitivas, adoptando un vocabulario computacional y apelando a mecanismos neurales (Chalmers, 1995). Así como difícil es un eufemismo para imposible, fácil no es precisamente la mejor descripción de la tarea que se encomienda a las neurociencias, sino que es la *attenuatio* para “empresa más grande que el hombre haya enfrentado, pero accesible”. Entre la extensa lista de problemas fáciles que Chalmers confecciona podemos destacar la habilidad para discriminar, la reportabilidad de los estados mentales, la atención y el control de la conducta. En todos los casos hay un aspecto cognitivo involucrado.

El problema difícil de la conciencia, por otra parte, es el problema de la experiencia. Si hay una pregunta que está más allá del alcance de las ciencias, dice Chalmers, esa es “¿a qué se debe el carácter fenoménico de nuestros estados mentales?”. ¿Qué procesos físicos darían lugar a algo como la conciencia fenoménica, al modo cualitativo propio y específico de nuestras experiencias? Podemos formular el argumento para el problema difícil -el argumento explicativo, según Chalmers (2003)- del siguiente modo:

² Es necesario aclarar, en relación a los términos “función” y “funcionalismo” (términos harto frecuentes en esta tesis), que los mismos están usados en un sentido amplio para señalar tan sólo que “la materia importa sólo por lo que la materia puede hacer. El funcionalismo, en un sentido amplio es tan ubicuo en ciencias que se ha convertido en equivalente a un supuesto reinante en toda práctica científica” (Dennett, 2005: p. 17).

- (1) Las consideraciones físicas puede explicar, a lo sumo, estructura y función.
 - (2) Explicar la estructura y la función no bastan para explicar la conciencia.
- Entonces,
- (3) Ninguna consideración física puede explicar la conciencia.

Mientras que una función cognitiva es precisamente eso, una función, de modo que para explicar tal función no tenemos que hacer más que señalar el mecanismo que la ejecuta y el modo en que lo hace, el problema de la experiencia y su emergencia excede al problema de la función de un mecanismo y su ejecución. Esto no significa que la conciencia no tenga una función, sino que tal función no agota el problema de la experiencia, del carácter subjetivo y cualitativo de la misma. Chalmers asegura que aún cuando logremos explicar todos los rasgos funcionales que pueden formar parte de un estado consciente, seguirá quedando sin explicar por qué ese estado está asociado a una expresión cualitativa particular, por qué va acompañado de una expresión fenoménica. Imaginemos, por ejemplo, que la conciencia tiene por función la integración de los estímulos: aún cuando logremos explicar tal integración no sabremos cómo los dispositivos materiales involucrados dieron lugar al aspecto cualitativo-fenoménico de la experiencia. Imaginemos ahora que la función de la conciencia es poner a disposición de los distintos subsistemas cognitivos aquellos contenidos procesados por un subsistema en particular. Aún cuando expliquemos el modo en que estos contenidos son procesados primero por un sub-sistema y luego puestos a disponibilidad global por la conciencia, seguiremos sin tener una explicación de cómo entran en ese escenario las propiedades fenoménicas de la experiencia.

1.3. Antecedentes del problema difícil: Dos argumentos antimaterialistas.

Las dos distinciones conceptuales hasta ahora desarrolladas -conciencia fenoménica/conciencia de acceso y problemas fáciles/problema difícil- no encajan como dos piezas de un mismo rompecabezas, son en realidad la misma pieza. En ambos casos se manifiesta la particularidad del problema de la *experiencia* -subjetiva, fenoménica- en contraste con aquellos otros fenómenos mentales que son

caracterizados en términos de funciones cognitivas, aún cuando puedan estar asociados a la experiencia consciente. En ambos casos las distinciones realizan un corte idéntico entre el tipo de fenómenos a explicar y las particularidades que implican dar cuenta de unos y otros. Esta correspondencia da cuenta del amplio predominio que estas dicotomías forjaron en el campo de las neurociencias desde el momento en que aparecieron. Por supuesto, tal predominio también está apoyado en una tradición filosófica que desde siempre ha prestado atención al misterio de la conciencia y su carácter fenoménico, así como a las particulares dificultades explicativas que ésta implica.

El problema difícil puede leerse como una reformulación de una designación anterior, introducida por Levine (1983), que habla de una “brecha explicativa” (*explanatory gap*) entre el tipo de explicaciones que requieren los estados subjetivos y el que nos permiten las ciencias y la consideración de *lo físico*. Un antecedente aún más lejano es el de Leibniz, quien en los albores del siglo XVIII, señalaba la imposibilidad de explicar la conciencia a partir de mecanismos -la brecha explicativa- del siguiente modo:

Nos vemos obligados, por lo demás, a confesar que la percepción, y lo que depende de ella, es inexplicable por razones mecánicas, es decir, por medio de figuras y movimientos. Supongamos pues la existencia de una máquina cuya estructura le permite pensar, sentir y ser capaz de percepción, y que sea posible ingresar en ella como en un molino. Y una vez supuesta, si la examinamos por dentro, no hallaremos sino unas piezas que trabajan unas sobre otras, pero nunca nada que explique una percepción. Así pues, habrá que buscar esa explicación en la substancia simple y no en el compuesto o en la máquina (Leibniz, 2005/1714, p. 40).

El experimento mental del gran molino, sugerido por Leibniz, ha cobrado numerosas formas recientemente: de sala china (Searle, 1980), nación china (Block, 1978) e incluso zombies (Chalmers, 1995). En todos los casos se destaca la aparente imposibilidad de reducir la conciencia a un conjunto de mecanismos y sus funciones, es decir, ilustran la brecha explicativa antes mencionada. Al mismo tiempo ponen de manifiesto una intuición generalizada según la cual la naturaleza de la conciencia viene dada por algún factor *especial*, no físico, y esta es la razón por la que es insalvable

a las explicaciones que se limitan a la consideración de mecanismos y funciones. Cuando nos preguntamos, entonces, si las neurociencias pueden explicar la conciencia nos estamos preguntando si puede haber una explicación *materialista* de la conciencia, es decir, que esté fundada en procesos físicos.

A continuación quisiera hacer referencia a dos argumentos que considero complementarios al problema difícil, y que definen un *vacío epistémico* a partir del cual se infiere un *salto ontológico* (Chalmers, 2003; Levine, 1983), por esta razón los llamaremos -para referirnos a ambos- argumentos *epistémicos* contra el materialismo. Si hacia el final de la tesis encontramos que estos argumentos siguen pareciendo convincentes entonces el problema difícil es tan difícil como pretende serlo. Si, en cambio, hemos logrado desarmarlos de modo tal que ya no resultan creíbles, entonces probablemente, aún cuando las teorías desarrolladas hasta el momento no sean lo suficientemente explicativas, podremos decir, al menos, que la ciencia está en condiciones potenciales de explicar la conciencia.

1.3.1 El argumento de la concebibilidad

En los últimos años se ha establecido una gran discusión en torno a la posibilidad de concebir *zombies filosóficos*: organismos que viven en un mundo idéntico al nuestro -cuya descripción partícula a partícula es idéntica a la nuestra-, y por lo tanto con los mismos estados cerebrales que los nuestros, pero que sin embargo no tienen ninguna experiencia como correlato de dichos estados. Los zombies son idénticos a nosotros, se comportan como nosotros, viven nuestra misma cotidianidad, se comunican, divierten e incluso afirman tener una conciencia fenoménica (cuya naturaleza también deben elucidar). La única diferencia es que ser zombie no se siente de ningún modo, para un zombie no hay nada que sea *ser cómo* (en referencia a Nagel) un zombie.

Un antecedente famoso de los zombies fue formulado por Locke (1975) y es la hipótesis del espectro invertido: “el mismo objeto podría producir en la mente de varios hombres diferentes Ideas al mismo tiempo; v.g. la *Idea*, que una Violeta produce en la mente de un hombre por sus ojos, mientras que al mismo tiempo es producida por un girasol en la de otro, y viceversa” (p. 89, mi traducción). Como estas hay otras formulaciones (Tye, 2013), aunque en este trabajo nos limitaremos a

considerar la versión de los zombies trabajada por Chalmers (1996) en el contexto propio de nuestra discusión. Empecemos por preguntarnos ¿Por qué es tan importante que podamos imaginar este tipo de casos?

Tal como Kripke defendiera en las célebres *Lectures* en Princeton publicadas en *Naming and Necessity* (1980), “la identidad no es una relación que se pueda mantener contingentemente entre objetos” (p.154). De acuerdo con Kripke sólo podemos predicar identidad en aquellos casos donde esta identidad sea necesaria; de hecho la necesidad de la identidad se sigue deductivamente de aceptar la identidad entre dos objetos, digamos x e y , y de aceptar que la identidad entre cada objeto y sí mismo es necesaria. Si lo pensamos en términos de mundos posibles decimos que podemos atribuir identidad cuando no podemos pensar un mundo posible donde dicha identidad no sea verdadera. En efecto estas consideraciones se cumplen para los nombres propios, como *phosphorus* y *hesperus*, y para las identidades *teóricas*, descubiertas por la ciencia, como el hecho de que el agua sea H₂O y el calor movimiento de partículas (Kripke, 1971).

La lección que obtenemos de Kripke es que si podemos imaginar un zombie filosófico, cuyo mundo sea materialmente idéntico al nuestro pero fenoménicamente distinto, entonces los zombies son *concebibles*, y si son concebibles hay un mundo posible donde la identidad entre nuestro mundo material y nuestro mundo fenoménico se rompe. Pero dado que las identidades no pueden ser contingentes (al menos no las que nos interesan), entonces no hay una identidad entre la realidad material y la realidad fenoménica. El argumento ligeramente formalizado quedaría así:

- (1) Los zombies filosóficos son concebibles.
- (2) Si son concebibles entonces son metafísicamente posibles.
- (3) Si es metafísicamente posible la existencia de zombies, la relación entre los estados fenoménicos y los hechos físicos es contingente.

Entonces,

- (4) La conciencia no puede ser idéntica a un hecho físico

1.3.2. El argumento del conocimiento

En el argumento explicativo del problema difícil se considera como una de las premisas que hay algo que no es funcional ni estructuralmente describable -la conciencia-. Una forma más general de poner esta idea es que nuestro conocimiento de los hechos físicos no nos conduce al conocimiento de los hechos fenoménicos, existiendo una brecha epistémica de la cuál se apela a una brecha ontológica entre uno y otro tipo de hechos (físicos y fenoménicos). Este modo de argumentar está bien ilustrado por el famoso pasaje de Jackson (1986) donde presenta el caso de la neurofisióloga Mary:

Mary es una científica brillante que es forzada, por la razón que fuera, a investigar el mundo desde una sala blanca y negro a través de un televisor en blanco y negro. Ella se especializa en neurofisiología de la visión y tiene, vamos a suponer, toda la información física que hay disponible acerca de lo que sucede cuando vemos tomates maduros, o el cielo, y usamos el término 'rojo', 'azul', y demás. Ella descubre, por ejemplo, qué combinación de ondas del cielo estimula la retina, y exactamente cómo esto produce a través del sistema nervioso central la contracción de las cuerdas vocales y la expulsión de aire de los pulmones que resulta en la expresión de la oración 'El cielo es azul'.... ¿Qué pasará cuando Mary sea liberada de su sala blanca y negra o tenga acceso a un televisor color? ¿Aprenderá algo o no? Parece obvio que ella aprenderá algo del mundo y de nuestra experiencia visual. Pero entonces es un hecho que su conocimiento previo era incompleto. Pero ella tenía *toda* la información física. *Ergo* hay algo más para tener que el conocimiento físico, y el fisicalismo es falso (p. 130, mi traducción).

La intuición generalizada parece sugerir, en acuerdo con Jackson, que no basta conocer todos los hechos del mundo físico para saber *a priori* cuál será la expresión cualitativa específica del *azul* cuando miramos al cielo una tarde despejada. Dado que no podemos conocer estas propiedades fenoménicas a partir del conocimiento de los hechos físicos, entonces tampoco podemos explicarlas a través de su estructura y función. El argumento se puede formalizar del siguiente modo:

(1) Mary tiene toda la información física concerniente a la visión humana del

color.

(2) Mientras está encerrada en su laboratorio blanco y negro Mary no tiene toda la información disponible del color.

Entonces

(3) No toda la información física del color es física.

Los tres argumentos desarrollados hasta el momento -explicativo, zombie y del conocimiento- dan forma a supuestas limitaciones de las ciencias para dar una respuesta al problema de la conciencia en términos materialistas. Además, todos estos argumentos, comparten una misma estructura lógica, ya que todos apelan a una brecha explicativa a partir de la cual se infiere una brecha ontológica (Chalmers, 2003). Por lo tanto podemos decir que:

(1) Hay una brecha epistémica entre las verdades físicas y fenoménicas.

(2) Si hay una brecha epistémica entre las verdades físicas y fenoménicas entonces hay una brecha ontológica.

Entonces,

(3) El materialismo es falso.

El hecho de que los tres argumentos epistémicos tengan la misma forma no sólo habla de un propósito y estrategia compartida, sino también de los desafíos y oportunidades que tiene por delante cualquier defensa de una posición materialista. Por ejemplo, una respuesta oportuna a uno de los argumentos seguramente nos dará una pista acerca de la premisa equivalente que debemos atacar en los argumentos restantes. Por el contrario, si atacando uno de los argumentos no podemos lidiar con los otros entonces es posible que haya que cambiar de estrategia.

1.4 Evidencia empírica: el caso del *overflow* fenoménico

Por último me gustaría considerar un argumento de carácter empírico, que se halla presente a lo largo de gran parte de la literatura científica, y apela al fenómeno del *overflow*, expresión por la que se designa el desborde del procesamiento fenoménico respecto del procesamiento cognitivo que facilita el

acceso. La dicotomía entre conciencia fenoménica y conciencia de acceso sugiere que ambos procesos funcionan de forma independiente, no supone una distinción meramente formal, sino que expresa una convicción acerca de los hechos del mundo. En este sentido la mejor evidencia para apoyar la distinción es encontrar los casos en los que se manifiesta sólo uno de los tipos de conciencia.

Hay amplia evidencia, que goza de un consenso casi absoluto, de que existen estados de acceso sin su contraparte fenoménica: procesamiento visual, memoria, control motor y toma de decisiones (Berlin, 2011). Sin embargo hay una discusión que divide las aguas: ¿pueden haber manifestaciones de la conciencia fenoménica a las que no tengamos acceso?

Existe una serie de experimentos que señalan que una parte de la experiencia fenoménica excede o desborda (en inglés *overflow*, término que usaremos en adelante) a los estados conscientes de acceso. El primero de los experimentos que apoyan la hipótesis del *overflow* fenoménico fue diseñado por Sperling (1960). En el experimento se presenta un conjunto de estímulos -letras dispuestas tal como señala la figura 1-, y se solicita al observador que brinde un reporte completo de su experiencia. Predeciblemente el reporte de los observadores sólo da cuenta de un número limitado de estímulos. Sin embargo el reporte del sujeto incluye la afirmación de que puede ver más de lo que puede reportar. Por ejemplo, sobre una presentación de doce letras dispuestas en tres filas de cuatro letras cada una, el sujeto puede reportar correctamente sólo una de estas líneas, y aún así afirma que ve más que esas cuatro letras que puede reportar. Esto parece confirmar que la experiencia desborda el acceso: “ver más de lo que es recordado implica que más información está disponible durante, y tal vez por un corto tiempo después de que el estímulo puede ser reportado” (Sperling, 1960, p. 1, mi traducción).

Para corroborar la intuición expresada por los sujetos, y determinar si efectivamente hay manifestaciones fenoménicas más allá de los límites impuestos por las capacidades cognitivas que nos brindan acceso (principalmente la atención y la memoria inmediata), Sperling pasa de un paradigma de reporte completo a uno de reporte parcial. En este caso se solicita al observador que responda sólo por una de las líneas de letras presentadas. Esta línea es sugerida por medio de un tono que suena con posterioridad a la presentación del estímulo y que funciona como señalador; por ejemplo, si suena un tono agudo el sujeto debe reportar la línea de

letras superior y si suena un tono bajo la línea inferior. Los resultados de estos experimentos muestran que los sujetos responden con acuracidad independientemente de la línea señalada, y, si tomamos el reporte parcial como una muestra aleatoriamente seleccionada por el tono, entonces debemos concluir que el porcentaje de información total disponible (no sólo aquella seleccionada para ser reportada) es mayor al que determina el límite de la memoria inmediata (Sperling, 1960). Esta información excedente es almacenada por un tipo de memoria que se denominó *icónica*.

F C H D
J R P O
D N B A

Figura 1. Modelo de tarjeta estímulo utilizada por Sperling. Los sujetos deben reportar la mayor cantidad de estímulos posibles durante el paradigma de reporte completo, pero sólo aquellos objetos especificados en el caso del reporte parcial.

Estos experimentos han sido reiterados recientemente en combinación con el paradigma de ceguera al cambio (Landman, Spekreijse y Lamme, 2003; Becker, Pashler y Anstis, 2000). Los experimentos de ceguera al cambio tienen por objetivo determinar la capacidad para detectar cambios en los estímulos presentados, tal como sucede con el clásico experimento del gorila, donde un primate atraviesa un grupo de personas que juegan con una pelota sin ser percibido en un número significativo de casos (Simons y Chabris, 1999). La combinación de la ceguera al cambio con el paradigma de reporte parcial se traduce en un experimento que solicita el reporte de la detección del cambio en *alguno* de los estímulos procesados. De manera análoga al experimento de Sperling se apela a una pista visual que se reproduce una vez que ha desaparecido el primer conjunto de estímulos (figura 2). Los resultados de los experimentos fueron consistentes con los de Sperling, y muestran que con anterioridad al acceso debe haber estados fenoménicos que expliquen cómo es que los sujetos pueden detectar el cambio con precisión ante un

número de estímulos que excede la capacidad de la memoria de trabajo. Siendo que los estímulos a los que se tiene acceso *efectivo* son limitados por la memoria de trabajo, es necesario suponer que existen estados de *eventual* o *potencial* acceso para explicar la arbitrariedad con que los mismos son accedidos o *escogidos* por los mecanismos cognitivos.

Tanto el experimento original de Sperling como las versión adaptadas a la ceguera al cambio toman los recaudos necesarios para evitar que haya algún mecanismo cognitivo particular que esté tomando parte en la determinación de estados que, de ser así, no serían propiamente fenoménicos. Recordemos que *identificar, atender* o *predecir*, junto con toda *conceptualización, memorización* o *vehiculación* de la información responden a formas de acceso, y nos desvían del *explanandum* propio del *problema difícil*: la conciencia de acceso. Ejemplos de ello son la determinación de Sperling por asegurar que los procesos atencionales no estén interfiriendo en el reporte de la línea sugerida por el tono (p. 24), y Landman, Sprekjeise & Lamme (2003) , que buscan *distraer* la atención, a través de pistas falsas, a líneas de letras distintas de aquellas que serán finalmente requeridas para el reporte por medio de la pista visual (pp. 158 - 160).

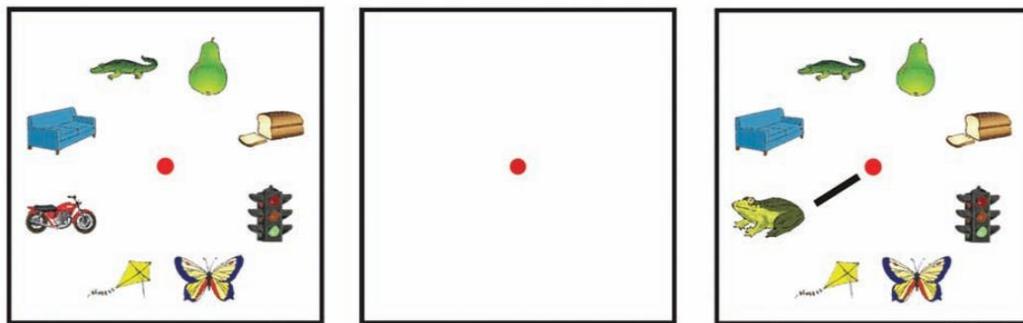


Figura 2. Ejemplo de los estímulos utilizados por Landman, et al. Las imágenes aparecen en orden de progresión, y la pista visible permite recuperar la información que permite detectar si hubo un cambio -o no- en el estímulo señalado.

Esta evidencia parece señalar la existencia del overflow fenoménico, lo cual nos permite inferir que la distinción entre conciencia fenoménica y conciencia de acceso no es una distinción meramente formal, sino que responde a la existencia de propiedades y/o entidades mentales distintas y en cierto grado autónomas. Si hubiéramos de formalizar este argumento, y su relación con el problema difícil diríamos algo así:

(1) Si existe el overflow (estados fenoménicos a los cuáles no hay acceso en primera persona), entonces el *explanandum* de la nueva ciencia de la conciencia es la conciencia fenoménica.

(2) Si el *explanandum* es la conciencia fenoménica en tanto distinta de los problemas de acceso, entonces el mismo constituye un *problema difícil*.

(3) El overflow existe, y esto se puede determinar experimentalmente.

Entonces,

(4) La nueva ciencia de la conciencia debe encargarse del estudio de la conciencia fenoménica.

Hasta aquí he introducido algunos de los conceptos más aceptados en relación al estudio de la conciencia por parte de las neurociencias. Nos hemos referido al modo en que se entiende la conciencia y a las dificultades que le son propias a tal definición. Este panorama conceptual nos ha puesto de cara a una supuesta brecha entre lo que pretendemos explicar y el tipo de explicaciones del que disponemos. Habiendo propiedades de la conciencia fenoménica distintas de aquellas que exceden a las propiedades que podemos considerar estructural y funcionalmente, se sigue que nos enfrentamos a un problema *difícil*.

El próximo capítulo está destinado a considerar la dificultad del problema difícil en relación a dos teorías neurocientíficas de la conciencia particulares: la *teoría del espacio de trabajo global* (GWST, por sus siglas en inglés) y la *teoría del procesamiento recurrente*.

Teorías neurocientíficas de la conciencia

Si bien las distinciones trazadas en el primer capítulo parecen suponer un futuro poco auspicioso para las neurociencias de la conciencia, hay un gran número de teorías que han pretendido realizar algún aporte para la comprensión del fenómeno de la conciencia. Chalmers (1998) reconoció, hace casi veinte años, una rica fauna de más de una veintena de mecanismos y procesos cerebrales que fueron postulados como un correlato de la conciencia. Ese número ha crecido sin ritmo hasta hoy, poniéndonos ante un panorama complejo y diversificado, que nos lleva a preguntarnos cómo se relaciona esta superabundancia teórica con el análisis escéptico de Block y Chalmers respecto a la posibilidad de éxito explicativo de la ciencia de la conciencia.

El objetivo principal de este capítulo es dimensionar el problema difícil en toda su problematicidad, y para ello es necesario evaluarlo en relación a las teorías particulares a las cuáles refiere. Dado que no podemos analizar todos los aportes teóricos efectuados en las últimas dos o tres décadas es necesario efectuar un recorte. Para ello hay tres consideraciones que me parecen relevantes y que pretendo desarrollar a continuación: (1) la distinción entre correlatos neurales de la conciencia (CNC) y teorías explicativas; (2) la distinción entre teorías disociativas y no disociativas; y (3) el supuesto del materialismo.

En cuanto a la primera distinción es cierto que la mayoría de las teorías no tiene una pretensión *explicativa* y se limitan a la búsqueda de *meros* correlatos neuronales. Crick y Koch (1990) adoptaron la expresión “correlato neural de la conciencia” (CNC) para referirse al “conjunto mínimo de mecanismos o eventos *suficientes* [la cursiva es añadida] para una experiencia consciente” (p?). Su contribución específica estaba dirigida a mostrar que la oscilación neuronal sincrónica entre los 40-70 hz es el (o uno de los) mecanismo a la base de toda experiencia consciente, pero lo que me interesa en este caso es la elección por parte de los autores del término “correlato”, en un claro intento de evitar un compromiso con una “relación de causalidad entre la ocurrencia de eventos mentales conscientes

y su asociada estructura fisiológica” (Koch, 2012, p. 25, mi traducción).

Dado que lo que me interesa es considerar las vías posibles para enfrentar al problema difícil, y defender la posibilidad de explicar la conciencia a partir de sus bases físicas -y no solamente correlacionarse-, las teorías que buscan un CNC sin asumir los desafíos explicativos no serán tenidas en cuenta. Esta tarea constituye un aporte valioso en relación a la identificación de los mecanismos y procesos específicos que participan en la actividad consciente, aún cuando tenga sus dificultades propias (como el llamado problema de la medición [Block, 2013]). Pese a ello no supone un desafío *explicativo* y por lo tanto tampoco supone un modo de evadir el problema difícil, razones por las cuales dejaremos de lado aquellas teorías comprometidas con la búsqueda del CNC pero que prescinden del cómo y el porqué de la conciencia. Esto no significa que no tenga en consideración la importancia de determinar el CNC, pero dado que la discusión que me interesa está atada a la posibilidad de *explicar* la conciencia dejaré de lado a las teorías que se limitan a la identificación de los CNC.

En segundo lugar me interesa evaluar las teorías en tanto disociativas o no disociativas (Kouider, Sid, de Gardelle, Sackur & Dupoux, 2010). Entre las primeras encontramos aquellas que llevan adelante una búsqueda de mecanismos y/o procesos *específicos* de la conciencia fenoménica, mientras que las segundas consideran que sea lo que sea la conciencia no puede buscarse más allá de los mecanismos cognitivos que proporcionan el acceso. Las teorías disociativas disocian la conciencia fenoménica de la conciencia de acceso, mientras que otras las identifican. Esta distinción, que Block (2014) también ha señalado en términos de teorías cognitivas vs. teorías no cognitivas, me parece particularmente relevante porque supone dos modos empíricamente excluyentes de encarar el problema de la experiencia consciente. Cualquier argumento que tengamos a favor de una de las líneas de investigación, sabremos que también funciona como un argumento en contra de la línea opuesta. Necesariamente resolver el meollo supone un avance para la investigación. A su vez, si primero recortamos las teorías que nos interesan en relación a sus pretensiones *explicativas*, parece lógico que el segundo recorte determine el *explanandum* de la pretendida explicación: Mientras que las teorías disociativas o no cognitivas intentarán explicar un fenómeno *sui generis*, posibilitado por mecanismos y procesos específicos, las teorías cognitivas o no disociativas

reducirán su explicación a otros problemas cognitivos o de acceso. Esto está directamente relacionado con el argumento del overflow fenoménico, ya que las teorías disociativas parecen apoyarse en evidencia de este tipo, defendiendo que hay estados fenoménicos sin acceso (tal como dijimos en el apartado 1.4). Por el otro lado, las teorías no disociativas aceptan que hay acceso sin conciencia fenoménica, pero niegan que pueda darse la inversa, es decir, conciencia fenoménica sin acceso. Toda experiencia consciente, para las teorías no disociativas, es una experiencia efectivamente accedida. Dicho de otro modo para estas teorías toda conciencia es cognitiva.

Otra forma de considerar la diferencia entre uno y otro tipo de teorías es que unas consideran la conciencia en tanto los efectos descendentes del procesamiento cognitivo (reacciones a estímulos, reportes verbales, efectos en la memoria, etc) o como las causas ascendentes que dan origen a la experiencia (actividad cerebral en una región cortical u otra). Las primeras son teorías cognitivas mientras las segundas son disociativas.

Por último, sólo consideraré aquellas teorías que dan continuidad al carácter naturalista y materialista de la investigación científica, algo que considero un rasgo fundamental y distintivo de las ciencias. Naturalista porque nadie diría que una explicación que apela a eventos sobrenaturales es científica, debemos poder reconstruir el fenómeno que ocupa nuestro estudio de alguna forma que lo haga a partir de las entidades y fuerzas conocidas y familiares de la naturaleza. En tal sentido, y como la etimología que remontamos a los griegos sugiere, la naturaleza es paradigmáticamente física, y lo físico no puede ser distinto de lo material. Si hay algo que no puede ser explicado en términos de las propiedades o el comportamiento de la materia, entonces este factor *emergente* constituye algún tipo de dualismo.

El requisito materialista explica por qué queda de lado en este capítulo el análisis de la propuesta de Chalmers (1995) para construir una teoría no reduccionista, ya que su “posición califica como una variedad de dualismo, en tanto postula propiedades básicas más allá de las propiedades invocadas por la física” (p. 210). Por supuesto que desde su perspectiva “it is an innocent version of dualism, entirely compatible with the scientific view of the world”, pero dicha compatibilidad descansa en un supuesto tan controvertido como aceptar propiedades “over and above” las propiedades físicas, y no hay consenso en absoluto acerca del tipo de

entidades no materiales que podemos incluir en nuestras explicaciones sin que las mismas dejen de explicar, que es justamente lo que está en juego en el problema difícil.

Una vez dicho esto ya hemos dejado de lado todas las teorías que persiguen un *mero* CNC y abandonan el desafío de explicar esos correlatos; hemos dejado de lado las teorías que no tienen ningún aporte que hacer en relación a la consideración de un mecanismo específico para la conciencia o uno reducible a los mecanismos cognitivos, y finalmente excluimos las teorías que no se apegan al materialismo. De la potencial lista inicial de teorías hay sólo dos que cumplen estos requisitos: la teoría del espacio de trabajo global (GWST, por sus siglas en inglés) y la teoría de la recurrencia. En el resto del capítulo me interesa presentar cada una de ellas para luego mostrar en qué sentido son incapaces de responder al problema difícil. De este modo pretendo evidenciar la necesidad de dar una respuesta al problema difícil si es que ha de tener sentido una ciencia explicativa de la conciencia, y me diferencio de aquellos autores que consideran que el problema difícil simplemente ha de ser ignorado hasta que el progreso científico muestre dónde estaba el error (Churchland; O'Hara y Scutt; en Shear, 1997).

2.1. Global Work-Space Theory (GWST)

Una de las primeras teorías de la conciencia, sino la primera, que surgió en el campo de las neurociencias fue la teoría del espacio de trabajo global, a través de Bernard Baars (1988). La teoría del GWS (Global Workspace) se construye desde una perspectiva arquitectónica que considera al cerebro como un “conjunto masivo de procesadores paralelos. En este sistema, coordinación y control pueden tener lugar por medio de un intercambio central, permitiendo a algunos procesadores -tales como el sistema sensorial del cerebro- distribuir la información al sistema como un todo” (Baars, 2005, p. 46). El rol de la conciencia en este esquema es el de soporte para permitir el acceso de subsistemas al contenido procesado por otros subsistemas (Baars, 1988, 2005).

Hay una metáfora recurrente que ilumina esta idea, se trata del *teatro* de la conciencia (Baars, 1997). Los módulos que procesan información están constituidos de múltiples redes neuronales que están activas y operando la mayor parte del

tiempo, pero la actividad de estas redes, que son el resultado de la transmisión de impulsos electroquímicos entre grupos de neuronas, suele permanecer inconsciente. La imagen del *teatro* sugiere que el cerebro es una suma de mecanismos que trabajan y operan en la *oscuridad*, pero que la conciencia actúa como un *foco* que permite *iluminar* algunos de esos contenidos y de este modo ponerlos a disposición del sistema en su totalidad. Esta arquitectura teatral permite que el cerebro opere en “orden a integrar, proveer acceso y coordinar el funcionamiento de un largo número de redes especializadas que de otro modo operarían autónomamente” (Baars, 2005, p. 48, mi traducción).

La propuesta inicial de Baars fue luego tomada por otros neurocientíficos que fueron añadiendo matices a la arquitectura operacional y teatral básica. Uno de los casos más representativos es el de Dehaene, quien ha realizado numerosos aportes para dar una interpretación basada en evidencia empírica del modo en que el foco de la atención se dirige hacia unos u otros contenidos, poniéndolos a disposición del sistema en su totalidad y dando lugar a los estados conscientes.

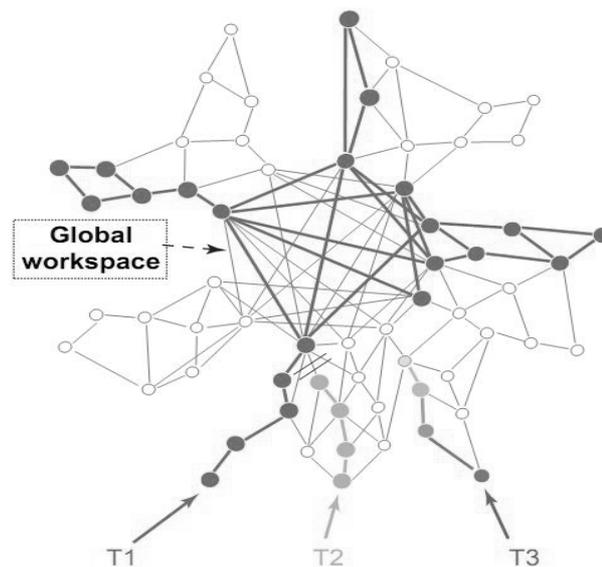


Figura 3. En la imagen puede observarse la distinción entre un espacio de trabajo global y el procesamiento modular. La oscuridad de las líneas indica el potencial acceso de la información del módulo al espacio global.

Dehane distingue, en base a patrones de conectividad neuronal distintivos, dos niveles que componen la arquitectura de la GWST (figura 3): en el nivel más bajo

se ubican el conjunto de las redes neuronales subcorticales y la mayor parte de las redes corticales. Cada una de estas redes está conformada por un número limitado de neuronas *locales* que se conectan entre sí, y pueden ser vistas como un núcleo de procesamiento automatizado, autónomo del resto y especializado en alguna tarea específica. El nivel alto se corresponde con el espacio de trabajo global. En este caso se trata de redes formadas por neuronas que se ubican en la corteza cerebral y son capaces de establecer conexiones de largo alcance, de modo que posibilitan la amplificación de la información procesada en un área particular al conjunto del sistema nervioso (Dehaene y Changeaux, 2005, pp. 910 – 911). Dehaene (y Naccache) sostiene que este nivel alto que constituye el espacio de trabajo y que posibilita la disponibilidad global de la información procesada por un subsistema específico, es lo que subjetivamente experimentamos como un estado consciente (2001, p. 1).

El vínculo entre las áreas locales y el espacio de trabajo global se da en un proceso de dos etapas: en un primer momento se observa la activación de redes a partir de la exposición a un estímulo y luego algunos de estos estímulos, los que más perduran, extienden su influencia a las distintas áreas corticales a través de mecanismos de propagación abajo-arriba (como la reverberancia) y de amplificación arriba-abajo (como la selección atencional) posibilitado por la actividad de largo alcance de las neuronas del espacio de trabajo global. Dehaene y Changeaux (2005) llaman a este proceso “ignición” o “encendido” (p. 921) y señala el momento en que un contenido correspondiente a un módulo local accede al espacio de trabajo global.

Los dos niveles de procesamiento descritos determinan la posibilidad que un contenido tiene de ser consciente y en base a ello se pueden distinguir tres niveles de accesibilidad. La información que es incorporada o amplificada por el espacio de trabajo global es consciente, pero esta es sólo una pequeña selección del total de información que se procesa en ese mismo momento en el cerebro. La posibilidad de que la información no consciente, que no forma parte del espacio global, acceda a dicho espacio determina los otros dos niveles de acceso: Es permanentemente inaccesible, o inconsciente, la información que no tiene posibilidad de acceso al espacio de trabajo, mientras que las redes que están en contacto con el espacio de trabajo y *podrían* ser amplificadas conforman un nivel *preconsciente* (Dehaene y Naccache 2001; Dehaene, Changeaux, Naccache, Sackur y Sergent, 2006).

2.1.1. Evidencia a favor de la GWST

Hay un gran número de estudios que parecen sugerir la necesidad de patrones de conectividad global para dar cuenta de la conciencia. Gran parte de esta bibliografía está indexada en Dehane & Naccache (2001). Los estudios de Casili *et al.* (2013), por ejemplo, han proporcionado evidencia que corrobora la dinámica de los procesos de ignición, mostrando los cambios pertinentes en la amplitud y la distribución de la actividad cerebral durante los estados conscientes a través del registro por EEG de la actividad electrofisiológica posterior a la estimulación magnética transcraneal (TMS).³

En una línea similar Laureys *et al.* (2011) han medido los patrones de activación neuronal ante distintos estímulos en pacientes en estado vegetativo (incapaces de tener experiencias conscientes), identificando el flujo sanguíneo a través de PET y midiendo el cambio en los potenciales eléctricos del cerebro por medio de EEG. Los resultados muestran que la activación cerebral se limita a las áreas sensoriomotoras afectadas directamente a través de estímulos sonoros, sin propagarse a otras áreas del cerebro. Al analizar pacientes con conciencia mínima, en cambio, se pudo determinar que la actividad cerebral que se corresponde con los mismos estímulos que se utilizaron en los pacientes vegetativos se propaga con mayor extensión, dando cuenta de la relación entre la amplitud de la actividad neuronal y los distintos grados de conciencia (Laureys *et al.*, 2011).

King *et al.* (2013) también han utilizado las señales EEG de pacientes sanos, vegetativos y en estado de conciencia mínima para identificar de qué modo la disponibilidad global de información permite indexar los distintos grados de conciencia. Para ello desarrollaron algoritmos que permiten identificar patrones de actividad eléctrica equivalente a lo largo de las distintas áreas del sistema nervioso. Bajo la suposición de que estos patrones equivalentes están asociados a la misma información, se pudo concluir que efectivamente había mayor información compartida -es decir, mismos patrones- en los pacientes sanos, luego en los pacientes en estado de conciencia mínima, y por último los pacientes en estado vegetativo, con

³ Si bien es cierto que estos desarrollos fueron considerados en el seno de la Teoría de la Integración de la Información de Tononi, Sitt *et al.* (2013) señalan que igualmente puede dar apoyo a la GWST.

el menor índice de información compartida.

Por otro lado hay una abundante cantidad de simulaciones computacionales que replican por este medio el comportamiento del sistema nervioso de modo acorde al sugerido por esta evidencia (Dehane y Changeaux, 2005). Algunos de estos modelos se corren con el objetivo de comprender, por ejemplo algunos fenómenos curiosos en el campo de la percepción, prototípicamente los casos de *attentional blink* [ceguera atencional] (Dehaene, 2003, p. 8520). Este fenómeno fue descrito por primera vez en 1992 (Raymond, Shapiro y Arnell) como la supresión de la atención visual por medio de la implementación del modelo experimental conocido como *rapid serial visual presentation* (RSVP – presentación visual serial rápida). De acuerdo con el protocolo del RSVP se presentan muchos objetos de manera sucesiva pero breve, y luego el sujeto debe ejecutar alguna tarea en relación a los estímulos, como identificar un color o número. Cuando los tiempos de presentación entre un objeto (T1) y otro (T2) oscilan entre 100 y 500 ms el sujeto parece disminuir o perder su atención visual. La GWST incorpora esta evidencia y da cuenta del fenómeno relacionándolo con la incapacidad del objeto (T2) para generar el estímulo que permita su reverberancia en el espacio de trabajo global.⁴ Todos los fenómenos vinculados a tipos de *blindsight* o visión ciega (inatencional, al cambio, por Continuous Flash Suppression -que es un mejoramiento del RSVP) responden a estructuras de simulación y explicación similares al caso abordado, lo que evidencia la potencia explicativa de la GWST. Pero ¿qué es lo que la teoría explica?

2.1.2. Lo que la GWST explica. Y lo que no.

Sin lugar a dudas la GWST brinda una potencial explicación de muchos fenómenos: explica cómo se da la integración cognitiva propia de los estados conscientes (...), da cuenta del modo en que el cerebro ejecuta actividades de modo inconsciente (...) y nos permite comprender los circuitos perceptivos y predecir su comportamiento, aún en casos inusuales como los de visión ciega.

La teoría del espacio de trabajo global explica muchas cosas pero no parece

4 En particular se atribuye a la incapacidad del estímulo para generar el potencial eléctrico indicado: parecería poder provocar ondas P1, P2 y N400, todas vinculadas al procesamiento semántico y visual arriba-abajo (Vogel, Luck & Shapiro, 1998), pero no P300, que serían las indicadas para generar la transición al espacio de trabajo.

explicar la conciencia, al menos en el sentido particular de conciencia que nos interesó en la sección anterior: no dice nada acerca de cómo es que tenemos estados fenoménicos. Sí señala y reconoce que los tenemos, afirma cuáles son y de qué nivel de la arquitectura dependen, pero no hay ninguna especificación de cómo esa arquitectura da lugar a estados fenoménicos. La teoría parece ser explicativa si nos referimos a la conciencia de acceso, pero se mantiene ajena al problema de la conciencia fenoménica.

La teoría del espacio de trabajo global es un caso claro y manifiesto del uso confuso, atestiguado por Chalmers y Block, entre conciencia de acceso y fenoménica. Cuando Baars presenta su teoría *cognitiva* de la conciencia (1988) uno ya puede esperar que vaya a concentrarse en la relación que la conciencia tiene con los procesos cognitivos, es decir, en la conciencia de acceso. Y eso es exactamente lo que sucede, aunque tal noble tarea está empañada por la pretensión explícita de comprender el aspecto experiencial de la conciencia que es impropio del tipo de teoría que se propone: “¿Qué es lo que hace de una teoría una teoría de la conciencia? En primer lugar es una teoría de la naturaleza de la experiencia” (Baars, 1988, p. 14, mi traducción).

Tal confusión persiste en el trabajo de Dehaene, quien oscila constantemente entre las referencias a la conciencia de acceso y la conciencia fenoménica. Afirma, por ejemplo:

Una teoría completa de la conciencia debería explicar por qué algunas representaciones cerebrales y cognitivas puede ser permanente o temporalmente inaccesibles a la conciencia, cuál es el rango posible de contenidos conscientes, como mapearlos en circuitos cerebrales específicos y si un mecanismo neural lo subyace. (Dehaene y Naccache, 2001, p. 1, mi traducción)

La arquitectura del espacio de trabajo global permite hipotetizar acerca del rol de un proceso como el de amplificación o reverberación de la información en el espacio de trabajo. Se puede suponer que tal nivel de procesamiento, el nivel global, permite una mayor variedad de procesos en relación al contenido amplificado, como su categorización semántica y perceptual, evaluación y disponibilidad para la memoria a largo plazo, etc. Todo esto forma parte de lo que Dehaene llama “una teoría

completa de la conciencia”, sin embargo nada de esto sirve para explicar por qué tenemos estados experienciales.

Dehaene y Naccache (2001) afirman, por ejemplo, que “la disponibilidad global de la información a través del espacio de trabajo es lo que experimentamos subjetivamente como un estado consciente” (p. 1, mi traducción), pero esto no constituye explicación alguna, sólo enuncia una correlación entre ciertos modos de procesamiento neurobiológico y estados conscientes, tal como si yo enunciara “cuando la marca del barómetro baja hay tormentas”: no estoy explicando -está claro que la caída en la marca del barómetro no explica por qué haya una tormenta-, sino que hay una mera correlación entre dos hechos.

La sugerencia de que la disponibilidad global está a la base de nuestra experiencia es reiterada en numerosas ocasiones, pero la naturaleza del vínculo entre estos hechos no siempre es precisada, no está claro si lo consideran explicativo, causal o meramente correlacional. Sírvanos otra cita para esclarecer el asunto:

Quando un pedazo de información, como la identidad de un estímulo, accede a un subconjunto suficiente de neuronas del espacio de trabajo su actividad se vuelve autosostenida y puede ser distribuída a través de conexiones de largo alcance a un gran número de áreas definidas, creando una disponibilidad global del estímulo dado, el cuál es entonces experimentado subjetivamente. (Dehane, 2003, p. 8520, mi traducción)

En este caso es manifiesta la pretensión explicativa, se explicita que la disponibilidad global de la información procesada a causa de un estímulo es responsable del surgimiento del carácter subjetivo y fenoménico de nuestra experiencia. Pero siguiendo a Chalmers siempre podemos volver a preguntarnos cómo y por qué esta disponibilidad es experimentada subjetivamente como consciente.

La confusión entre los usos de conciencia de acceso y la noción de conciencia fenoménica por parte de Dehaene se vuelve evidente cuando propone usar la distinción entre los tres niveles de accesibilidad -asociados a la capacidad de la información para acceder al espacio de trabajo- para dar cuenta de las diferencias entre la conciencia fenoménica y de acceso. Dice:

Sería sorprendente si estas distinciones bastaran para capturar las intuiciones detrás de las definiciones de conciencia fenoménica (P) y de acceso (A) de Block [...]. Lo que Block ve como una diferencia en la esencia podría ser meramente una diferencia cualitativa debido a una discrepancia entre el tamaño de la información potencialmente accesible [...] y la escasez de información que puede ser reportada de hecho en un momento dado (Dehane, 2003, p. 30, mi traducción).

Pero nada parece indicar por qué la diferencia cuantitativa entre bits de información procesada redundaría en una diferencia de naturaleza, como parecen constituir los estados subjetivos. Los estados fenoménicos implican nuevas propiedades que no están en todo procesamiento de información, y la teoría del espacio de trabajo no afirma nada de por qué la mayor accesibilidad posibilita estados conscientes. Una computadora procesa información y no posee ni experimenta estados subjetivos.

Para no dar lugar a interpretaciones equivocadas: no estoy diciendo que la conciencia no requiera de una articulación de mecanismos similar a la que propone la teoría del espacio de trabajo, pero esto no es explicar la conciencia. Si una arquitectura tal fuera *suficiente* para exhibir conciencia, entonces ¿por qué las simulaciones diseñadas y corridas replicando una arquitectura del tipo de la GWST no tienen estados fenoménicos? ¿o es sólo que no lo sabemos? Una respuesta intuitiva y realista a este problema es apelar al límite de procesamiento, ya que no hay una computadora que procese la cantidad de datos que procesa nuestro cerebro. Pero entonces, si hay distintas magnitudes de información compartida, ¿cuál es el límite entre la información inconsciente o preconscious y la información consciente? Cualquiera sea el límite que fijemos podemos llevar a su defensor por una pendiente resbaladiza, restando de bit en bit de información, hasta que lo hayamos convencido de que todo procesamiento de información, independientemente de su magnitud, debería implicar estados conscientes, una forma de panconciencia que parece demasiado encontrada con la intuición.

Cada vez que se reitera que una diferencia cuantitativa determina la aparición de estados fenoménicos en un sistema u organismo, cada vez que se intenta aclarar que una diferencia de magnitudes es la responsable de la aparición de nuevas propiedades -cualidades, subjetivas-, cada vez se hace más evidente la ausencia de una explicación que dé cuenta de la naturaleza de la experiencia.

Una vez más: es probable que la arquitectura del trabajo de espacio sea relevante para explicar lo que nos interesa explicar de la conciencia, es probable también que sólo ante grandes -inmensas- cantidades de información emerjan estados conscientes, pero aún no tenemos una explicación de por qué estos elementos son importantes para explicarlos. Por todo esto decimos que no hay una explicación de la conciencia, no de la conciencia fenoménica.⁵

2.2. Teoría de la conciencia como procesamiento recurrente

Lamme (2006) toma como caso particular de conciencia la conciencia visual.⁶ Supongamos -al menos por el momento- que este modo de proceder es aceptable. Entonces la pregunta es ¿cómo llegamos a tener una experiencia visual consciente? De acuerdo con Lamme (2006) los múltiples niveles de procesamiento pueden ser considerados en dos etapas: barrido proalimentativo [feedforward sweeping] y procesamiento recurrente. En un primer momento, cuando una imagen impacta en la retina es procesada rápidamente a través de los sucesivos niveles de la corteza visual primaria V1, hacia las áreas corticales visuales extraestriadas V2, V3, V4 Y V5/MT, y hasta alcanzar la corteza motora y la región prefrontal, por medio de conexiones proalimentativas, como lo indica la figura 4 (Lamme, 2006, p. 495; 2010, p. 212). Luego, en un segundo momento, aparecen dinámicas de procesamiento recurrente de la información que permiten el intercambio de información entre las distintas áreas del cerebro por medio de conexiones horizontales -hacia otras áreas- y retroalimentativas -llegando nuevamente hasta V1-.

5 Hay aún más razones por las que la teoría del espacio de trabajo global no da respuestas al problema de la conciencia. Hay problemas conceptuales propios de la teoría, por ejemplo, no queda claro si la disponibilidad global posibilita la conciencia o si la conciencia posibilita la disponibilidad o si una y otra son lo mismo. También hay problemas en relación a la evidencia: numerosa bibliografía cuestiona los índices para la medida de la conciencia (Irvine, 2013; Freeman, 2007) así como el instrumental científico y los dispositivos de recabación de datos, en particular aquellos que involucran métodos por imagen (Bechtel & Stufflebeam, 2001; Klein, 2010; Ramsey et al., 2009; Sarty, 2007), a punto tal que hay quien calificó de nueva frenología al intento de localizar en el cerebro procesos cognitivos (Uttal, 2001).

6 Este recorte del estudio de la experiencia consciente al estudio de la conciencia visual- ha sido una estrategia frecuente, que también fue inaugurada por Crick y Koch (1990) en los primeros estudios dirigidos a identificar el CNC. Entre las razones que arguyen para adoptar una estrategia tal dicen que “el sistema visual tiene varias ventajas para la investigación de las bases neuronales de la conciencia. A diferencia del lenguaje, el sistema visual es bastante parecido entre los humanos y los primates más desarrollados. Hay mucho trabajo experimental al respecto, tanto de psicofísicos como de neurocientíficos. Además creemos que el sistema visual es más fácil de estudiar que los aspectos más complejos de la conciencia asociados a la autoconciencia. De ahora en adelante, entonces, discutiremos la conciencia en su aspecto visual” (p. 267, mi traducción).

El barrido proalimentativo y el procesamiento recurrente son efectuados cada uno en dos subetapas: de procesamiento superficial y de procesamiento profundo, donde se distingue entre el procesamiento sin y con atención del estímulo respectivamente. La diferencia neural entre uno y otro tipo radica principalmente en el destino de los estímulos, ya que al prescindir de la atención éstos no alcanzan las áreas superiores de procesamiento (barrido proalimentativo), o no son puestos a disposición de otras áreas ni replicados en las conexiones reentrantes, por lo que se mantienen *locales* (procesamiento recurrente); mientras que la atención de los estímulos permite que éstos prosperen hacia las áreas superiores y se vuelvan *globales* (Lamme, 2010, p. 213).

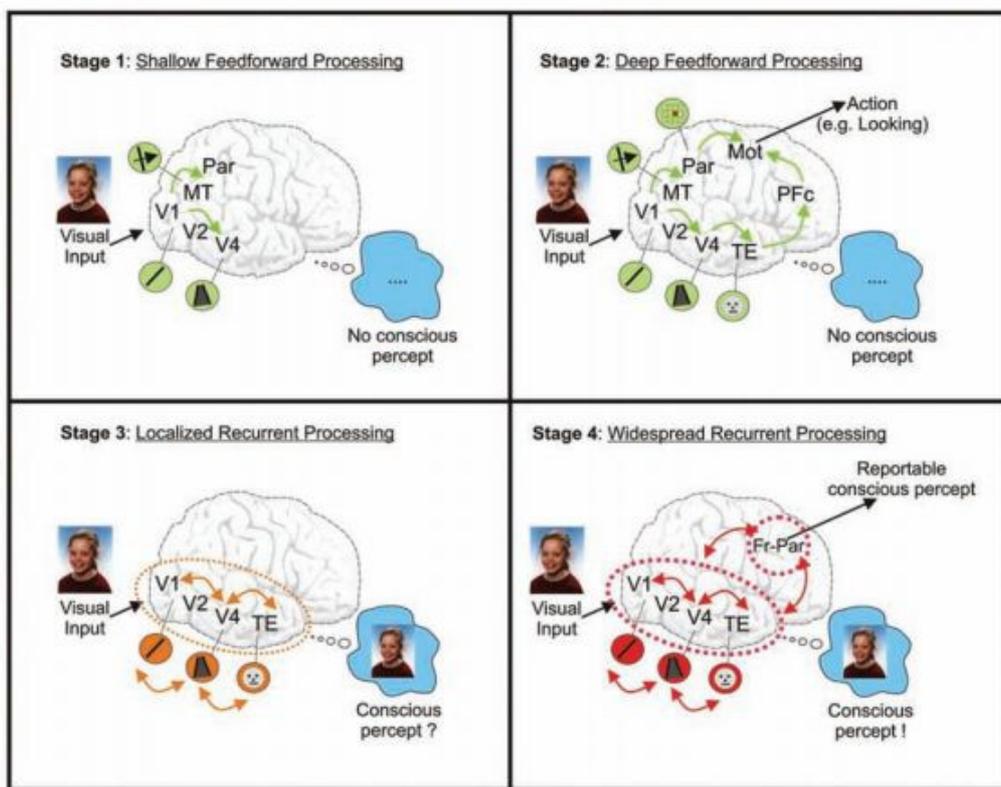


Figura 4. Arriba las dos etapas de barrido proalimentativo, abajo las correspondientes al procesamiento recurrente

Según Lamme (2006) los mecanismos de procesamiento recurrente son los encargados de generar estados fenoménicos, de modo que podemos “define consciousness as recurrent processing” (2006, p. 499). En particular basta el procesamiento recurrente local (sin atención) para que ocurran estados fenoménicos, aún cuando el agente no pueda dar cuenta de ellos.

La importancia del procesamiento recurrente para la comprensión de los estados conscientes implica desestimar que un área particular del cerebro pueda ser la responsable de la conciencia, sino que el proceso, independientemente del área, constituye el mecanismo de la conciencia. La activación de una zona neural x , desde esta perspectiva, es insuficiente en tanto no incluya procesamiento recurrente. Las distintas áreas involucradas determinan los rasgos específicos de los estados conscientes, así como los distintos comportamientos posibles en relación a esos estados, pero el ingrediente determinante de la conciencia es la recurrencia (Lamme, 2006, p. 499).

Hay evidencia que parece apoyar el rol primordial de la recurrencia en el procesamiento visual. Boehler, Schoenfeld, Heinze y Hopf (2008) utilizaron el registro de la señal magnetoencefalográfica (MEG), que provee una resolución temporal prácticamente ilimitada, para “proveer evidencia de que la conciencia de un estímulo se correlaciona con modulaciones rápidas en la actividad recurrente” (p. 8742, mi traducción). Los resultados indicaron que los casos en los que los sujetos manifiestan estar conscientes del estímulo (reportando su orientación entre cuatro opciones -arriba, abajo, derecha, izquierda-) se corresponden con procesos de recurrencia en el área V1. La modulación de estos procesos, además, se da con una velocidad que nos lleva a descartar la influencia de mecanismos de atención top-down. Mientras estos últimos se comienzan a manifestar a los 150 ms de presentados los estímulos, los procesos identificados por Boehler *et al.* aparecían en tan sólo 100-120 ms (p. 8745). Esto muestra que, tal como sugiere Lamme, “(i) el procesamiento recurrente en V1 se correlaciona con la conciencia visual y (ii) que la atención y la conciencia involucran operaciones de procesamiento recurrente distintas” (Boehler *et al.*, 2008, p. 8742, mi traducción).

Por otro lado distintos modelos computacionales del procesamiento de la información visual han mostrado una mayor eficiencia si se apela a arquitecturas recurrentes (Jehee, Lamme, y Roelfsema, 2007; Thielscher y Neumann, 2008).

Un aporte novedoso de Lamme acerca de la relación entre la evidencia y los procesos conscientes tiene que ver con la inversión particular que opera. Lamme sugiere apelar a una “definición neural” de la conciencia, es decir, debemos abandonar la definición de conciencia como un concepto meramente psicológico, del cual sólo el agente puede dar cuenta, para complementarlo con una definición que

contemple los mecanismos neurales que definen la conciencia (pp. 494 - 500). Este nuevo rol definitorio de ciertos mecanismos neurales -en Lamme el procesamiento recurrente- responde, por un lado, a la dificultad para verificar la experiencia consciente sin ninguna función cognitiva interviniente, y por el otro a la imposibilidad de fundar la ciencia en una el reporte de la primera persona.

La necesidad de trascender la primera persona radica en el hecho de que los que sujetos mienten, se equivocan, e incluso pueden estar incapacitados para dar la respuesta pertinente. Por ejemplo, queremos saber si ciertos pacientes en coma o estado vegetativo tienen experiencias conscientes, pero ¿cómo preguntamos? Supongamos que un sujeto padece de agenesia del cuerpo calloso, es decir que ha nacido sin el conjunto de fibras nerviosas que posibilitan la comunicación entre los hemisferios cerebrales izquierdo y derecho. Si presentamos un estímulo visual al campo visual derecho de un paciente en estas condiciones entonces responderá sin reparos que puede ver el estímulo, pero si en cambio presentamos el mismo estímulo en su campo visual izquierdo dirá no haberlo visto, aún cuando pueda dibujarlo con su mano izquierda (ligada al hemisferio derecho). ¿Podemos concluir a partir de esta situación que uno u otro hemisferio forman parte del correlato neural de la conciencia? ¿O en realidad estamos hablando del correlato de la memoria, la atención, las capacidades lingüísticas o las funciones de toma de decisiones asociadas?

Supongamos, por ejemplo, que el mecanismo *x* es un fuerte candidato para ser considerado el CNC. Supongamos también que la investigación adopta algún sistema de medición de la conciencia en base a la conducta o al reporte verbal del sujeto acerca de sus estados conscientes como la “Signal Detection Theory” [SDT], desarrollada por Green y Swets (1966), que ha sido utilizada ampliamente en la psicología como método para la discriminación de estímulos. Ahora bien, cada vez que el sujeto pueda reportar estados conscientes respondiendo a una pregunta o resolviendo una tarea deberíamos monitorear sus estados cerebrales y así correlacionar su activación con la información que brinda el agente sobre sus estados conscientes. Pero ¿cómo saber que estamos midiendo el correlato neural de la conciencia y no de las capacidades que le permitieron *hablar* sobre ese estado, o de la *memoria* que le permite responder sobre un evento ya sucedido, o del sistema ejecutivo que le permite *comportarse* en relación a una experiencia consciente

determinada?

Tras la sugerencia de Lamme para “definir” la conciencia en términos de recurrencia, no es necesario pedir un reporte en primera persona sino que podemos buscar la activación del mecanismo de la conciencia que hayamos acordado como CNC. Pero si bien es cierto que a partir de esta estrategia podemos prescindir del reporte del agente, también es cierto que para ello es necesario acordar en cuál sea el CNC que es justamente lo que está en cuestión, constituyendo una circularidad en el argumento (Kouider, Sid, de Gardelle, Sackur & Dupoux, 2010, p. 304).

2.2.1. Explicatividad del procesamiento recurrente.

El intento de Lamme incluye la designación de un CNC específico, el procesamiento recurrente, y al mismo tiempo es considerado por él mismo un intento propiamente explicativo. Reconoce que “la cuestión clave es hasta que punto el procesamiento recurrente puede ser considerado el ingrediente [...] que explica la fenomenalidad” (2010, p. 216, mi traducción), y se compromete con una posición particular: “el procesamiento recurrente tiene un alto poder explicativo respecto de importantes características de la percepción consciente, en tanto hay una fuerte homología entre la estructura integrada de la percepción y la estructura del procesamiento recurrente” (2010, p. 217, mi traducción). Hay varios aspectos del procesamiento recurrente que, según Lamme, nos permiten dar cuenta de la naturaleza de la conciencia. Por ejemplo, el procesamiento recurrente posibilita la amplificación de la información entre las áreas que procesan los distintos rasgos de la experiencia visual, lo cual nos ayuda a comprender el *agrupamiento* de dicha información bajo una aparente unidad (2006, p. 497).

Por otro lado el procesamiento recurrente, en contraposición al barrido proalimentativo, parece satisfacer la regla de Hebb. Prescindiendo de las sutilezas técnicas de dicha regla podemos señalar su asunción básica:

Cuando un axón de la célula *A* está lo suficientemente cerca como para excitar una célula *B* y repetidamente o persistentemente toma parte en su activación, un proceso de crecimiento o un cambio metabólico tiene lugar en una o ambas células, de modo tal que la eficiencia de *A*, como célula que dispara a *B*, ha crecido (Hebb, 2002/1949,

p. 50, mi traducción).

Con esta regla podemos predecir y explicar procesos de plasticidad neuronal a través de la activación simultánea de las neuronas pre y postsináptica. La plasticidad neuronal es considerada la base neuronal del aprendizaje y la memoria, y de ahí su importancia para la comprensión de los procesos cognitivos. Esto es relevante para nuestra discusión porque la plasticidad neuronal es más efectiva cuando depende de los ritmos oscilatorios situados en el rango gamma, que a su vez dependen de conexiones reentrantes como las que suceden durante el procesamiento recurrente (Lamme, 2010, p. 218). En palabras de Lamme (2006) “los estímulos que evocan RP [procesamiento recurrente] cambian tu cerebro, mientras que los estímulos que sólo evocan activación proalmentativa no tienen mayor impacto” (p. 499, mi traducción).

Sin embargo ni el agrupamiento ni el aprendizaje son elementos propiamente fenoménicos, y explican algo distinto de lo que habíamos fijado como el objeto del problema difícil, la conciencia fenoménica. Estos elementos pueden estar vinculados, en el mejor de los casos, a la estructura de la experiencia (como defiende Lamme), pero explicar la estructura de la conciencia no es explicar las propiedades fenoménicas de la conciencia.

Lo que discuto no es si Lamme responde con éxito o no al problema difícil, sino que pongo en cuestión que su propuesta siquiera apunte o esté dirigida al objeto de dicho problema. Por supuesto que Lamme acepta que “el ‘problema difícil’ no se resuelve sólo con esto” (2006, p. 500, mi traducción), pero al mismo tiempo sostiene que su aporte “es más prometedor de encontrar la solución que las perspectivas estándar del CNC” (2006, p. 500, mi traducción), y esto último es lo que yo niego. La teoría sugerida por Lamme no explica por qué o cómo habría de surgir la conciencia fenoménica, y se limita a analizar algunos rasgos del mecanismo de recurrencia para asociarlos a rasgos estructurales de la conciencia (como el agrupamiento) o con elementos que están directamente relacionados al acceso -en contraposición a la conciencia fenoménica- como la memoria y el aprendizaje.

Según hemos podido observar ninguna de las teorías presentadas hasta aquí responden al problema difícil en la forma en que Chalmers (1995) lo presenta. Siguiendo esta línea de pensamiento tanto la teoría del espacio de trabajo global

como la teoría de la conciencia como recurrencia fallan en identificar la conciencia fenoménica como *explanandum* de sus explicaciones, y lo suplantán con un problema de atención, memoria o estructura de la experiencia, en ambos casos. El problema difícil se planta con firmeza ante las respuestas de los teóricos, y supone un límite para la posibilidad de explicar la naturaleza de la conciencia a partir de la identificación y comprensión de los mecanismos neuronales que la acompañan.

En el próximo capítulo me propongo desarrollar un argumento con el objetivo de superar el problema difícil y así dar a la nueva ciencia de la conciencia un lugar distinto -y superador- del de encontrar los correlatos materiales de algo que es propiamente fenoménico, para pasar a considerar la posibilidad de dar una explicación genuina de la experiencia consciente a partir de la sola consideración de entidades y propiedades materiales.

La estrategia eliminativista

Muchos filósofos han pensado que la mejor forma de hacer frente al problema difícil y mantener a la conciencia dentro del ámbito de estudio de las ciencias es diluyendo tal problema a través de una vía eliminativista. Las posiciones eliminativistas en general proceden argumentalmente señalando que la referencia de una denominación x es en realidad ilusoria. Allí donde alguien utiliza un término para referirse a algo, el eliminativista sostiene que en realidad no hay nada para referir. La historia de las ciencias está plagada de entidades cuyos conceptos se evidenciaron espurios y sus referencias vacías. Entre los casos paradigmáticos podemos contar al flogisto y el éter. Tal vez algún día la ciencia, dicen los eliminativistas, nos muestre que allí donde señalamos la conciencia fenoménica no hay nada que se corresponda con ella, y que ésta designación es el resultado de tomar por verdadera alguna realidad ilusoria. De esta forma, sin conciencia fenoménica, tal vez el problema difícil deje de ser tal. Pero ¿qué significa eliminar el concepto de conciencia?

Parece inevitable acusar de locura a quien desconociera su propia experiencia consciente, o a quien pretendiera que su experiencia no es desde un punto de vista particular, el suyo propio. Con seguridad lo que entendemos por conciencia irá cambiando a medida que vayamos conociendo más sobre ella pero esto no significa que la mejor estrategia sea eliminar el concepto de conciencia. Lo que hay que eliminar para dar una respuesta satisfactorio al problema difícil, de acuerdo con lo que argumentaré en el presente capítulo, es *cierta* forma de entender la conciencia o ciertas propiedades usualmente atribuidas a la experiencia consciente. En particular, lo que hay que eliminar es la conciencia entendida como *qualia*.

A continuación presentaré la estrategia argumentativa desarrollada por Dennett contra la existencia de los *qualia* (1988, 1991, 2005) para iluminar de qué modo se puede construir, a partir de allí, una opción que haga del problema difícil un problema (al menos) tratable. Defenderé, por tanto, un eliminativismo de los *qualia* con perspectivas a eliminar el problema difícil. Para eso será necesario aclarar las

especificidades filosóficas de la noción de qualia; mostrar cuál es la relación entre esta noción y la de conciencia fenoménica; evaluar y comentar los puntos centrales del argumento de Dennett; y por último mostrar cómo esta estrategia nos permite cuestionar la legitimidad del problema difícil.

3.1 ¿Qué son los qualia?

La literatura científica y filosófica ha acuñado una gran variedad de términos para hacer referencia al carácter fenoménico de la experiencia, uno de ellos es el de *qualia*. De acuerdo con el artículo de Tye (2013) para la entrada “qualia” en la *Stanford Encyclopedia of Philosophy*:

Yo soy el sujeto de un estado mental con rasgos distintivos y subjetivos Hay algo que para mi estar en ese estado, tiene alguna fenomenología particular. Los filósofos a menudo usan el término ‘qualia’ (singular ‘quale’) para referir a los aspectos fenomenales, introspectivamente accesibles, de nuestra vida mental (p. 1, mi traducción).

Bajo esta descripción los qualia no son algo distinto de aquello que venimos refiriendo como conciencia fenoménica, sino que ambas designaciones hacen alusión a las *cualidades* de *nuestra* experiencia, es decir a su carácter fenoménico y subjetivo. Tye (2013) también señala que “en este sentido amplio del término es difícil negar que hayan qualia” (p. 1, mi traducción) porque, claro, esto parece implicar la negación de algo que se nos impone como evidente: que tenemos estados experienciales subjetivos, con un carácter cualitativo definido, que damos en llamar qualia (o conciencia fenoménica). Sin embargo esto es justamente lo que Dennett se propone hacer su artículo de 1988, *Quining Qualia*, donde reconoce que su propósito es “subversivo. Quiero [Dennett] arrojar una idea que, de alguna forma u otra, es ‘obvia’ para la mayoría de la gente--científicos, filósofos, gente ordinaria” (p. 1, mi traducción).⁷

Si bien hay usos más restringidos y técnicos del concepto de qualia (Tye, 2013), Dennett (1988) está interesado en lo que él considera la noción nuclear y

⁷ Una línea argumentativa muy similar -sino idéntica- desarrolla en *Consciousness Explained* (1991) y *Sweet Dreams* (2005).

preteórica, previa a y compartida por los refinamientos posteriores. En este sentido la propuesta de Dennett es radical, ya que no pretende eliminar cierta noción de qualia, o poner en evidencia las vaguedades y ambigüedades entre distintas concepciones acerca de los qualia, sino que apunta al corazón de su definición, a aquellos elementos sin los cuáles no se puede hablar de qualia en absoluto. Según la tradición los qualia son (1) inefables, (2) privados, (3) directa o inmediatamente aprehendidos en la conciencia e (4) intrínsecos (Dennett, 1988).

Hablamos de (1) inefabilidad para señalar la imposibilidad para verbalizar o vehicular lingüísticamente estas propiedades cualitativas de nuestra experiencia. ¿Cómo podemos poner en palabras la apariencia cualitativa determinada del particular verde grisáceo que exhiben las nubes antes de la caída de granizo? ¿Qué parte del lenguaje debo usar para comunicar la intensidad y naturaleza propia del dolor que me aqueja en este momento en la muñeca, producto de tanta escritura?

Cuando decimos que los estados cualitativos de la experiencia son (2) privados decimos que ningún test objetivo, psicológico o meramente conductual puede revelarlos y que es imposible cualquier tipo de comparación interpersonal en el modo en que se presentan tales estados.

Por otro lado la aprehensibilidad *inmediata* de los contenidos fenoménicos en la conciencia (3) refiere al carácter privilegiado de la primera persona: las cosas se me presentan bajo ciertas cualidades, y estas cualidades son aprehendidas de modo directo, sin mediación conceptual alguna, se presentan como datos puros e incuestionables para el agente de tal experiencia. Puede ser que los mismos no se correspondan con ningún elemento externo, conformen una experiencia “ilusoria” del mundo, y sin embargo las cualidades siguen estando allí para el sujeto.

Sin embargo ninguno de estos puntos es tan importante para la necesidad de efectuar el giro eliminativista como el hecho de que los qualia sean de carácter intrínseco (4). Dennett entiende por intrínsecas aquellas cualidades “atómicas e inanalizables”, ‘simples’ u ‘homogéneas’” (Dennett, 1988, p. 1, mi traducción). Efectivamente, si consideramos de este modo las propiedades intrínsecas, es inevitable pensar que aquello que describimos como intrínseco será inasible para las ciencias: no podemos nombrarlo por inefable, señalarlo por privado, ni conocerlo por medio de sus causas, efectos o relaciones ya que son *inanalizables*, y “no hay nada de lo que agarrarse cuando tratamos de describir tales propiedades a alguien que no

conoce la instancia particular en cuestión” (Dennett, 1988, p. 1, mi traducción).

Si bien definir la noción de intrinsicalidad en relación a los fenómenos mentales es una tarea que ha presentado serias dificultades a los teóricos (Lewis, 1983; Chisholm, 1976; Weatherson & Marshall, 2014), el significado que le otorga Dennett está próximo al canónico en la literatura sobre el tema: “algo tiene propiedad intrínsecas en virtud de lo que ese algo mismo es, y de nada más. No así para las propiedades extrínsecas, donde algo bien podría tener sus virtudes en virtud de lo que otra cosa es” (Lewis, 1983, p. 197, mi traducción).⁸ Basta quedarnos con la primer parte de la cita para obtener la definición de intrinsicalidad,

- ◆ “F es una propiedad intrínseca sii, necesariamente, cualquier cosa que sea F es F en virtud del modo en que sí misma, y nada distinta de sí, es” (Weatherson & Marshall, 2014, p. 1, mi traducción).

De acuerdo con esta enunciación los qualia son lo que son en virtud del modo en que ellos mismo, y nada distinto de ellos, son. No son lo que son en virtud de la materia y sus propiedades, ni son lo que son en relación por los inputs que los generan ni los outputs que producen. La definición de Lewis y la interpretación de Weatherson & Marshall parecen coincidir con la primitividad señalada por Dennett.

Dennett también apela a una metáfora que ayuda a esclarecer la noción de intrinsicalidad: Hay quienes consideran que el valor del dólar es intrínseco, en oposición al valor de las otras monedas, que parece ser relacional y que se mide,

⁸ Esta cita se encuentra inserta en un contexto mayor, donde Lewis (1983) afirma que: “Una oración o enunciado o proposición que adscribe propiedades intrínsecas a algo es enteramente acerca de esa cosa; mientras que una adscripción de propiedades extrínsecas a algo no es enteramente acerca de esa cosa, sino que bien podría ser acerca de un todo más grande que incluye aquello a lo que referimos como parte. Algo tiene propiedades intrínsecas en virtud del modo en que eso mismo, y nada más, es. No así para las propiedades extrínsecas, que pueden ser parte del objeto en virtud del modo en que algo distinto de sí es. Las propiedades intrínsecas de algo dependen solamente de ese algo, mientras que las propiedades extrínsecas pueden depender, total o parcialmente, de algo más. Si algo tiene una propiedad intrínseca, entonces también la tiene cualquier duplicado perfecto de ese algo; mientras que los duplicados situados en diferentes contextos tendrán diferentes propiedades extrínsecas” (p. 197, mi traducción).

Tal como Weatherson & Marshall (2014) han señalado, cada una de éstas comparaciones da lugar a un concepto distinto de intrinsicalidad. En la presente tesis aceptaré como correcta, siguiendo al mencionado trabajo (Weatherson & Marshall, 2014), la distinción que predomina en el trabajo de Lewis y a las que parece reducir el resto, que es la distinción que citamos en el cuerpo del trabajo: “algo tiene propiedad intrínsecas en virtud de lo que ese algo mismo es, y de nada más. No así para las propiedades extrínsecas, donde algo bien podría tener sus virtudes en virtud de lo que otra cosa es” (Lewis, 1983, p. 197, mi traducción)

justamente, por contraste con el valor intrínseco del dólar. Por supuesto que esto no es más que una ilusión. El mismo error nos acomete, según Dennett, cuando consideramos los qualia: decir que un estado fenoménico es intrínseco implica afirmar que puede ser considerado con independencia de cualquier otra entidad o propiedad. Una vez más se enfatiza el hecho de que los qualia son propiedades *sui generis* de la experiencia y constituyen unidades atómicas e inanalizables.

Si los qualia pueden caracterizarse de este modo, entonces la relación entre problema difícil y las propiedades intrínsecas de los qualia es evidente. El carácter inefable, privado y subjetivo dan lugar a una brecha epistémica que últimamente se corresponde con el salto ontológico que supone la intrinsicalidad de los qualia.

3.2. Qualia y conciencia fenoménica

Antes de proceder al argumento de Dennett en contra de los qualia es necesario mostrar que podemos atribuir, a los defensores del concepto de conciencia fenoménica y del problema difícil, una noción de propiedades fenoménicas tan cargada como la que suponen los qualia tal como se los ha presentado. Dennett considera que la intrinsicalidad (junto a las otras propiedades) forman parte, como hemos señalado con anterioridad, de la noción nuclear y preteórica de qualia. Sin embargo Chalmers y Block en ningún momento apelan a dicha noción, e incluso han intentado desprenderse de la misma, por lo tanto es razonable preguntarnos si el argumento que presentaremos más adelante ha de servir para el concepto de conciencia sostenido por estos autores, o si el mismo se mantiene ajeno a nuestras consideraciones.

Dennett afirma que ~~“el término 'fenomenal' no significa nada obvio y no tendencioso para mí, y parece sospechosamente como un gesto en la dirección que lleva vuelta a lo inefable, privado”~~ (p. 1, mi traducción). Referirse a cualquier propiedad en términos de fenoménica es *siempre* una referencia a los qualia. Frankish ha argumentado con más detenimiento respecto de la dificultad para considerar la idea de qualia con independencia de estas propiedades canónicas. En *Quining Diet Qualia* (2011) niega que pueda haber una noción deflacionada de qualia -los *diet qualia*- que refiera exclusivamente al carácter fenoménico de la experiencia, sin comprometerse con la intrinsicalidad o la inefabilidad, por ejemplo. En

contraposición a esta noción *diet* están los *classic qualia*, que reúnen las propiedades canónicas, y los *zero qualia* que se corresponden con un concepto deflacionado que denota las “propiedades de la experiencia que nos disponen a juzgar que las experiencias tienen propiedades cualitativas introspectivas que son intrínsecas, inefables y subjetivas” (Frankish, p. 669, mi traducción). El autor discute varios caminos a partir de los cuáles se podrían definir los *diet* qualia, concluyendo que todos los casos son reducibles al conjunto de los *classic* qualia o de los *zero* qualia, y, por lo tanto, que la noción de *diet* qualia es espuria.

En nuestro caso, antes que una discusión conceptual útil para todos los casos, debería bastar con indagar a qué se refieren Chalmers y Block cuando hablan de conciencia fenoménica.

Block (1995) presenta el concepto de conciencia fenoménica asumiendo que **“la parte controversial es que yo tomo las propiedades conscientes fenoménicas como distintas de cualquier propiedad funcional, cognitiva o intencional”** (p. 230, mi traducción). En efecto su afirmación es controversial, aún cuando esté “tratando duramente de limitar la controversialidad de sus asunciones” (p. 230, mi traducción), ya que en esa mera asunción se ha comprometido con todo aquello que puede comprometerse el más fiel defensor de los qualia, en especial con la intrinsicalidad de las propiedades “fenoménicamente conscientes”.

~~Las propiedades fenoménicas son distintas de cualquier elemento cognitivo, lo cual implica que no son representacionales, no son intencionales, no tienen un contenido y no son funcionales, razón por la cual no tienen ningún rol en el resto de los procesos mentales ni son el resultado de ninguna disposición personal o subpersonal.~~ Si esto no es ser intrínseco entonces ¿qué lo será? Al mismo tiempo Block toma la definición que hace Nagel (1974) de los estados fenoménicamente conscientes en tanto hay algo que se siente al estar en ese estado, definición que lo compromete con la propiedad subjetiva de los estados conscientes. Es cierto que no hay referencias a la privacidad ni a la inefabilidad en el trabajo que hemos tomado de referencia, sin embargo habiendo aceptado que hay aspectos fenoménicos intrínsecos y subjetivos difícilmente pueda evitar comprometerse con las propiedades restantes. Igualmente debemos destacar que la propiedad realmente problemática de los qualia, que según argumentaré deriva en el problema difícil, es la intrinsicalidad, a la cuál Block adhiere por su definición de conciencia fenoménica en

contraposición a la conciencia de acceso.

La cercanía del concepto de quale con el de conciencia fenoménica es aún más claro si se lo lee a través de Chalmers (1995), quien admite que “a veces términos como ‘conciencia fenoménica’ y ‘qualia’ también son usados, pero encuentro más natural hablar de ‘experiencia consciente’ o simplemente ‘experiencia’” (p. 201, mi traducción). Chalmers no da ninguna razón ni justificación por la que sea más natural uno u otro término, por lo cuál podemos suponer que la suya es una elección arbitraria y que no responde a una diferencia en el contenido de los conceptos. El concepto de experiencia de Chalmers -que habíamos identificado como equivalente a la conciencia fenoménica de Block en el primer capítulo- es, por lo tanto, sinónimo con el de qualia. Esto también se evidencia cuando presenta al problema difícil como aquél cuya resolución es distinta de la resolución de cualquiera de los problemas fáciles, e.e. distinta de de todo problema que pueda ser descrito en términos de estructura y función lo cual supone que la conciencia fenoménica es algo distinto de cualquier disposición conductual en sentido amplio y está definida en virtud de sí misma, por lo que tampoco la consideramos representacional. Esto, que parecen ser meras manchas de nuestra percepción, los *raw feels* o los *sense data* de los empiristas clásicos ingleses, el mito de lo dado sellarsiano, es necesariamente inefable, subjetivo, privado, inmediato y, principalmente, intrínseco.

Bajo el supuesto de que podemos identificar el concepto de conciencia fenoménica con el de qualia podemos pasar al argumento eliminativista.

3.3. Quining⁹ Qualia

La estrategia de Dennett está dirigida a depositar el peso de la prueba en los defensores de la existencia de los qualia, mostrando las inconsistencias y paradojas que se desprenden del concepto mismo de qualia. Dennett asegura que se puede objetar cualquier intento argumentativo de los *qualiófilos* en virtud del grupo aporético de propiedades que se asocian a los qualia. Para mostrar esto apela, en *Quining Qualia* (Dennett, 1988), a 15 *bombas de intuición*¹⁰ (independientes aunque

9 “Quinear’ [...]: Dengar resueltamente la existencia o importancia de algo real o significativo” (Dennett, 1988, p.1, mi traducción).

10 Intuiciones 1 y 2 están construidas para señalar de qué modo se llega a la noción de qualia; intuiciones 3 a 6 ponen en cuestión el acceso -subjetivo e intersubjetivo, en primera y en tercera persona- a los qualia; intuiciones 7 a 12 están dirigidas, principalmente (junto a algunas referencias

relacionadas) que pretenden poner en evidencia la inconcebibilidad de propiedades como los qualia. Algunas de estas intuiciones fueron retomadas en otros de sus trabajos (1991, 2005) con ligeras modificaciones. A continuación me concentraré en algunas de ellas, en particular en aquellas que atacan la noción de intrinsicalidad, que, según hemos dicho, es central para hacer a un lado el obstáculo que supone el problema difícil.

Consideremos en primer lugar la intuición del bebedor de cerveza experimentado:

La cerveza sabe diferente al bebedor de cerveza experimentado. Si la cerveza fuera a tener siempre el mismo sabor que la primera vez, nunca me hubiera vuelto un bebedor de cerveza! O para poner el mismo punto desde el otro lado, si mi primer sorbo de cerveza me hubiera sabido del mismo modo que el sorbo que tomé recién, nunca habría mejorado mi gusto en primer lugar! Habría amado el primer sorbo tanto como el que acabo de disfrutar (Dennett, 1988, p. 1, mi traducción).

Dennett retrata una situación muy común: muchas veces podemos apreciar una evolución en nuestro gusto. En este caso se trata de la cerveza, pero podemos pensar también en bebedores de vino, whisky o café. ¿Qué es lo que cambia cuando podemos nos convertimos en *catadores*? Las propiedades de la cerveza han de ser las mismas, sin embargo su degustación frecuente cambia a tal punto que aquello que provocaba rechazo provoca luego placer. Si el gusto de la cerveza fuera una propiedad intrínseca entonces tales reacciones deberían variar aún cuando uno pueda reconocer que el gusto es el mismo, pero de hecho el gusto ha cambiado. Por supuesto, en algún sentido continúa siendo el mismo: no sabe a frutillas ni a mermelada, aún sabe a cerveza. Pero no es exactamente el mismo gusto, y sólo a raíz de que el gusto es distinto al de aquel primer sorbo, el bebedor puede disfrutarlo. En tanto el cambio en la reactividad garantiza el cambio en el gusto decimos que los qualia en este caso no son intrínsecos, ya que no se explican en virtud de lo que ellos mismos son sino en relación a las disposiciones afectivas que tenemos respecto de ellos (Dennett, 1988).

a la immediatez), contra la noción de intrinsicalidad; nociones 13 a 16 presentan un modo de interpretar (o al menos uno de individuar) el carácter cualitativo de la experiencia en términos de propiedades informacionales.

La *educación* del gusto es similar al entrenamiento que uno puede adquirir en relación a los otros sentidos, como lo indica la intuición de *la cuerda de guitarra*:

Toca el mi grave abierto, y escucha cuidadosamente el sonido. [...] Ahora toca la cuerda abierta de nuevo y cuidadosamente apoya un dedo suavemente sobre el octavo traste para crear un 'armónico'. [...] Pero luego, en una tercera ejecución al aire uno puede escuchar, con sorprendente nitidez, el armónico que fue aislado en la segunda ejecución (Dennett, 1988, p. 1, mi traducción).

Si bien esta intuición está originalmente dirigida contra la supuesta homogeneidad e inefabilidad de los qualia también muestra que los sonidos no son intrínsecos. ¿Cuál es en este caso el quale correspondiente al sonido de la cuerda E *al aire* en la primera ejecución? El E ejecutado antes de la cuerda no puede ser, porque entonces el quale correspondiente no sería atómico y estaría compuesto de partes; además, y más importante, si habíamos dicho que para ser intrínseco algo tiene que ser definido en virtud de sí mismo y de nada distinto de sí, este sonido no es intrínseco ya que puede ser definido a partir de sus partes.¹¹

La última intuición que me limitaré a considerar contempla un experimento de inversión del campo visual por medio del uso de lentes inversores:

Después de usar lentes de inversión por varios días los sujetos tuvieron una adaptación sorprendentemente exitosa. Supongamos que los presionamos con la pregunta: '¿Tu adaptación consiste en una reinversión del campo visual, o en dar vuelta el resto de la mente en un juego de compensaciones?' Si ellos dudaran ¿podríamos insistir en que tiene que haber una respuesta correcta, aún si no pudiéramos decir con confianza cuál es? (Dennett, 1988: p. 1, mi traducción)

Si a la inversión del campo visual le acompaña la adaptación motora progresiva, esto supone que también en dicha adaptación opera un proceso de inversión de modo tal que hay una sincronización entre percepción y comportamiento, es decir que ha operado algún tipo de inversión también en las disposiciones conductuales (Stratton,

¹¹ Solo salvamos la intrinsicalidad si afirmamos que la experiencia del sonido de la cuerda E al aire está compuesta por los quale componentes, identificados a partir de la ejecución del armónico, pero en tal caso la experiencia primera del sonido estaba compuesta por sonidos indisintiguibles para el oyente, lo cuál parece contradecir el carácter subjetivo e inmediato de los qualia.

1896, 1897). Esto es así a punto tal que en los experimentos de Stratton, pioneros en el asunto, se describe un fenómeno de reacomodamiento del campo visual, e.e. aquello que se ve patas para arriba de pronto se ve en su orientación normal. Imaginemos que la totalidad de las capacidades disposicionales de un sujeto, y no sólo unas pocas, se acopla con el fenómeno de inversión visual, entonces ¿podría el sujeto referir a inversión alguna? Ni él mismo, ni una tercera persona podrá referir a una inversión del campo visual, ya que tanto su comportamiento, como sus respuestas estarán alteradas de modo tal que será imposible hablar de la normal dirección de las cosas en contraste con una supuesta inversión. Sin lugar a dudas esto apoyaría la tesis de no intrinsicalidad de los elementos de nuestra experiencia defendida por Dennett, ya que, aplicando la misma receta que con las intuiciones anteriores, implica definir una propiedad de la experiencia -en este caso la orientación del campo visual- en relación a las disposiciones conductuales (lingüísticas, corporales e incluso aquellas que determinan el procesamiento de la imagen retinal en nuestro ojo), y no en virtud de sí misma.

Pero no todos estarán conformes con este argumento. Alguien puede afirmar que nunca estuvo en cuestión la intrinsicalidad del posicionamiento del campo visual, el cual no es un quale de la experiencia, sino que es un elemento representacional/cognitivo y que por lo tanto no constituye un aporte en contra de la intrinsicalidad de los qualia. Supongamos que hay algún modo válido de argumentar a favor de esto (Dennett lo desestima, yo también). En tal caso me gustaría apelar a la experiencia cotidiana para para argumentar en relación a los colores lo mismo que Dennett argumenta para el posicionamiento del campo visual:

Cualquier persona acostumbrada al uso frecuente de lentes de sol polarizados alguna vez habrá experimentado *olvidar* que llevaba los lentes puestos. En estos casos se tiene la sensación de que uno ve tal como vería si no llevara puesto lente alguno, en eso reside la confusión. Incluso, tras el uso sostenido de los lentes de sol, el sujeto puede notar que si intenta imaginar la posible opacidad tonal a la que se encuentra sometido por los lentes oscuros, no podrá más que sorprenderse ante la falta de una referencia visual que le permita estimar el cambio tonal. Si hay algún tipo de información que pueda servir para construir tal referencia, esta información puede ser caracterizada funcional o disposicionalmente, pero si sólo tomamos la información perceptiva no podremos dar cuenta de un cambio en la tonalidad de los

colores. Por ejemplo, podremos estimar el cambio tonal rememorando situaciones anteriores en las que comparé lo que veía con o sin lentes, puedo apelar a mis conocimientos sobre mis colores y saber que el tono será más oscuro, en gama de grises, si es que los lentes son sólo polarizados, etc.

Los momentos en que la experiencia del color bajo el uso de lentes oscuros es experimentada como *equivalente* a la experiencia del color sin lentes polarizados sólo son posibles en tanto aceptemos que la experiencia del color no depende de lo que los colores en sí mismo sean, sino que estos son lo que son en virtud del entramado conductual y disposicional en el que se hayan insertos, paradigmáticamente de los órganos que los procesan. Esto implica que nuestra experiencia no puede estar constituida por propiedades que sean intrínsecas como los qualia. A su vez hemos dicho que los qualia designan las propiedades fenoménicas e introspectivas de la conciencia, siendo una denominación indistinta la de qualia o la de conciencia fenoménica. Por transitividad, negar la existencia de los qualia significa negar la conciencia fenoménica.

Pero ¿No habíamos dicho que la idea de negar la conciencia es una mala idea? No sólo parece implicar la negación de algo *evidente* como la experiencia consciente, sino que además parece dejarnos sin un *explanandum* para la nueva ciencia de la conciencia. Debemos contemplar tal posibilidad.

3.3. ¿Eliminando la conciencia?

Hay quienes han afirmado que Dennett pretende eliminar la conciencia. Voorhees (2000), por ejemplo, afirma con severidad:

Daniel Dennett es el Diablo... No hay un testigo interno, no hay un intérprete semántico, y ningún yo distinto de un abstracto “Centro de Gravedad Narrativa” que no es nada más que una conveniente ficción... Para Dennett no es el caso del Emperador que no tiene ropas. Son más bien las ropas las que no tienen Emperador (pp. 55–56, mi traducción).

Sin embargo estas lecturas facilistas han dado lugar a una seria confusión en torno al programa de Dennett, que si bien eliminativista, no pretende negar un hecho

evidente, la conciencia. En palabras del mismo Dennett (1988): “Todo lo real tiene propiedades, y dado que yo no niego la realidad de la experiencia consciente, doy por hecho que la experiencia consciente tiene propiedades” (p. 1, mi traducción). Del mismo modo en *Dulces Sueños* (2005), que constituye una actualización de la teoría de la conciencia presentada en *Consciousness Explained* (1991), Dennett se preocupa por aclarar la situación nuevamente: “Yo no sostengo, por supuesto, que la conciencia humana no exista, sólo que no es lo que la gente suele creer que es” (p. 71, mi traducción). Pero entonces ¿qué es lo que Dennett elimina? y más importante aún ¿qué es la conciencia? Prosigamos con la cita anterior de *Quining Qualia*:

Doy por hecho que la experiencia consciente tiene propiedades. Doy por hecho además que cada estado consciente de las personas tiene propiedades en virtud de las cuales esos estados tienen el contenido experiencial que tienen. Esto es decir, cuando alguien experimenta algo como siendo de un modo u otro, esto es verdad en virtud de alguna propiedad de algo que está sucediendo con él, pero estas propiedades son tan distintas de las propiedades imputadas tradicionalmente a la conciencia que sería un grosero error llamar a alguna de ellas los largamente buscados qualia. Se supone que los qualia son propiedades especiales, en algún modo difícil de definir. [...] La experiencia consciente no tiene propiedades que sean especiales en ninguno de los sentidos en que se supone que los qualia son especiales (Dennett, 1988, p. 1, mi traducción)

En particular, como ya hemos destacado, la condición intrínseca de los qualia es una de las propiedades que hace de ellos algo *especial e inanalizable* en términos de función y estructura. Teniendo esto en mente podemos leer la cita anterior del siguiente modo: **tenemos experiencias conscientes, pero las propiedades de esta experiencia no pueden considerarse intrínsecas, aisladas de toda reacción psicológica previa así como del procesamiento de información a través de los mecanismos cognitivos**. Siendo los qualia entidades o propiedades no relacionales “eludirán todo tipo de análisis funcional amplio” (Dennett, 2005, p. 176, mi traducción) en virtud de una maniobra meramente definatoria y sin que hayan verdaderas razones para creer en la existencia de entidades o propiedades tales. El objetivo fundamental de

Dennett, por lo tanto, es eliminar la experiencia consciente *qua* experiencias fundada en propiedades intrínsecas, no relacionales, últimas e inanalizables.

3.4. Eliminando el problema difícil

Hemos señalado en el primer capítulo que las distinciones entre problema difícil y problema fácil, y entre conciencia fenoménica y conciencia de acceso son equivalentes. En ambos casos se supone que hay una parte de la conciencia que es independiente de cualquier mecanismo cognitivo, de cualquier mecanismo de acceso, y que constituye el *explanandum* propio de cualquier ciencia de la conciencia. Este modo de formular ambas distinciones implica que las propiedades -fenoménicas- de nuestra conciencia, independientes del contexto psicológico y fisiológico en el cuál se hallan insertos, son intrínsecas. Pero siendo propiedades fenoménicas e intrínsecas podemos identificarlas con la noción clásica de qualia. Paradójicamente, según hemos acordado con Dennett, los qualia pueden ser eliminados; en particular podemos eliminarlos de nuestra descripción del mundo mental en tanto propiedades que no son definidas en virtud de nada más que de sí mismas (intrínsecas). De esto se sigue que la conciencia fenoménica, como algo distinto de todo mecanismo de acceso, y que el problema difícil, en tanto independiente de todo problema fácil, son inexistentes; y siendo inexistente uno de los términos se disuelven las dicotomías.

Ya sin la distinción entre conciencia fenoménica y de acceso no hay dificultades para repartir dificultades entre problemas distintos, sino que todo se reduce al mismo problema (fácil): identificar y explicar los mecanismos que posibilitan la ejecución de las distintas funciones cognitivas, algunas de las cuáles tendrán el rasgo característico de expresarse de modo consciente. Así puestas las cosas el eliminativismo deviene en identitarismo: no es que los mecanismos cognitivos *causan* la conciencia, sino que *son* la conciencia. Por eso podemos negar la segunda premisa del argumento explicativo, por la cual existe algo que no es explicable en términos de *estructura* y *función*. Al haber reducido nuestro *explanandum* a los mecanismos cognitivos, o -lo que es lo mismo- a la conciencia de acceso, no hay nada que explicar que quede por fuera de la estructura y la función.

Desde esta perspectiva la discusión sobre la legitimidad del problema *difícil* descansa en una disputa ontológica en torno a la existencia o no de los qualia: de la

eliminación de los mismos se sigue la eliminación de la noción equivalente de *conciencia fenoménica* y de aquí la eliminación del *problema difícil*. Esta manera de enfocar la discusión supone, entonces, la disolución de las dos distinciones referidas, y presenta ventajas importantes respecto de otros modos de enfrentar el *problema difícil*, que se limitan a la consideración del problema sin cuestionar el supuesto *explanandum* del problema, e.e. la conciencia fenoménica.

Hasta aquí hemos presentado el concepto de qualia, establecido su identidad con el concepto de conciencia fenoménica, y mostrado como tras su eliminación (vía argumento tomado de Dennett) podemos proceder a la eliminación del problema difícil. En el próximo capítulo quiero defender que esta doble eliminación, este modo particular de atacar el problema difícil, presenta dos ventajas importantes: (1) permite discutir los argumentos *epistémicos* que dan apoyo al problema difícil, es decir, nos brinda una heurística filosófica para el análisis de los argumentos antimaterialistas; y (2) permite decantarnos por ciertas teorías neurobiológicas antes que otras, es decir, nos brinda una heurística científica para decidir entre teorías rivales.

Ventajas de la estrategia eliminativista

La estrategia eliminativista es una de tantas formas de enfrentar el problema difícil. Tiene de particular que lo hace a partir de la eliminación de aquello que se pretende sea el *explanandum* de la ciencia que enfrente tal problema: la conciencia fenoménica. Es posible eliminar el problema difícil sin poner en cuestión la existencia de la conciencia fenoménica, o al menos es un camino que muchos han intentado (Churchland, 1997; Hardcastle, 1997; O'hara & Scott, 1997). En este capítulo quiero mostrar que la estrategia de Dennett es la mejor opción porque además de abordar el problema difícil nos permite dar tratamiento a los argumentos *epistémicos* en contra del materialismo, al mismo tiempo que brinda una heurística de investigación para decidir entre teorías empíricas rivales.

4.1. Estrategias conceptuales para los argumentos *epistémicos*

Según hemos afirmado en el primer capítulo los argumentos epistémicos (el argumento del conocimiento y el argumento *zombie*), que anteceden y complementan el argumento explicativo -que da lugar al problema difícil-, comparten con éste una misma estructura lógica. Esto es una razón suficiente para esperar que una buena respuesta al problema difícil implique también alguna sugerencia acerca de cómo atacar problemas que pueden ser enunciados con premisas equivalentes.

La existencia de un vínculo entre la estrategia eliminativista y los argumentos *epistémicos* ya había sido sugerido, aunque no desarrollado, por Levine, quien sobre el final de su trabajo *Materialism and Qualia: The Explanatory Gap* (1981) se limita a señalar que

hay sólo un modo que puedo ver de escapar a este dilema y permanecer materialista. Uno debe negar, o disolver, la intuición que yace a la base del argumento. Esto implica, creo, tomar más una línea eliminativista respecto a los *qualia* de lo que muchos filósofos materialistas están listos para tomar (pp. 360-361, mi traducción).

El dilema del que habla Levine refiere a la supuesta contingencia de los enunciados que expresan identidades psico-físicas tal como lo formulara Kripke en *Naming and Necessity* (1972). Como destacamos en el primer capítulo esta contingencia está dada por la *concebibilidad* de dos mundos materialmente idénticos -partícula por partícula-, donde las realidades fenoménicas sean distintas, lo cuál no es más que otra forma de plantear el argumento zombie.

4.1.1. Matando zombies

Recordemos que, de acuerdo con este argumento, no se puede expresar la identidad entre los estados fenoménicos porque es concebible, y por lo tanto metafísicamente posible, que existan mundos materiales idénticos cuyos correlatos fenoménicos sean distintos. Dada esta posibilidad la identidad entre estados fenoménicos y materiales es contingente, violando el requerimiento de necesidad para las identidades teóricas elaboradas por las ciencias. En otras palabras: si podemos imaginar un zombie, es porque es *concebible* un mundo físicamente idéntico al nuestro, pero fenoménicamente distinto, lo cual sería imposible si hubiera una identidad entre el mundo material y el fenoménico.

El lugar que la necesidad y la posibilidad metafísica -contrapuestas a la necesidad y la posibilidad física- ocupan en estos argumentos ha sido ampliamente discutido (Dennett, citar más). Naturalmente, una cosa es aceptar que los argumentos conceptuales tienen un lugar propio en el desarrollo de las ciencias, y me parece muy de atender cuando por esa vía se declama la *imposibilidad* de un concepto, entidad o propiedad; sin embargo algo muy distinto es decretar la *posibilidad* de un concepto, entidad o propiedad, y sólo por ello considerar oportuno su inclusión dentro del edificio de las ciencias. Aún así es pertinente analizar el argumento con más detenimiento, especialmente habiendo una segunda parte de la cita de *Levine* que señala la relación entre “intuición subyacente al fundamento del argumento [de la concebibilidad]” y “la línea eliminativista con respecto a los qualia”. Nosotros ya hemos tomado la línea eliminativista, pero ¿cuál es la intuición de los argumentos que abogan por la brecha explicativa que queda expuesta? Ésta no es otra que la intuición *zombie*, según la cual las propiedades cualitativas de la

experiencia pueden ser imaginables con -relativa- independencia de su soporte material. Esta intuición (antimaterialista, por cierto) también es la intuición de que las propiedades de la experiencia son intrínsecas. La propiedad *azul*, por ejemplo, es intrínseca: el modo en que se experimenta dicha propiedad al mirar, por ejemplo, el cielo en una tarde despejada, no está determinado por la configuración cerebral, ya que no hay nada en dicha configuración que nos diga cuál será el modo específico en que el quale *azul* será experimentado, e incluso puedo concebir dos mundos materialmente idénticos, donde el mismo sujeto tenga la experiencia de *azul* ante distintos estímulos. Pero eliminando la existencia de propiedades intrínsecas como los qualia, tal como sugirió Levine, podemos contradecir la intuición -que después de todo es sólo eso, una intuición- y afirmar que no es *concebible* que en dos mundos idénticos partícula por partícula, el mismo sistema nervioso pueda ser soporte de propiedades experienciales distintas. En otras palabras podemos comprometernos con que los aspectos funcionales y estructurales de una configuración física *x* determinan la experiencia particular del azul del cielo una tarde despejada, porque *son* esa experiencia.

La eliminación de la conciencia fenoménica como algo distinto a la conciencia de acceso puede considerarse, entonces, la introducción de una nueva intuición materialista: *Todo lo que hay que saber del modo en que experimentamos el azul es lo que sabemos del azul cuando nuestro mecanismo visual está procesando el color azul.*

Si bien las intuiciones operan de forma muy particular en el intelecto de las personas, quiero ofrecer una razón por la que la intuición inicial -antimaterialista- puede merecer revisión

Pongamos a prueba nuestra imaginación: pensemos en un viaje en el tiempo. En principios no parece un desafío demasiado grande, cierro los ojos y me imagino entrando en una máquina blindada, en un lugar que parece un gran laboratorio del futuro y escoltado por científicos turbados... Entro, de afuera cierran la puerta herméticamente y por la pequeña ventana alguien levanta los pulgares. De repente un sinfín de destellos eléctricos, ruido ensordecedor y una sensación de congelamiento en mi imagen, en mi cuerpo, punto a partir del cuál todo empieza a aclarar. Cuando miro a través de la pequeña ventana un grupo de científicos me reciben excitados en un laboratorio que no se veía tan distinto del anterior, pero que

evidenciaba los 100 años de tiempo al futuro que había viajado.

¿He imaginado realmente un viaje en el tiempo? ¿Puedo decir que a partir del relato es *concebible* un viaje en el tiempo? Para imaginar dicho viaje en el tiempo de un modo que sea argumentativamente válido hay que preguntarse cosas como “¿qué es lo que sucede con el tiempo mientras yo hago ese viaje?” o “¿qué dispositivo es el que permite la operación de la máquina?”, debería tener al menos una concepción de tiempo que haga posible decir que un sujeto *viaja en el tiempo*. Tal tarea tal vez no sea imposible, pero requiere mucho más esfuerzo que simplemente imaginarse a uno en dos instantes sucesivos con almanaques atrás indicando fechas distintas.

De acuerdo con esto existen casos en los que *aparentemente* nos imaginamos algo, y sin embargo no es así. Esto puede ser lo que nos pasa cuando imaginamos un zombie: tal vez no nos imaginamos en absoluto un zombie. Tendríamos que imaginarnos, en primer lugar, el sistema nervioso que procesa la información, ¿pero eso es algo que podemos hacer? Es probable que aún cuando creamos sinceramente poder concebir un zombie nos falte el conocimiento necesario para saber qué implicancias tiene el funcionamiento del sistema nervioso, con sus miles de millones de neuronas, redes, neurotransmisores y toda la mar en coche. La intuición materialista, entonces, nos lleva a ver como posibilidad que si tuviéramos en cuenta todas las leyes y entidades físicas que participan en el modo en que el sistema nervioso da lugar a la experiencia ya no podríamos imaginarnos dos mundos materialmente idénticos, cuyos correlatos fenoménicos fueran distintos.

La imaginación basta para responder sobre la *necesidad* de muchas cosas. Por ejemplo, a la pregunta “¿es posible un círculo cuadrado?” la imaginación no alcanza siquiera a concebir qué es lo que se está preguntando. ¿Qué sería un círculo cuadrado? Si intentamos imaginar un círculo cuadrado tropezamos siempre con algún tipo de líneas que viola la definición de círculo o de cuadrado. De esto obtenemos, apelando a la concebibilidad, que ni siquiera es posible un cuadrado. Sin embargo no logro imaginar qué sería imaginar -valga la redundancia- un cableado quasi infinito entre en un enorme número de componentes y subcomponentes. Imaginar un zombie filosófico es una tarea imposible para nuestra imaginación, y en tanto no podemos dimensionar dicha tarea, especificando los elementos que hay que hacer *concebibles*, tampoco tiene sentido afirmar que sea una posibilidad imaginar que los mismos estados físicos no dan lugar a los mismos estados fenoménicos.

4.1.2. Lo que Mary sabe

La hipótesis del identitarismo, construida a partir de la eliminación de la conciencia fenoménica, nos permite desconocer la brecha epistémica sugerida por Chalmers como primer premisa del argumento epistémico: A partir de esta hipótesis debemos pensar que los zombies tal vez no sean concebibles, o que, contrario a lo que sugiere el argumento explicativo, no hay nada que no pueda ser descrito en términos estructurales y/o funcionales. En efecto, si las propiedades *fenoménicas* no son intrínsecas y si la experiencia consciente *es*, como afirma Dennett, la disponibilidad general de la información procesada por los mecanismos cognitivos, tal vez no hayan razones para pensar que hay una brecha epistémica.

Nos queda por considerar el particular caso de Mary, la científica que nunca salió de una habitación en blanco y negro (es decir, que nunca vio el color) y que sin embargo puede conocer todo sobre nuestro cerebro. Este argumento también se sostiene en una intuición acerca de la brecha explicativa, ya que supone que una vez que Mary salga de su incolora habitación al mundo colorido descubrirá que hay más información acerca de los colores de la que ella podía obtener considerando los hechos físicos. Sin embargo, al igual que hemos hecho con los argumentos anteriores, debemos pensar que esta brecha responde a una intuición que no está necesariamente justificada.

La eliminación de propiedades intrínsecas efectuada por Dennett implica que todas las propiedades fenoménicas de la conciencia se pueden conocer a partir de los datos del mundo físico. De acuerdo con esto Mary, teniendo un conocimiento omnisciente de los procesos físicos, debería poder saber cómo se ve el color rojo. Quien no considere esto intuitivo lo invito a recordar la exagerada que nos asistió al considerar concebibles a los zombies podría estar operando en este caso: ¿realmente podemos imaginar qué implica conocer todas las leyes físicas? ¿el modo particular en que opera cada elemento de nuestro cerebro y todas sus implicancias? De no ser así no podemos decir que logramos imaginar lo que sería conocer todos los hechos físicos sin conocer los fenoménicos. Tal vez lo que puedo hacer es imaginar *sólo* una parte del fenómeno.

Hay un elemento que se reitera en cada uno de los argumentos anti-

materialistas, la intuición acerca del vínculo entre la realidad material y fenoménica parece preceder la construcción de los argumentos. En un sentido los argumentos constituyen una petición de principios: Mary podrá prever la experiencia del color rojo siempre que hayamos convenido que conocer los estados físicos basta para conocer los estados cualitativos de la experiencia, y no podrá hacerlo si creemos que los estados conscientes son independientes de la configuración física. Esto, a su vez, nos lleva a cuestionar la legitimidad con la que Chalmers reconstruye los argumentos epistémicos y el argumento explicativo, especialmente cuando afirma que los argumentos constituyen una apelación a una brecha epistémica para luego inferir un salto ontológico. En las intuiciones desarrolladas parece que de las suposiciones ontológicas se sigue una brecha epistémica.

Finalmente, si reparamos que la estructura de los argumentos anti-materialistas tiene la forma de un condicional del tipo ‘Si hay una brecha epistémica entonces hay una brecha ontológica’ (Chalmers, 2003), el eliminativismo presentado permite, por aplicación del *modus tollens*, e.e. negando que haya una brecha ontológica, negar que haya una brecha epistémica.

4.2. Teorías rivales y evidencia empírica

Supongamos que, contrariamente a la estrategia que he desarrollado, nos conformamos con la eliminación del problema difícil sin cuestionar el lugar de la *conciencia fenoménica* o los qualia en nuestra ontología. Supongamos que el problema difícil es una ilusión originada en la imposibilidad para determinar las variables en juego para el surgimiento de los estados conscientes (después de todo las neurociencias son una disciplina muy joven aún). En este caso todos los problemas atinentes a la naturaleza de nuestros estados mentales, incluso los conscientes, pueden ser explicados en términos de función y estructura, es decir, no hay más que problemas *fáciles*, entre los que ahora también está la conciencia fenoménica. Obviamente, **la noción de conciencia fenoménica en este caso será necesariamente deflacionada, ya que para constituir un problema fácil al menos debe ser descriptible en términos de función y estructura, lo que la hace relacional y, por supuesto, no intrínseca.**

Nos vemos llevados a elegir entre dos opciones: (1) hay una conciencia

fenoménica explicable en términos de función y estructura, de manera que sólo hay que encontrar el mecanismo correcto para explicarla; y (2) no hay una conciencia fenoménica en tanto tal, sólo hay conciencia de acceso y por lo tanto no hay un mecanismo específico que da lugar a la conciencia fenoménica, sino que los distintos mecanismos cognitivos operan de modo que *constituyen* estados conscientes.

Estas dos opciones se corresponden con las teorías de tipo disociativista, en el primer caso, y con las teorías que rechazan dicha estrategia en el segundo. Ambas teorías son igualmente viables con la sola eliminación del problema difícil. Dicha eliminación *tal vez* baste para romper el vacío epistémico entre los estados conscientes fenoménicos y *algo más* en virtud de lo cual esos estados se puedan comprender. Pero este *algo más* no está especificado por la sola eliminación del problema difícil. Mientras que la opción (2) correspondiente a la propuesta de Dennett es viable, también lo es la opción (1), por la cual las teorías disociativas pueden explicar los *-diet-* qualia en términos de algún mecanismo específico, distinto de todos los mecanismos de acceso (como sucede con el procesamiento recurrente en el caso de Lamme).

Pero **la eliminación del problema difícil a partir de la eliminación de los qualia (o de la conciencia fenoménica como estados cuyas propiedades son independientes de los estados de acceso) implica que hay una identificación estricta entre acceso y conciencia: los mecanismos cognitivos son la conciencia.** De este modo el argumento eliminativista -de la conciencia fenoménica y no sólo del problema difícil- se constituye en una heurística de la investigación empírica y nos permite escoger entre vías alternativas -y contradictorias- como lo son las teorías disociativistas, que postulan un mecanismo específico para la conciencia, y aquellas teorías que se niegan a este postulado. En particular, si aceptamos que no hay una conciencia en los términos en los que la hemos presentado inicialmente, nos vemos forzados a escoger por las teorías que limitan su *explanandum* a los estados conscientes a los estados *accedidos*.

Dennett (2005), muy consecuentemente, da apoyo explícito a la teoría del GWS, que antes se hallaba confinada a explicar la conciencia de acceso por no poder dar una respuesta a qué sea la conciencia, y que ahora puede considerarse explicativa de la conciencia en tanto y en cuanto explica aquellos mecanismos de cognición y percepción que se corresponden con la experiencia consciente. De hecho el modelo

de conciencia de los *múltiple drafts* (1991) elaborado por Dennett (1991, 2005) comparten un mismo trasfondo teórico con la teoría del espacio de trabajo global bajo el cuál se puede considerar la continuidad entre ambos proyectos.

Sin embargo esto parece contradecir la evidencia del *overflow* fenoménico, citada en el apartado 1.4, según la cuál pueden identificarse *residuos* fenoménicos independientes del acceso por parte de los sujetos. Esta evidencia aún necesita ser discutida para poder dar crédito a la estrategia doblemente eliminativista presentada.

4.2.1. El *overflow* fenoménico revisado

En el primer capítulo hemos comentado dos experimentos que dan lugar a la especulación en relación a la existencia de estados fenoménicos que escapan al acceso cognitivo. El experimento de Sperling (1960) presenta un conjunto de letras que excede el límite de la memoria inmediata y muestra que los sujetos pueden acceder a cualquiera de ellas si se utiliza una pista aún cuando hayan desaparecido los estímulos. Las versiones más recientes de este experimento (Landman, Spekreijse y Lamme, 2003; Becker, Pashler y Anstis, 2000) también utilizan un número de estímulos mayor al que determina la memoria de trabajo o inmediata y muestra que los sujetos pueden identificar los cambios en cualquiera de ellos si los mismos son señalados con una pista, aún cuando hayan desaparecido los estímulos. Tanto Sperling (1960) como las modificaciones experimentales que le sucedieron aciertan en mostrar que los sujetos pueden manipular una cantidad de estímulos superior a la que estipula la memoria de trabajo, sin embargo me parece objetable que de ello se infiera la existencia de estados fenoménicos que exceden al acceso por parte del sujeto. Esta crítica puede ser abordada desde: (1) la evidencia empírica que contradice los experimentos mencionados, así como aquellas interpretaciones alternativas posibles de estos experimentos; y (2) las consideraciones en relación a la supuesta naturaleza de la conciencia fenoménica, su supuesta relación con la conciencia de acceso, y el lugar del problema difícil en la explicación científica.

En relación a la evidencia empírica es de destacar evidencia reciente que muestra que la confianza de los sujetos en que su experiencia es más rica de lo que pueden reportar puede deberse a un sesgo. De Gardelle, Sackur y Kouider (2009)

han mostrado que los sujetos creen tener una experiencia rica de los estímulos, aún cuando los mismos son alterados de modo drástico. En el experimento se presentaron tarjetas que combinan letras y pseudoletas -i.e. letras rotadas- a modo de estímulo, pero los sujetos no han podido identificar las pseudoletas en sus reportes (p.). La confusión de las pseudoletas con letras reales se debe, según los autores, al carácter parcialmente ilusorio de la experiencia por fuera del foco de la atención. Los autores también consideran que la persistencia en el reporte de ver sólo letras, aún cuando los sujetos saben que se están usando como estímulos pseudoletas, se debe al importante rol que juegan los sistemas de predicción en interacción con la información parcialmente accesible (p. 569).

Los defensores del overflow pueden argumentar que las letras rotadas aún pueden ser percibidas como letras, por ejemplo podemos suponer que existe un mecanismo específico que detecta letras, así como hay uno que detecta rostros, y defender que este se activa aún cuando las letras están al revés (del mismo modo que se activaría el mecanismo neural que se activa al detectar rostros si ve un rostro al revés). Sin embargo, para explicar de este modo el fenómeno de detección de las letras y pseudoletas debemos considerar la mediación de un mecanismo postcategorial, que se activa tras la identificación del estímulo. Esto no sólo es contradictorio con los tiempos de presentación de los estímulos -que son presentados por debajo del tiempo requerido para el direccionamiento atencional- (Coltheart, 1980), sino que además implica un mecanismo de acceso, del cuál podemos dar cuenta funcionalmente y para el cuál no necesitamos en absoluto apelar a estados fenoménicos. Siempre se mantendrá “controversial si una gran cantidad de información que está disponible sin ser reportable refleja procesamiento consciente o inconsciente” (Kouider, *et. al.*, 2010, p. 301, mi traducción).

Sin embargo, aún cuando hubiera otro modo de discutir el experimento de de Gardelle, Sackur y Kouider, éste no deja de aportar una nueva hipótesis para interpretar la evidencia en torno al overflow, ya que nos sugiere como argumento que la confianza de los sujetos en relación a la riqueza de su experiencia puede ser ilusoria. Apelar a una ilusión para interpretar la evidencia empírica es significativo porque supone que la experiencia depende en gran medida de los mecanismos cognitivos que la generan, en particular los mecanismos de predicción y expectación que intervienen en la interpretación relativamente pobre de los estímulos para

construir una experiencia más homogénea. Estos mecanismos son mecanismos cognitivos, de acceso, y por lo tanto no es necesario suponer la existencia de un mecanismo específico para la conciencia. Tal como hemos anticipado la forma en la que se interpreta la evidencia es transversal a la distinción entre teorías de la conciencia disociativas y cognitivas.

En 1980 Coltheart ya había cuestionado que los métodos utilizados permitieran estudiar que la memoria icónica pudiera entenderse a partir de la persistencia de la *experiencia* visual. En su trabajo considera tres sentidos en los cuales “un estímulo visual puede persistir psicológicamente por algún tiempo después de que el estímulo ha desaparecido” (p. 183, mi traducción): (1) En un sentido neural, cuando decimos que persiste la actividad neuronal del sistema visual; (2) en un sentido fenomenológico, por el cual decimos que el estímulo aún es *visible*; y (3) en un sentido informacional, al referirnos a la persistencia de la información acerca de las propiedades visuales del estímulo. Experimentos como el de Sperling sólo dan cuenta del overflow en caso de que podamos referirnos a la persistencia en un sentido fenomenológico, sin embargo de acuerdo con Coltheart sólo necesitamos considerar la persistencia de la *información* visual, para dar sentido al experimento:

Los experimentos sobre memoria icónica descritos, y otros como ellos, afirman que la esencia del concepto es que refiere a un almacenamiento de la información visual en descomposición. Estos experimentos no requieren que esta información sea visible (aunque podría serlo), o que sus bases neurales sean actividad persistente en el sistema visual (aunque también podría) (Coltheart, 1980, p. 188, mi traducción).

Coltheart acepta que el experimento de Sperling muestra dos tipos de memoria a corto plazo, una de corta capacidad, que perdura más pero que limita el acceso de los sujetos (memoria inmediata); y otra memoria de larga capacidad, que sin embargo decae de forma demasiado rápida para que los sujetos puedan acudir a ella para efectuar un reporte extensivo (memoria icónica), pero no acepta que esto implique una persistencia de la *experiencia consciente visual*. Su principal argumento surge de considerar que la persistencia en un sentido fenomenológico tiene dos propiedades: “su duración está inversamente relacionada con la duración del estímulo, y su duración está inversamente relacionada con la luminancia del

estímulo” (p. 207, mi traducción) pero que la memoria icónica no reacciona de acuerdo con estas dos formas, sino que su duración se mantiene estable ante las alteraciones en la duración y la luminosidad de los estímulos.

Di Lollo (1977) también afirma que los patrones de duración de la memoria icónica, independientes de la duración del estímulo “sugiere fuertemente que la memoria icónica podría ser vista con más propiedad como el producto de procesos neurales de duración finita más que como el almacenamiento sensorial cuyos contenidos comienzan a decaer cuando la estimulación termina” (p. 243, mi traducción). Mientras que Coltheart se inclina por identificar la persistencia psicológica de las propiedades visuales en términos de información, Di Lollo toma el camino de identificarla con los procesos neuronales, pero en ambos casos se mantiene la prescindibilidad de la persistencia fenomenológica para interpretar los resultados de Sperling. Pero ambos señalan lo mismo que Philips (2011): “no tenemos razones para negar que la performance en los estudios de detección del cambio de Landman and Sligte’s resulta simplemente de la persistencia informacional no consciente y subpersonal. Correlativamente, no tenemos razones para postular una conciencia rica preservada en una forma frágil de memoria explícita” (p. 407, mi traducción)

Es cierto que los argumentos de Coltheart y Di Lollo no son concluyentes, y no excluyen la posibilidad de que haya conciencia visual mientras se ejecuta la memoria icónica, pero el punto es que tampoco hay razones para identificar estos dos fenómenos. En efecto: ¿por qué la información disponible por los sujetos del experimento de sperling debe ser consciente, y no simple procesamiento de la información por parte de los mecanismos cognitivos? La única respuesta que parece encontrar es la autoridad del reporte en primera persona que señala que la experiencia es más rica que aquello que se puede reportar.

Por último me gustaría citar en apoyo a la hipótesis de la ilusión como explicación del reporte fenoménico cierta evidencia en relación a la ceguera al cambio en las zonas de visión periférica. En el ejemplo de la figura 5, tomada de Cohen y Dennett (2011), se muestran dos imágenes, una de una escena natural, y la otra de la misma escena pero con los colores y las formas de la periferia degradados y alterados. Cuando estas imágenes son presentadas en una rápida sucesión los observadores, fijando su atención en el centro de la imagen, son incapaces de

detectar la diferencia entre una y otra imagen, preservando la percepción de la escena natural, sin las alteraciones. Esto sugiere, una vez más, una participación activa de la atención y de los mecanismos de expectación en relación a la percepción periférica, o, dicho de otro modo, que hay un factor ilusorio en la percepción de los sujetos.



Figura 5. La fotografía muestra dos escenas. La de la derecha presenta una degradación en sus formas y colores respecto de la de la izquierda. Si se cambia una imagen por otra mientras el sujeto atiende al centro estos cambios pasan desapercibidos.

4.2.2. El experimento perfecto

La evidencia citada, así como las interpretaciones alternativas a la que ésta nos conduce, también nos permite considerar un modo más agresivo de atacar el argumento del overflow. Si antes dijimos que los experimentos de Sperling y otros nos permitían identificar una memoria de procesamiento con características particulares, pero equivocaban el camino al identificar esta memoria con propiedades fenoménicas, ahora me interesa considerar de un modo más general en qué medida podemos decir que hay estados fenoménicos no accedidos y que aún así pueden ser identificados. En particular me propongo defender que **“parece imposible mapear estados neuronales y fenoménicos sin confiar en el acceso”** (Kouider, *et. al.*, 2010, p. 304, mi traducción), apoyándome en el **experimento perfecto** (Cohen y Dennett, 2011).

El experimento perfecto busca un aislamiento perfecto entre el supuesto mecanismo de la conciencia visual para los colores y el resto de los mecanismos de acceso. Imaginemos que por medio de una implementación quirúrgica aislamos el área que permite la conciencia del color. En tal caso, ante un estímulo el sujeto podrá

identificar todos los rasgos de su experiencia excepto el color, y de esta manera reportará una experiencia sin color, ya que el área que procesa la visión del color está aislada de las áreas que permiten el reporte verbal o conductual, incluso está aislada de toda memoria de trabajo que permita retener el estímulo. Debemos aceptar que “las teorías que postulan una disociación entre conciencia y función asumirán necesariamente que los participantes del ‘experimento perfecto’ están conscientes del color de la manzana pero simplemente no pueden acceder a la experiencia” (Cohen y Dennett, 2011, p. 362, mi traducción). Esto nos lleva a preguntarnos: ¿cuánto podemos saber de la conciencia fenoménica independientemente del acceso?

El único modo que tenemos de acceder a estados fenoménicos, subjetivos y privados, es a través del reporte por parte del sujeto de dichos estados, de dichas experiencias. ¿Qué sería decir que un estado fenoménico no es accedido por el sujeto, que el sujeto tiene una experiencia que no puede identificar? Este problema es el mismo que motivó (como ya hicimos notar en el capítulo 2) a Lamme a buscar una definición neural de la conciencia. Sin embargo esta es una estrategia inválida, ya que para definir la conciencia a partir de un marcador neural necesitamos tener razones para escoger tal marcador y no otro, y esto es algo que sólo podemos conseguir a través del reporte de sujetos que indiquen tener o no tener conciencia acerca de algo.

Imaginemos el procedimiento estándar para identificar los correlatos neurales de la conciencia: presentamos distintos estímulos a un sujeto y pedimos su reporte acerca de lo que percibe conscientemente mientras prestamos atención al escaneo por fMRI de sus estados cerebrales. Cada vez que el sujeto afirma tener una experiencia del estímulo no sabremos decidir si aquellas zonas que se activan se corresponden con el mecanismo de la conciencia fenoménica o el mecanismo del reporte verbal o reporte conductual si ponemos botones en vez de pedir palabras. Si el sujeto afirma haber tenido la experiencia tampoco podemos descartar la atención, o con la memoria inmediata que almacena y sostiene la información que se procesa.

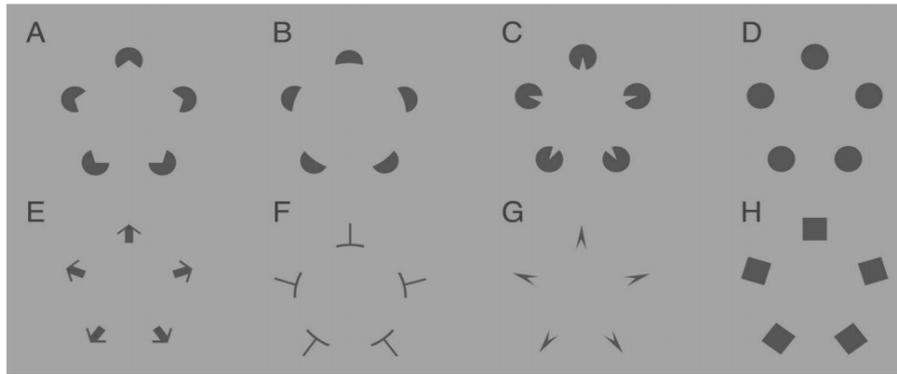


Figura 6. Arriba a la izquierda la figura de Kanizsa, que genera la ilusión de estar viendo un rombo. El resto de las figuras son ejemplos de estímulos de control de los autores.

Lamme (2012) discute que este experimento sea perfecto, de hecho no muestra lo que un experimento debería mostrar: que cierta área de procesamiento se corresponde con la conciencia fenoménica y no con algún mecanismo de acceso. En su breve respuesta argumenta que para ello el experimento perfecto debería identificar mecanismos neuronales asociados a la percepción, y que si estos “mecanismos permanecen cuando los mecanismos neurales de acceso son eliminados, uno debe concluir que la conciencia sin acceso existe” (Lamme, 2012, p. 138, mi traducción). En esta dirección Vandenbroucke, Fahrenfort, Sligte y Lamme (2014) lograron identificar un marcador neural para la detección de la figura Kaniza (figura 6) que se activa con independencia de si la misma es o no accedida por el sujeto. Esto es tomado por ellos como un indicador de actividad perceptiva aún cuando otros indicadores de acceso -como el de atención- están ausentes. Lamme parece entender que el problema que Dennett ve está en la imposibilidad efectiva para aislar distintas áreas, pero el verdadero problema radica en la imposibilidad para distinguir entre los mecanismos de acceso y el supuesto mecanismo de la conciencia ya que cualquier interacción del paciente con el medio implica la activación de algún sistema cognitivo. En palabras de Cohen y Dennett (2012):

Sin estos mecanismos un paciente no podría retener las instrucciones en la memoria, atender las palabras que decimos por encima de los sonidos del scanner ni decidir imaginar el estímulo que corresponde a una respuesta prediseñada. Sin estos mecanismos no hay razón para creer que este paciente sea consciente de algo (p. 139, mi traducción)

Cuando Vandenbroucke, Fahrenfort, Sligte y Lamme inhiben los mecanismos de acceso, entonces no tenemos modos de cerciorarnos de que la activación de un marcador neuronal cualquiera va acompañado de una experiencia consciente. De hecho la ilusión que surge de la figura de Kaniza puede ser *cognitiva* y no *meramente* perceptual, caso en el cual podremos detectar la activación de patrones idénticos tanto si la ilusión es cognitiva y consciente, como si es cognitiva y no manifiesta el aspecto cualitativo de la experiencia fenoménica. Afirman que “a través de la inferencia perceptual, el input visual pasa de ser fragmentado y sinsentido a objetos delimitados y escenas” (p. 955, mi traducción) pero lo que está en cuestión es que haya *algo* fragmentado y sin sentido a lo cuál los mecanismos de inferencia perceptual tienen acceso, o si dichos mecanismos dan lugar -en su conjunción con otros tantos (o no)- a la conciencia. Cómo saberlo es imposible en tanto cualquier intento por determinar ese *excedente* vendrá de la mano de algún tipo de operacionalización y funcionalización.

Otra forma de entender las implicancias del experimento perfecto es notando que si la conciencia puede ser explicada tanto como el resultado de las causas ascendentes (neuronales), como el resultado descendente de funciones cognitivas de alto orden (memoria, atención, reporte, etc), sólo el segundo de los permite eludir la crítica del experimento perfecto, ya que si la conciencia fuera una capacidad psicológica autónoma explicable en términos de causas ascendentes, no por ello dejaremos de necesitar las funciones de alto orden que evidencien la hipótesis a confirmar (que tal o cual mecanismo es el encargado de la conciencia -fenoménica).

4.2.3. Overflow y subjetividad

Hay una última razón que creo relevante considerar para desestimar la evidencia empírica que señala la existencia de overflow fenoménico, y la misma se deduce de la consideración del carácter subjetivo del overflow. Nagel (1974) es quien elabora la ya citada fórmula para definir la conciencia: “Fundamentalmente un organismo tiene estados mentales conscientes si y sólo si hay algo que es como ser ese organismo - algo que es como ser para ese organismo [...]./ Podríamos llamar a

esto el carácter subjetivo de la conciencia” (p. 436, mi traducción). En efecto decimos que alguien es consciente de algo cuando ese alguien se *siente* en un estado particular, y no sólo *está* en ese estado. Tener una experiencia consciente radica en que alguien es el sujeto de esa experiencia. A menos que uno se incline por considerar dualismos anacrónicos, donde las propiedades cualitativas de la experiencia tengan existencia incluso con independencia del sujeto, entonces necesita considerar al sujeto de una experiencia consciente para afirmar que hay una experiencia consciente.

Tanto Block (1995) como Chalmers (1995) aceptan el carácter subjetivo de los estados fenomenológicos. El primero afirma que “lo que hace a un estado fenoménicamente consciente es que hay algo que es ‘como’ [Nagel, 1974] ser o estar en ese estado” (p. 228, mi traducción). El segundo, en la misma tónica, dice: “El problema realmente difícil de la conciencia es el problema de la experiencia. [...] Como Nagel (1974) ha puesto, hay algo que es como ser un organismo consciente. Este aspecto subjetivo es la experiencia” (p. 436, mi traducción). Nótese las sendas referencias a Nagel y podremos notar que el rol de la subjetividad en la consideración de los estados conscientes es una cuestión que requiere mucho menos discusión que la consideración del lugar de la intrinsicalidad en relación a esos estados. Pero ¿por qué es tan importante la discusión de la subjetividad para considerar el overflow?

Coltheart (1980) ya había notado, tras distinguir la información visual, procesada subliminalmente, y la persistencia visual que “los experimentos sobre memoria icónica requieren que el sujeto procese información visual después de que el estímulo visual ha terminado; por contraste, los experimentos sobre persistencia visible dependen de las experiencias de sensación visual del sujeto después de que el estímulo ha desaparecido” (p. 188, mi traducción).

En efecto, sería un resultado sumamente paradójico -es cierto que esto no significa que ello sea imposible- encontrarnos con que hay estados conscientes que no son accedidos por el sujeto, porque un estado consciente se caracteriza por la relación de acceso por parte del sujeto: alguien debe ser el sujeto de esa experiencia, sentirse *de algún modo*, ante el estímulo presentado. No hay alternativa posible, un estado fenoménico no puede ser procesado subliminalmente, ¿o cuál sería la distinción en tal caso entre esos estados y los estados cognitivos?

~~En suma podemos decir que la estrategia eliminativista permite discutir el argumento explicativo que da lugar al problema difícil negando la brecha ontológica ('la conciencia fenoménica posee rasgos no funcionales') y asumiendo una intuición materialista a partir de la cuál puede replicar aquellas intuiciones que conforman las premisas de los argumentos antimaterialistas.~~

~~Asimismo desde una posición eliminativista podemos discutir el argumento del overflow fenoménico y sugerir una interpretación consistente con una teoría cognitiva de la conciencia, al mismo tiempo que ofrecer un argumento que parece dejar fuera de las ciencias la posibilidad de determinar el overflow. En relación a este punto me gustaría aclarar que aún cuando he defendido tal posición, el estado actual de la cuestión amerita dejar abierta la posibilidad de que nuevos diseños experimentales nos den una mejor aproximación a la cuestión. Más que determinar con absoluta certeza el rol de la cognición en la conciencia, he pretendido seguir a Brown (2014) cuando afirma que "no necesitamos aspirar a una meta así inalcanzable. Deberíamos preguntarnos si, dentro de los confines de los estándares científica mente aceptables de evidencia, el balance de la misma favorece a una u otra teoría" (p. 3, mi traducción).~~

Desafíos para el eliminativismo

5.1. El nuevo *explanandum* de la ciencia de la conciencia

Hasta ahora hemos mostrado que la mejor forma de eliminar el *problema difícil* es eliminando la conciencia fenoménica en tanto distinta de la conciencia de acceso. De acuerdo con esta visión la tarea de las neurociencias es abordar el estudio de los mecanismos cognitivos, que Block incluye entre los mecanismos de acceso y Chalmers entre los problemas fáciles. En efecto, Dennett (2005) afirma que los problemas a abordar “tienen explicaciones mundanas, que no requieren revoluciones en la física ni novedades emergentes. Sucumben, con mucho esfuerzo, a los métodos standard de las ciencias cognitivas” (p. 74-75, mi traducción) y defiende que el problema difícil es reducible a la suma de problemas sencillos. Esto es consecuente con el identitismo que predica, por el cual la conciencia no es el resultado del procesamiento cognitivo, sino que *es* dichos mecanismos. Sin embargo puede parecer que al reducir el problema difícil de la conciencia a la suma de los problemas fáciles y aceptar la identidad entre ambos estamos eliminando también la conciencia como un *explanandum* independiente de cada mecanismo cognitivo. El eliminativismo debe dar cuenta en qué sentido se debe interpretar esta identidad, de lo contrario el argumento eliminativista implica la eliminación de la experiencia consciente en tanto un fenómeno particular (algo a lo que Dennett y yo nos oponemos). En este sentido el principal problema del eliminativismo es que no hay consenso en los términos en los que se puede identificar la experiencia consciente como *explanandum* de la ciencia de la conciencia.

5.2. Evidencia de procesamiento cognitivo inconsciente

Además de la simple intuición de que somos conscientes, y de que ser conscientes designa un estado particular y distintivo, hay otra razón por la que no podemos aceptar la identificación entre conciencia de acceso y conciencia, o entre procesamiento cognitivo y conciencia. Hay múltiple evidencia que señala que

podemos ejecutar un gran número de procesos cognitivos de alto nivel de modo inconsciente: procesamiento semántico (Van Opstal *et al.*, 2011), procesamiento e integración de la información visual (Mudrik *et al.*, 2011), condicionamiento afectivo y toma de decisiones (Berlin, 2011), etc. Esta evidencia pone en cuestión constantemente la línea divisoria entre lo consciente y lo inconsciente, pero además supone que un mecanismo cognitivo puede ejecutar funciones complejas y de alto nivel sin necesidad de que los agentes sean conscientes de ello. Esto es muy distinto del argumento zombie, aunque éste puede ayudar a identificar la dificultad: Mientras al apelar a la concebibilidad de los zombies ponemos en juego la posibilidad de que en un mundo gemelo los sujetos no tengan experiencias conscientes, en este caso está en cuestión el hecho empírico de que los mismos aparatos cognitivos parecen poder operar con o sin la experiencia consciente.

Podemos resolver operaciones matemáticas y no ser conscientes, podemos procesar el lenguaje sin ser conscientes (y de hecho alguno de los niveles de procesamiento nunca llegarán a ser conscientes), podemos discriminar propiedades *visibles* de los objetos sin que podamos verlos, podemos jugar ajedrez sin ser conscientes -bueno, tal vez no nosotros pero sí las computadoras-, y una larga lista de etcéteras capacidades cognitivas que pueden ser procesadas sin ser conscientes.

5.2.1. Evidencia de procesamiento semántico inconsciente de números

Van Opstal, *et al.* (2011) muestran que existen capacidades cognitivas que trabajan de manera eficiente e inconsciente en el procesamiento de símbolos numéricos, sin que seamos capaces de percibirlos (de acuerdo con el reporte en primera persona). Más precisamente, la investigación indica que las personas podemos realizar cálculos aproximativos, o procesar información proporcionada a través de símbolos numéricos sin tener una experiencia de dichos símbolos. En el experimento que permitió evidenciar esto se presentaron sucesivamente a los participantes dos conjuntos de números, cada conjunto con cuatro dígitos en simultáneo (figura 7). Mientras el segundo de los conjuntos es el objeto visible de la percepción, el primero es inaccesible gracias a la aplicación de técnicas de masking. De acuerdo a las instrucciones del experimento se indicaba a los participantes responder si la suma de los cuatro dígitos del conjunto perceptible de números era

mayor o menor a veinte, mientras que otra variación del mismo experimento pedía indicar si el promedio de los dígitos era mayor o menor a 5. Sorprendentemente, en los casos en los que la respuesta correcta acerca del conjunto de números encubiertos coincidía con aquella que se seguía de los números mostrados, el tiempo de respuesta era menor que en aquellos casos en los cuáles el conjunto de dígitos enmascarados conllevaba una respuesta distinta del conjunto que era percibido conscientemente. Estos resultados no pueden ser sino una prueba de que inconscientemente procesamos el contenido propio de algunos símbolos, en este caso numéricos. Veamos otro caso.

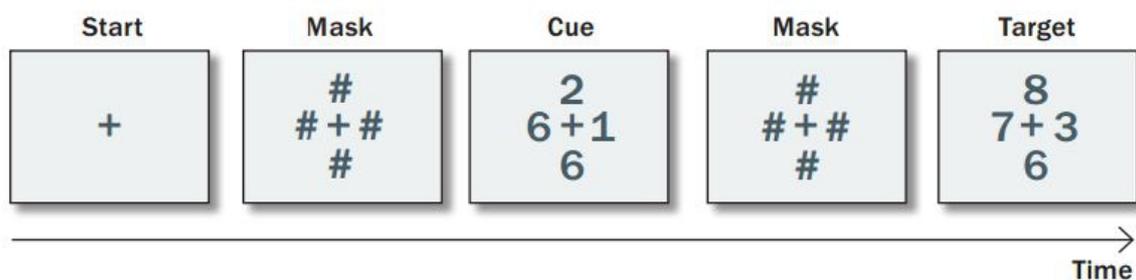


Figura 7. Se puede apreciar el tipo de estímulos y la secuencia experimental con la que Van Opstal, et. al., obtuvieron sus resultados.

5.2.2. Integración inconsciente de la información visual

Mudrik, *et al.* (2011) intentan combatir una vieja suposición acerca del rol de la conciencia: que ésta constituye un elemento necesario para la integración de la información. En su trabajo buscan identificar aspectos del procesamiento visual que se ejecutan sin necesidad de la conciencia, en particular si es posible la integración de los objetos visuales *significativos* en una escena, donde ni el objeto ni la escena son conscientemente percibidos. Para tal fin una serie de patrones coloridos aleatorios y rápidamente cambiantes es proyectada en un ojo (*masking por continuous flash suppression [CFS]*), mientras que la fotografía de una persona realizando una tarea es expuesta en el otro ojo. El primero de estos estímulos suprime al segundo por cierto tiempo, al cabo del cual la imagen pasa a ser visible. Las imágenes presentaban la particularidad de insertar objetos significativamente extraños para el contexto. La figura 8 incluye las imágenes utilizadas en el

experimento: tableros de ajedrez en vez de hornos, raquetas como flechas, y sandías que reemplazan pelotas de basketball. La hipótesis del experimento es que las imágenes con un objeto inusual para la escena alcanzarían el umbral de la conciencia en tiempos distintos a aquellas imágenes de mayor congruencia. Efectivamente, tras realizar los experimentos, pudieron constatar que había una diferencia promedio de 14 centésimas (2:50 y 2:64 segundos) entre el tiempo que requería la imagen incongruente para alcanzar el umbral de la conciencia y la imagen congruente. La mayor rapidez con la que la primera se hacía visible para los participantes señala que de algún modo la incongruencia entre el objeto y la escena fue percibida de modo inconsciente, aún cuando no podamos decir cómo.



Figura 8. Se pueden apreciar las fotografías estímulo utilizadas para determinar la integración inconsciente de los elementos visuales de una imagen. De cada par, la imagen de la izquierda alcanzó el umbral con mayor rapidez que la imagen de la derecha.

En experimentos similares los mismos autores presentaron -CFS mediante imágenes de rostros familiares y rostros desconocidos. En estos casos los resultados mostraron que la imagen de una cara conocida y afectivamente significativa alcanza el umbral de la conciencia con mayor rapidez que la imagen de un rostro desconocido.

5.2.3. Condicionamiento afectivo e inconsciente de la conducta

Otro caso interesante es el de procesamiento afectivo inconsciente, documentado por Winkielman, Berridge y Wilbarger (2005). Los autores presentaron a los sujetos experimentales -por medio de técnicas de masking, de modo que los sujetos no accedieran a ellas conscientemente- caras enojadas y felices, con el objetivo de evaluar si las mismas, siendo procesadas subliminalmente, podían influenciar la conducta posterior. Para ello, luego de presentado el estímulo, se intentaba determinar cómo reaccionan los sujetos antes el consumo de bebidas,

tomando como parámetros la cantidad de líquido ingerido y la preferencia cualitativa en relación a órdenes preestablecidos por el sujeto. De acuerdo con los resultados cuando el estímulo es un rostro feliz el consumo aumenta y la elección se inclina hacia las de mayor preferencia; mientras que las elecciones se invierten cuando el estímulo es un rostro enojado.

Berlin (2011) es una referencia obligada para acceder a un gran número de experimentos como los citados, donde se intenta determinar los límites y las capacidades de los sistemas cognitivos, o de acceso, para procesar información de manera inconsciente. Sólo quiero agregar un experimento citado allí, que también evidencia el comportamiento condicionado por el procesamiento inconsciente afectivo. En dicho experimento se trabajó con pacientes incapaces de discriminar gustos y sabores, se prepararon tres soluciones, una salada, otra a base de jugo de limón y una tercera dulce. Cuando estas bebidas fueron ingeridas por los participantes todos las describieron como sabiendo del mismo modo, pero cuando fueron ofrecidas de modo simultáneo todos inclinaron su elección por las soluciones dulces. Estos estudios también muestran que el valor afectivo de estímulos que no son conscientemente percibidos, y no producen respuesta afectiva consciente, aún puede motivar la conducta.

5.3. La distinción entre los procesos conscientes e inconscientes

Estos experimentos muestran, por un lado, que no podemos identificar, tal como intuíamos, la conciencia con los mecanismos cognitivos; ya que evidentemente hay situaciones en las que afirmamos que cierto estado de cosas es procesado inconscientemente y otras en las que el procesamiento es consciente, y en ambas el mismo mecanismo desencadena las acciones. Todo esto pone en evidencia la necesidad de encontrar un criterio que distinga la conciencia del inconsciente. Para Dennett (1991) identificar un límite claro entre el procesamiento consciente y el inconsciente es una cuestión compleja y en gran medida arbitraria (p. 125), lo cual no esconde la necesidad de encontrar y clarificar la naturaleza de tal límite.

El mismo Dennett y los teóricos de la GWST apelan al espacio de trabajo la diferencia depende de los niveles de procesamiento de la información de los que participan los contenidos. Los contenidos procesados modularmente, por *alguno* de

los mecanismos cognitivos específicos, preservan su anonimato, mientras que aquellos que son incorporados a la red global de trabajo, e.e. que están disponibles para el *conjunto* de las funciones cognitivas, se expresan de manera consciente. En el contexto de la GWST sólo la disponibilidad *global* de la información cognitiva es un estado consciente:

La tesis propuesta a consenso no es que la disponibilidad global causa algún efecto posterior -algo así como el soplo que aviva el fuego de los qualia, que así obtienen ingreso al Teatro Cartesiano, o algo así- sino que es, en sí mismo, un estado consciente (Dennett, 2005, p. 134, mi traducción).

Sin embargo, aún cuando Dennett haya disuelto el problema difícil, cuando la pregunta “¿por qué y cómo surge la experiencia consciente?” pueda ser respondida en términos científicos (apelando a estructuras y funciones) queda preguntarnos en qué sentido la disponibilidad global de la información puede ser la respuesta adecuada. Si así fuera con seguridad podríamos identificar por qué el procesamiento modular no alcanza para producir conciencia, o por qué algunas arquitecturas del tipo del GWS (con jerarquías, módulos y núcleos procesadores) no alcanzan a ser conscientes. ¿O si lo son?

Es cierto que ya no hay que dar cuenta de propiedades *fenoménicas*, pero sí esperamos que una teoría explicativa de la conciencia nos proporcione un criterio para distinguir los estados conscientes de los estados inconscientes. Si toda la información procesada responde a una misma naturaleza es difícil determinar por qué tenemos una experiencia tan distinta respecto de otros sistemas que también procesan e integran información. La imposibilidad para distinguir sistemas conscientes e inconscientes es el gran riesgo para aquellas posiciones que descansan en la consideración de un único tipo de procesamiento de la información, ya que es muy fácil construir una pendiente resbaladiza por la que todo se vuelva inconsciente.

Dennett, empero, considera que su propuesta tiene el elemento necesario para dar cuenta de la particularidad de la experiencia, a través de la explicación de la subjetividad en tanto ilusión constituida en el seno del espacio de trabajo global:

La reacción más natural a esta propuesta es una franca incredulidad: parece estar

dejando afuera el elemento más importante-el Sujeto! [...] El error detrás de esta extraviada objeción es no notar que la Primera Persona ya ha sido de hecho incorporada en los multifacéticos efectos posteriores a la influencia política que se obtiene en las distintas competencias (p. 143, mi traducción)

Cada disposición conductual, pensando tanto en las capacidades psicológicas de alto orden como en los procesos y estímulos de nivel más bajo, forma parte de una multiplicidad de sistemas y subsistemas, donde el “centro de gravedad” (Dennett, 1991) de la información procesada en la sala de cómputos constituye la ilusión del teatro cartesiano. Sin embargo la GWST esto no parece ser explicativo en absoluto, sino más bien la traducción a un nuevo grupo de metáforas que parecen ser más elocuentes y estar en un vocabulario que nos hacen imaginable el rumbo que debería tomar la investigación, pero que en absoluto son explicativos. ¿Por qué la suma de las disposiciones conductuales, aún cuando compartieran información y se refieran las unas a las otras, estableciendo relaciones jerárquicas y de alta complejidad organizacional, habría de llevarnos a la experiencia consciente? ¿No sería de lo más sencillo hacer un programa así? El mero acceso a múltiples niveles de información no parece darnos más que una supercomputadora.

Dennett no *apela* al sujeto para dar cuenta de la experiencia consciente, sino que lo define como el todo resultante de la suma de las partes -modulares, cognitivas, inconscientes. La relación entre ambas es de constitución, del mismo modo que muchos perros hacen una jauría, pero ninguno de los perros es la jauría. Con la conciencia pasa lo mismo, aún cuando muchos estados informacionales se procesan de modo inconsciente, la suma de ellos -en particular de aquellos que tienen relación con el espacio global de trabajo- somos nosotros y nosotras, personas conscientes. Este es el problema del tópico, en este caso de la conciencia, pero el problema del tópico se usa para referir a cualquier relación entre las partes y el todo, donde el todo exhibe propiedades que no exhiben las partes.

Dennett (2005) afirma que “hay muchos modelos de sistemas que responden a *versiones* de la cuestión del tópico de manera no controvertida, y todos son computacionales” (p. 12, traducción propia), y en efecto casi cualquier instrumento electrónico actual ejemplifica el modo en que los componentes están limitados en sus capacidades respecto del todo. El ejemplo reiterativo es, siempre, la computadora,

“cuyas partes no saben nada de ajedrez, pero puede derrotarte al ajedrez con una impactante confiabilidad” (Dennett, 2005, p. 12, mi traducción).

Es cierto que el problema del t3pico es un problema resoluble funcionalmente, y que apelar a 3l no es, *en principios*, controversial. Para discutir esto me gustar3a tomar el ejemplo que propone Hofstadter (1979) en *G3del, Escher, Bach*. Imaginemos una colonia de hormigas. La colonia manifiesta propiedades y capacidades particulares, distintas de las que atribuimos a las hormigas que forman parte de dicha colonia. Cada hormiga, por ejemplo, tiene un grado de libertad para la acci3n lo suficientemente amplio como para que atendiendo al comportamiento de cada individuo nos veamos impedidos de identificar los trazos, caminos y t3neles de la colonia. Del mismo modo para *interactuar* con la colonia no nos basta interactuar con una hormiga particular, al hacerlo seguramente la hormiga se desconcierte, o trate de seguir su camino, pero no ser3 hasta que hayamos obtenido un embotellamiento, algo que excede totalmente a la hormiga particular, que empecemos a ver que un trazo comienza a despoblarse. Arrastrar una hormiga tampoco forzar3 nuevos caminos, y s3lo marc3ndolos -es decir, sin necesidad de actuar con ninguna hormiga particular- podremos observar un comportamiento inducido en la colonia, y a trav3s de este en las hormigas particulares. Este ejemplo ilustra perfectamente por qu3 algunos sistemas, como las colonias de hormigas, al ser considerados como un todo poseen propiedades que no son adscribibles a sus partes. Pero el ejemplo no nos dice por qu3 es an3logo al caso de la conciencia. **Podemos identificar y en gran medida estudiar numerosos sistemas de este tipo, pero eso s3lo no hace que la conciencia sea uno de esos sistemas.**

Las distintas versiones del problema del t3pico no parecen exhibir un comportamiento del que llamamos propiamente consciente. Estoy dispuesto a aceptar, incluso, que el comportamiento de la colonia es equiparable al dominio de alg3n tipo de lenguaje (Hofstadter, 197, p. 315), pero tener un lenguaje, una inteligencia, habilidades pr3cticas no hace a un ser consciente, 3o s3? El eliminativismo que hemos defendido hasta ahora implica eso, pero no podemos apelar al eliminativismo ya que 3ste supone al problema del t3pico como camino para la explicaci3n de la experiencia consciente. Por otro lado tener una experiencia consciente, a3n cuando 3sta quede reducida a la suma de funciones cognitivas m3s simples, implica una serie de propiedades distintas de las propiedades de las partes

-el procesamiento del lenguaje, por caso-. Elaborar una serie de criterios no está dentro de mis objetivos, pero una vez más podemos apelar a la noción de Nagel, por la cual tener o estar en un estado consciente implica que dicho estado se *siente* de algún modo *para* el sistema. Difícilmente diríamos que la colonia se *siente* de un modo u otro en relación a sus propiedades, o que una computadora está *jugando* al ajedrez (de seguro nos está ganando, pero difícilmente esté jugando), o que la Galaxia de Andrómeda sea consciente de su espiralidad (si así fuera seguro preferiría verse más delgada, como una galaxia elíptica y no espiral). De hecho ninguno de los sistemas conocidos que ejemplifican el problema del tópico son sistemas de los que predicamos conciencia.

Otro problema que se presenta a la hora de considerar la explicación de la conciencia a partir del problema del tópico es que el holismo (“el todo es más que la suma de las partes”) incluye la aparición de propiedades emergentes. En un sentido amplio “la ‘emergencia’ describe el conjunto de propiedades nuevas que surgen cuando un nivel más alto de complejidad es formado a partir de componentes de menor complejidad, donde aquellas propiedades no están presentes” (Luisi, 2006, p. 112, mi traducción). Pero en este caso ¿cuáles serán las nuevas propiedades que tienen los estados conscientes? Cuando pensamos en los sistemas-*tópico* conocidos estas nuevas propiedades no son problemáticas porque generalmente nos referimos a propiedades relativamente simples y mensurables (la dirección de los caminos en contraste con la dirección que toma cada hormiga individualmente, en el caso de la colonia; la disposición de los cuerpos que contiene, en el caso de la galaxia), sin embargo cuando nuestra nueva propiedad es el sujeto parece que estamos sugiriendo una nueva propiedad un tanto más sofisticada.

Estas preguntas y razonamientos no pretenden que las dificultades señaladas sean inabordables, aunque sí son orientativas respecto del tipo de desafíos que aún no hemos ni empezado a responder y que son necesarios tratar para dar forma a la ciencia de la conciencia. Sin embargo la dificultad que algunas de ellas sugiere que identificar la experiencia consciente en tanto fenómeno particular de estudio puede resultar algo más complejo que simplemente identificarla con el conjunto de los sistemas cognitivos y capacidades psicológicas que nos son propias. Con esto no quiero decir que debemos renunciar al materialismo, ni que debemos reconsiderar propiedades fenoménicas, pero sí que tal vez debemos aceptar que la *naturaleza* de

la conciencia es distinta -no tan distinta- a la de la información eléctrica en mi cabeza o en tu computadora.

CONCLUSIONES

A lo largo de este trabajo he pretendido argumentar contra la idea de que hay alguna propiedad intrínseca de la experiencia consciente que nos impide explicarla en términos de estructura y función. Esto es lo que Chalmers (1995) llama el problema difícil. En el capítulo 1 hemos mostrado que la distinción que él propone entre problemas fáciles y problema difícil se corresponde con la otra de Block, entre conciencia de acceso y conciencia fenoménica. También hemos visto que estas formas de concebir la conciencia están sustentadas en argumentos filosóficos que atacan al materialismo y de evidencia empírica que parece apoyar las dicotomías mencionadas. La fuerza de las intuiciones que subyacen a este modo de concebir la conciencia se pone en evidencia al usarlas para analizar dos de las teorías de mayor circulación académica, la teoría del espacio de trabajo global, defendida por Dehaene y Dennett principalmente, y la teoría de la conciencia como recurrencia neuronal, propuesta por Lamme. Ninguna de ellas logra dar respuesta a las inquietudes sugeridas por el problema difícil y fracasan en su intento explicativo.

Sin embargo también hemos visto que se puede tomar una estrategia doblemente eliminativa: respecto de la conciencia fenoménica y del problema difícil. Si los argumentos que presenté en el capítulo 3 son correctos podemos quinear la conciencia fenoménica -en tanto intrínseca- y de este modo dejar atrás al problema difícil. De este modo no quedan más que tareas sencillas para explicar la conciencia. Entre **las p** principales ventajas de la estrategia eliminativista están (1) la heurística que nos brinda para dar respuesta a los argumentos antimaterialistas, socavando las intuiciones/premisas que los sostienen (ni los zombies son tan concebibles ni es tan inconcebible que alguien adquiriera conocimientos de la experiencia a través del correcto uso del conocimiento de la física); y (2) la potencialidad para generar una heurística empírica de investigación, favoreciendo a las teorías cognitivas por sobre aquellas disociativas, así como brindando una interpretación plausible del supuesto overflow fenoménicos. Respecto a las teorías disociativas, sus problemas fueron expuestas al tratar la teoría de la recurrencia (apartado 2.2), al considerar la

imposibilidad de una noción *diet* de conciencia fenoménica (apartado 3.2) y al discutir la evidencia del overflow a ella asociada (apartado 4.2). De todo ello se sigue que no puede haber experiencia consciente por fuera del procesamiento cognitivo (mientras que sí podemos identificar casos del procesamiento cognitivo que no consideramos conscientes).

Por último también hice manifiestas mis preocupaciones respecto a la dificultad que sucede al eliminativismo para identificar de modo positivo el *explanandum* propio de una ciencia de la conciencia, tanto desde el nivel neuronal como en relación a las capacidades psicológicas de alto nivel involucradas. Expuse argumentos filosóficos así como evidencia empírica que parecen generar dificultades a la hora de pensar la conciencia reducida a conciencia de acceso -o a los mecanismos cognitivos.

Sin embargo estos son desafíos que se presentan como resolubles a través de una mayor indagación empírica, que invitan a nuevas líneas de investigación y aportes que permitan avanzar en nuestra comprensión de la conciencia. En este sentido ninguna de las críticas y advertencias contra el simplismo al considerar la experiencia consciente como *explanandum* de una ciencia de la conciencia es equiparable a la limitación obscena que impone el problema difícil a la potencialidad explicativa de las ciencias.

En este sentido podemos resaltar la necesidad de que futuras investigaciones aborden el problema del overflow y añadan nueva evidencia empírica que permita complejizar las discusiones y -por qué no- dar mayor apoyo a las interpretaciones aquí defendidas. Del mismo modo es necesario profundizar en nuestra comprensión del modo en que funcionan los mecanismos cognitivos, si es que éstos son los componentes que hacen posible la conciencia. Este punto es fundamental si tenemos en cuenta que investigaciones recientes (Irvine, 2013) ha puesto en cuestión no ya la posibilidad de medir la conciencia, sino la posibilidad de construir medidas libres de sesgo que nos permitan describir el funcionamiento de las capacidades cognitivas.

Muchas otros tópicos de investigación están quedando sin tratar (el problema de la medida, los tantos algoritmos que pretenden predecir e identificar la medida, incontable y novedosa evidencia que nos acerca a una lectura neural de nuestra experiencia consciente, consideraciones en torno a la distinción entre conciencia transitiva e intransitiva, A.I., la conciencia en el reino animal, etc.), sin embargo la

legitimidad de todas las discusiones por venir dependen en gran medida de la problemática que he intentado abordar. Si la conciencia queda al margen del tratamiento funcional parece inevitable que las neurociencias se parezcan siempre a una nueva frenología (Uttal, 2001).

Sin embargo espero haber mostrado que esa situación está directamente vinculada a una concepción cargada -y no en un buen sentido- de la noción de conciencia fenoménica. Prescindiendo de ella podemos destinar nuestros esfuerzos a la empresa del estudio de la conciencia con el optimismo de saber que tenemos en nuestras manos un fenómeno real, y no una ilusión. El problema difícil, entonces, no es más que el resultado de una posición que se socava a sí misma, que crea una entidad que, por ficticia, no puede explicar.

REFERENCIAS

- Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge University Press.
- Baars, B. J. (1997). *In the theater of consciousness: The workspace of the mind*. Oxford University Press.
- Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Progress in Brain Research*, 150, pp. 45–53. Elsevier.
- Becker, M. W., Pashler, H., & Anstis, S. M. (2000). The role of iconic memory in change-detection tasks. *Perception*, 29(3), 273–286.
- Bennett, M. R., Hacker P.M.S. (2003). *Philosophical Foundations of neuroscience*. Oxford: Blackwell.
- Berlin, H. A. (2011). The neural basis of the dynamic unconscious. *Neuropsychanalysis*, 13(1), 5–31.
- Block, N. (1978). Troubles for functionalism. *Minnesota Studies in the Philosophy of Science* 9, 261-325.
- Block, N. (1995). How many concepts of consciousness? *Behavioral and Brain Sciences*, 18(02), 272–287.
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9(2), 46–52.
- Block, N. (2007). Comparing the Major Theories of Consciousness. En Michael Gazzaniga (ed.), *The Cognitive Neurosciences IV*. 1111-1123.
- Block, N. (2007a). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30, 5-6.
- Block, N. (2007b). Wittgenstein and Qualia. *Philosophical Perspectives*, 21(1), 73–115.
- Block, N. (2008). Consciousness and Cognitive Access. *Proceedings of the Aristotelian Society*, 108, 289–317. doi:10.1111/j.1467-9264.2008.00247.x
- Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*, 15(12), 567–575. doi:10.1016/j.tics.2011.11.001
- Block, N. (2014). Rich conscious perception outside focal attention. *Trends in*

- Cognitive Sciences*, 18(9), 445–447. doi:10.1016/j.tics.2014.05.007
- Block, N., Carmel, D., Fleming, S. M., Kentridge, R. W., Koch, C., Lamme, V. A. F., Rosenthal, D. (2014). Consciousness science: real progress and lingering misconceptions. *Trends in Cognitive Sciences*, 18(11), 556–557. doi:10.1016/j.tics.2014.09.004
- Boehler, C. N., Schoenfeld, M. A., Heinze, H.-J., & Hopf, J.-M. (2008). Rapid recurrent processing gates awareness in primary visual cortex. *Proceedings of the National Academy of Sciences*, 105(25), 8742–8747.
- Brown, R. (2014). Consciousness doesn't overflow cognition. *Frontiers in Psychology*, 5. doi:10.3389/fpsyg.2014.01399
- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
- Chalmers, D. J. (1998). On the Search for the Neural Correlate of Consciousness. *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*, 2, 219.
- Chalmers, D. J. (2003). Consciousness and its place in nature. *Blackwell guide to the philosophy of mind*, 102–142.
- Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, 15(8), 358–364. doi:10.1016/j.tics.2011.06.008
- Cohen, M. A., & Dennett, D. C. (2012). Response to Fahrenfort and Lamme: defining reportability, accessibility and sufficiency in conscious awareness. *Trends in Cognitive Sciences*, 16(3), 139–140. doi:10.1016/j.tics.2012.01.002
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, 27(3), 183–228.
- Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, 2, 263–275.
- De Gardelle, V., Sackur, J., & Kouider, S. (2009). Perceptual illusions in brief visual presentations. *Consciousness and Cognition*, 18(3), 569–577.
- Dehaene, S., & Changeux, J. P. (2005). Ongoing Spontaneous Activity Controls Access to Consciousness: A Neuronal Model for Inattentional Blindness. *PLoS Biology*, 3(5), e141. doi:10.1371/journal.pbio.0030141
- Dehaene, S., & Changeux, J. P. (2011). Experimental and Theoretical Approaches to

- Conscious Processing. *Neuron*, 70(2), 200–227.
doi:10.1016/j.neuron.2011.03.018
- 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.
doi:10.1016/j.tics.2006.03.007
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1), 1–37.
- Dehaene, S., Sergent, C., & Changeux, J.P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Sciences*, 100(14), 8520–8525.
- Dennett, D. C. (1988). Quining qualia. *Consciousness in Modern Science*. Consultado de <http://cogprints.org/254>
- Dennett, D. C. (1991). *Consciousness explained*. London: Back Bay Books
- Dennett, D. C. (1996). Facing backwards on the problem of consciousness. *Journal of Consciousness Studies*, 3(1), 4–6.
- Dennett, D. C. (2005). *Sweet dreams: Philosophical obstacles to a science of consciousness*. MIT press.
- Di Lollo, V. (1977). Temporal characteristics of iconic memory. *Nature*, 267(5608), 241–243.
- Frankish, K. (2012). Quining diet qualia. *Consciousness and Cognition*, 21(2), 667–676.
- Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, 31–52.
- Hebb, D. O. (2002/1949). *The organization of behavior: A neuropsychological theory*. Psychology Press.
- Hofstadter, D. R. (1979). *Goedel, Escher, Bach: an eternal golden braid*. New York: Random House-
- Huang, L., Treisman, A., & Pashler, H. (2007). Characterizing the Limits of Human Visual Awareness. *Science*, 317(5839), 823–825. doi:10.1126/science.1143515
- Irvine, E. (2013). *Consciousness as a scientific concept: A philosophy of science perspective* Springer.

- Jackson, F. (1982). Epiphenomenal qualia. *The Philosophical Quarterly*, 127–136.
- Kirk, R. (2003). Zombies. Disponible en <http://stanford.library.usyd.edu.au/archives/spr2012/entries/zombies/>
- Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, 11(1), 16–22. doi:10.1016/j.tics.2006.10.012
- Koch, C. (2012). *Consciousness: Confessions of a romantic reductionist*. MIT press.
- Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), 857–875. doi:10.1098/rstb.2007.2093
- Kouider, Sid, de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences*, 14(7), 301–307. doi:10.1016/j.tics.2010.04.006
- Kripke, S. (1971). Identity and necessity. *Perspectives in the Philosophy of Language*, 93-126.
- Kripke, S. (1980). *Naming and Necessity*, rev. ed. Cambridge, MA: Harvard.
- Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, 7(1), 12–18.
- Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), 494–501. doi:10.1016/j.tics.2006.09.001
- Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, 1(3), 204–220. doi:10.1080/17588921003731586
- Lamme, V. A., Supèr, H., Landman, R., Roelfsema, P. R., & Spekreijse, H. (2000). The role of primary visual cortex (V1) in visual awareness. *Vision Research*, 40(10), 1507–1521.
- Landman, R., Spekreijse, H., & Lamme, V. A. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, 43(2), 149–164.
- Laureys, S., & Tononi, G. (Eds.). (2011). *The neurology of consciousness: cognitive neuroscience and neuropathology*. Academic Press.
- Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64(4), 354-361.
- Lewis, D. (1983). Extrinsic properties. *Philosophical Studies*, 44(2), 197–200.

- Locke, J., (1975) *An Essay concerning Human Understanding*. New York: Oxford.
- Luisi, P. L. (2006). *The emergence of life: from chemical origins to synthetic biology*. Cambridge University Press.
- Mudrik, L., Breska, A., Lamy, D., & Deouell, L. Y. (2011). Integration without awareness: Expanding the limits of unconscious processing. *Psychological Science*, 22(6), 764–770. doi:10.1177/0956797611408736
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, 435-450.
- O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(05), 939–973.
- Phillips, I. B. (2011). Perception and iconic memory: what Sperling doesn't show. *Mind & Language*, 26(4), 381–411.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink?. *Journal of Experimental Psychology: Human perception and performance*, 18(3), 849.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(03), 417-424.
- Shear, J. (1997). Explaining consciousness: The hard problem. London: MIT press.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, 28(9), 1059-1074.
- Sitt, J. D., King, J. R., Naccache, L., & Dehaene, S. (2013). Ripples of consciousness. *Trends in Cognitive Sciences*, 17(11), 552–554.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74(11), 1.
- Stratton, G. M. (1897). Vision without inversion of the retinal image. *Psychological Review*, 4(4), 341.
- Stratton, G. M. (1896). Some preliminary experiments on vision without inversion of the retinal image. *Psychological Review*, 3(6), 611.
- Tye, M. (2013) Qualia, En The Stanford Encyclopedia of Philosophy, URL = <<http://plato.stanford.edu/archives/fall2013/entries/qualia/>>.
- Uttal, W. R. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain*. The MIT Press.
- Van Opstal, F., de Lange, F. P., & Dehaene, S. (2011). Rapid parallel semantic processing of numbers without awareness. *Cognition*, 120(1), 136–147.

doi:10.1016/j.cognition.2011.03.005

Vandenbroucke, A. R. E., Fahrenfort, J. J., Sligte, I. G., & Lamme, V. A. F. (2014). Seeing without knowing: Neural signatures of perceptual inference in the absence of report. *Journal of Cognitive Neuroscience*, 26(5), 955–969.

doi:10.1162/jocn_a_00530

Weatherson, B. & Marshall, D. (2014) Intrinsic vs. Extrinsic Properties, *En The Stanford Encyclopedia of Philosophy*, URL = <<http://plato.stanford.edu/archives/fall2014/entries/intrinsic-extrinsic/>>.

Winkielman, P. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin*, 31(1), 121–135.

doi:10.1177/0146167204271309