

Autoconciencia prerreflexiva: dos argumentos contra las teorías representacionales de la conciencia

Israel Grande-García
Instituto de Investigaciones Filosóficas, UNAM

Abstract

Representational theories of consciousness try to explain phenomenal experience by the use of terms without implying consciousness. The two main theories are representationalism and higher-order representational theories (HOR). The key argument for the first one is the transparency of experience according to which, it is impossible to discern, via introspection, any intrinsic properties of an experience of x that are not experienced as properties of x . Representationalism states that our experiences show qualitative features that are identical to the properties things are represented as having, and therefore, phenomenal character of experience *is* identical to its representational content. On the other hand, HOR theories account for conscious states in terms of higher-order representation, that may take the form of a quasi-perception or a thought that it's directed to a first-order mental state. In this paper I put forward two arguments against representational theories, on the basis of a theory which affirms that every conscious experience has an implicit pre-reflective self-consciousness of some kind, that it is characterized of having a for-me-ness aspect (sense of ownership). The first argument is against representationalism. The argument asserts that this sense of ownership (1) is an intrinsic aspect of experience, and (2) is discernible via introspection; therefore, the phenomenal character of experiences is *not* identical to its representational content, and consequently, representationalism is false. The second argument is the problem

of targetless higher-order states. Suppose that a subject *S* has a HOR of being in a pain state *p*, but without having *p*. According to HOR theories, *p* is non-conscious, because it doesn't exist. But, *S*'s having HOR about *p*, will cause that *S* would have a phenomenal experience of *p*. Nevertheless this result in incoherent. For pre-reflexive self-consciousness theory, on the other hand, *p* will not be a conscious state if it doesn't belong to *S*, nevertheless *S* *does* have consciousness of *p* (being *p* different from pain, *e.g.*, a believe, a mental imagery) therefore this sense of ownership is what makes *p* a phenomenal experience.

Keywords: phenomenal consciousness, pre-reflective self-consciousness, sense of ownership, representationalism, higher-order representation.

Resumen

Las teorías representacionales de la conciencia intentan dar cuenta de la experiencia fenoménica con términos mentales sin implicar la conciencia. Las dos principales son el representacionalismo y las teorías de representación de orden superior. El principal argumento para la primera es el de la transparencia de la experiencia, según el cual es imposible discernir, vía introspección, las propiedades intrínsecas de una experiencia de *x* que no se experimenten como propiedades de *x*. El representacionalismo sostiene que nuestras experiencias muestran rasgos cualitativos que son idénticos a las propiedades de los objetos que nuestra experiencia representa y que, por lo tanto, el carácter fenoménico de la experiencia *es* idéntico a su contenido representacional. Las teorías de representación de orden superior o HOR (por *higher-order representation*), por su parte, explican los estados conscientes en términos de una representación de orden superior, que puede tomar la forma de una cuasi percepción o un pensamiento que va dirigido a un estado mental de primer orden. En este artículo presento dos argumentos contra las teorías representacionales desde una teoría que establece que toda experiencia consciente lleva implícita un tipo de autoconciencia prerreflexiva caracterizada por tener un aspecto por tener un aspecto: ser característicamente mía o para mí (sentido de pertenencia). El primer argumento es contra el representacionalismo. El argumento sostiene que este sentido de pertenencia (1) es un rasgo intrínseco de la experiencia y (2) es discernible por medio de la introspección; por lo tanto, el carácter fenoménico de la experiencia *no* es idéntico a su contenido representacional

y por consiguiente, el representacionalismo es erróneo. El segundo argumento es el problema de los estados de orden superior sin un objetivo al cual dirigirse (*i.e.*, un estado mental de primer orden). Supongamos que un sujeto S tiene una representación de orden superior (HOR) de encontrarse en un estado de dolor *d*, pero sin tener *d*. Según las teorías HOR, *d* no es consciente porque no existe. Pero, como S tiene una HOR acerca de *d* entonces tendrá una experiencia fenoménica de *d*. No obstante este resultado es incoherente. Para la teoría de la autoconciencia prerreflexiva, *d* no es consciente si no le pertenece a S, pero como S *sí* tiene conciencia de *d* (siendo *d* otro tipo de estado diferente al dolor, *e.g.*, una creencia, una imaginación mental), entonces este sentido de pertenencia es lo que hace que *d* sea una experiencia fenoménica.

Palabras clave: conciencia fenoménica, autoconciencia prerreflexiva, sentido de pertenencia, representacionalismo, representación de orden superior.

Introducción

La conciencia es una propiedad de algunos estados y sucesos mentales que implica la instanciación de *propiedades fenoménicas*. Estas propiedades caracterizan cómo es ser un sujeto (*e.g.*, cómo es ser yo mismo ahora) o cómo es estar en un estado mental (*e.g.*, cómo es escuchar una melodía). La experiencia de escuchar una melodía es entonces un estado fenoménicamente consciente. Cuando haya algo que es como ser un sujeto, ese sujeto tendrá propiedades fenoménicas específicas. Cuando haya algo que es como encontrarse en un estado mental, ese estado tendrá propiedades fenoménicas específicas. Se suele decir también que cómo es tener una experiencia consciente hace referencia a su *carácter fenoménico*. Thomas Nagel así definía a la conciencia: “fundamentalmente un organismo tiene estados mentales conscientes si y sólo si hay algo que es cómo *ser* ese organismo, algo que es cómo ser *para* ese organismo. [...] Podemos llamar a esto el carácter subjetivo de la experiencia” (1974: 436).

Conciencia como adjetivo, es decir “consciente”, se aplica a los sistemas cognitivos y a sus estados. En este sentido, David Rosenthal (1993, 1997) ha hecho una útil distinción entre:

- a) *Conciencia de criatura*. Los humanos y otras especies animales son criaturas conscientes, mientras que las plantas y las piedras no lo son. Así, ser consciente en este sentido es lo opuesto a estar en un sueño profundo o en coma. Describimos a una persona o a un animal como consciente si está despierta, alerta y se encuentra con capacidad de sentir y responder a su entorno.
- b) *Conciencia estado*. Generalmente los estados y sucesos sensorio-perceptivos, como los dolores y las experiencias visuales y auditivas, son estados y sucesos conscientes, pero también pueden no serlo como en la percepción subliminal, en la visión periférica o en la visión ciega. Los estados afectivos como la alegría, la tristeza, el enojo, suelen ser también conscientes. Otros estados y sucesos mentales como pensamientos, recuerdos, creencias y deseos, también pueden ser conscientes, aunque muchas veces estos estados y sucesos mentales ocurren sin que nos percatemos de ellos.
- c) *Conciencia intransitiva*. Decir que un sistema cognoscitivo es consciente *simpliciter*, significa decir que carece de objeto al cual dirigirse. Entonces la expresión “*x* es consciente” se refiere al uso intransitivo de conciencia.
- d) *Conciencia transitiva*. Cuando un sistema cognitivo siente algo o piensa en alguien, decimos que éste está consciente *de* algo o alguien. Entonces la expresión “*x* es consciente *de y*” se refiere al uso transitivo de conciencia.

Podemos dividir las teorías de la conciencia en teorías reduccionistas y no reduccionistas. Las teorías reduccionistas intentan dar cuenta de la conciencia en términos físicos (neuronales, computacio-

nales) o en términos mentales (representacionales, cuasi percepción, pensamiento de orden superior) que no presuponen a la conciencia. La mayoría de los teóricos abogan por un enfoque reduccionista; sólo algunos defienden un enfoque no reduccionista. Al afirmar que la conciencia no se puede explicar en términos físicos, tratan a la conciencia como un rasgo fundamental del mundo, vinculada a los más elementales niveles estructurales del universo físico.

Los otros enfoques reduccionistas que intentan dar cuenta de la conciencia con otros conceptos mentales sin implicar términos fenoménicos son las teorías representacionales (el representacionalismo y las teorías de representación de orden superior). El representacionalismo sostiene que la conciencia es reducible a un tipo de representación, a saber: la naturaleza de los estados y sucesos conscientes es agotada por la manera en la cual estos representan el mundo, *i.e.*, por su contenido representacional. Las teorías de representación de orden superior por su parte, explican los estados y sucesos conscientes en términos de una representación de orden superior que puede tomar la forma de una cuasi percepción o un pensamiento, que va dirigido a un estado de primer orden.

Desde hace unos diez años, un grupo de teóricos, principalmente filósofos, han estado desarrollando un nuevo enfoque —fundamentalmente por su descontento hacia el representacionalismo y las teorías de orden superior— cuya tesis básica es que todo estado o evento consciente se puede explicar en términos de un tipo de autoconciencia; que toda experiencia consciente lleva implícita un tipo de autoconciencia primitiva, prerreflexiva.

En este artículo voy a hacer una crítica a las teorías representacionales de la conciencia, desde este nuevo enfoque que ahora se conoce como *enfoque autorrepresentacional de la conciencia* (Kriegel y Williford, 2006). Uno de los defensores de este enfoque (Zahavi, 2004: 74) opina que debería mejor hablarse en términos de “auto-presencia”, “auto-presencia” o “auto-manifestación”, en lugar de “auto-

representación”, porque el conocimiento que tenemos de nuestras propias experiencias parece tener una inmediatez presentacional, que no queda captada por la noción de *re*-presentación. De hecho, dice Zahavi, en la mayoría de los casos mis experiencias se presentan ante mí, en lugar de re-presentarse. Voy a referirme entonces a este enfoque como *teoría de la autoconciencia prerreflexiva de la conciencia*.

I. Teorías representacionales de la conciencia

1. Representacionalismo fuerte

En la filosofía occidental ha existido una añeja tradición que considera que los contenidos de la mente se dividen en dos categorías generales: 1) los estados mentales intencionales y 2) los estados mentales fenoménicos no intencionales. La intencionalidad es la propiedad de los estados mentales de ser acerca de algo o de estar dirigidos a algo. En algún sentido podríamos identificar los estados intencionales con las representaciones mentales dado que, por ejemplo, el sentido en el cual mi deseo de que mañana será un día soleado es *acerca* del día, parece estar muy relacionado con el sentido de que mi deseo *representa* el día soleado (Seager, 1999: 132).¹ En contraste, los estados no

¹ Uno de los dictaminadores de este artículo me señaló que no era del todo correcto identificar la intencionalidad de los estados mentales con las representaciones. La noción de “representación mental” es, posiblemente y en primera instancia, un constructo teórico de la ciencia cognitiva (Pitt, 2008: §1) y, como tal, es un concepto básico de la teoría computacional de la mente, de acuerdo con la cual los estados y sucesos cognoscitivos están constituidos por la ocurrencia, transformación y almacenamiento de estructuras informativas (representaciones) de uno u otro tipo. Estas representaciones son estructuras que poseen propiedades semánticas (*e.g.*, contenido, referencia, condiciones de verdad, valor de verdad, etc.) y, en tanto

intencionales o no representacionales, no llevan un contenido, no representan, por lo cual son puramente fenoménicos, *i.e.*, son *qualia*.

El representacionalismo prescinde de esta extendida convicción filosófica de que los contenidos de la mente se dividen en las dos habituales categorías básicas, estados intencionales y estados fenoménicos no intencionales, y va a proponer, en cambio, que la última categoría se incluye de alguna manera, es una especie de o es agotada por la primera categoría. En otras palabras, el representacionalismo intenta romper con la tradición y considera que todos y cada uno de los estados mentales conscientes, siempre representan algo, es decir, siempre son estados representacionales.

Dado que no basta con identificar la conciencia con la representación, ya que puede haber representaciones que no son conscientes (*e.g.*, las imágenes en una pantalla de computadora, un mapa, etc.), lo que se requiere es una teoría representacional que establezca *qué* tipo de representaciones son los estados mentales conscientes y *cómo* y *por qué* este tipo de representaciones se hacen conscientes.

Hay muchos tipos de representacionalismo. Una división que se ha establecido es entre el representacionalismo débil y el representacionalismo fuerte. De acuerdo con el representacionalismo débil,

tales, se pueden entender no sólo en términos computacionales. Así, las representaciones mentales han sido un tema filosófico muy antiguo que se puede rastrear al menos hasta Aristóteles. Una teoría representacional de la mente toma como punto de partida los estados mentales de sentido común (sensaciones, percepciones, pensamientos, creencias, deseos, etc.) y se dice que estos estados son estados intencionales. La teoría representacional de la mente define entonces los estados mentales intencionales como relaciones con representaciones mentales y explica la intencionalidad de aquellos en términos de las propiedades semánticas de las representaciones mentales. Por ejemplo, la creencia de que Superman se debilita con la kriptonita, está apropiadamente relacionada a una representación mental cuyo contenido es la siguiente situación descrita por la oración: "Superman se debilita con la kriptonita".

todos los estados mentales son representacionales, pero que algunos de estos estados mentales tienen ciertas propiedades no intencionales, a saber, *qualia*. Otra forma de expresar el representacionalismo débil consiste en decir que la naturaleza intencional de ciertos estados mentales conscientes no agota su carácter consciente o fenoménico (Crane, 2003: 45). Otra manera de enunciar esta tesis es que no toda diferencia fenoménica o consciente en los estados mentales es una diferencia intencional o representacional. En general este enfoque se defiende describiendo maneras en las cuales el carácter fenoménico de una experiencia puede cambiar pero sin que haya cambio en el contenido representacional o casos en los cuales el carácter fenoménico sigue siendo el mismo pero el contenido representacional cambia.

A diferencia del representacionalismo débil, el representacionalismo fuerte (a partir de aquí RF) hace una afirmación adicional: el carácter cualitativo de nuestros estados mentales fenoménicos *consiste en* el contenido representacional de tales estados. Muchos teóricos representacionalistas defienden el RF (Dretske, 1995, 2003; Tye, 1995, 2000, 2009) y dado que éste *reduce* el carácter fenoménico al contenido intencional, también se le ha llamado representacionalismo reduccionista (Chalmers, 2004: 161-5). El RF también se describe como una forma de externismo sobre los *qualia*, ya que sostiene que los *qualia son* las propiedades representadas que tienen los objetos externos. Esta posición teórica que ubica los *qualia* en el mundo externo, permitiría rechazar la idea de estos con cualidades privadas de la experiencia interna, haciendo posible su naturalización desde un fisicismo sobre la conciencia.

El principal argumento a favor del RF proviene de una observación fenomenológica conocida como *transparencia de la experiencia*. Suponga que usted está viendo una manzana roja. Usted es consciente de la manzana. Ahora se le pide que ponga atención no a la manzana, sino a su propia *experiencia* de la manzana. ¿Qué sucede? ¿De qué se hace consciente ahora? Lo que ocurre, *ex hypothesi*, es que usted sigue

viendo la manzana roja, sigue siendo consciente de la manzana y no de otra cosa intermedia entre la manzana y su experiencia.

Este argumento ha sido articulado de manera influyente por Gilbert Harman, aunque ya había sido discutido por G. E. Moore. De acuerdo con Harman, al percibir un objeto no somos conscientes de las propiedades intrínsecas de nuestra experiencia (*i.e.*, la representación o el vehículo representacional del objeto) sino de las propiedades del objeto que experimentamos (*i.e.*, el objeto representado). De acuerdo con este argumento, vemos los objetos “a través” de nuestros estados perceptuales e incluso no nos damos cuenta de que nos encontramos *en* un estado perceptual. Dice Harman:

Cuando usted ve un árbol, no experimenta rasgos como rasgos intrínsecos de su experiencia. Mire un árbol e intente situar su atención a los rasgos intrínsecos de su experiencia visual. Predigo que se va a encontrar con que los únicos rasgos en los que sitúa su atención serán los rasgos del árbol presentado, incluyendo los rasgos relacionales del árbol “desde aquí” (1990: 39).

¿Por qué entonces la transparencia de la experiencia se utiliza como argumento a favor del RF? Volviendo al ejemplo de la manzana roja, cuando usted pone atención a su experiencia, lo único que su introspección le revela, *prima facie*, son los rasgos del cielo azul *representados* por su experiencia. En otros términos, las únicas propiedades de nuestra experiencia consciente a las que tenemos acceso introspectible son sus propiedades representacionales. ¿Por qué? La respuesta, dice Tye:

[...] es que sus experiencias perceptuales no tienen rasgos *introspectibles* sobre y arriba de aquellos rasgos implicados en sus contenidos intencionales. Entonces el carácter fenoménico de tales experiencias [...] es idéntico a, o está contenido dentro de, sus contenidos intencionales (1995: 136).

2. Teorías de representación de orden superior de la conciencia

Hay una clase de teorías representacionales que se ha desarrollado independientemente del representacionalismo fuerte que revisamos anteriormente. De acuerdo con esta clase de teorías, un estado mental es un estado mental consciente si es el objeto de un estado representacional de segundo orden u orden superior, por lo que podemos llamar a esta clase de teorías, teorías de representación de orden superior de la conciencia o, en breve, teorías HOR (por *higher-order representation*). Las teorías HOR se dividen a su vez en: 1) *teorías de experiencia de orden superior* o teorías HOE (por *higher-order experience*), y 2) *teorías de pensamiento de orden superior* o teorías HOT (por *higher-order thought*).

De acuerdo con las teorías HOE (Armstrong, 1981; Lycan, 1996), un estado mental es consciente si es el objeto de una metarrepresentación que puede consistir en una cuasi-percepción, un sentido interno o una monitorización, que se dirige hacia los estados mentales de primer orden u orden inferior sobre algún estado de cosas ϵ , por lo cual a la teoría se le conoce también en ocasiones como *teoría de percepción de orden superior* o *teoría del sentido interno*. Esta representación de segundo orden opera de un modo similar a como lo hacen nuestros sistemas perceptivos externos, pero “está dirigida internamente y no externamente” (Churchland, 1988: 74) y es una metarrepresentación similar a un “sentido”, análogo a un “ojo de la mente”, un mecanismo de “escáner interno” (Lycan, 1995: 250) que monitoriza los estados mentales de primer orden acerca de ϵ .

Por su parte, las teorías HOT establecen que la metarrepresentación es una abstracción más que una reflexión directa y toma la forma de un *pensamiento de segundo orden* u *orden superior* acerca de los estados mentales de primer orden u orden inferior acerca de ϵ .

La teoría HOT ha sido desarrollada por David Rosenthal (1997, 2002) y ha sido defendida por otros autores como Peter Carruthers (2000) y Rocco Gennaro (1996). Un *pensamiento* es “cualquier estado intencional episódico con una actitud mental asertiva” (Rosenthal, 1997: 751, n. 43; 2002: 410). Un pensamiento es entonces, en el sentido de Rosenthal, lo que los filósofos de la mente llaman por lo general una “creencia ocurrente”. Un pensamiento de orden superior o HOT es un pensamiento acerca de algún estado mental. Estar *consciente de algo*, *transitivamente* consciente de algo o percatarse de algo, es estar “en un estado mental cuyo contenido pertenece a ese algo” (Rosenthal, 1997: 737), por ejemplo, cuando uno tiene un pensamiento acerca de algo o una percepción de algo. Los *estados sensoriales* son estados como experiencias visuales o experiencias de dolor, que involucran cualidades sensoriales, cuando estos estados son fenoménicamente conscientes, es decir, tienen cualia. Los estados conscientes son siempre *intransitivamente conscientes*.

Para resolver el problema de qué es lo que hace que un estado mental sea consciente, Rosenthal propone que los estados mentales son conscientes cuando están acompañados por un estado de orden superior, a saber, un pensamiento. He aquí la propuesta de Rosenthal:

Según este modelo, somos conscientes de algo cuando tenemos un pensamiento sobre ese algo. Entonces un estado mental será consciente si va acompañado por un pensamiento sobre ese estado. El acaecimiento de tal pensamiento de orden superior (HOT) nos hace conscientes del estado mental; de esta manera el estado del que somos conscientes es un estado consciente. De forma similar, cuando no ocurren tales HOT, no nos percatamos de encontrarnos en el estado mental en cuestión y entonces el estado no es un estado consciente. Entonces, el núcleo de la teoría es que un estado mental es un estado consciente cuando y sólo cuando, va acompañado por un HOT apropiado (1997: 741).

II. Autoconciencia prerreflexiva *versus* las teorías representacionales de la conciencia

1. Autoconciencia prerreflexiva y sentido de pertenencia

Aunque hay diferentes versiones de la teoría de la autoconciencia prerreflexiva de la conciencia (Kriegel y Williford, 2006), la idea principal, puesta del modo más directo, es que toda experiencia consciente lleva implícita una forma de autoconciencia prerreflexiva. Llamaré por simplicidad a la tesis general de este enfoque, tesis de la autoconciencia prerreflexiva implícita *simpliciter* o simplemente APR.

El enfoque APR es un intento de dar cuenta del carácter fenoménico o subjetivo de los estados mentales, partiendo principalmente de bases fenomenológicas. Por ejemplo, los teóricos APR nos invitan a tomar como punto de partida cualquier experiencia consciente, digamos una experiencia perceptual. El objeto de mi experiencia perceptual es intersubjetivamente accesible, es decir, es público. En cambio, mi experiencia perceptual *per se*, sólo me es accesible a mí y no puede ser compartida por otra persona. Esto es, aunque dos personas pueden compartir un mismo tipo de estado mental, vía algún proceso empático o lo que sea, en principio ninguna persona puede tener o experimentar *mis* estados mentales.

Entonces, de acuerdo con los teóricos APR, la experiencia consciente se caracteriza por tener un aspecto de primera persona, un componente, rasgo o cualidad de ser característicamente mía o para mí² y que este aspecto es el que hace subjetiva o fenoménica la experiencia (Gallagher y Zahavi, 2008: 50-1). Este aspecto es lo que los neurobiólogos y científicos cognoscitivos llaman el *sentido de perte-*

² Los teóricos, la mayoría anglosajones, usan los términos ingleses *meness*, *myness* o *for-me-ness*, para referirse a este aspecto de que la experiencia es para mí o mía.

nencia (véase *e.g.*, Gallagher, 2007: 2; Gray, 2004: 264-6; Haikonen, 2007: 191; Martin, 1995; Roessler y Eilan, 2003; Synofzik, Vosgerau y Newen, 2008) el cual está íntimamente relacionado con el *sentido de agencia*.³

Cuando tengo una experiencia consciente (perceptual, emocional, etc.) desde la perspectiva de la primera persona, “la experiencia en cuestión se da inmediata y no inferencialmente como *mía*. No escruto *primero* una percepción o sensación de dolor y *entonces* la identifico como *mía*” (Thompson y Zahavi, 2007: 76). Dice Kriegel, uno de los defensores del enfoque APR:

Este carácter subjetivo o rasgo de ser para mí [*for-me-ness*], es ciertamente un fenómeno elusivo, pero está presente en cualquier experiencia consciente. De hecho, su presencia parece ser una condición para cualquier fenomenalidad: resulta difícil darle sentido a la idea de una experiencia consciente que no tenga este *for-me-ness* para ella. Si no tuviera este *for-me-ness*, sería solamente un estado subpersonal, un estado que tiene lugar *en* mí pero que no es *para* mí en el sentido pertinente. Tal estado subpersonal no parece contar como una experiencia consciente (2007: 46).

Ahora, si este sentido de pertenencia es fundamental para que una experiencia sea una experiencia consciente, los teóricos APR dan un paso más y van a sugerir entonces que la autoconciencia es pertinente para tener un mejor entendimiento de la conciencia fenoménica. Por ejemplo, de acuerdo con Zahavi (2005):

³ En la clínica neuropsiquiátrica se han descrito dos fenómenos que ejemplifican claramente la diferencia entre estos dos sentidos (agencia y pertenencia): el síndrome de *mano alienada* y el síndrome de *mano anárquica*. El fenómeno de la mano alienada, que por lo general afecta la mano izquierda, se caracteriza por la experiencia de que la mano no pertenece al paciente. En el fenómeno de la mano anárquica, por otra parte, la mano afectada (contralateral a la lesión), ejecuta movimientos aparentemente propositivos y dirigidos hacia una meta, pero que el paciente afirma no haber iniciado o controlado por voluntad propia.

[...] la experiencia se da (al menos tácitamente) como *mi* experiencia, como experiencia por la que *yo* atravieso o tengo vivencia. Todo esto sugiere que la experiencia de primera persona me presenta un acceso inmediato a mí mismo y que la conciencia (fenoménica) implica una forma (mínima) de autoconciencia. Puesto de otra forma, la autoconciencia se puede ver como una condición necesaria para la conciencia fenoménica. A menos que un proceso mental sea autoconsciente no habrá nada que sea como pasar por este proceso y por lo tanto no puede ser un proceso fenoménicamente consciente (16).

2. Autoconciencia prerreflexiva versus el representacionalismo fuerte

Siguiendo a Janzen (2006b: 332-335) y Legrand (2007: 585), podemos usar la tesis APR *simpliciter* como contraejemplo para el representacionalismo fuerte (RF). El blanco del ataque es el argumento de la transparencia. Primero, basándonos en Janzen (2006b) y en Kind (2003: 230, 2007: 418) podemos distinguir entre *Transparencia fuerte* y *Transparencia débil*.

Transparencia fuerte (TF): Es imposible discernir, vía introspección, las propiedades intrínsecas de una experiencia de x que no se experimenten como propiedades de x .

Transparencia débil (TD): Es difícil, pero no imposible, discernir, vía introspección, las propiedades intrínsecas de una experiencia de x que no se experimenten como propiedades de x .

Esta distinción es importante por dos razones. La primera es que la TF, pero no la TD, motiva el RF defendido por Dretske y Tye, entre otros. Recordemos que para el RF el carácter fenoménico es igual al contenido representacional. Pero si nuestras experiencias muestran rasgos cualitativos que no se experimentan como idénticos a los rasgos de los objetos que experimentamos (*i.e.*, si la TF es falsa) entonces el carácter fenoménico no puede ser idéntico al contenido represen-

tacional, ya que este último es, digamos, un producto causal directo de los rasgos de los objetos que experimentamos, en el sentido de que la presencia de estos rasgos “activa” los procesos sensoriales internos que generan el contenido representacional.

La segunda razón es que los defensores de la TF presentan el argumento de la transparencia en contraposición a la teoría de los datos sensoriales (*e.g.*, Harman, 1990: 39; Tye, 2000: 45-46). La teoría de los datos sensoriales es una teoría *mediacional* de la percepción de acuerdo con la cual los datos sensoriales median nuestra percepción de los objetos del mundo externo. Sin embargo, los defensores del RF sostienen que la introspección *no* nos revela ninguna de estas supuestas entidades mentales mediacionales; de hecho, la introspección, nos dicen los teóricos RF, no revela rasgos distintivos de la experiencia por encima de los rasgos de los objetos que experimentamos. A partir de esto los teóricos representacionistas concluyen que la teoría de los datos sensoriales es falsa.

Ahora, el teórico representacionista debe evitar apelar a la TD porque ésta minaría el RF. Si la experiencia de un objeto externo fuera sólo débilmente transparente, entonces podríamos (al menos en principio) evitar ver a través de dicha experiencia, y al hacerlo seríamos introspectivamente conscientes de algunas de sus propiedades intrínsecas y que van más allá de su contenido representacional. Esto significaría que en el carácter fenoménico de la experiencia forma parte las propiedades intrínsecas de la experiencia misma, en lugar de sólo aquello que es representado. De ser así, fallaría la reducción del carácter fenoménico de la experiencia al contenido intencional y por lo tanto el representacionismo fuerte sería falso. Entonces, el representacionista debe apelar a la TF para defender su teoría representacional. Pero aun la TF no es vulnerable a ciertos contraejemplos.

Recordemos que la TF dice que es imposible discernir, vía introspección, las propiedades intrínsecas de una experiencia de x que no se experimenten como propiedades de x . Para que el contraejemplo

sea exitoso debe ser, entonces, 1) intrínseco y 2) discernible por medio de la introspección. APR *simpliciter* reúne justamente estos dos criterios.

En primer lugar, que toda experiencia consciente implica alguna forma de autoconciencia es un hecho introspeccionable. Janzen (2006b: 326) nos invita a llevar a cabo el siguiente ejercicio: mientras usted lee estas palabras cambie su atención de las letras impresas a su acto mismo de percibir las; ahora tiene el siguiente reto: negar que usted no puede experimentar el acto de ver como suyo, es decir, ¿podría usted negar que el acto de percibir no le “pertenece”?

Se podría objetar diciendo que al estar implícitamente consciente de uno mismo viendo, oyendo o lo que sea, en realidad uno no está consciente de uno mismo viendo, sino que más bien uno se percata *inconscientemente* de uno mismo viendo. Sin embargo, esta imagen de la conciencia es inaceptable bajo fundamentos fenomenológicos, a saber, que toda experiencia consciente implica un sentido de pertenencia o *for-me-ness*. Mi experiencia de esta página, aunque no esté pensando en ella, va acompañada por tal sentido de pertenencia, de ser mía, *i.e.*, en ella va implícita alguna forma de autoconciencia y ésta no es una percatación inconsciente, sino algo que *yo* experimento.

Cómo es estar enamorado de mi novia es algo muy diferente de cómo es oler la tierra mojada cuando está por llover, que a su vez difiere de cómo es ver el cielo azul o de imaginarme acostado en la playa. Sin embargo a pesar de ser tan diferentes estas experiencias fenoménicas, el punto es que se experimentan por un sujeto. Estas experiencias *me* presentan a *mí* diferentes objetos intencionales. No sólo estoy fenoménicamente en contacto con las propiedades de estos diversos objetos intencionales, si el lector me permite decirlo de este modo, sino que estos objetos y sus propiedades están ahí para mí y las experiencias son de hecho, mías. Y este sentido de pertenencia, este *for-me-ness*, permanece constante, cualesquiera que sean los objetos

intencionales, por eso no tendría sentido *reducir* este sentido a los rasgos cualitativos de los objetos que uno experimenta.

Todo lo anterior se puede presentar sucintamente de la siguiente manera:

- 1) Es imposible discernir, vía introspección, las propiedades intrínsecas de una experiencia de x que no se experimenten como propiedades de x (ésta es la transparencia fuerte);
- 2) nuestras experiencias muestran rasgos cualitativos que son idénticos a los rasgos de los objetos que nuestra experiencia representa; por lo tanto,
- 3) el carácter fenoménico de la experiencia *es* idéntico a su contenido representacional (ésta es la tesis del representacionalismo fuerte); sin embargo,
- 4) la experiencia consciente tiene un sentido intrínseco de pertenencia, un componente, cualidad o rasgo de ser característicamente mía o para mí; este aspecto es el que hace que la experiencia sea fenoménica o que tenga una cualidad de “cómo es” (*what it's like*) estar en ella;
- 5) este sentido de pertenencia o *meness* es a) un rasgo intrínseco de la experiencia y b) discernible por medio de la introspección; por lo tanto,
- 6) (1)-(3) son falsos: el carácter fenoménico de la experiencia *no* es idéntico a su contenido representacional y por lo tanto la reducción del carácter fenoménico a su contenido representacional falla.

3. Autoconciencia prerreflexiva versus las teorías de representación de orden superior

Las teorías HOR de la conciencia han recibido diferentes críticas (*e.g.*, Aquila, 1990; Byrne, 1997; Drestke, 1995: cap. 4, esp. 103-116;

Seager, 1999: cap. 3). Una de las principales objeciones a las teorías HOR y en particular a la teoría HOT (Byrne, 1997: 119-124; Caston, 2002: 780-781; Kriegel, 2007: 49; Janzen, 2006a: 49-53; Levine, 2001: 108-109; Neander, 1998: 418-421; Rowlands, 2001: 92-93; Seager, 1999: 72-84), es el problema de los estados de orden superior sin un objetivo al cual dirigirse. De acuerdo con la teoría HOT de Rosenthal, una condición de adecuación para una explicación de la naturaleza de la conciencia, es que la teoría proporciona una explicación plausible de la naturaleza de la conciencia intransitiva estado y que la mejor manera de satisfacer este desiderátum es en términos de la conciencia transitiva de criatura (*e.g.*, Rosenthal, 1997: 737). Ahora, explicar no circularmente la conciencia intransitiva estado en términos de la conciencia transitiva de criatura, implica estipular que la conciencia transitiva y la conciencia intransitiva son dos estados *differentes*. Empero, las teorías HOR no salen libradas al intentar explicar esta diferencia. Si el estado transitivo Mt y el estado intransitivo Mi fuesen estados distintos y si, como Rosenthal ha argumentado (1997: 744) el objetivo —*i.e.*, el estado mental de primer orden u orden inferior— al cual va dirigido Mt no necesita implicarse al causar Mt , entonces Mt puede ocurrir sin Mi y viceversa. Si Mi ocurre sin Mt entonces Mi es un estado mental no consciente, pero si Mt ocurre sin Mi , entonces el sujeto de Mi tiene lo que Rosenthal llama un estado consciente “confabulado” (2002: 415-417). Los estados conscientes confabulados son, básicamente, HOT sin un objetivo; ocurren cuando el sujeto tiene un HOT en el sentido de encontrarse en un estado mental en particular, pero sin que aquel estado exista. Sin embargo esto conduce a un dilema. Por un lado, la teoría HOT no puede decir que el estado consciente confabulado es un genuino estado consciente, ya que el requisito de que debe existir un estado de orden inferior no se cumple porque este último no existe. Y por otro lado, la teoría se compromete al decir que *parece* que el sujeto S tiene un estado consciente, ya que el solo hecho de tener un HOT en el sentido de en-

contrarse en un estado mental sin que este último exista, es suficiente para que S crea que se encuentra en ese estado mental. De hecho, para S no hay ninguna diferencia fenoménica entre la experiencia verídica y la confabulada.

Supongamos que un sujeto S, en el tiempo t , tiene un HOT de encontrarse con un fuerte dolor de muelas d , pero que S no tiene d . Según la teoría HOT (lo mismo se aplica para la HOE o de sentido interno) d no es consciente, simple y sencillamente porque d no existe. Pero, como S tiene un HOT acerca de d , entonces S tendrá una experiencia fenoménica de d . Es más, *ex post facto*, S buscará aliviar su dolor ¡sin tener ningún dolor! Desde mi punto de vista, éste es un resultado contraintuitivo para las teorías HOR porque es a todas luces una incoherencia.

Por supuesto, un teórico HOR tendría algunas respuestas a esta objeción. Una respuesta sería decir que el estado metarrepresentacional M^* no tiene objetivo y que la propiedad de ser consciente, aunque no sea instanciada por el estado primario M, es, no obstante, instanciada por M^* . Sin embargo, según la teorías HOR, M^* es un estado *no* consciente.⁴ Decir que M^* instancia la propiedad de ser

⁴ Con esta premisa, derivada de la teoría HOT, algunos autores (*e.g.*, Drummond, 2006: 202-203; Janzen, 2006a: 51-53; Kriegel, 2007: 49-50; Textor, 2006: 411-412; Zahavi, 2005: 25-26) han generado otra objeción a las teorías HOR. El argumento es que la tesis de que un estado mental es un estado mental consciente si es el objeto de un estado representacional de orden superior, conduce a una regresión al infinito: si un estado mental M es consciente porque es el objeto de un estado metarrepresentacional M^* , entonces para que este segundo estado M^* sea consciente, se requiere que sea el objeto de otro estado mental metarrepresentacional M^{**} y así *ad infinitum*. La réplica de los teóricos HOR, sin embargo, es que el argumento anterior comete petición de principio y que la salida al argumento de la regresión al infinito se detiene fácilmente si se acepta la existencia de estados mentales no conscientes. Esta es por ejemplo la respuesta que dan los teóricos HOT para quienes el pensamiento de segundo orden no tiene por qué ser consciente.

consciente sería decir que, *ceteris paribus*, M^* es consciente y no consciente, lo cual sería incoherente.

Otra respuesta consistiría en afirmar que no hay estados meta-representacionales sin un objetivo. Pero aun cuando esto se podría probar (lo cual resulta muy difícil de imaginar cómo sería el caso), fácilmente podríamos concebir una situación contrafáctica en la cual *sí* ocurren estados meta-representacionales sin un objetivo.

Al no postular estados meta-representacionales, este tipo de problemas no se presentan para la teoría de la autoconciencia prerreflexiva. Volvamos al dolor de muelas de S. S cree tener d en t y además tiene la experiencia fenoménica de encontrarse en d , pero no tiene d . Para el teórico APR d es otro tipo de estado mental diferente de los dolores, por ejemplo, una creencia o un estado de imaginación mental, como un estado de simulación mental y S es consciente de ese estado porque le pertenece y este sentido de pertenencia es lo que hace que d sea una experiencia fenoménica.

Ahora, Rosenthal arguye que para que un estado mental sea consciente, no es suficiente que seamos no inferencialmente conscientes de ese estado, sino que también debemos ser conscientes de ser nosotros mismos de estar en ese estado mental. “Sólo si nuestro pensamiento es acerca de uno mismo como tal y no sólo acerca de alguien que resulta ser uno mismo, entonces el estado mental será consciente” (1997: 750). Para ponerlo en otros términos, no es suficiente explicar cómo un estado mental se hace consciente, sino también explicar cómo ese estado es *mi* estado, un estado en el que *yo* me encuentro. ¿Por qué? Porque este sentido de ser para mí o mío, es un elemento crucial para que un estado sea consciente. De no tener este sentido de pertenencia entonces el estado sería un estado mental no consciente. No obstante al postular la existencia de estados mentales

Dice Rosenthal: “la introspección ocurre cuando hay un pensamiento de tercer orden que hace consciente al pensamiento de segundo orden” (1997: 745). Pero este pensamiento de tercer orden no es consciente.

metarrepresentacionales *no* conscientes para evitar una regresión al infinito (véase la nota 4), las teorías HOR fallan al dar cuenta de este hecho fenomenológico.

Conclusiones

He presentado la teoría de la autoconciencia prerreflexiva de la conciencia como un avance con respecto a las teorías representacionales de la conciencia, estableciéndola como contraposición para estas teorías. Ahora, la mayor parte de los defensores de la teoría APR parten de un vocabulario puramente mentalista tomando como base las ideas de la tradición fenomenológica en filosofía. Sin embargo, esto no impide que la tesis de la autoconciencia prerreflexiva pueda ser reducida en términos físicos. Recientemente algunos teóricos han estado desarrollando el enfoque APR desde la perspectiva de la ciencia cognoscitiva situada (*embodied*) por lo que proponen que la autoconciencia prerreflexiva es en realidad una forma de autoconciencia corporal (Bermúdez *et al.*, 1995; Gallagher, 2005; Gallagher y Zahavi, 2008; Legrand, 2005, 2006, 2007; Thompson, 2007; Wider, 1997). Algunos investigadores en neurociencia cognoscitiva también han formulado algunas hipótesis para establecer cuáles pueden ser las bases cerebrales de la autoconciencia prerreflexiva (Baars *et al.*, 2003; Newen y Vogeley, 2003; Seeley y Sturm, 2007).

Aunque el enfoque APR se ha venido desarrollando desde hace poco, lo importante es que este enfoque ha combinado las ideas de dos tradiciones en filosofía que, a primera vista, parecen inconmensurables, la tradición fenomenológica y la tradición analítica, para dar cuenta de un fenómeno único: la conciencia.

Agradecimientos

Quiero agradecer a la doctora María Christiansen Renaud (Departamento de Filosofía, Universidad de Guanajuato) por todas sus aten-

ciones y por su invitación a participar en la revista *Valenciana*. También un agradecimiento a los dos dictaminadores anónimos, quienes revisaron una versión previa de este trabajo. Un reconocimiento especial a la señorita Vannesa Erica Tizatl Martínez por todo su aliento y apoyo.

Bibliografía

- Aquila, Richard, 1990, "Consciousness and higher-order thoughts: Two objections", en *American Philosophical Quarterly*, núm. 27, pp. 81-87.
- Armstrong, David Malet, 1981, *The Nature of Mind*, The Harvester Press, Sussex.
- Baars, Bernard, Thomas Z. Ramsøy y Steven Laureys, 2003, "Brain, conscious experience and the observing self", en *Trends in Neurosciences*, núm 26, pp. 671-675.
- Bermúdez, José Luis, Anthony Marcel y Naomi Eilan (eds.), 1995, *The Body and the Self*, MIT, Cambridge.
- Byrne, Alex, 1997, "Some like it HOT: Consciousness and higher-order thoughts", en *Philosophical Studies*, núm. 86, pp. 103-129.
- Carruthers, Peter, 2000, *Phenomenal Consciousness: A Naturalistic Theory*, Cambridge University Press, Nueva York.
- Caston, Victor, 2002, "Aristotle on consciousness", en *Mind*, núm. 111, pp. 751-815.
- Chalmers, David John, 2004, "The representational character of experience", en Brian Leiter (ed.), *The Future for Philosophy*, Oxford University Press, Oxford, pp. 153-181.
- Churchland, Paul M., 1988, *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind*, MIT, Cambridge.
- Crane, Tim, 2003, "The intentional structure of consciousness", en Smith, Quentin y Aleksandar Jokic (eds.), *Consciousness: New Philosophical Perspectives*, Oxford University Press, Oxford, pp. 33-56.
- Dretske, Fred, 1995, *Naturalizing the Mind*, MIT, Cambridge.
- _____, "Experience as representation", en *Philosophical Issues*, núm. 13, pp. 67-82.
- Drummond, John, 2006, "The case(s) of (self-)awareness", en Kriegel, Uriah y Kenneth Williford (eds.), *Self-Representational Approaches to Consciousness*, MIT, Cambridge, pp. 199-220.

- Gallagher, Shaun, 2005, *How the Body Shapes the Mind*, Oxford University Press, Oxford.
- _____, 2007, “The natural philosophy of agency”, en *Philosophy Compass*, núm. 2, pp. 1-11.
- _____ y Dan Zahavi, 2008, *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*, Routledge, Londres.
- Gennaro, Rocco J., 1996, *Consciousness and Self-Consciousness: A Defense of the Higher-Order Thought Theory of Consciousness*, John Benjamins, Philadelphia.
- Gray, Jeffrey, 2004, *Consciousness: Creeping Up On the Hard Problem*, Oxford University Press, Oxford.
- Haikonen, Pentti O., 2007, *Robot Brains: Circuits and Systems for Conscious Machines*, John Wiley & Sons, West Sussex.
- Harman, Gilbert, 1990, “The intrinsic quality of experience”, en James E. Tomberlin (ed.), *Philosophical Perspectives, 4: Action Theory and Philosophy of Mind*, Ridgeview Publishing, Atascadero, pp. 31-52.
- Janzen, Greg, 2006a, “Phenomenal character as implicit self-awareness”, en *Journal of Consciousness Studies*, núm. 13, pp. 44-73.
- _____, 2006b, “The representational theory of phenomenal character: a phenomenological critique”, en *Phenomenology and the Cognitive Sciences*, núm. 5, pp. 321-339.
- Kind, Amy, 2003, “What’s so transparent about transparency?”, en *Philosophical Studies*, núm. 115, pp. 225-344.
- _____, 2007, “Restrictions on representationalism”, en *Philosophical Studies*, núm. 134, pp. 405-427.
- Kriegel, Uriah, 2007, “Philosophical theories of consciousness: contemporary western perspectives”, en Zelazo, Philip David, Morris Moscovitch y Evan Thompson (eds.), *The Cambridge Handbook of Consciousness*, Cambridge University Press, Nueva York, pp. 35-66.
- _____ y Keneth Williford (eds.), 2006, *Self-Representational Approaches to Consciousness*, MIT, Cambridge.
- Legrand, Dorothee, 2005, “Le soi corporel”, en *L'Évolution Psychiatrique*, núm. 70, pp. 709-719.
- _____, 2006, “The bodily self: The sensori-motor roots of pre-reflective self-consciousness”, en *Phenomenology and the Cognitive Sciences*, núm. 5, pp. 89-118.

- _____, 2007, “Pre-reflective self-as-subject from experiential and empirical perspectives”, en *Consciousness and Cognition*, núm. 16, pp. 583-599.
- Levine, Joseph, 2001, *Purple Haze: The Puzzle of Consciousness*, Oxford University Press, Oxford.
- Lycan, William G., 1995, “A limited defence of phenomenal information”, en Thomas Metzinger (ed.), *Conscious Experience*, Imprint Academic, Reino Unido, pp. 243-258.
- _____, 1996, *Consciousness and Experience*, MIT, Cambridge.
- Martin, M.G.F., 1995, “Bodily awareness: A sense of ownership”, en José Luis Bermúdez, Anthony Marcel y Naomi Eilan (eds.), *The Body and the Self*, MIT, Cambridge, pp. 267-289.
- Nagel, Thomas, 1974, “What is it like to be a bat?”, en *The Philosophical Review*, núm. 83, pp. 435-450.
- Neander, Karen, 1998, “The division of phenomenal labor: A problem for representational theories of consciousness”, en *Philosophical Perspectives*, núm. 12, pp. 411-434.
- Newen, Albert y Kai Vogeley 2003, “Self-representation: Searching for a neural signature of self-consciousness”, en *Consciousness and Cognition*, núm. 12, pp. 529-543.
- Pitt, David, “Mental representation”, en Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, en línea: <http://plato.stanford.edu/entries/mental-representation>.
- Roessler, Johannes y Naomi Eilan (eds.), 2003, *Agency and Self-Awareness: Issues in Philosophy and Psychology*, Oxford University Press, Oxford.
- Rosenthal, David M., 1993, “Higher-order thoughts and the appendage theory of consciousness”, en *Philosophical Psychology*, núm. 6, pp. 155-166.
- _____, 1997, “A theory of consciousness”, en Ned Block, Owen Flanagan y Güven Güzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*, MIT, Cambridge, pp. 729-753.
- _____, 2002, “Explaining consciousness”, en David J. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*, Oxford University Press, Nueva York, pp. 406-421.
- Rowlands, Mark, 2001, *The Nature of Consciousness*, Cambridge University Press, Cambridge, 2001.
- Seager, William, 1999, *Theories of Consciousness: An Introduction and Assessment*, Routledge, Londres.

- Seeley, William W. y Virginia E. Sturm, 2007, "Self-representation and the frontal lobes", en Bruce L. Miller y Jeffrey L. Cummings (eds.), *The Human Frontal Lobes: Functions and Disorders*, The Guilford Press, Nueva York, pp. 317-334.
- Synofzik, Matthis, Gottfried Vosgerau y Albert Newen, 2008, "I move, therefore I am: a new theoretical framework to investigate agency and ownership", en *Consciousness and Cognition*, núm.17, pp. 411-424.
- Textor, Mark, 2006, "Brentano (and some neo-Brentanians) on inner consciousness", en *Dialectica*, núm. 60, pp. 411-432.
- Thompson, Evan, 2007, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, The Belknap Press of Harvard University Press, Cambridge.
- _____ y Dan Zahavi, 2007, "Philosophical issues: phenomenology", en Philip David Zelazo, Morris Moscovitch y Evan Thompson (eds.), *The Cambridge Handbook of Consciousness*, Cambridge University Press, Nueva York, pp. 67-87.
- Tye, Michael, 1995, *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, MIT, Cambridge.
- _____, 2000, *Consciousness, Color, and Content*, MIT, Cambridge.
- _____, 2009, *Consciousness Revisited: Materialism without Phenomenal Concepts*, MIT, Cambridge.
- Wider, Kathleen V., 1997, *The Bodily Nature of Consciousness: Sartre and Contemporary Philosophy of Mind*, Cornell University Press, Ithaca.
- Zahavi, Dan, 2004, "Back to Brentano?", en *Journal of Consciousness Studies*, núm 11, pp. 66-87.
- _____, 2005, *Subjectivity and Selfhood: Investigating the First-Person Perspective*, MIT, Cambridge.