

LOS ENFOQUES AUTORREPRESENTACIONALES DE LA CONCIENCIA¹

Cette conscience (de) soi, nous ne devons pas la considérer comme une nouvelle conscience, mais comme *le seul mode d'existence qui soit possible pour une conscience de quelque chose*. De même qu'un objet étendu est contraint d'exister selon les trois dimensions, de même une intention, un plaisir, une douleur ne sauraient exister que comme conscience immédiate (d') eux-mêmes.

JEAN-PAUL SARTRE (1943: 20)

En la filosofía occidental ha existido una añeja tradición que considera que los contenidos de la mente se dividen en dos categorías generales: (i) los estados mentales intencionales y (ii) los estados mentales fenoménicos no-intencionales. Franz Brentano distinguía entre actos mentales y contenidos mentales. De acuerdo con él, el contenido mental es un *contenido intencional*, es decir, los estados mentales (los procesos) son siempre acerca de algo, están siempre dirigidos a algo. Este algo puede no existir, por lo que se dice que los estados mentales también pueden representar *inexistentes intencionales*.

Este rasgo de ser *acerca de algo* es lo que se conoce como *intencionalidad*. Según Brentano, la *intencionalidad* es el rasgo definitorio de los fenómenos mentales y es lo que hace la diferencia entre lo mental y lo físico.

¹ Reseña del libro de Uriah Kriegel y Kenneth Williford (eds.) (2006), *Self-Representational Approaches to Consciousness*, Cambridge, Estados Unidos, MIT Press, vi + 561 p.

Ahora, otra forma de referirse a los estados intencionales² en vista de llevar un contenido es cuando se dice que los estados mentales *representan* algún estado de cosas. Así, mi creencia de que el cielo es azul es *acerca del* cielo o, en otros términos, *representa* el cielo. En contraste, los estados no-intencionales o no-representacionales, no llevan un contenido, no representan, por lo cual son puramente fenoménicos, *i.e.*, son *qualia*.

En la filosofía de la mente contemporánea se ha desarrollado una concepción con el nombre de *teorías representacionales de la mente y la conciencia* (en breve, *representacionismo*), que intenta romper con la tradición y considera que todos y cada uno de los estados mentales, siempre representan algo, es decir, siempre son estados representacionales.

Hay dos principales clases de teorías representacionales: (i) *teorías representacionales de primer orden* (RPO), y (ii) *teorías representacionales de orden superior* (ROS). La meta de las teorías RPO³ es caracterizar las propiedades fenoménicas de la experiencia en términos de los contenidos representacionales de la misma. Así, la diferencia entre una experiencia de amarillo y una experiencia de azul se explica como una diferencia en las propiedades representadas en cada caso (*i.e.*, las propiedades de las superficies para reflejar la luz). De igual modo, la diferencia entre un dolor y una comezón se explica en términos representacionales diciendo que ésta radica en las propiedades o tipos de perturbaciones que se representan como localizadas en diversas regiones del cuerpo de una persona. Así, de acuerdo con las teorías RPO, una experiencia fenoménica está

² Se dice que los estados intencionales o representacionales también son actitudes proposicionales, porque expresan una *actitud* hacia una proposición específica (Churchland, 1988: 63) e incluyen estados mentales como deseos, creencias, miedos, etcétera, y su rasgo común es que se los podemos atribuir a los sujetos utilizando enunciados de la forma 'S \emptyset que p', donde 'S' es una persona, un animal no-humano o un sistema artificial, 'p' es una proposición acerca de un estado de cosas en el mundo y \emptyset es cualquier verbo de una actitud proposicional. Así, decimos por ejemplo que 'Pedro cree que lloverá', 'María quiere ir a la fiesta', etcétera.

³ En su forma moderna la teoría de representación de primer orden, o simplemente representacionismo, ha sido desarrollada por Fred Dretske (1995) y Michael Tye (1995).

*posicionada (poised)*⁴ para tener un impacto en las creencias y en los procesos de razonamiento práctico de los sujetos para guiar su conducta.

Las teorías ROS establecen, en cambio, que para experimentar un estado de cosas ϵ , se requiere un estado representacional de segundo orden u orden superior acerca del estado representacional de primer orden u orden inferior acerca de ϵ .

Las teorías ROS se dividen a su vez en: (i) *teorías de experiencia de orden superior* (EOS),⁵ y (ii) *teorías de pensamiento de orden superior* (POS).⁶ De acuerdo con las teorías EOS, un estado mental es consciente si es el objeto de una representación de segundo orden u orden superior. Esta metarrepresentación puede consistir en una *cuasipercepción*, un sentido interno o un monitoreo, acerca de los estados mentales de primer orden u orden inferior respecto de algún estado de cosas ϵ , por lo cual a la teoría también se le conoce, 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 perceptuales externos, pero “está dirigida internamente y no externamente” (Churchland, 1988: 74), y es una metarrepresentación de *grano fino* en tanto es densa, indescriptible y concreta, *i.e.*, no es una abstracción, sino más bien un *sentido*, análogo a un *ojo de la mente*, un mecanismo de *escáner interno* que monitorea los estados mentales de primer orden acerca de ϵ .

Por su parte, las teorías POS 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 . Así, según las teorías POS un estado mental es consciente si va acompañado por una de ellas.

En *Self-Representational Approaches to Consciousness*, los filósofos Uriah Kriegel y Kenneth Williford tienen como objetivo mostrar una teoría alternativa de la conciencia que, según sus propias palabras, “ha ido ganando

⁴ Tye, 1995: 236. De acuerdo con Tye: “un estado mental es fenoméricamente consciente sólo en caso de que tenga un CPANI —un Contenido Posicionado, Abstracto, No conceptual e Intencional adecuado—”(1998/2004: 658).

⁵ Véanse, Armstrong, 1981; Churchland, 1988; Lycan, 1995.

⁶ Véanse, Carruthers, 2007; Rosenthal, 1997.

visibilidad pero sigue estando [...] sin ser discutida” (p. 1). A esta teoría la llaman teoría autorrepresentacional de la conciencia (ARC) y en breve sostiene que un estado mental es consciente si y sólo si se representa a sí mismo del modo correcto.

La teoría ARC combina algunos elementos de las teorías RPO y ROS, pero evita las dificultades que éstas enfrentan. El atractivo inicial de dichas teorías es que abren el camino para explicaciones reductivas de la conciencia en términos representacionales no fenoménicos e incluso en términos físicos, dejando lugar a explicaciones fisicalistas de la conciencia.

Sin embargo, las teorías RPO y ROS enfrentan algunas dificultades. Además de los contraejemplos que se les han lanzado, tienen algunos problemas principales. En el caso de la teoría RPO, desde el supuesto de que cualquier estado de cosas ϵ pueda ser representado consciente o inconscientemente, el problema es que no queda claro cómo es que la sola representación de ϵ puede hacer que el estado representado sea un estado consciente.

Las teorías ROS superan este problema introduciendo un estado representacional de orden superior de los estados conscientes. Así, lo que hace consciente a un estado representacional es justo el hecho de que la representación es en sí misma representada.

Pero la introducción de un estado de orden superior puede conducir a algunos problemas. En particular, las representaciones de orden superior pueden representarse de forma equívoca, como sus contrapartes de primer orden. Además, pueden manifestar de modo erróneo no sólo las propiedades de un estado de cosas ϵ , sino también su existencia misma. De esto se sigue que las teorías ROS sin un ϵ al que se dirijan, pueden resultar en una impresión subjetiva de encontrarse en un estado consciente sin realmente estarlo, *i.e.*, un sujeto puede no tener un estado consciente en el tiempo t , pero tiene cierta experiencia fenoménica de que sí lo está.

La teoría ARC evita estas dificultades y, al igual que las teorías ROS, puede dar cuenta de la diferencia entre estados representacionales conscientes e inconscientes del mismo ϵ , pero a diferencia de las teorías ROS, en la teoría ARC los estados inconscientes también son autorrepresentacionales como los conscientes. Entonces, cuando tengo una experiencia de ver el cielo azul, me encuentro en un estado que representa tanto un ϵ azul como a sí mismo.

En su introducción, Kriegel y Williford (p. 3) desarrollan un breve argumento para la teoría ARC, como sigue: (1) los estados conscientes son estados de los que uno está consciente de encontrarse en ellos; (2) los estados de los que uno está consciente de encontrarse en ellos son estados representados; por lo tanto, (3) los estados conscientes son estados representados; (4) cuando un estado es representado lo es por sí mismo o por otro diferente; pero, (5) hay buenas razones para pensar que un estado consciente no es representado por uno distinto; por lo tanto, (6) hay buenas razones para pensar que un estado consciente es representado por sí mismo.

(1) a (3) hacen un primer paso de la teoría ARC al excluir las teorías RPO y conducir a la noción de que los estados conscientes implican la conciencia que tiene el sujeto de ellos, y de (4) a (6) forman un segundo paso que excluye las teorías ROS y conduce a la teoría ARC.

Según Kathleen Wider, la teoría ARC y la idea general de que los estados conscientes tienen un aspecto autorreferencial, ya fue defendida por René Descartes, John Locke, Immanuel Kant, y más tarde por Brentano y la tradición fenomenológica, en especial Sartre.⁷ Pero no fue sino hasta muy recientemente que los filósofos analíticos han atendido a la teoría ARC.⁸

El texto de Kriegel y Williford puede considerarse entonces como un trabajo pionero que pretende no sólo desarrollar la teoría ARC, sino capturar la atención de los filósofos analíticos que han desatendido a esta teoría.

Pasando al libro como tal, éste se divide en cuatro partes. La primera contiene capítulos que defienden alguna forma de la teoría ARC. La segunda, comprende apartados en contra de ella. Los capítulos de la tercera parte exploran ciertas conexiones entre la teoría y algunas cuestiones filosóficas. La última está formada por dos capítulos de dos figuras importantes.

La primera parte abre con un capítulo de Robert Van Gulick, quien presenta su *teoría de los estados globales de orden superior de la conciencia*,

⁷ Cfr., Wider, 1997: 7-39. Caston sostiene incluso que una forma de esta teoría ya está presente en Aristóteles (2002).

⁸ Por ejemplo, Anderson y Perlis, 2005; Hossack, 2003; Kriegel, 2003; Legrand, 2005; Zahavi, 2005.

una versión no fuerte de la teoría ARC en la que los estados conscientes son característicamente autorrepresentacionales, pero no de modo necesario o, incluso, universal. En el siguiente capítulo, Terry Horgan, John Tienson y George Graham hacen notar que ciertas formas de escepticismo global acerca de nuestra vida consciente no tienen la misma sujeción en nosotros como sus contrapartes en el mundo externo y proponen que la única explicación plausible de esto implica el supuesto de que las experiencias conscientes son autorrepresentacionales. Después, Kathleen Wider argumenta en favor de (i) la teoría ARC y (ii) que la naturaleza esencialmente afectiva de las experiencias conscientes es la fuente de su carácter autorrepresentacional. Andrew Brook ofrece después una interpretación de la teoría kantiana de la conciencia y en los dos últimos capítulos de esta parte, los editores defienden dos versiones de la teoría ARC: una versión fuerte (Williford) y una débil (Kriegel).

La segunda parte empieza con un capítulo de Joseph Levine, quien arguye que ninguna versión de la teoría ARC puede cerrar la brecha explicativa entre los estados físicos y la experiencia consciente e incluso capturar la naturaleza o fuente de tal brecha. En el capítulo siguiente, John J. Drummond defiende una concepción husserliana de la conciencia contra la versión de Brentano de la ARC. Posteriormente, Rocco Gennaro defiende una versión débil de la ARC contra una versión más fuerte de la misma. Christopher Hill en el siguiente capítulo argumenta en favor de las teorías RPO de la conciencia perceptual que no implican ninguna referencia a una estructura autorrepresentacional. Dan Zahavi termina la segunda parte con una exposición de la teoría ARC en la tradición fenomenológica y concluye con la afirmación de que la teoría no hace justicia a la fenomenología de la experiencia, lo cual sugiere una concepción puramente intrínseca de la conciencia.

En el primer capítulo de la tercera parte, Peter Carruthers sugiere que las experiencias conscientes tienen una función autorrepresentacional, pero que las creencias y deseos no tienen este papel funcional y, por lo tanto, no son estados conscientes. Robert Lurz propone, contra Carruthers, un modelo de *monitorización de orden mismo* para dar cuenta de creencias y deseos como estados conscientes. Después, Jason Ford y David Woodruff Smith exploran la posibilidad de naturalizar y desmitificar la autorrepresentación implícita en la conciencia apoyándose en la literatura referente

a la atención. En el penúltimo capítulo de este apartado, Tomis Kapitan argumenta: (i) que los pensamientos y experiencias deícticas implican una forma de autoconciencia, y que, entonces, (ii) si todos los estados conscientes están deícticamente mediados, la autoconciencia es ubicua en nuestra vida consciente. Keith Lehrer sugiere después que la fuente de lo problemático de los estados conscientes es su *contenido lúcido*, y que entonces una explicación satisfactoria de esta lucidez representacional involucrará la tesis de que estos estados implican una clase particular de autorrepresentación.

La última parte contiene los dos capítulos más extensos. En el primero, David Rudrauf y el neurólogo portugués Antonio Damasio proponen una teoría de la conciencia en la que están íntimamente conectadas la afectividad y la subjetividad. Por último, Douglas R. Hofstadter sugiere que en la esencia de la conciencia hay algo análogo a un *bucle extraño* de autorreferencia matemática, sugerido por los teoremas de incompletitud de Kurt Gödel.

Self-Representational Approaches to Consciousness es, sin duda alguna, una importante y fundacional obra que no debe perder de vista el interesado en el estudio de la conciencia. El texto, sin embargo, es de carácter básicamente filosófico y las implicaciones de la teoría autorrepresentacional para otros modelos teóricos —en particular en neurociencias y ciencias cognitivas— aún están por determinarse. A pesar de la creciente literatura en favor de la teoría ARC hacen falta exámenes críticos que establezcan sus defectos y limitaciones. Por ejemplo, si queremos estudiar las bases neuronales de los estados conscientes autorrepresentacionales y asumimos una concepción de identidad mente-cerebro, ¿qué significa entonces que un estado cerebral se autorrepresente o se autorrefiera? De igual manera, si se asume una explicación causal de la representación, ¿cómo puede un estado mental autorrepresentarse si no puede causarse a sí mismo?

BIBLIOGRAFÍA

Anderson, Michael L. y Donald R. Perlis (2005), "The roots of self-awareness", *Phenomenology and the Cognitive Sciences*, vol. 4, pp. 297-333.

- Armstrong, David M. (1981), *The Nature of Mind*, Ithaca/Nueva York, Estados Unidos, Cornell University Press.
- Carruthers, Peter (2007), "Higher-order theories of consciousness", en Max Velmans y Susan Schneider, *The Blackwell Companion to Consciousness*, Oxford, Reino Unido, Blackwell, pp. 277-286.
- Caston, Victor (2002), "Aristotle on consciousness", *Mind*, vol. 111, pp. 751-815.
- Churchland, Paul (1988), *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind*, Cambridge, Estados Unidos, MIT Press.
- Dretske, Fred (1995), *Naturalizing the Mind*, Cambridge, Estados Unidos, MIT Press.
- Hossack, Keith (2003), "Consciousness in act and action", *Phenomenology and the Cognitive Sciences*, vol. 2, pp. 187-203.
- Kriegel, Uriah (2003), "Consciousness as sensory quality and as implicit self-awareness", *Phenomenology and the Cognitive Sciences*, vol. 2, pp. 1-26.
- Legrand, Dorothée (2005), "Le soi corporel", *L'Évolution Psychiatrique*, vol. 70, pp. 709-719.
- Lycan, William (1995), "Consciousness as internal monitoring", *Philosophical Perspectives*, vol. 9, pp. 1-14.
- Rosenthal, David M. (1997), "A theory of consciousness", en Ned Block, Owen Flanagan y Güven Güzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*, Cambridge, Estados Unidos, MIT Press, pp. 729-753.
- Sartre, Jean-Paul (1943), *L'être et le néant: Essai d'ontologie phénoménologique*, París, Francia, Gallimard.
- Tye, Michael (1998/2004), "Précis of *Ten problems of consciousness*", en John Heil (ed.), *Philosophy of Mind: A Guide and Anthology*, Oxford, Reino Unido, Oxford University Press, pp. 657-664.
- Tye, Michael (1995), "A representational theory of pains and their phenomenal character", *Philosophical Perspectives*, vol. 9, pp. 223-239.
- Wider, Kathleen V. (1997), *The Bodily Nature of Consciousness: Sartre and Contemporary Philosophy of Mind*, Ithaca/Nueva York, Estados Unidos, Cornell University Press.
- Zahavi, Dan (2005), *Subjectivity and Selfhood*, Cambridge, Estados Unidos, MIT Press.

ISRAEL GRANDE-GARCÍA*

D. R. © Israel Grande-García, México, D.F., julio-diciembre, 2007.

* Universidad Autónoma Metropolitana-Iztapalapa, israel.grandegarcia@gmail.com