

DOI: 10.22201/iifs.18704905e.2018.16

## RESPUESTA A SEBASTIÁN: EN DEFENSA DE LA NATURALEZA CONSCIENTE DE LAS CREENCIAS DE PRIMERA PERSONA

JAVIER VIDAL  
Universidad de Concepción (Chile)  
fravidal@udec.cl

RESUMEN: En esta revista, argumenté en favor de la naturaleza consciente de las creencias de primera persona (2014). En un artículo más reciente, Miguel Ángel Sebastián (2017) trata de mostrar que tal argumento no es sólido. Aquí procedo a lidiar con sus tres críticas principales. En primer lugar, sostengo que el argumento es válido si conocer una creencia de primera persona consiste en creer que uno la tiene. Defiendo también que no hay ninguna regresión al infinito de creencias conscientes de orden superior que sea dañina. En segundo lugar, argumento que la noción de conciencia en juego es aquella que es relevante en la discusión acerca de la naturaleza de la conciencia y del inconsciente. Por último, reviso la justificación del principio (SC\*) para mostrar que no prejuzga la cuestión.

PALABRAS CLAVE: conciencia, accesibilidad, creencia de orden superior, creencia inconsciente, dinámica cognitiva

SUMMARY: In this journal, I argued for the conscious nature of first-person beliefs (2014). In a more recent paper, Miguel Ángel Sebastián (2017) tried to show that such an argument is unsound. Here I address his three core criticisms. Firstly, I maintain that my argument is valid if to know a first-person belief consists in believing one has it. Likewise, I defend that no damaging infinite regression of conscious higher-order beliefs is involved. Further, I argue that the notion of consciousness at stake is the relevant notion for discussion about the nature of consciousness and the unconscious. Finally, I examine the justification for the principle (SC\*) in order to show that it does not beg the question.

KEY WORDS: consciousness, accessibility, higher-order belief, unconscious belief, cognitive dynamics

En esta revista publiqué un artículo (2014) en el que argumentaba que las creencias con un contenido de primera persona, o creencias *de se*, son necesariamente conscientes. A este respecto puede hablarse de la naturaleza consciente de las creencias de primera persona. En un artículo más reciente en esta revista, Miguel Ángel Sebastián (2017) cuestiona el argumento por no ser sólido. Sebastián presenta un abanico de consideraciones, algunas de las cuales él mismo admite que no tienen fuerza contra mi argumento, como ocurre con la argumentación de Williamson contra la luminosidad de la mayor

parte de los estados mentales (2017, p. 49). En este artículo sólo me centraré en las objeciones dirigidas al valor probativo y a la relevancia de mi argumento. En primer lugar, Sebastián establece que si me comprometo con la idea de que saber implica creer, entonces el argumento que él bautiza como CONOCERME, que es una parte del argumento principal, resulta ser inválido o introducir una dañina regresión al infinito. También el principio (SC\*), que, aplicado a las creencias de primera persona, lleva de CONOCERME a la conclusión de que tales creencias son necesariamente conscientes, introduciría una regresión dañina (2017, pp. 49–51). En segundo lugar, considera que la noción de conciencia que tendría sentido aquí no es la noción relevante en la discusión de la teoría de la conciencia en términos de creencias o pensamientos de orden superior, que en la versión de Rosenthal yo había criticado, ni de la teoría freudiana como una explicación de ciertos comportamientos en términos de estados intencionales inconscientes (2017, pp. 52–53, 57). Por último, Sebastián sostiene que no ofrezco ninguna justificación en favor del principio (SC\*), o bien la justificación que él propone como reconstrucción de mi pensamiento prejuzgaría la cuestión (2017, pp. 53–57).

A continuación responderé a estas críticas en el orden recién expuesto. En la primera sección mostraré, de entrada, que CONOCERME no es un argumento inválido si saber implica creer o, más bien, si en este caso particular saber consiste en creer. Además, veremos que, aunque es cierto que tanto CONOCERME como (SC\*) introducen una regresión al infinito, esa regresión no es dañina según cierta concepción de las creencias de orden superior y de las creencias conscientes. En la segunda sección utilizaré, en relación con el término “conciencia”, la distinción entre fijar la extensión de un término de clase natural y determinar lo que es esencial para ser un elemento en esa extensión, con el fin de argumentar que la noción de conciencia empleada aquí sí es la noción relevante tanto en la discusión de la teoría de Rosenthal como en la discusión de la teoría freudiana del inconsciente. En la tercera sección, revisaré la justificación en favor del principio (SC\*) en términos de una dinámica cognitiva dependiente entre creencias para mostrar, en conclusión, que en ningún caso puede hablarse de prejuzgar la cuestión.

## I

Mediante el subargumento que Sebastián denomina CONOCERME, empecé por establecer en mi artículo la existencia de una conexión necesaria entre tener una creencia de primera persona y saber que

uno mismo tiene esa creencia. En efecto, supongamos que un individuo  $X$  tiene una creencia con un contenido de primera persona de la forma:

(1) Yo  $\Phi$ .

En términos del uso del cuasiíndexico “él mismo”, una adscripción de esa creencia será de la forma:

(2)  $X$  cree que él mismo  $\Phi$ .

Ahora bien, la referencia de primera persona puede caracterizarse por referirse con éxito a uno mismo *como uno mismo* de tal modo que, en el caso de la creencia, el individuo que la tiene no puede ignorar la identidad que guarda con el individuo acerca del cual es esa creencia. En otras palabras, no puede ser que  $X$  ignore que el individuo acerca de quien cree algo es él mismo. A este respecto podríamos decir, como señala Sebastián, que las creencias de primera persona son inmunes al error por fallo en la referencia (IER).<sup>1</sup> Así, sostuve crucialmente que de (2) se sigue:

(3)  $X$  sabe que la creencia de que alguien  $\Phi$  es acerca de quien cree que alguien  $\Phi$ , es decir, él mismo.

Por último, de (3) se sigue trivialmente:

(4)  $X$  sabe que él mismo cree que él mismo  $\Phi$ .

Según Sebastián, CONOCERME es un argumento inválido porque el paso de (2) a (3) no se justifica. Más aún, si saber implica creer, eso significaría que como resultado de tener la creencia de que él mismo  $\Phi$ ,  $X$  también tendría la creencia de que  $p$ , donde  $p$  es la proposición de que *la creencia de que alguien  $\Phi$  es acerca de quien cree que alguien  $\Phi$ , es decir, él mismo*. Pero Sebastián sostiene que la idea de que  $X$  no puede ignorar que  $p$ , que es lo que establece IER, admite una lectura débil según la cual  $X$  no puede tener la creencia de que no  $p$  y una lectura fuerte según la cual  $X$  no puede no tener la creencia de que  $p$  (que, como parece plausible, sería una creencia que satisfaría las condiciones para el conocimiento). Concluye entonces

<sup>1</sup> Adviértase que no es simplemente que la referencia esté garantizada, como podría sugerir la terminología de Sebastián, sino que  $X$  no puede ignorar la identidad que guarda con el referente. Pero mantendré esa terminología.

que la lectura débil es suficiente para dar cuenta de IER sin que, no obstante, justifique el paso de (2) a (3) (2017, p. 50).

Empezaré por señalar que en este artículo asumiré, por mor de la discusión, que en (3) saber consiste en creer bajo condiciones que, en este caso, son adecuadas. Téngase en cuenta que eso no equivale a decir que saber implique creer ni que el conocimiento sea analizable en términos de creencia. Puede ser, por ejemplo, que los estados perceptivos sean estados de conocimiento sin que haya ninguna creencia involucrada en la percepción. Mi argumento será simplemente asumir que, en el caso particular de (3), saber sí consiste en tener una creencia.

Ahora recordemos por un momento una versión de la historia de Edipo, según la cual Edipo cree que el hijo de Layo es culpable de su muerte sin saber que él mismo es el hijo de Layo. Esta historia muestra que la creencia de que el hijo de Layo es culpable no es inmune al error por fallo en la referencia. Es cierto que, cuando Edipo tiene la creencia de que el hijo de Layo es culpable, el análogo mental del término “hijo de Layo” refiere a Edipo. Más aún, puesto que es Edipo quien tiene la creencia, resulta que mediante el análogo mental de “hijo de Layo” Edipo se refiere reflexivamente a él mismo. Sin embargo, de acuerdo con esta versión, Edipo debe ignorar que *la creencia de que el hijo de Layo es culpable es acerca de quien cree que el hijo de Layo es culpable, es decir, él mismo*. Digamos que ésta es la proposición  $q$ . Según la lectura débil que propone Sebastián, la ignorancia que tendríamos aquí es equivalente al hecho de que Edipo tiene la creencia errónea de que no  $q$ . En efecto, si no poder ignorar que  $q$  corresponde a la imposibilidad de tener la creencia errónea de que no  $q$ , entonces ignorar de hecho que  $q$  corresponde a tener la creencia errónea de que no  $q$ .<sup>2</sup> Pero se trataba de que Edipo no sabe (o no cree) que él mismo es el hijo de Layo, no de que cree que él mismo no es el hijo de Layo. No he incluido esta creencia en mi versión de la historia (que, desde luego, puede estar presente en otra versión). Por lo tanto, no estaría justificado atribuir a Edipo la creencia de que no  $q$ . Así, para explicar el error por fallo en la referencia, estaríamos atribuyendo a Edipo una creencia de orden superior que de hecho Edipo no tiene en este caso. En cambio, este problema no surge cuando introducimos la lectura fuerte. Según la lectura fuerte, la ignorancia de Edipo

<sup>2</sup> Desde luego, hacer equivalentes “ignorar que  $q$ ” y “tener la creencia errónea de que no  $q$ ” no resiste mucho análisis si se considera en sí mismo. Como es obvio, puedo ignorar que se da algo sin tener la creencia errónea de que no se da.

consiste simplemente en no saber que la creencia de que el hijo de Layo es culpable es acerca de quien cree que el hijo de Layo es culpable, es decir, él mismo. En otras palabras, la cuestión es simplemente que Edipo no tiene la creencia de que  $q$ . Podría decirse a este respecto que la lectura débil de Sebastián introduce creencias de orden superior en el lugar equivocado. Más aún, resulta que Edipo debe tener la creencia de que no  $q$ , cuando Sebastián había cuestionado precisamente que en el caso de las creencias de primera persona uno tuviera que formarse una creencia de orden superior, lo que, según él, uno podría perfectamente no hacer.

Considérese también el siguiente argumento. Cuando Edipo tiene la creencia de que el hijo de Layo es culpable puede o no saber que él mismo es el hijo de Layo. Podría tener la creencia de que el hijo de Layo es culpable y, además, saber que él mismo es el hijo de Layo. En cualquier caso, por el hecho mismo de tener la creencia de que el hijo de Layo es culpable, Edipo no sabe que  $q$  (es decir, no sabe que esa creencia es acerca de él mismo). Pero tenemos una explicación de lo que Edipo necesita saber en primer lugar para llegar a saber que  $q$ . A Edipo le falta saber que él mismo es el hijo de Layo para llegar a saber que la creencia de que el hijo de Layo es culpable es acerca de quien cree que el hijo de Layo es culpable, es decir, él mismo. Sin embargo, la situación es distinta cuando Edipo tiene la creencia de que él mismo es culpable. Según el planteamiento de Sebastián, por el hecho mismo de tener la creencia de que él mismo es culpable, Edipo no sabe que *la creencia de que alguien es culpable es acerca de quien cree que alguien es culpable, es decir, él mismo*.<sup>3</sup> Digamos que ésta es la proposición  $r$ . El problema aquí es que no parece que falte ningún conocimiento previo para llegar a saber que  $r$ . No hay nada equivalente al conocimiento de que él mismo es el hijo de Layo para que Edipo llegue a saber que  $r$ .<sup>4</sup> Sebastián puede insistir en que, a pesar de que no falte ningún conocimiento previo para llegar a saber eso, Edipo podría simplemente no formarse la creencia de

<sup>3</sup> Por cierto, téngase en cuenta lo extraño que resulta decir que así como Edipo no sabe que la creencia de que el hijo de Layo es culpable es una creencia acerca de él mismo, tampoco sabe que la creencia de que él mismo es culpable es una creencia acerca de él mismo.

<sup>4</sup> Una cuestión distinta es que, de acuerdo con la historia, Edipo primero tiene la creencia de que el hijo de Layo es culpable y llega a creer (o saber) que él mismo es culpable mediante el conocimiento de que él mismo es el hijo de Layo. Por lo tanto, para llegar a creer (o saber) que él mismo es culpable se requería este conocimiento previo. Pero aquí se trata de que una vez que Edipo tiene la creencia de que él mismo es culpable no es necesario ningún otro conocimiento previo para llegar a saber que  $r$ .

que  $r$  cuando tiene la creencia de que él mismo es culpable. Pero esta respuesta no resulta satisfactoria en términos epistémicos. Tenemos una explicación según la cual Edipo no sabe que la creencia de que el hijo de Layo es culpable es una creencia acerca de él mismo si no tiene el conocimiento previo de que él mismo es el hijo de Layo. Es legítimo concluir que Edipo sabe por defecto que la creencia de que él mismo es culpable es una creencia acerca de él mismo porque no le falta ningún conocimiento previo para llegar a saberlo. Pero si saber consiste en creer en este caso, entonces Edipo tuvo que formarse la creencia de que  $r$ .

Concluyo entonces que el paso de (2) a (3) está justificado. A continuación Sebastián establece que si el paso de (2) a (3) está justificado, se produce una dañina regresión al infinito de creencias de orden superior (2017, p. 50). Veamos primero que, en efecto, se produce una regresión al infinito. Como vimos, de (3) se sigue trivialmente (4). Pero si saber consiste en creer en este caso, entonces de (4) se sigue que  $X$  cree que él mismo cree que él mismo  $\Phi$ . Esta creencia de segundo orden es a su vez una creencia de primera persona, de manera que, por aplicación de CONOCERME, se sigue que  $X$  tiene una creencia de tercer orden. Pero esta creencia de tercer orden es a su vez una creencia de primera persona, de manera que, por aplicación de CONOCERME, se sigue que  $X$  tiene una nueva creencia de orden superior, y así *ad infinitum*. Por si esto fuera poco, Sebastián considera que mi propuesta en favor de la naturaleza consciente de las creencias de primera persona introduciría una dañina regresión al infinito de creencias conscientes (2017, pp. 50–51). Veamos primero otra vez que efectivamente se trata de una regresión al infinito. En mi argumento empleo el siguiente principio fundamental, que discutiré en la última sección:

(SC\*) Si (si  $X$  cree que  $\alpha$ , entonces  $X$  sabe que él mismo cree que  $\alpha$ ), entonces (si  $X$  cree que  $\alpha$ , entonces  $X$  cree conscientemente que  $\alpha$ ).

Es una consecuencia de CONOCERME que las creencias de primera persona satisfacen el antecedente de (SC\*). Ahora bien, acabamos de ver que si  $X$  tiene una creencia de primera persona se sigue, por aplicación reiterada de CONOCERME, que  $X$  tiene una infinidad de creencias de orden superior. Por lo tanto, habrá un conjunto infinito de creencias de orden superior que satisfacen el antecedente de (SC\*). Se sigue entonces, por aplicación reiterada de (SC\*), que  $X$  tiene una infinidad de creencias conscientes.

Argumentaré ahora, frente a Sebastián, que ninguna de estas regresiones es dañina según cierta concepción de las creencias de orden superior y de las creencias conscientes. Para ello me serviré del planteamiento de Shoemaker (2009) en un contexto diferente. De entrada, Shoemaker sostiene que algunas creencias permanentes [*standing*], sean o no sean de primera persona, son conscientes, en el sentido de que uno está dispuesto a asentir mental o verbalmente a su contenido cada vez que se plantea la cuestión de su verdad; asimismo, son creencias que están disponibles para guiar el razonamiento y la acción. Considérese, por ejemplo, la creencia que la mayor parte de los latinoamericanos tienen de que Santiago es la capital de Chile. En primer lugar, ésta es una creencia permanente porque no se manifiesta, en todo momento dado, a través de un pensamiento o juicio con ese contenido. Pero es una creencia consciente, o de acceso consciente, porque su contenido se afirmará cada vez que uno se ponga a considerar si es o no es verdadera, además de estar disponible para ser la premisa de un razonamiento e influir racionalmente en el modo como uno actúa (si uno lo estima relevante para la consecución de los propios objetivos) (2009, p. 40).

Shoemaker establece luego que si uno tiene una de tales creencias permanentes de que  $\alpha$ , entonces tendrá la creencia de segundo orden de que él mismo tiene esa creencia de que  $\alpha$ . Su idea es, en primer lugar, que en el caso de tales creencias permanentes conscientes la cuestión de si uno cree o no cierta proposición es “transparente” a la cuestión de si la proposición es o no es verdadera (2009, pp. 41, 48). Por ejemplo, la cuestión de si uno cree o no que Santiago es la capital de Chile es transparente a la cuestión de si Santiago es o no es la capital de Chile. Más aún, parece incoherente, o irracional, que uno asienta a una proposición (de primer orden) y a la vez se rehúse a asentir a la proposición de que uno mismo cree esa proposición (de primer orden) (2009, p. 36). Pero podría objetarse que el hecho de que una creencia permanente de que  $\alpha$  sea consciente en el sentido introducido antes es suficiente para que uno asienta a la proposición de que uno mismo cree que  $\alpha$  si surge esta última cuestión. La disposición a ese asentimiento podría explicarse así sin necesidad de postular la existencia de una *creencia* de segundo orden cuyo contenido es que uno mismo cree que  $\alpha$ .<sup>5</sup> En este punto es donde Shoemaker introduce una concepción de las creencias de orden

<sup>5</sup> La objeción podría plantearse en términos de la distinción entre tener una creencia disposicional (o permanente) y tener una disposición a creer. Desde luego, estoy dispuesto a asentir tanto a la proposición de que Santiago es la capital de Chile como a la proposición de que Santiago no es la capital de Tanzania. Pero podría

superior que vuelve estéril la objeción. Como es evidente, no es sólo la creencia permanente de que  $\alpha$ , en su modalidad consciente, lo que explica la disposición a ese asentimiento: se requiere, además, poseer los conceptos de creencia y de uno mismo (o de primera persona), así como cierto grado de racionalidad. Propone entonces que todo eso, junto con el hecho de tener una creencia permanente consciente de que  $\alpha$ , *constituye* el hecho de tener la creencia de segundo orden cuyo contenido es que uno mismo cree que  $\alpha$  (Shoemaker 2009, p. 42). A continuación Shoemaker se da cuenta de que si esta creencia de segundo orden es a su vez consciente, como parece serlo,<sup>6</sup> entonces se sigue, por el razonamiento anterior, que uno tiene una creencia de tercer orden, y así *ad infinitum*. Pero concluye que, en términos de su concepción de las creencias de orden superior, esta amenaza de regresión no es dañina:

Tener la creencia de segundo orden requiere la posesión de conceptos, de creencia y de uno mismo, que van más allá de los requeridos por la posesión de la creencia de primer orden, pero para la posesión de creencias de órdenes más altos no se requieren conceptos adicionales. Si tener la creencia de segundo orden de que uno tiene una creencia disponible [consciente] de primer orden sólo es tener esa creencia disponible [consciente] junto con los conceptos de creencia y de uno mismo y un grado adecuado de racionalidad, parece que tener esa creencia de segundo orden debería contar como tener la creencia de tercer orden de que uno tiene esa creencia, que esto a su vez debería contar como tener la creencia de cuarto orden, y así *ad infinitum*. (Esto tiene la consecuencia de que creencias con diferentes contenidos pueden tener la misma base metafísica.) (2009, pp. 42–43; la traducción es mía.)

decirse que mientras que en el primer caso tengo una creencia permanente, en el segundo caso sólo tengo una disposición a creer (Audi 1994, Lycan 1986). La objeción sería entonces que, más que tener una creencia permanente de segundo orden, tengo solamente una disposición a creer que creo una proposición (de primer orden). Pero, una vez que introducimos la concepción de Shoemaker de las creencias de orden superior, como veremos ahora, no es más problemático decir que tengo una creencia permanente de segundo orden que decir que sólo tengo una disposición a tener esa creencia.

<sup>6</sup> Parece serlo, puesto que, según el razonamiento anterior, la creencia de segundo orden satisface uno de los criterios para ser una creencia consciente, a saber, uno está dispuesto a asentir a su contenido. Pero entonces también la creencia de tercer orden será una creencia consciente. En efecto, sería irracional que uno asintiera a la proposición que es el contenido de la creencia de segundo orden, mientras se rehúsa a asentir a la proposición de que uno cree ese contenido. Este proceso se repite *ad infinitum*.

Con esto pretendo ilustrar, a través del planteamiento de Shoemaker, de qué modo una regresión al infinito de creencias de orden superior y de creencias conscientes puede ser metafísicamente inocua. En el caso de mi propuesta, concebida ahora *à la* Shoemaker, esa base metafísica común a un conjunto infinito de creencias de orden superior sería una creencia de primera persona (de primer orden) más la posesión del concepto de creencia y cierto grado de racionalidad. Pues me parece que es plausible entender la conexión necesaria establecida por CONOCERME entre tener una creencia de primera persona (de primer orden) y tener una creencia de segundo orden acerca de ella en términos de la idea de constitución que usa Shoemaker:<sup>7</sup> tener la creencia de primera persona (de primer orden), más la posesión del concepto de creencia y cierto grado de racionalidad, constituye el hecho de tener la creencia de segundo orden. Esto significa que tener la creencia de segundo orden *sólo es* tener la creencia de primera persona (de primer orden) más el concepto de creencia y cierto grado de racionalidad. Considérese la siguiente comparación. Supongamos que aceptamos el axioma de la lógica epistémica (o doxástica) según el cual si alguien tiene la creencia de que  $\alpha \rightarrow \beta$  y la creencia de que  $\alpha$ , entonces tiene la creencia de que  $\beta$ . Parece plausible sostener que, en ese caso, tener la creencia de que  $\beta$  sólo es tener la creencia de que  $\alpha \rightarrow \beta$  y la creencia de que  $\alpha$  (más poseer cierto grado de racionalidad). Ahora bien, también se sigue, por aplicación de CONOCERME, que uno tiene una creencia de tercer orden. Pero dado que ya no se requiere la posesión de ningún concepto adicional para tener creencias de órdenes más altos, resulta que tener la creencia de segundo orden contará como tener la creencia de tercer orden de que uno tiene esa creencia, que esto a su vez contará como tener la creencia de cuarto orden, y así sucesivamente.<sup>8</sup>

<sup>7</sup> Recordemos que si saber consiste en creer en este caso, por aplicación de CONOCERME se sigue que si uno tiene una creencia de primera persona (de primer orden), entonces tendrá también una creencia de segundo orden acerca de ella. Téngase en cuenta que el planteamiento de Shoemaker es más débil: no sostiene que si uno tiene una creencia permanente (de primer orden), entonces tendrá también una creencia de segundo orden acerca de ella. La creencia permanente debe ser, en primer lugar, una creencia disponible. Pero no todas las creencias permanentes son creencias disponibles (2009, p. 40).

<sup>8</sup> Una alternativa que consideré en mi artículo es sostener que la regresión no es dañina porque no hay independencia de rol funcional entre una creencia de primera persona y la creencia (o el conocimiento) de que uno tiene esa creencia (2014, n. 8). En palabras de Shoemaker, se trataría de que no hay independencia en la dinámica cognitiva de ambas creencias. No voy a repetir aquí el argumento, dado que a

Adviértase que, a diferencia de Shoemaker, no trato de justificar la idea de que uno tiene una creencia de segundo orden en términos de la noción de constitución. Para él, la noción de constitución permite explicar por qué uno tiene una creencia de segundo orden, y no nada más una creencia de primer orden que satisface ciertas condiciones, cuando está dispuesto a asentir a la proposición de que uno tiene esa creencia de primer orden. Sin embargo, en principio no podría excluirse, como me señaló uno de los árbitros, que tener una creencia de primera persona (de primer orden), junto con el concepto de creencia y cierto grado de racionalidad, sólo constituya el hecho de tener una *disposición* a creer que uno tiene esa creencia de primera persona (de primer orden). Pero mi afirmación de que uno tiene una creencia de segundo orden más que una disposición a tener esa creencia es independiente de la noción de constitución: es simplemente que, puesto que por CONOCERME uno *sabe* que tiene la creencia de primera persona (de primer orden) —donde saber consiste en creer—, se sigue que uno tiene una creencia de segundo orden más que una disposición a tenerla. Una vez que no está en duda que la conexión necesaria establecida por CONOCERME es entre tener una creencia de primera persona (de primer orden) y tener una creencia de segundo orden acerca de ella, parece legítimo, como resultado de entender dicha conexión necesaria en términos de la noción de constitución, aplicar esta noción a la creencia de segundo orden para luego mostrar que la regresión al infinito de las creencias de orden superior no es dañina.

Como vimos, mi propuesta también implica que si uno tiene una creencia de primera persona, entonces tiene una infinidad de creencias conscientes. A Sebastián este resultado le parece inaceptable para cualquier sentido relevante de “conciencia”. Por mi parte sostendré, en la siguiente sección, que el sentido que introduce Shoemaker forma parte del significado de “conciencia”. Como es obvio, la respuesta que daré ahora sobre la regresión al infinito de creencias conscientes no resultará plenamente satisfactoria mientras no argumente en favor de ese punto. Supongamos entonces, provisionalmente, que la idea de conciencia en términos de disponibilidad es relevante en la discusión. Esto significa, en particular, que tener una creencia consciente consiste en estar dispuesto a asentir a su contenido<sup>9</sup> (además de ser una creencia que está disponible para guiar el razonamiento y la

Sebastián no le parece convincente (2017, n. 12). En la sección III me ocuparé de este asunto en relación con la justificación del principio (SC\*).

<sup>9</sup> Para ser más exacto, debería decir que tener una creencia consciente es tener una creencia con cierta propiedad *categorica* por la que uno está dispuesto a asentir

acción).<sup>10</sup> Esta versión de la propuesta establece, por aplicación del principio (SC\*), que si uno tiene una creencia de primera persona (de primer orden), entonces esa creencia es consciente en el sentido de estar dispuesto a asentir a su contenido. Pero también se sigue, por aplicación de (SC\*), que la creencia de segundo orden, que uno tiene como consecuencia de tener esa creencia de primera persona (de primer orden), es consciente en el sentido de estar dispuesto a asentir a su contenido, y así sucesivamente. A este respecto no encuentro problemático decir que uno tiene una infinidad de creencias de orden superior que son conscientes en el sentido de estar dispuesto a asentir a sus contenidos. Olvidémonos por un momento del argumento en favor de la naturaleza consciente de las creencias de primera persona. Es difícil poner a discusión que tengamos una infinidad de disposiciones al asentimiento. En efecto, consideremos el simple hecho de que uno está dispuesto a asentir a una proposición de primer orden. Parece irracional, como sostiene Shoemaker, que uno asienta a una proposición y a la vez se rehúse a asentir a la proposición de que uno mismo cree esa proposición (de primer orden). Esto implica que uno está dispuesto a asentir a la proposición de que uno mismo cree esa proposición (de primer orden). Pero por el mismo razonamiento uno estará dispuesto a asentir a la proposición de que uno cree que uno mismo cree esa proposición (de primer orden), y así sucesivamente. En cierto modo puede decirse que mi argumento establece que, en el caso de una creencia de primera persona, tenemos necesariamente una infinidad de creencias de orden superior que son conscientes en el sentido de que tenemos necesariamente una infinidad de disposiciones al asentimiento, algo que, como acabo de mostrar, es un hecho en cualquier caso.

## II

Antes de pasar a discutir la justificación del principio (SC\*), Sebastián establece que la noción de conciencia que introduce Shoemaker, o yo mismo, no es la noción relevante en las teorías que explican la

a su contenido. Expongo esta idea en la próxima sección. Considérese aquí que de esto no se sigue que a una infinidad de creencias conscientes le corresponda una infinidad de tales propiedades categóricas. Pues, si una infinidad de creencias de orden superior está fundamentalmente *constituída* por una creencia de primera persona (de primer orden), entonces todas esas creencias de orden superior tendrán en común *la* propiedad categórica de esa creencia de primera persona (de primer orden). Esa única propiedad categórica explicará, en relación con cada creencia de orden superior, que uno esté dispuesto a asentir a su contenido.

<sup>10</sup> Por simplicidad, hablaré sobre todo de la disposición a asentir.

conciencia fenoménica en términos de pensamientos o creencias de orden superior, como la teoría de Rosenthal (Sebastián 2017, pp. 52–53).<sup>11</sup> De entrada, es obvio que los estados mentales permanentes no son conscientes en sentido fenoménico. Pero supongamos que explicar la conciencia fenoménica de un estado mental ocurrente en términos de una creencia de orden superior es dar una explicación en términos de la disponibilidad de tal estado para el asentimiento así como para guiar el razonamiento y la acción. Puesto que se trata de un estado mental ocurrente, esa disponibilidad debe consistir en la difusión actual del contenido de ese estado en el espacio o memoria global de trabajo (Block 1997, 2007), como sostienen las teorías cognitivas de la conciencia.<sup>12</sup> Pero esta noción funcional de acceso consciente o de acceso cognitivo no puede aplicarse a los estados mentales permanentes. Como estado permanente, mi creencia de que Santiago es la capital de Chile no tiene un contenido que esté actualmente difundido en el espacio de trabajo. Eso explicaría precisamente que no haya conciencia fenoménica en este caso. Por ello mismo, la noción funcionalmente relevante de acceso consciente no puede aplicarse a las creencias de primera persona consideradas como estados permanentes. Sebastián concluye entonces que la noción de conciencia que puede aplicarse en estos casos no es relevante en la discusión.

Empiezo por advertir que el propósito de la teoría de Rosenthal no es sólo explicar la conciencia *fenoménica* en términos de creencias de orden superior. Entre los estados mentales ocurrentes cuya concien-

<sup>11</sup> En realidad, Rosenthal habla de *pensamientos* de orden superior. Puede afirmarse que, como tal, una creencia es un estado permanente que se manifiesta mediante pensamientos o juicios. Por lo tanto, no habría creencias ocurrentes. Sin embargo, he preferido mantener uniforme la terminología. Desde luego, si no hubiera creencias ocurrentes, sería necesario añadir una versión de CONOCERME y de (SC\*) para los pensamientos o juicios. En particular, en correspondencia con (SC\*) para las creencias como estados permanentes, se trataría del principio:

(SC\*+) Si (si  $X$  piensa o juzga que  $\alpha$ , entonces  $X$  sabe que él mismo piensa o juzga que  $\alpha$ ), entonces (si  $X$  piensa o juzga que  $\alpha$ , entonces  $X$  piensa o juzga conscientemente que  $\alpha$ ).

<sup>12</sup> En el planteamiento original de Block, la teoría de Rosenthal corresponde a una explicación, más que en términos de una conciencia de acceso como mera disponibilidad, en términos de una conciencia de *monitoreo* por la que un estado mental ocurrente es objeto de otro estado mental (Block 2001, p. 205). Pero una vez que introdujo (Block 1997, 2007) una noción específica de acceso consciente como difusión *actual* en el espacio global de trabajo, me parece que la conciencia de monitoreo podría considerarse como acceso consciente en este sentido.

cia la teoría debe explicar se encuentran los estados intencionales, como las creencias, y los estados propiamente fenoménicos, como las experiencias visuales (1997, p. 729; 2005 [1986], pp. 23–25). Según Rosenthal, un estado fenoménico tiene propiedades cualitativas (o cualidades mentales) que hacen que haya algo para el sujeto que sea como encontrarse en ese estado cuando ese estado va acompañado por una creencia de orden superior. Pero las creencias ocurrentes no tienen propiedades cualitativas que puedan traducirse en fenomenología: en este sentido, no hay algo para el sujeto que sea como tener una creencia o un pensamiento.<sup>13</sup> Ésta es una de las razones para decidir entre una teoría en términos de creencias de orden superior y una teoría en términos de percepciones de orden superior. En efecto, si se tratase de percepciones, entonces los estados de orden superior tendrían propiedades cualitativas que harían que hubiera algo para el sujeto que fuera como encontrarse en un estado de orden superior cuando ese estado va acompañado a su vez por un estado de un orden más alto aún (en la introspección). Pero el hecho es que no parece haber tal cosa (1997, pp. 740–741; 2005 [1999], pp. 161–162; 2005 [2001], pp. 104–106; 2002, pp. 657–658). En relación con la noción funcional de acceso consciente, resulta que la disponibilidad de una creencia ocurrente consciente para el asentimiento así como para guiar el razonamiento y la acción no tiene asociada ninguna fenomenología. A este respecto no se diferencia de una creencia permanente que esté también disponible. Sin duda, la diferencia está en que en el primer caso, pero no en el segundo, esa disponibilidad consiste en la difusión actual del contenido de la creencia en el espacio de trabajo. Mi punto aquí es sencillamente que no es correcto argumentar que la noción de conciencia que puede aplicarse a las creencias de primera persona, como estados permanentes, no es la noción relevante en la teoría de Rosenthal porque tales creencias no son conscientes en sentido fenoménico.

Ahora bien, Sebastián muestra acertadamente, como yo había establecido en mi artículo (2014, p. 49), que la naturaleza consciente

<sup>13</sup> A pesar de esto, podría plantearse si, para Rosenthal, no hay algo así como una fenomenología *cognitiva*, es decir, una fenomenología no sensorial propia de las creencias, pensamientos o juicios. Parece que no, pues usa el término “fenomenología” y otros relacionados sólo para hablar de los estados sensoriales conscientes. Sin duda, esta posición es contraria a mucha literatura reciente sobre la materia. Incluso se ha argumentado que la teoría de la conciencia de Rosenthal en términos de pensamientos de orden superior implica que los pensamientos conscientes tienen fenomenología cognitiva (Brown y Mandik 2012). Aquí no puedo entrar en esta discusión.

de las creencias y de otros estados intencionales de primera persona es incompatible con ciertos casos freudianos paradigmáticos (2017, p. 54). Por ejemplo, no podría ser que un hombre tuviera la creencia inconsciente de que su pareja le es infiel como resultado de reprimir un deseo hacia otra persona. Pero, es claro que el contrapunto de una creencia inconsciente en sentido freudiano no es una creencia *ocurrente* consciente. Se supone que la atribución de esa creencia inconsciente, como un estado permanente, explicaría ciertos comportamientos del hombre hacia su pareja que no pueden explicarse en términos de la disponibilidad del contenido de esa creencia para el asentimiento o como guía del razonamiento y la acción. Por ello mismo, las creencias inconscientes en sentido freudiano contrastan con las creencias cuyo contenido está permanentemente disponible. La creencia permanente (de primera persona) de que la pareja de uno le es infiel contrastaría, por ejemplo, con la creencia permanente (de primera persona) de que uno vive en Concepción. Desde luego, esta última puede convertirse en una creencia ocurrente consciente que explique ciertos comportamientos, pero la acción por la cual uno se compra un pasaje Concepción-Santiago (y no uno Valdivia-Santiago) se puede explicar, en la mayoría de los casos, en términos de la disponibilidad permanente del contenido de esa creencia. Esto es lo que no sucede con la explicación del comportamiento en términos de una creencia inconsciente. Veámoslo de este otro modo. Si sólo se tratara de que uno tiene una creencia inconsciente en el sentido de que el contenido de esa creencia no está actualmente difundido en el espacio de trabajo, entonces el concepto psicoanalítico de inconsciente no sería relevante en la explicación del comportamiento. Puesto que las creencias permanentes no tienen un contenido que esté actualmente difundido, resulta que todas las creencias permanentes serían inconscientes en sentido freudiano. Ni la creencia de que la pareja de uno le es infiel ni la creencia de que uno vive en Concepción son creencias de acceso consciente, entendido como difusión actual en el espacio de trabajo. Por lo tanto, no hay ninguna diferencia entre ambas creencias respecto de esta noción funcional de acceso consciente que pueda explicar una diferencia en el comportamiento. Está claro entonces que la noción de conciencia que va de la mano con el concepto psicoanalítico de inconsciente debería poder aplicarse a las creencias de primera persona que son estados permanentes.

Téngase en cuenta, sin embargo, que Sebastián concluye que la noción de conciencia que empleo no es la noción de conciencia relevante en la discusión de la teoría de la conciencia en términos de creencias o pensamientos de orden superior ni de la teoría freudiana

del inconsciente (2017, p. 57). Pero acabamos de ver que es poco discutible que ésa sea la noción de conciencia propia de la teoría freudiana del inconsciente. Su artículo debería apuntar más bien a la idea de que, mientras que esa noción de conciencia es la noción relevante en la discusión de los casos freudianos de creencias inconscientes de primera persona, no sería la noción relevante en la discusión de la teoría de Rosenthal.<sup>14</sup> Esto parece implicar que el término “conciencia” (o “estado consciente”) tendría al menos dos significados distintos: por un lado, estaría el significado según el cual los estados ocurrentes conscientes son estados de acceso consciente en el sentido de disponibilidad como difusión actual en el espacio de trabajo y, por otro lado, estaría el significado según el cual los estados permanentes conscientes son estados de acceso consciente en el sentido de disponibilidad que introduce Shoemaker o, como podría decirse, de disponibilidad como *potencial* difusión en el espacio de trabajo (Dehaene y Naccache 2001; Block 1997, p. 160; 2007, pp. 491–493).<sup>15</sup> Es evidente que sería mejor, con vistas a una explicación unificada de la conciencia, si el término “conciencia” tuviera un solo significado, de tal manera que la misma noción tuviera aplicación en la discusión de la teoría freudiana y en la de la teoría de Rosenthal. Ahora argumentaré que en efecto es así utilizando la distinción, para los términos de clase natural, entre fijar la extensión de un término y determinar lo que es esencial para ser un elemento en esa extensión (Kripke 1980, pp. 54–59; Putnam 1975, pp. 223–235).

Consideremos, en primer lugar, cómo podría utilizarse esta distinción en el planteamiento de Rosenthal. Como vimos, entre los estados ocurrentes conscientes distingue estados intencionales y estados fenoménicos. Podría establecerse entonces que, dado que un estado fenoménico es aquel que se caracteriza por haber algo para el sujeto que es como encontrarse en ese estado, nosotros fijamos la extensión del término “conciencia” por referencia a ese conjunto de estados ocurrentes conscientes que son los estados fenoménicos. Pero eso no implica que sólo los estados fenoménicos tengan las propiedades esenciales para ser un elemento en esa extensión. Más aún, la investigación posterior podría revelar que tener fenomenología no es una

<sup>14</sup> El problema es que, después de decir que sólo considerará la noción de acceso consciente como difusión actual en el espacio de trabajo (2017, p. 53), señala que mi propuesta constituye una amenaza para los casos freudianos (p. 54). Pero la idea de que las creencias *ocurrentes* de primera persona son necesariamente conscientes no afectaría en nada a los casos freudianos.

<sup>15</sup> Block 2007 habla de “accesibilidad cognitiva” en sentido estrecho y en sentido amplio.

propiedad esencial para ser un elemento en esa extensión. Es el hecho de que los estados fenoménicos fijen la extensión lo que generaría la ilusión de que tener fenomenología es la propiedad esencial que estamos buscando. En particular, podría descubrirse que la propiedad esencial para ser un estado ocurrente consciente es tener una creencia de orden superior acerca de ese estado (cuando esa creencia no es el resultado de un proceso inferencial). En relación con la noción funcional de acceso consciente, diríamos que la propiedad esencial consiste en la disponibilidad del contenido de ese estado como difusión actual en el espacio de trabajo. Desde luego, todos los estados fenoménicos tienen esa propiedad esencial según una teoría de la conciencia fenoménica en términos de creencias de orden superior (o cualquier otra teoría cognitiva), pero ahora resulta que, según Rosenthal, la extensión del término “conciencia” también estaría constituida por los estados intencionales que, carentes de fenomenología, se acompañan de una creencia de orden superior de tal modo que su contenido está actualmente difundido en el espacio de trabajo. Así que, a pesar de fijar la extensión del término “conciencia” por referencia a los estados fenoménicos, esa extensión comprende tanto los estados fenoménicos como los estados intencionales que tienen la propiedad esencial para ser un estado ocurrente consciente.<sup>16</sup>

Pues bien, me parece que los resultados de la investigación psicoanalítica,<sup>17</sup> y los de la investigación reciente sobre los prejuicios o sesgos implícitos (Brownstein y Saul 2016),<sup>18</sup> muestran que lo

<sup>16</sup> Al utilizar la distinción entre fijar la extensión y determinar lo que es esencial para ser un elemento en esa extensión me he inspirado en el uso que le da Rosenthal en otro contexto, donde sostiene que fijamos la extensión del término “mental” (o “estado mental”) por referencia a los estados conscientes. Pero eso no implica que sólo los estados conscientes tengan las propiedades esenciales para ser un elemento en esa extensión. Esa extensión comprende los estados que tienen propiedades cualitativas, como es propio de los estados sensoriales, y los estados que tienen intencionalidad. Pero adviértase que ser un estado sensorial no es ser un estado fenoménicamente consciente, pues las propiedades cualitativas sólo se traducen en fenomenología cuando el estado sensorial va acompañado por una creencia de orden superior; es decir, hay estados sensoriales inconscientes. Por lo tanto, la conciencia no es una propiedad esencial para ser un elemento en la extensión del término “mental”. Por otra parte, tampoco los estados ocurrentes intencionales son necesariamente conscientes (1997, pp. 733–734; 2005, pp. 177–178; 2005 [1991], p. 138; 2005 [1986], pp. 44–45).

<sup>17</sup> En efecto, no estoy cuestionando el poder explicativo del concepto psicoanalítico de inconsciente ni, en particular, que podamos tener creencias permanentes de naturaleza inconsciente. En mi artículo usé el ejemplo de creer inconscientemente que comer es obsceno (2014, pp. 43–44).

<sup>18</sup> Se trata de actitudes racistas, sexistas o discriminatorias de algún tipo que negamos sinceramente tener, lo que significa que su contenido no está disponible

que es esencial para ser un elemento en la extensión del término “conciencia” es que un estado tenga una propiedad por la que su contenido está disponible tanto para el asentimiento como para guiar el razonamiento y la acción, *esté o no esté actualmente difundido en el espacio de trabajo*. Esa propiedad esencial será una propiedad categórica que explica, por un lado, que un estado tenga el potencial para difundir su contenido en el espacio de trabajo (propiedad disposicional) y, por el otro, que el contenido de un estado esté actualmente difundido en el espacio de trabajo (propiedad categórica a su vez) cuando se dan ciertas condiciones adicionales. Frente a Rosenthal, de aquí se sigue, obviamente, que la extensión del término comprende estados ocurrentes y estados permanentes. Pero no veo que pueda negarse este resultado so pena de prejuzgar la cuestión. Supongamos que la extensión del término “conciencia” se fija por referencia a los estados fenoménicos. Como planteamos, es el hecho de que los estados fenoménicos fijen la extensión lo que generaría la ilusión de que tener fenomenología es la propiedad esencial para ser un elemento en esa extensión. Pero los estados fenoménicos son estados mentales ocurrentes. Podría plantearse entonces que es el hecho de que los estados fenoménicos fijen la extensión lo que también genera la ilusión de que sólo los estados mentales ocurrentes pueden ser elementos en esa extensión. Pues, podría ser que, como sostengo, la investigación haya revelado que la propiedad esencial para ser un estado consciente, cuya naturaleza aún puede estar por descubrir, no discrimina entre estados mentales ocurrentes y estados mentales permanentes. No veo por qué deberíamos aceptar que tener fenomenología no es una propiedad esencial para ser un estado consciente y que, sin embargo, una investigación no pueda revelar que ciertos estados permanentes son estados conscientes (en contraste con los estados permanentes de carácter inconsciente). Téngase en cuenta, por lo demás, que esta propuesta no es criticable sobre la base de que la extensión del término “conciencia” ahora dejaría fuera ciertos estados conscientes que cualquier propuesta debería incluir. Si la conciencia fenoménica puede explicarse en términos de difusión actual en el espacio de trabajo, entonces los estados fenoménicos tienen la propiedad esencial en cuestión. Eso sucede también con los

para el asentimiento. Pero la evidencia experimental muestra que tenemos tales actitudes, que, por lo tanto, deben darse en un nivel inconsciente; téngase en cuenta que “implícito” aquí no se usa en el sentido en el que, en cierta literatura sobre las creencias como estados permanentes, se habla de creencias implícitas para referirse a las creencias que, sin ser consideradas explícitamente, tienen un contenido que está disponible para el asentimiento (Gertler 2011).

estados intencionales cuyo contenido está actualmente difundido en el espacio de trabajo.

Aunque ésta es la posición que defiendo, alguien que insista en la diferencia entre conciencia fenoménica y acceso consciente podría argumentar del siguiente modo. Dejando de lado la conciencia fenoménica de los estados sensoriales conscientes, se trataría de usar la distinción entre fijar la extensión del término “acceso consciente” y determinar lo que es esencial para ser un elemento en esa extensión. Adviértase primero que, como suele aceptarse, los estados fenoménicos son estados de acceso consciente (Tye 2003, pp. 122–125) incluso si la conciencia fenoménica no puede explicarse en términos de difusión actual en el espacio de trabajo o de alguna noción relacionada (Chalmers 1996, p. 26; 1997). La idea entonces es que fijamos la extensión del término “acceso consciente” por referencia a los estados mentales ocurrientes, tanto fenoménicos como intencionales, cuyo contenido está actualmente difundido en el espacio de trabajo. Eso podría generar la ilusión de que sólo los estados mentales ocurrientes pueden ser elementos en esa extensión, pero la investigación considerada líneas antes mostraría que lo que es esencial para ser un elemento en esa extensión es que un estado mental tenga cierta propiedad categórica por la que su contenido está disponible tanto para el asentimiento como para guiar el razonamiento y la acción, esté o no esté actualmente difundido en el espacio de trabajo. Esto implica que los estados mentales permanentes, como las creencias permanentes, pueden tener la propiedad esencial para ser un estado de acceso consciente. El término “acceso consciente” no tendría, por lo tanto, dos significados sino uno solo cuya extensión comprende tanto los estados ocurrientes (fenoménicos e intencionales) como los estados permanentes que tienen la misma propiedad categórica esencial.

En mi artículo argumenté que las creencias de primera persona son necesariamente conscientes sin discriminar entre tales creencias como estados ocurrientes y como estados permanentes. El análisis anterior apoya la idea de que la noción de conciencia empleada en el principio (SC\*) no sólo es la noción relevante para establecer que las creencias de primera persona, como estados permanentes, son necesariamente conscientes, sino que también es la noción relevante para establecer que las creencias de primera persona, como estados ocurrientes, son necesariamente conscientes. Pero esta última aplicación es justo lo que se necesita en la discusión de la teoría de Rosenthal que abordé en mi artículo.

## III

Por último discutiré la crítica de Sebastián según la cual no ofrezco una justificación satisfactoria del principio (SC\*), o bien simplemente prejuizo la cuestión (2017, pp. 54–57). Mi argumentación partía de examinar un caso en el que uno tiene una creencia inconsciente, la creencia de que comer es obsceno, a pesar de llegar a saber, a través de la terapia psicoanalítica, que uno tiene esa creencia. Que la creencia sigue siendo inconsciente se pone de manifiesto en el hecho de que uno aún no estaría dispuesto a asentir al contenido de esa creencia. Desde luego, uno estaría dispuesto a asentir a la proposición de que uno mismo cree que comer es obsceno (que es el contenido de una creencia de segundo orden) después de la revelación psicoanalítica, pero parece que, al menos en una primera fase, uno aún no estaría dispuesto a asentir a la proposición de que comer es obsceno (que es el contenido de la creencia de primer orden).<sup>19</sup> La explicación de este hecho sería que la creencia de que comer es obsceno tiene una dinámica cognitiva o rol funcional, en la vida mental de un sujeto, que es independiente de cualquier conocimiento, o creencia de segundo orden, que uno pudiese adquirir después. Por ello mismo, la creencia de segundo orden también tendrá una dinámica cognitiva independiente en relación con la creencia de que comer es obsceno. En ese contexto, definí una creencia inconsciente en términos de la existencia de una dinámica cognitiva independiente entre cierta creencia y el conocimiento (o creencia) de que uno mismo tiene esa creencia. Por lo tanto, una creencia consciente se definía por no existir esa dinámica cognitiva independiente (2014, pp. 45–46). La justificación del principio (SC\*) procedía entonces del siguiente modo: si existe una conexión necesaria entre tener cierta creencia y saber (o creer) que uno tiene esa creencia (antecedente de (SC\*)), entonces no hay una dinámica cognitiva independiente, lo que, por la definición de creencia consciente, implica que esa creencia es consciente (consecuente de (SC\*)) (2014, pp. 46–47).

Consideraré brevemente algunos aspectos de esta argumentación. En primer lugar, parece poco discutible que si existe una conexión necesaria entre tener cierta creencia (de primer orden) y tener una creencia de segundo orden acerca de ella, ambas creencias no tienen

<sup>19</sup> En realidad, sólo introduje esta consideración en los términos específicos de una disposición a la expresión *verbal*: uno aún no estaría dispuesto a expresar de manera espontánea su creencia (de primer orden) mediante una emisión sincera de “Comer es obsceno” (2014, p. 44).

una dinámica cognitiva independiente en la vida mental de un sujeto. Más aún, si esa conexión necesaria se entiende en términos de la noción de constitución que introduce Shoemaker, de tal modo que tener la creencia (de primer orden) constituye el hecho de tener la creencia de segundo orden, esa conclusión parece trivial (Shoemaker 2009, p. 42). Esto significaría que tendríamos una justificación inmediata del principio (SC\*) en términos de la definición de creencia consciente que usé en mi artículo. Ahora bien, por lo que he venido diciendo me parece mejor reemplazar la definición en términos de dinámica cognitiva dependiente por la definición en términos de disposición al asentimiento.<sup>20</sup> Esto significa que la justificación del principio (SC\*) ya no es tan inmediata. Ahora es necesario justificar el paso de una dinámica cognitiva dependiente entre cierta creencia (de primer orden) y una creencia de segundo orden a la conclusión de que esa creencia es consciente en el sentido de estar dispuesto a asentir a su contenido. Aquí es donde Sebastián cuestiona, hasta cierto punto, la justificación del principio (SC\*) (2017, pp. 54–55).<sup>21</sup>

Empezaré por señalar que no se trata de que si *cualesquiera* creencias tienen una dinámica cognitiva dependiente, entonces (algunas de) esas creencias son conscientes en el sentido de que uno está dispuesto a asentir a su contenido. Queda claro que esto no es así. En efecto, supongamos que si alguien tiene la creencia de que  $\alpha \rightarrow \beta$  y la creencia de que  $\alpha$ , entonces tiene la creencia de que  $\beta$ . Supongamos también que, en ese caso, tener las creencias de que  $\alpha \rightarrow \beta$

<sup>20</sup> Desde el planteamiento introducido en la sección anterior, debería hablarse más bien de la definición en términos de la propiedad categórica por la que uno está dispuesto al asentimiento, pero por simplicidad me expresaré de ese modo, como he venido haciendo en este artículo. Ahora bien, me parece plausible que esa propiedad categórica sea, en primer lugar, la propiedad por la que existe una dinámica cognitiva dependiente entre cierta creencia y una creencia de orden superior, lo cual implicaría a su vez que uno esté dispuesto a asentir a su contenido. A este respecto, no se trataría tanto de reemplazar la definición cuanto de completarla con una referencia a la disposición al asentimiento.

<sup>21</sup> En realidad, Sebastián sólo cuestiona la justificación de este paso para un sentido del término “conciencia”. Desde luego, él sostiene que no hay tensión conceptual en la idea de que una creencia sea inconsciente a pesar de que exista tal dinámica cognitiva dependiente, pero sólo está pensando en una creencia inconsciente en el sentido de que su contenido no esté actualmente difundido en el espacio de trabajo (o no tenga conciencia fenoménica). En efecto, no hay tensión conceptual a este respecto. Eso no excluye que esa creencia deba ser consciente en el sentido de tener la propiedad categórica por la que su contenido está disponible para el sujeto, esté o no esté actualmente difundido en el espacio de trabajo. Por lo tanto, su cuestionamiento es más bien, otra vez, un cuestionamiento de la noción de conciencia en juego.

y de que  $\alpha$  (más cierto grado de racionalidad) constituyen el hecho de tener la creencia de que  $\beta$ . Por lo que hemos venido diciendo, se sigue trivialmente que existe una dinámica cognitiva dependiente entre esas creencias. Pero, como es obvio, no tendría sentido concluir que (algunas de) esas creencias son conscientes en el sentido de estar dispuesto a asentir a su contenido. Es perfectamente posible que la dinámica cognitiva de esas creencias se desarrolle como parte de la vida mental *inconsciente* del sujeto. Ciertamente, cualquier cambio en la vida mental del sujeto que se produzca por las creencias de que  $\alpha \rightarrow \beta$  y de que  $\alpha$  contará como un cambio que se produce también por la creencia de que  $\beta$ ; sin embargo, la cuestión es que no hay nada por lo que esa dinámica cognitiva dependiente deba tener efectos en el nivel consciente.

El caso es diferente cuando se considera la justificación del principio (SC\*). Resulta que en este caso se establecerá una dinámica cognitiva dependiente entre la creencia de que  $\alpha$  y una creencia, la creencia de segundo orden, cuyo rol funcional es justamente *informar de la posesión de la creencia de que  $\alpha$* . Desde luego, una creencia de segundo orden puede informar de la posesión de una creencia (de primer orden) sin que esta última sea consciente en el sentido de estar dispuesto a asentir a su contenido. Eso es lo que sucede en el caso de alguien que cree que comer es obsceno pero que sólo a través de la terapia psicoanalítica llega a tener la creencia de que él mismo cree que comer es obsceno. Pero en el caso en consideración existe una dinámica cognitiva dependiente en virtud de la cual cualquier cambio en la vida mental del sujeto que se produzca por la creencia de que  $\alpha$  contará como un cambio que se produce también por la creencia de que uno mismo cree que  $\alpha$ . En los términos informacionales que estoy empleando, esto significa, en particular, que, a diferencia de lo que sucede en el caso psicoanalítico, no puede transmitirse la información de que  $\alpha$  a los distintos sistemas “consumidores” de información —que incluyen los mecanismos motores— sin transmitirse también la información de que uno mismo cree que  $\alpha$ . Ahora bien, parece que informar que uno mismo cree que  $\alpha$  *al* informar que  $\alpha$  es poner la información de que  $\alpha$  a disposición del sujeto. Lo que muestra el caso psicoanalítico es precisamente que la (pseudo)información de que comer es obsceno no está disponible para el sujeto justo cuando esa información puede transmitirse a los sistemas “consumidores” de información sin transmitirse por ello la información de que uno mismo cree eso. No veo cómo se puede explicar el caso psicoanalítico y sostener a la vez que la información de que

$\alpha$  podría no estar disponible para el sujeto a pesar de que ambos elementos de información no pueden separarse.

Téngase en cuenta, además, que no forma parte de la explicación que, cuando existe una dinámica cognitiva dependiente, la creencia de segundo orden es consciente en el sentido de estar dispuesto a asentir a su contenido. En el caso psicoanalítico no es relevante si la creencia de segundo orden es o no es consciente. De hecho, a través de la terapia psicoanalítica el sujeto llega a tener la creencia consciente de que él mismo cree que comer es obsceno. En efecto, adquirió esa creencia de segundo orden mediante un proceso consciente reflexivo. Por lo tanto, la dinámica cognitiva independiente es lo único que explica que la creencia de que comer es obsceno no sea consciente. Entonces, aunque una creencia de segundo orden no fuera consciente, haría la tarea de volver consciente la creencia de que  $\alpha$  como resultado de una dinámica cognitiva dependiente entre ambas creencias.<sup>22</sup>

Termino esta sección argumentando que, sea o no sea satisfactoria la justificación del principio (SC\*) que ofrezco, no hay ningún sentido relevante en el que se prejuzgue la cuestión, como sostiene Sebastián (2017, pp. 55–57). No prejuzgo la cuestión en el sentido de que, en la derivación de (SC\*) a partir de su antecedente, no trato de pasar directamente de la afirmación de que uno sabe que uno mismo cree que  $\alpha$  a la afirmación de que la creencia de que  $\alpha$  está disponible para el sujeto, que es precisamente lo que está en necesidad de justificación. Este punto es más claro aquí que en el artículo anterior, pues he asumido explícitamente que, en este caso, saber consiste en creer. Pero incluso así puede argumentarse que se prejuzgaría la cuestión si tratara de pasar directamente de la afirmación de que uno cree que uno mismo cree que  $\alpha$  a la afirmación de que la creencia de que  $\alpha$  está disponible para el sujeto. Sin embargo, no he propuesto en absoluto esa transición directa (que en la reconstrucción que presento a continuación sería pasar directamente de (iv) a (vi)). Eso significaría

<sup>22</sup> Debe distinguirse con claridad este planteamiento de la teoría de Rosenthal según la cual una creencia es consciente cuando va acompañada por una creencia *inconsciente* de orden superior. En la propuesta de Rosenthal una creencia de orden superior no vuelve consciente otra creencia como resultado de una dinámica cognitiva dependiente entre ellas. Otra cosa es que una dinámica cognitiva dependiente sea la consecuencia de que la creencia de orden superior vuelva consciente esa creencia. En cualquier caso, para Rosenthal no existe una relación de constitución entre tener una creencia y tener una creencia de orden superior acerca de ella, puesto que se trata de “existencias separadas”. Este punto es esencial en la atribución de una regresión dañina a la teoría de Rosenthal tanto en mi crítica (2014, pp. 49–51), como en la que Sebastián (2018) formuló recientemente.

olvidarse de que la justificación de (SC\*) se da por la existencia de una dinámica cognitiva dependiente entre ambas creencias. En vez de repetir mis palabras concluyo con una reconstrucción de la justificación del principio (SC\*):

- (i) Si  $X$  cree que  $\alpha$ , entonces  $X$  sabe que él mismo cree que  $\alpha$  (asunción).
  - (ii)  $X$  cree que  $\alpha$  (asunción).
  - (iii)  $X$  sabe que él mismo cree que  $\alpha$  (de (i) y (ii)).
  - (iv)  $X$  cree que él mismo cree que  $\alpha$  (de (iii): saber consiste en creer).
  - (v) La creencia de que  $\alpha$  y la creencia de que uno mismo cree que  $\alpha$  tienen una dinámica cognitiva dependiente (de (i)–(iv)).
  - (vi) La creencia de que  $\alpha$  está disponible para  $X$  (de (v)).
  - (vii)  $X$  cree conscientemente que  $\alpha$  (de (vi) y la definición de creencia consciente).
  - (viii) Si  $X$  cree que  $\alpha$ , entonces  $X$  cree conscientemente que  $\alpha$  (de (ii)–(vii)).
- 
- (ix) Si (si  $X$  cree que  $\alpha$ , entonces  $X$  sabe que él mismo cree que  $\alpha$ ), entonces (si  $X$  cree que  $\alpha$ , entonces  $X$  cree conscientemente que  $\alpha$ ) (de (i)–(viii)).

En mi artículo (2014) argumenté, a través de CONOCERME, que en el caso de las creencias con un contenido de primera persona puede descargarse (i), lo que implica, a través de (SC\*), que las creencias de primera persona son necesariamente conscientes. En este artículo creo haber contribuido a mostrar que ese argumento es sólido.<sup>23</sup>

## BIBLIOGRAFÍA

Audi, R., 1994, “Dispositional Beliefs and Dispositions to Believe”, *Noûs*, vol. 28, no. 4, pp. 419–434.

<sup>23</sup> Para la realización de este trabajo, agradezco el apoyo otorgado por el Gobierno de Chile mediante el Proyecto Fondecyt no. 1140395. Con Miguel Ángel Sebastián tengo una deuda especial e impagable por haberme dado la oportunidad de revisar y fortalecer los fundamentos de mi argumentación. Expreso también gratitud a los comentarios de un árbitro de la revista.

- Block, N., 2007, “Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience”, *Behavioral and Brain Sciences*, vol. 30, no. 4, pp. 481–499 y 530–548.
- , 2001, “Paradox and Cross Purposes in Recent Work on Consciousness”, *Cognition*, vol. 79, nos. 1–2, pp. 197–219.
- , 1997, “Biology versus Computation in the Study of Consciousness”, *Behavioral and Brain Sciences*, vol. 20, no. 1, pp. 159–166.
- Brown, R. y P. Mandik, 2012, “On Whether the Higher-Order Thought Theory of Consciousness Entails Cognitive Phenomenology, or: What Is It Like to Think that One Thinks that P?”, *Philosophical Topics*, vol. 40, no. 2, pp. 1–12.
- Brownstein, M. y J. Saul (comps.), 2016, *Implicit Bias and Philosophy*, Oxford University Press, Oxford.
- Chalmers, D., 1997, “Availability: The Cognitive Basis of Experience”, *Behavioral and Brain Sciences*, vol. 20, no. 1, pp. 148–149.
- , 1996, *The Conscious Mind*, Oxford University Press, Oxford.
- Dehaene, S. y I. Naccache, 2001, “Towards Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework”, *Cognition*, vol. 79, nos. 1–2, pp. 1–37.
- Gertler, B., 2011, “Self-Knowledge and the Transparency of Belief”, en A. Hatzimoysis (comp.), *Self-Knowledge*, Oxford University Press, Oxford, pp. 125–145.
- Kripke, S., 1980, *Naming and Necessity*, Harvard University Press, Cambridge, Mass. [Versión en castellano: *El nombrar y la necesidad*, 2a. ed. rev., trad. Margarita M. Valdés, Instituto de Investigaciones Filosóficas-UNAM, 2005.]
- Lycan, W.G., 1986, “Tacit Belief”, en R.J. Bogdan (comp.), *Belief*, Oxford University Press, Oxford, pp. 61–82.
- Putnam, H., 1975, “The Meaning of ‘Meaning’”, en *Mind, Language, and Reality*, Cambridge University Press, Cambridge, pp. 215–271.
- Rosenthal, D.M., 2005, *Consciousness and Mind*, Oxford University Press, Oxford.
- , 2002, “How Many Kinds of Consciousness?”, *Consciousness and Cognition*, vol. 11, pp. 653–665.
- , 2001, “Introspection and Self-Interpretation”, *Philosophical Topics*, vol. 28, no. 2, pp. 201–233; reimpresso en Rosenthal 2005, pp. 103–131 (a cuya paginación remito en este texto).
- , 1999, “Sensory Quality and the Relocation Story”, *Philosophical Topics*, vol. 26, no. 1–2, pp. 321–350; reimpresso en Rosenthal 2005, pp. 149–174 (a cuya paginación remito en este texto).
- , 1997, “A Theory of Consciousness”, en N. Block, O. Flanagan y G. Güzeldere (comps.), *The Nature of Consciousness: Philosophical Debates*, The MIT Press, Cambridge, Mass., pp. 729–753.
- , 1991, “The Independence of Consciousness and Sensory Quality”, en E. Villanueva (comp.), *Consciousness: Philosophical Issues I*, Ridge-

- view Publishing Company, Atascadero, California, pp. 15–36; reimpresso en Rosenthal 2005, pp. 135–148 (a cuya paginación remito en este texto).
- Rosenthal, D.M., 1986, “Two Concepts of Consciousness”, *Philosophical Studies*, vol. 49, no. 3, pp. 329–349; reimpresso en Rosenthal 2005, pp. 21–45 (a cuya paginación remito en este texto).
- Sebastián, M.A., 2018, “Drop It Like It’s HOT: a Vicious Regress for Higher-Order Thought Theories”, *Philosophical Studies*, DOI: 10.1007/s11098-018-1078-7.
- , 2017, “Hay muchas cosas que creo de mí mismo (consciente e inconscientemente) sin saber que las creo”, *Crítica. Revista Hispanoamericana de Filosofía*, vol. 49, no. 146, pp. 37–60.
- Shoemaker, S., 2009, “Self-Intimation and Second-Order Belief”, *Erkenntnis*, vol. 71, no. 1, pp. 35–51.
- Tye, M., 2003, *Consciousness and Persons*, The MIT Press, Cambridge, Mass.
- Vidal, J., 2014, “Creencia de primera persona, conciencia y la paradoja de Eroom”, *Crítica. Revista Hispanoamericana de Filosofía*, vol. 46, no. 138, pp. 37–64.

*Recibido el 22 de abril de 2018; revisado el 23 de agosto de 2018; aceptado el 27 de agosto de 2018.*