teorema
Vol. XXX/2, 2011, pp. 129-144
ISSN: 0210-1602
[BIBLID 0210-1602 (2011) 30:2; pp. 129-144]

ARTÍCULOS

Pensamientos-matriz, ¿una solución compatibilista a los casos de traslado? [1]

Silvia Andrés Balsera

ABSTRACT

The purpose of this paper is to examine some arguments offered by P. Ludlow against the incompatibilist challenge represented by the switching cases. His strategy consist in modifying the relationships between thoughts and their contents in order to consider each thought as a matrix of several contents. It will be argued that Ludlow can avoid Boghossian's conclusion at the expense of defending a non externalist account of our thoughts.

KEYWORDS: *Externalism, Internalism, Compatibilism, Switching Cases, Mental Content.*

RESUMEN

El propósito de este artículo es examinar algunos de los argumentos ofrecidos por P. Ludlow para combatir el desafío incompatibilista de los casos de traslado. La estrategia de este autor consiste en modificar las relaciones entre los pensamientos y sus contenidos a fin de que, en lo sucesivo, un pensamiento sea considerado una matriz de varios contenidos. Se defenderá que Ludlow consigue evitar la conclusión de Boghossian sólo a costa de mantener una concepción no externista de nuestros pensamientos.

PALABRAS CLAVE: *externismo, internismo, compatibilismo, casos de traslado, contenido mental*

I. INTRODUCCIÓN

El debate sobre la posibilidad de conjugar el externismo psicológico con el acceso privilegiado a los propios contenidos mentales ha ocupado un

lugar muy destacado en las preocupaciones de los filósofos de la mente en los últimos años. De entre las distintas estrategias que se han diseñado para precisar la clase de incompatibilidad que se presume entre ambas tesis, el artículo presentado por P. Boghossian en (1989) constituye una de las propuestas más interesantes al respecto. Nuestro objetivo en este trabajo es hacernos cargo de los argumentos que P. Ludlow desarrolla en (1995) y (2004) con el fin de neutralizar el argumento de Boghossian y defender, con ello, que es posible mantener, simultáneamente, el externismo y el acceso privilegiado. Se concluirá que la propuesta de Ludlow, pese a que se presenta a sí misma como favorable al externismo compatibilista, no es, en realidad, una genuina propuesta externista e, incluso, atentaría contra rasgos básicos del acceso privilegiado.

II. El debate sobre el incompatibilismo y los casos de traslado

Antes de introducir el debate sobre la compatibilidad entre el externismo y el acceso privilegiado –debate que constituye el contexto de los casos de traslado y de las réplicas de Ludlow a los mismos–, resulta conveniente describir la problemática ante la que se sitúan el internismo y el externismo como posiciones enfrentadas. Dicha cuestión concierne a las condiciones de identidad de nuestros pensamientos. Por "pensamiento" entenderemos aquellas actitudes mentales (creencias, deseos, temores, esperanzas, etc.) dotadas de contenido. El contenido de un pensamiento es aquello que especificamos haciendo uso de la denominada "cláusula-que", es decir, la expresión lingüística que comienza por el nexo "que" en nuestros informes psicológicos tales como "Ana piensa que el oro es valioso". Usualmente se considera, además, que los conceptos involucrados en un contenido determinan la identidad de éste y que, a su vez, el contenido de un pensamiento determina la identidad de dicho pensamiento.

Pues bien, retomando la cuestión principal, el desacuerdo entre internistas y externistas se explica de la siguiente manera. Mientras que los primeros consideran que los elementos que fijan la identidad de un pensamiento son todos ellos factores internos a la mente de los individuos, esto es, factores con respecto a los cuales los individuos poseen pleno acceso epistémico, los externistas mantienen que en ocasiones es necesario contar con elementos que el individuo podría no conocer para identificar uno de sus pensamientos, ya sean elementos sociales (como los usos lingüísticos comunitariamente asociados a un término) o naturales (como los géneros naturales presentes en el entorno de los sujetos).

La dificultad más importante que presenta el externismo es, según se alega, su difícil compatibilidad con el acceso privilegiado que un individuo posee con respecto a sus propios pensamientos. Se piensa que si la identidad del pensamiento de que P que un sujeto mantiene en un momento dado depende de condiciones que pueden residir más allá de su conocimiento, parece

que dicho sujeto no será capaz de saber con la inmediatez y fiabilidad requeridas que el pensamiento que mantiene es el pensamiento de que P, más bien que otro distinto (de que Q o de que R).

Uno de los planteamientos más sugerentes que se han elaborado para explotar la presunta incompatibilidad entre la individuación externista de nuestros pensamientos y el privilegio epistémico del auto-conocimiento es el que ponen en juego los "casos de traslado inadvertido". Aclararemos que aunque la formulación inicial de los mismos fue ofrecida en (1988) por Burge, gran valedor del externismo en su variante social, aquí sólo tomaremos en consideración la argumentación de Boghossian en (1989).

Para empezar, diremos que los casos de traslado se originaron como una prolongación del experimento mental de la Tierra Gemela ideado por Putnam en (1975)[2]. En ellos se nos pide suponer que $Óscar_1$, el habitante de la Tierra, es trasladado mientras duerme a la Tierra Gemela[3]. Puede pensarse que este individuo no advertirá que dicho traslado ha tenido lugar, dado que, si nos atenemos a los términos del experimento original de Putnam, la única diferencia que existe entre su entorno de origen y el nuevo entorno de acogida escapa a su conocimiento. En particular, él no sabe que el líquido denominado "agua" en un entorno y otro es esencialmente distinto, es decir, que en la Tierra ese líquido se compone de moléculas de H_2O mientras que en la Tierra Gemela se compone de moléculas de XYZ, aun cuando una y otra sustancia resultan ser cualitativamente idénticas (ambas carecen de color, de olor, de sabor, etc.). Según continúa el relato, $Óscar_1$ debe permanecer en la Tierra Gemela el tiempo suficiente para que, de acuerdo con el externismo, las condiciones de su nuevo entorno figuren entre los determinantes de sus contenidos mentales. De este modo, tras admitir que los líquidos a los que nuestro protagonista aplica el término "agua" en un entorno y otro son esencialmente distintos, el *dictum* externista establece que los pensamientos que $Óscar_1$ expresaba *entonces*, en la Tierra, y *ahora*, en la Tierra Gemela, mediante el uso de dicho término son pensamientos diferentes. Lo interesante es advertir que el cambio que se produce en sus propias creencias le resultará tan inadvertido a $Óscar_1$ como el traslado mismo. Dado el carácter externo a la mente de las condiciones de identidad de sus pensamientos, podría producirse un cambio en estos últimos –como el que se deriva del traslado– sin que el individuo tenga noticia de dicha variación [Burge (1988), p. 652]. La posibilidad de una alteración inconsciente de nuestros pensamientos es el punto de partida del cuestionamiento de Boghossian, cuyo fin es determinar de qué manera esta ignorancia conlleva atentar contra el acceso privilegiado a nuestros contenidos mentales. En concreto, el argumento que este autor plantea en (1989) es el siguiente:

(M1) Si un individuo S no olvida nada, lo que sabe en t_1 debe saberlo en t_2.

(M2) En el caso de traslado S no ha olvidado nada.

(M3) Dado el externismo, S no sabe que *P* en t_2.

Por tanto, S no sabe que *P* en t_1.

La idea general que subyace al argumento es que la posibilidad de una variación inadvertida en los contenidos mentales de un sujeto que se deriva del externismo puede generalizarse hasta alcanzar un escepticismo global sobre los propios pensamientos. Si la primera premisa de este argumento no es más que una "perogrullada" sobre el funcionamiento de la memoria [Boghossian (1989), p. 23], la segunda premisa se limita a establecer que el cambio en los pensamientos de $Óscar_1$ no puede explicarse en términos de olvido. La tercera premisa recoge la consecuencia externista de que al producirse un cambio en los factores del entorno que contribuyen a individuar los pensamientos relevantes de nuestro protagonista, $Óscar_1$ no puede mantener en su nuevo entorno los pensamientos que mantenía en su entorno de origen. Por todo ello, si este individuo no ha olvidado nada entre t_1 y t_2; si esto implica, además, que lo que sabía en t_1 también debería saberlo en t_2 y aceptamos, como nos muestra el caso de traslado, que en t_2 $Óscar_1$ no sabe qué pensaba en t_1, entonces no queda más remedio que concluir que, en realidad, $Óscar_1$ no conocía su pensamiento *P* en t_1.

Boghossian, de este modo, se apoya en la falta de habilidad para advertir un cambio entre los contenidos mentales pasados y los presentes que se desprende del examen externista de los casos de traslado para mostrar que el externismo psicológico socava el carácter privilegiado del conocimiento de nuestros propios pensamientos.

III. Las propuestas de Ludlow frente a los casos de traslado

No obstante, la opinión más generalizada entre los partidarios del externismo es que el argumento de Boghossian no tiene nada de definitivo. En esta sección vamos a examinar las respuestas que Ludlow ofrece en dos artículos, (1995) y (2004), con el objetivo de contrarrestar la fuerza del argumento incompatibilista. La relación entre estos dos trabajos es clara, puesto que en (2004) Ludlow sostiene que su propuesta de (1995) era correcta aunque no había ido lo suficientemente lejos [Ludlow (2004), p. 421]. Esto nos permite decir que la solución que se plantea en este último trabajo sigue la misma línea argumental que (1995) aunque radicaliza sus consecuencias.

Para empezar diremos que la estrategia de Ludlow en (1995) se apoya en una auto-denominada "explicación externista de los contenidos de la memoria" combinada con el modelo de auto-conocimiento que Burge desarrolla en (1988). Esta explicación pretende demostrar que la primera premisa del

argumento de Boghossian, (M1), es una premisa falsa. De este modo, el modelo de Burge, de gran éxito entre los partidarios del compatibilismo[4], permite asegurar la infalibilidad en el conocimiento de los pensamientos presentes. En primer lugar, se propone distinguir en el conocimiento de nuestros pensamientos dos tipos de juicios, a saber, los juicios de primer orden, es decir, los pensamientos que conocemos reflexivamente, y los juicios de segundo orden, que tienen como contenido los juicios de primer orden. Pongamos como ejemplo el siguiente informe:

(1) Pienso que creo que el agua moja

En (1) puede distinguirse un juicio de primer orden, "creo que el agua moja", y un juicio de segundo orden que versa sobre el primero: "*Pienso que* creo que el agua moja". En segundo lugar, se sostiene que el juicio de primer orden queda *incluido* en el juicio de segundo orden como su contenido, lo cual supone que los conceptos que figuran en el juicio de primer orden son también los conceptos que figuran en el juicio de segundo orden. De esta manera, tras producirse un cambio en las condiciones externas que supone un cambio en la identidad de los pensamientos de primer orden, aquellos juicios de segundo orden mediante los cuales conocemos los primeros también verán alterada su identidad. Por ello, si Óscar$_1$ fuera transportado a la Tierra Gemela, entonces el pensamiento reflexivo que formularía acerca de la creencia que pasa a expresar mediante una proferencia como "el agua es húmeda" no contendría ya el concepto de agua, como ocurría en la Tierra, sino otro distinto, el concepto de agua-gemela o, como lo denominaremos a partir de ahora, el concepto de agua$_2$.

Una vez explicado en lo esencial el modelo de Burge, expondremos el contenido esencial de la "explicación externista de la memoria" que constituye el otro puntal del argumento de Ludlow en (1995):

(EEM) Las condiciones de identidad (externas en parte) de los contenidos que figuran en los actos de rememoración son aquellas que pertenecen al entorno en que se lleva a cabo dicho acto.

Aplicaremos este principio al caso de traslado para ver cómo funciona. Si en la Tierra, en el tiempo t_1, Óscar$_1$ pensaba *que el agua moja* (puesto que, por hipótesis, ese es el contenido individuado por las condiciones terrestres), ¿qué es lo que pensará en t_2, tras sufrir el traslado, cuando diga "en mi infancia pensaba que el agua moja"? Atendiendo a (EEM), lo que Óscar$_1$ piensa en t_2 es *que el agua$_2$ moja*. Serán las condiciones del nuevo entorno las que individuarán el contenido rememorado en t_2 y, haciendo uso del modelo basado en la inclusión, diremos que en ese instante del tiempo este individuo conoce su contenido de *que el agua$_2$ moja* porque sus juicios reflexivos también que-

dan individuados por las nuevas condiciones externas. Ludlow usa este resultado para impedir el establecimiento del condicional que figura en la premisa (M1) del argumento de Boghossian. Así, un individuo puede no olvidar nada entre t_1 y t_2, y, sin embargo, no saber en t_2 lo que sabía en t_1 porque, al producirse un cambio en las condiciones externas de individuación, es incapaz de acceder al mismo pensamiento que pensó en el pasado.

Hay que decir, sin embargo, que esta estrategia de Ludlow recibió severas críticas por parte de otros partidarios del compatibilismo y él mismo reconoció en (2004) su insatisfacción con la solución propuesta[5]. Por esta razón, nuestro autor ha optado por presentar una argumentación más radical. El primer aspecto de la misma consiste en seguir manteniendo (EEM). El segundo pasa por describir la relación entre un pensamiento y su contenido de una manera muy distinta a la usual. Ludlow sostiene que nuestros pensamientos quedarían mejor descritos como pensamientos-matriz que contienen, cada uno de ellos, una serie de contenidos indexados en el tiempo [Ludlow (2004), p. 424]. De esta manera, el contenido que en cada instante del tiempo $Óscar_1$ podría traer a la memoria será distinto en función del entorno en el que se produzca la rememoración, puesto que la identidad del contenido dependerá, según (EEM), de los factores relevantes del entorno en el que se efectúe la rememoración mencionada. Sin embargo, siempre que se trate de contenidos pertenecientes a un mismo pensamiento-matriz, entonces podríamos sortear la conclusión de Boghossian de que $Óscar_1$ no conoce sus pensamientos: $Óscar_1$ es perfectamente capaz de recordar en t_2 el pensamiento que tuvo en t_1, al contrario de lo que se afirma en (M3), aunque lo haga otorgando a ese pensamiento un contenido distinto, *siempre y cuando ambos contenidos lo sean de un mismo pensamiento-matriz*.

Ludlow refuerza su propuesta indicando que "los informes de creencia se construyen con miras al beneficio del oyente H, para ayudar a H a construir una teoría sobre la vida mental de un agente relativa a los intereses y preocupaciones de S" [(2004), p. 421]. Es decir, que la función de las atribuciones mentales consiste, básicamente, en volver inteligible para un espectador el comportamiento de un individuo y, por ello, si dos atribuciones distintas para dos espectadores independientes explican igualmente bien la vida mental de un sujeto de atribución, entonces serán dos atribuciones verdaderas.

Para verlo podemos inspirarnos en un ejemplo del propio Ludlow [(2004), p. 429]. Pensemos en una conversación telefónica entre $Óscar_2$, residente en la Tierra Gemela, con $Pedro_1$, habitante de la Tierra. Imaginemos, además, que $Juan_2$, compañero de piso de $Óscar_2$ en la Tierra Gemela, escucha dicha conversación. En el transcurso de la misma, $Óscar_2$ dice a su interlocutor "todo está lleno de agua". ¿Cómo interpretarán $Pedro_1$ y $Juan_2$ a $Óscar_2$? Según dice Ludlow, es natural suponer que cada uno de ellos atribuirá contenidos distintos a $Óscar_2$. Mientras que $Pedro_1$ dirá que $Óscar_2$ tiene una creencia sobre el agua, $Juan_2$ le adscribirá un contenido sobre el $agua_2$.

Lo interesante es que, de acuerdo con nuestro autor, ambos están en lo cierto con respecto al pensamiento de $Óscar_2$, puesto que dicho pensamiento consiste en una matriz de contenidos, entre los que se encuentran aquellos contenidos que $Pedro_1$ y $Juan_2$ le atribuyen. En palabras de Ludlow: "no existe una correspondencia uno-a-uno entre pensamientos y contenidos […] los pensamientos están individuados por una serie de contenidos, no un contenido único, y […] los contenidos variarán a lo largo del tiempo" [(2004), p. 422]. Los espectadores pueden informar sobre la vida mental de un individuo adscribiéndole distintos contenidos –en función del entorno y el tiempo de la atribución–. Aún más, un individuo podría reflexionar sobre sus propias creencias y atribuirse a sí mismo un contenido que no coincidiera con ninguno de los contenidos que otros informantes le atribuyen sin que ello suponga que dicho individuo desconoce lo que cree. Para ello, bastaría con que cualquiera de los contenidos (incluido el contenido auto-atribuido) pertenezca al mismo pensamiento-matriz.

La nueva propuesta de Ludlow pretende poner fin a las inquietudes incompatibilistas. Si en virtud de (EEM), el contenido del pensamiento que $Óscar_1$ tuvo en t_1 cambia al rememorarlo en t_2, ello no implica que este individuo sea incapaz de conocer su pensamiento pasado y, por tanto, que carezca de acceso privilegiado a sus pensamientos. En la medida en que ambos contenidos lo son de un mismo pensamiento-matriz, $Óscar_1$ puede rememorar el pensamiento que mantuvo en t_1, aunque, eso sí, dotándole de un contenido distinto.

IV. Pensamientos-matriz o cómo no argumentar a favor del compatibilismo

Una vez descrita la propuesta de Ludlow nos gustaría realizar un análisis crítico de la misma, para lo cual, primeramente, debemos indagar en torno a la naturaleza de las proposiciones-matriz. En especial, nos interesaría descubrir qué clase de contenidos pueden atribuirse con verdad a un sujeto que mantiene un determinado pensamiento *P*. De esta investigación se desprenderá, según esperamos, que la nueva propuesta de Ludlow ha sorteado de manera ingeniosa el problema de Boghossian, pero a costa de dejar de lado el externismo.

En cuanto a la pregunta por la naturaleza de las proposiciones-matriz, no apreciamos demasiadas indicaciones positivas en el artículo de Ludlow de (2004). Según señalamos, la única descripción que este autor ofrece presenta a un pensamiento-matriz *P* como conteniendo una multiplicidad de contenidos indexados en el tiempo. De esta manera, si $Óscar_1$ tuvo en t_1 el pensamiento *P* con el contenido c_1, ello no impide que en t_2 un espectador diga con verdad que su pensamiento de $Óscar_1$ en t_1 tenía el contenido c_2. Según esta

idea, parece que las condiciones de verdad de los informes psicológicos que distintos hablantes atribuyen en diferentes entornos a Óscar$_1$ son tales que si uno de ellos le adscribe (al menos) uno de los contenidos albergados por el pensamiento-matriz estará realizando una atribución psicológica verdadera.

Con esta afirmación, Ludlow modifica una de las premisas del debate entre externistas e internistas, a saber, que la identidad de un pensamiento depende de la identidad de su (único) contenido. Ahora bien, debe decirse que esta premisa goza de una plausibilidad muy intuitiva, en el sentido de que nos permite conectar el debate filosófico sobre la identidad de nuestros pensamientos con nuestros criterios ordinarios de atribución de los mismos. De este modo, en nuestras prácticas habituales de atribución coincidimos en que dos individuos que comparten un lenguaje *piensan lo mismo, es decir, tienen la misma creencia*, cuando expresan uno de sus pensamientos usando las mismas palabras. Así, supongamos que dos individuos dicen, delante del escaparate de una joyería, "el platino es más caro que la plata". Si ambos se refieren a las mismas cosas haciendo uso de los mismos términos, nuestro dictamen habitual es que uno y otro expresan el mismo pensamiento porque piensan *lo* mismo, donde este "lo" especifica el contenido de sus creencias. Ahora bien, la idea de los pensamientos-matriz de Ludlow tendría como primera consecuencia que, a partir de ahora, la identidad de un pensamiento dependerá, no ya de la identidad de su único contenido, sino de la multiplicidad de contenidos que alberga.

La segunda pregunta que nos interesa plantear es qué deben tener en común dos contenidos, c_1 y c_2, para formar parte del pensamiento-matriz *P*. A fin de resolver esta cuestión, creemos conveniente retomar el problema que intenta resolver Ludlow al presentar su propuesta. De lo que se trata es de admitir un rango de variación entre los contenidos atribuidos y/o rememorados tal que las diferencias entre ellos no afecten a la identidad del pensamiento que los contiene. Sólo así podremos mantener que las diferencias en el contenido no le impiden a Óscar$_1$ recordar el *mismo* pensamiento-matriz que mantuvo en el pasado.

Lo último que necesitamos para despejar la incógnita es atender al siguiente aspecto de los casos de traslado. El hecho de que Óscar$_1$ no advierta que en el nuevo entorno su proferencia "el agua es húmeda" ha pasado a expresar un contenido distinto del que expresaba en la Tierra sólo se explica si *el nuevo contenido es, por lo que a él respecta, completamente indistinguible del contenido que expresaba en la Tierra mediante el uso de la misma expresión*. La ignorancia del cambio de sus contenidos se debe a que Óscar$_1$ se encuentra en t_2 en el mismo estado interno, de carácter cualitativo o fenomenológico que se encontraba en t_1. Esta identidad viene propiciada por el hecho de que, *ex hypothesi*, el agua y el agua$_2$ son líquidos que comparten todas sus propiedades superficiales y cuyo factor diferenciador le resulta desconocido a Óscar$_1$.

Algo similar ocurriría si Óscar$_1$ fuera trasladado sin su conocimiento a un tercer entorno, la Tierra Gemela-bis, donde todo es idéntico a la Tierra excepto un aspecto, a saber, que el líquido inodoro, incoloro e insípido que corre por sus ríos y lagos, líquido que todos los habitantes de este nuevo entorno denominan "agua", es, en realidad, esencialmente distinto al agua y tiene, por ejemplo, la composición química ABC. En este caso, si un habitante de la Tierra Gemela-Bis, José$_3$, dijera de Óscar$_1$ "él cree que el agua es un líquido", atribuyéndole con ello un contenido –relativo al agua$_3$– distinto del que le atribuían a este hablante Pedro$_1$ y Juan$_2$, esta atribución sería, según las pautas de Ludlow, igualmente apropiada para caracterizar el pensamiento de Óscar$_1$. Lo que vemos es que sean cuales sean los contenidos que se le atribuyan a Óscar$_1$ servirán para capturar su pensamiento *P* si se trata de contenidos idénticos *por lo que a él respecta*.

Para alcanzar mayor precisión, haremos uso del término "contrapartida epistémica" propuesto por Sawyer en (1999):

> (Contrapartida epistémica): los hechos h y h' son contrapartidas epistémicas para un sujeto S en un tiempo t si y sólo si sitúan a S en el mismo estado cualitativo en t [Sawyer (1999), p. 361].

Sawyer afirma que no sólo los hechos, sino también los eventos u objetos, pueden ser contrapartidas epistémicas. Lo esencial es que, en todo caso, ha de tratarse de entidades distintas, tales que dan lugar a un mismo estado cualitativo en S en un tiempo t. Y S estará situado en estados cualitativos idénticos con respecto a h y h' si desde su punto de vista no hay ninguna diferencia de carácter cualitativo entre ambas entidades. El ejemplo que Sawyer nos ofrece de contrapartidas epistémicas es, justamente, el del agua y el agua$_2$. Visto esto, podemos redescribir el *quid* de los casos de traslado diciendo que Óscar$_1$ no advierte la diferencia en los contenidos (de un mismo pensamiento) por la sencilla razón de que cuando los piensa se encuentra en el mismo estado cualitativo al remitir ambos contenidos a contrapartidas epistémicas.

En función de esto, podemos decir que la réplica de Ludlow al argumento de Boghossian –la cual establece que los contenidos que Pedro$_1$, Juan$_2$ y José$_3$ le atribuyen a Óscar$_1$ son contenidos que caracterizan el *mismo* pensamiento– pasa por aceptar el siguiente criterio de identidad para el pensamiento-matriz *P*:

> (Condiciones de identidad de los pensamientos-matriz): Un pensamiento *P* de un sujeto S que tenga como contenido c_1 y remita al estado de cosas e_1 albergará sólo aquellos contenidos (c_2, c_3, etc.) que remitan a estados de cosas (e_2, e_3, etc.) que sean para S en t contrapartidas epistémicas de e_1.

En otras palabras, los diferentes contenidos (c_1, c_2, etc.) que distintos hablantes podrían atribuir a Óscar$_1$ capturarán con verdad su pensamiento siempre y cuando se trate de contenidos que remitan a contrapartidas epistémicas, es decir, a estados de cosas (que el agua sea húmeda, que el agua$_2$ sea húmeda, etc.) que den lugar a la misma fenomenología o estado cualitativo en Óscar$_1$. Hemos extraído, así, el criterio de identidad básico para un pensamiento-matriz. Sin embargo, el texto de Ludlow no profundiza tanto como sería deseable en la cuestión de la relación existente entre esta clase de pensamientos y su contenido. En lo que sigue, ofreceremos dos interpretaciones de dicha relación que presuponen el criterio de identidad mencionado. Según pensamos, ninguna de estas interpretaciones será demasiado útil para los propósitos del externismo compatibilista, dado que una y otra resultan difícilmente conjugables con exigencias más básicas del externismo. Si estuviéramos en lo correcto, nuestro autor tendría que enfrentarse a una consecuencia contraria a sus intenciones, puesto que (EEM), la tesis propuesta en (1995), constituye para Ludlow una tesis externista fuerte y las modificaciones de (2004) llevan aún más lejos las conclusiones externistas de su anterior trabajo.

Así, en primer lugar, podríamos preguntarnos si es que acaso cada pensamiento-matriz está *compuesto* por una serie de contenidos cualitativamente idénticos para el sujeto. Esta concepción de las relaciones entre los contenidos y los pensamientos-matriz es apropiada para dar cuenta de las condiciones de verdad de las atribuciones psicológicas en los términos de Ludlow: el informe psicológico será verdadero siempre y cuando se especifique al menos uno de los contenidos (o bien c_1, o bien c_2 o bien c_3, etc.) que componen el pensamiento-matriz de un individuo. Sin embargo, creemos que esta concepción resulta difícilmente compatible con la lectura externista usual del experimento mental de la Tierra Gemela. De acuerdo con ésta, el pensamiento que Óscar$_1$ expresa al afirmar "el agua es húmeda" es *que el agua es húmeda* y no cualquier otro (que el agua$_2$ es húmeda o que el agua$_3$ es húmeda). Es decir, el externista saca partido de la intuición de que este individuo pretende referirse a *ese* líquido y no a otro distinto cuando dice de él que posee la propiedad de la humedad. Pero la tesis de Ludlow tendría como consecuencia que cuando Óscar$_1$ emite la proferencia "el agua es húmeda" ante una cantidad de H_2O puede decirse con verdad de él que ha atribuido la propiedad de la humedad no sólo a ese líquido, sino a cualquiera de los líquidos que él denominaría "agua" (o bien al H_2O, o bien al XYZ, o bien al ABC, etc.), dado que todos ellos constituyen contrapartidas epistémicas para él en ese momento.

Asimismo, creemos que esta interpretación de la relación entre un pensamiento-matriz y sus contenidos tendría consecuencias un tanto contraintuitivas en una situación como la siguiente. Pensemos en un individuo que descubre que en un momento dado ha tomado dos clases de cosas distintas por una y la misma, que es precisamente lo que le ocurre a Óscar$_1$ con respecto al agua y al agua$_2$. Este sujeto advertiría que ha permanecido en un error al

tomar como idénticas (porque *así se lo han parecido*) dos sustancias diferentes. De este modo, si él hubiera hablado en t_1 de una sustancia A creyendo que se trata de la sustancia B y en un momento posterior, t_2, se le informa de que ambas sustancias son diferentes, cabe pensar que este sujeto reconocería que *lo* que él pensó (que B es P) en t_1 es diferente de *lo* que expresó de manera efectiva su proferencia (que A es P). Justamente, la confusión de base (al ser A y B contrapartidas epistémicas en un momento dado para este individuo) explica su error, esto es, el hecho de que no haya dado a conocer su verdadero pensamiento. Si un segundo hablante hubiera escuchado su proferencia, habría estado legitimado, sin duda, para atribuirle el pensamiento de que A es P, ateniéndose a criterios públicos sobre el significado de sus palabras. Pero esto no significa que nuestro protagonista no tenga derecho a admitir su error y aclarar que el pensamiento que se le ha atribuido es *distinto* del pensamiento que quiso expresar en realidad[6].

Ahora bien, este fenómeno es difícil de explicar desde las propuestas de Ludlow. Y es que, atendiendo a su tesis, si el observador le hubiera atribuido a nuestro protagonista el pensamiento de que A es P en t_1, habría informado *con verdad* acerca del pensamiento (matriz) que nuestro protagonista quería expresar en este momento. La razón es que en t_1 A y B son contrapartidas epistémicas para nuestro individuo, dado lo cual *que A es P* y *que B es P* habrían sido contenidos incluidos dentro del mismo pensamiento-matriz, según el criterio de Ludlow. No obstante, hemos convenido en que nuestro protagonista, una vez que es consciente de su error, se negaría a admitir que ambas atribuciones son verdaderas. O, dicho de otro modo, aunque el observador no estableció su informe de manera arbitraria, puesto que se apoyó en la interpretación convencional de la proferencia emitida ("A es P"), esto no evita que la única atribución que captura el pensamiento que él quiso expresar en realidad, sea distinta de aquella que realizó el observador.

Una vez dicho esto, podemos ofrecer una segunda interpretación de las tesis de Ludlow sobre la relación que existe entre un pensamiento-matriz y sus contenidos. De este modo, se diría, un pensamiento-matriz es simplemente un *pensamiento descriptivo* que contiene las cualidades comunes a todos los contenidos que alberga. Así, puesto que $Óscar_1$ se encuentra en el mismo estado cualitativo con respecto a ambos líquidos, en un entorno y otro él tendría el mismo pensamiento puramente descriptivo. Este pensamiento podría capturarse de múltiples maneras; por ejemplo, diciendo que $Óscar_1$ piensa que el agua es húmeda (como diría un habitante de la Tierra) o que piensa que el $agua_2$ es húmeda (como diría un habitante de la Tierra Gemela), etc. Lo fundamental es que todas estas especificaciones atribuyen a $Óscar_1$ contenidos indistinguibles para él mismo, en la medida en que remiten a contrapartidas epistémicas que dan lugar a un mismo pensamiento-matriz descriptivo. La idea de Ludlow es, como vimos, que podríamos mantener que el traslado inadvertido no afecta de ningún modo al conocimiento que un individuo tiene

de su pensamiento, así concebido. Y es que una vez que se produce el traslado a la Tierra Gemela, Óscar$_1$ será capaz de recordar el mismo pensamiento-matriz (descriptivo) que expresó en la Tierra al decir "el agua es húmeda", a pesar de que lo rememore especificando un contenido distinto para el mismo. Así, dada la variación en las condiciones externas del entorno en que se produce la rememoración, por (EEM), el contenido recordado será un contenido sobre el agua$_2$, y no sobre el agua$_1$, que fue el contenido original del pensamiento terrestre. Sin embargo, mientras se trate de contenidos involucrados en un *mismo* pensamiento-matriz podremos decir que Óscar$_1$ recuerda el pensamiento matriz (descriptivo) que pensó en la Tierra y negaremos, con ello, la premisa (M3) del argumento de Boghossian, a saber, que este individuo deja de saber en la Tierra Gemela el pensamiento que mantenía en la Tierra.

Para ver en qué sentido su propuesta, así interpretada, entra en tensión con el externismo, conviene reparar en que la utilidad del experimento mental de la Tierra Gemela para los propósitos externistas reside en el hecho de que nos invitaba a concluir que Óscar$_1$ y Óscar$_2$, dos *réplicas internas* que comparten su historia interior cualitativa o fenomenológica, tienen, no obstante, *pensamientos distintos* en función de una diferencia en sus entornos externos. Ahora bien, vemos que Ludlow ha optado por describir nuestros pensamientos como pensamientos descriptivos que albergan múltiples contenidos para sortear las conclusiones de Boghossian. La identidad de un pensamiento-matriz vendrá dada, no ya por un contenido, sino por el mínimo descriptivo que comparten todos los contenidos que remiten a contrapartidas epistémicas para S en un momento t. O, dicho de otra manera, un mismo pensamiento-matriz admitirá una multiplicidad de contenidos siempre y cuando estos respeten la historia fenomenológica o cualitativa de un individuo. La conclusión que nos interesa extraer es que, para Ludlow, *todas las réplicas internas de Óscar$_1$ en otros mundos posibles (como Óscar$_2$) serán individuos que mantienen el mismo pensamiento cuando hacen uso del término "agua"*. La razón es que, aun cuando la diferencia en el líquido designado de ese modo en cada entorno dé lugar a diferencias en el contenido, estas diferencias no afectan a la identidad del pensamiento-matriz *P*, ya que los contenidos de todas las réplicas internas de Óscar$_1$ respetan, por hipótesis, su historia fenomenológica.

En este punto tenemos todos los elementos que nos permiten ver la tensión existente entre el externismo y la segunda interpretación de la propuesta de Ludlow. En efecto, debemos considerar que cuando Óscar$_1$ profiere la oración "el agua carece de sabor" está expresando un pensamiento-matriz *P* determinado por su historia fenomenológica, esto es, por ciertas cualidades descriptivas. Pero, si esto es así, el pensamiento que Óscar$_1$ expresa al proferir la oración "el agua carece de sabor" es el mismo pensamiento que expresarían mediante la misma proferencia todas sus réplicas internas en otros mundos posibles, incluido, por supuesto, Óscar$_2$. De todo esto debe concluirse que, para Ludlow, a pesar de que existe una diferencia en los contenidos

de $Óscar_1$ y $Óscar_2$, *esta diferencia es completamente irrelevante desde un punto de vista psicológico porque no impide que ambos tengan el mismo pensamiento*. Pero si aceptamos, con Ludlow, que los estados mentales de ambos individuos son idénticos porque $Óscar_1$ y $Óscar_2$ comparten una misma historia interna (fenomenológica o cualitativa), con ello estaremos extrayendo una conclusión del experimento mental de la Tierra Gemela enteramente favorable al internismo.

Por último, no deseamos concluir este trabajo sin señalar lo siguiente. Aun cuando se admitiera que la solución de este autor puede bloquear el argumento de Boghossian en particular, pensamos que su propuesta atenta contra un aspecto básico del privilegio epistémico del auto-conocimiento. Y es que como hemos indicado previamente, la concepción del acceso privilegiado que se desprende de la propuesta de Ludlow le arrebataría al sujeto la posibilidad de aclarar que él pensó una cosa, y no otra distinta, tras advertir que ha cometido errores empíricos comunes. De la misma manera que la confusión entre el agua y el $agua_2$ se explica por el parecido superficial entre ambos líquidos y la falta de información sobre su distinta composición química, encontramos que un hablante puede confundir con frecuencia cualesquiera sustancias que guarden similitudes superficiales entre sí, pero que no quepa considerar como pertenecientes a una misma clase.

La carencia de información empírica puede dar lugar, como se vio, a que dos sustancias distintas constituyan contrapartidas epistémicas para un individuo en un momento dado. Pero, una vez subsanada dicha carencia, debería reconocerse el derecho del individuo a defender que él *sabía* lo que estaba pensando en aquellos casos en los que era claro que su pretensión era referirse a (pensar sólo en) una de las sustancias y no referirse a (o pensar en) la otra. Por esta razón, deberíamos rechazar la tesis de Ludlow de que todo informe que atribuya a un individuo cualquier contenido que remita a una contrapartida epistémica en un momento dado puede caracterizar con verdad el pensamiento de dicho individuo, bien porque cualquiera de esos contenidos están incluidos en su pensamiento, bien porque ese pensamiento consiste en un mínimo descriptivo común a los contenidos de esa clase[7].

En último término, lo que defendemos es que el hecho de que un individuo esté confundiendo dos clases de cosas no parece invalidar la intuición de que en muchas ocasiones sólo hay una atribución psicológica correcta, siendo así que el testimonio del individuo goza de una autoridad especial al respecto.

VI. Conclusiones

A lo largo de este artículo hemos tratado de demostrar que, con el fin de esquivar el argumento de Boghossian en (1989), Ludlow ha elaborado una nueva caracterización de las relaciones pensamiento-contenido que asegurar

que $Óscar_1$ conoce el mismo pensamiento *P* tanto en t_1 como en t_2, a pesar de que toma en consideración contenidos distintos en cada momento. Si, como pide Ludlow, todas las rememoraciones futuras de *P* fueran rememoraciones que remiten al mismo pensamiento, entonces $Óscar_1$ poseería un acceso privilegiado a *P*, admitiendo que puede conocer *a priori* y con autoridad los distintos contenidos que, en el fondo, lo son del *mismo* pensamiento. Ludlow califica su solución de "radical" y "provocativa", en la medida en que nos invita a dejar de lado uno de los presupuestos fundamentales de las discusiones sobre la identidad metafísica de los pensamientos, a saber, que la identidad de un pensamiento depende de la identidad de su (único) contenido. Sin embargo, nuestro examen nos ha permitido concluir que Ludlow, en realidad, sustituye este criterio de identidad por otro de carácter claramente internista: un pensamiento-matriz *P* albergará aquellos contenidos que remiten a contrapartidas epistémicas, es decir, contenidos determinados por la historia interna fenomenológica del individuo. De aplicar este criterio, bajo cualquiera de las dos interpretaciones que hemos ofrecido, vimos que, en primer lugar, se derivan consecuencias bastante alejadas de los dictámenes externistas usuales y, en segundo lugar, se desatiende a un rasgo fundamental del privilegio epistémico del auto-conocimiento.

Departamento de Lógica y Filosofía de la Ciencia
Universidad Complutense
Ciudad universitaria, E-28040 Madrid (España)
E-mail: s.andres@filos.ucm.es

NOTAS

[1] Este trabajo ha sido realizado en el marco del proyecto FFI2008-03092 financiado por el Ministerio de Ciencia e Innovación. Agradezco sinceramente las interesantes sugerencias de los evaluadores, que, sin duda alguna, han contribuido a mejorar el presente trabajo.

[2] A pesar de que el experimento original de Putnam, expuesto en su trabajo de (1975), está planteado para defender una posición externista en semántica y no en psicología, suele considerarse que dicho experimento admite ciertas modificaciones que lo hacen apto para extraer conclusiones favorables al externismo psicológico.

[3] En lo que sigue, adoptaremos la convención de acompañar el nombre de los protagonistas del experimento de Putnam (1975) con un subíndice numérico. Los habitantes de la Tierra llevarán el número 1, mientras que los habitantes de la Tierra Gemela llevarán el número 2. Si se requiriera considerar otros entornos, continuaremos progresivamente con la numeración. Haremos lo mismo con los líquidos relevantes: "$agua_1$" o simplemente "agua" para el H_2O y "$agua_2$" para el XYZ.

[4] Cf. Davidson (1990), p. 271; Heil (1988), p. 244; Davies (2000), p. 391; Gibbons (1996), p. 290; Bernecker (1996), p. 338; Ludlow (1996), p. 312 y Shoemaker (1994), p. 260.

[5] Brueckner, por ejemplo, alega en (1997) que la premisa del argumento de Boghossian que ataca Ludlow en (1995) es, en realidad, (M2). Según Brueckner, si el pensamiento rememorado en t_2 es un *pensamiento distinto* del pensamiento mantenido en t_1 –en función de la variación en los entornos en que Óscar$_1$ está inserto en cada momento–, entonces lo que ocurre es que este individuo no puede recordar qué pensamiento mantenía en el pasado y, consiguientemente, se ha producido es un fallo en el funcionamiento de su memoria, justo lo que se negaba en (M2).

[6] Desde luego, alguien podría replicar que esta clase de situaciones no pueden darse en el escenario del experimento mental de la Tierra Gemela, puesto que las sustancias A y B relevantes (el H_2O y el XYZ) llevan el mismo nombre, "agua", y no se sabe *aún* que son distintas. Sin embargo, el externista alegaría que esto no invalida en absoluto la situación de fondo, que es la misma en ambos casos: hay dos sustancias que son, de hecho, diferentes, aunque lleven el mismo nombre; si uno o más sujetos creen que son la misma entonces están equivocados y el descubrimiento de la diferencia que existe entre ambas llevaría a los sujetos a la situación de admitir ese error.

[7] Podemos explicar esta idea retomando el ejemplo anterior. Así, Ludlow dictamina que las atribuciones realizadas por Pedro$_1$ y Juan$_2$ a Óscar$_1$ que son, respectivamente, *que Óscar$_1$ piensa que el agua es húmeda* y *que Óscar$_1$ piensa que el agua$_2$ es húmeda*, son dos atribuciones que, al especificar contenidos incluidos en un mismo pensamiento-matriz, resultan igualmente apropiadas para caracterizar con verdad el estado mental de Óscar$_1$. Ahora bien, si nuestro protagonista pretendía referirse sólo al agua$_2$ y la creencia que él quería expresar era un pensamiento *concreto* acerca de esta sustancia *en particular* cabe considerar que, tras haber recibido la información empírica relevante, Óscar$_1$ podrá escoger justificadamente la atribución correcta (la de Juan$_2$) como la única que capturó verdaderamente su estado mental. Así, de las dos atribuciones que recibe Óscar$_1$, sólo una captura con verdad el contenido real de su pensamiento y mantener lo contrario sería atentar contra el propio testimonio de Óscar$_1$ y, consiguientemente, contra su privilegio epistémico para seleccionar justificadamente el contenido específico de su pensamiento.

REFERENCIAS BIBLIOGRÁFICAS

BOGHOSSIAN, P. (1989), "Content and Self-Knowledge", en Ludlow, P. y Martin, N. (1998), pp. 149-173.

BOGHOSSIAN, P. y PEACOCKE, C. (eds.) (2000), *New Essays on the A Priori*, Oxford, Clarendon Press.

BRUECKNER, A. (1997), "Externalism and Memory", en Ludlow, P. y Martin, N. (1998), pp. 319-331.

BURGE, T. (1988), "Individualism and Self-Knowledge", *Journal of Philosophy*, vol. 85, pp. 649-663.

DAVIDSON, D. (1990), "Epistemology Externalized", en Davidson (2001), pp. 205-220. (Traducción castellana: "La externalización de la epistemología", en Davidson,

D., *Subjetivo, intersubjetivo, objetivo*, Madrid, Cátedra, 2003, pp. 265-279. Citado por la versión castellana).

— (2001), *Subjective, Intersubjective, Objective*, Oxford, Oxford University. (Traducción castellana: *Subjetivo, intersubjetivo, objetivo*, Madrid, Cátedra, 2003).

DAVIES, M. (2000), "Externalism and Armchair Knowledge", en Boghossian, P. y Peacocke, C. (2000), pp. 384-414.

FALVEY, K. y OWENS, J. (1994), "Externalism, Self-Knowledge, and Skepticism", *Philosophical Review*, vol. 103, pp. 107-137.

GIBBONS, J. (1996), "Externalism and Knowledge of Content", *The Philosophical Review*, vol. 105, pp. 287-310.

HEIL, J. (1988), "Privileged access", *Mind*, vol. 47, pp. 238-51.

LUDLOW, P. (1995), "Social Externalism, Self-Knowledge, and Memory", en Ludlow, P. y Martin, N. (1998), pp. 307-310.

— (2004), "What Was I Thinking? Social Externalism, Self-Knowledge and Shifting Memory Targets", en Schantz (2004), pp. 419-425.

LUDLOW, P. y MARTIN, N. (eds.) (1998): *Externalism and Self-Knowledge*, Stanford, CSLI Publications.

PESSIN, A. y GOLDBERG, S. (eds.) (1996), *The Twin Earth Chronicles: Twenty Years of Reflection on Hilary Putnam's 'The meaning of 'meaning'*, Londres, Sharpe.

PUTNAM, H. (1975), "The Meaning of 'Meaning'", en Pessin, A. y Goldberg, S. (1996), pp. 3-52. (Traducción castellana en: L. M. Valdés (ed.), *La búsqueda del significado*, Madrid, Tecnos, 1991, pp. 131-193).

SAWYER, S. (1999), "An Externalist Account of Introspective Knowledge", *Pacific Philosophical Quarterly*, vol. 80, pp. 358-378.

SCHANTZ, R. (ed.) (2004), *The Externalist Challenge*, Berlin, Walter De Gruyter.

SHOEMAKER, S. (1994), "Self-Knowledge and Inner Sense", *Philosophy and Phenomenological Research*, vol. 54, pp. 249–314.