

(Pre-print. Texto original castellano del capítulo que, traducido al portugués, fue publicado en el volumen *Mente, linguagem e mundo*, São Paulo, 2010)

REFLEXIONES SOBRE ANTIINDIVIDUALISMO Y AUTOCONOCIMIENTO

Carlos J. Moya. Universitat de València

Carlos.Moya@uv.es

El debate sobre las relaciones entre el antiindividualismo o externismo¹ y el autoconocimiento empieza a tener ya una larga historia, desde que Michael McKinsey publicase, en 1991, su artículo “Anti-Individualism and Privileged Access”.² Desde entonces han aparecido varios libros colectivos, formados por artículos previamente publicados en revistas, y siguen apareciendo nuevas contribuciones a la discusión, centrada sobre todo en la cuestión de la compatibilidad o incompatibilidad entre ambas doctrinas, el externismo y el autoconocimiento.

El problema del autoconocimiento se extiende mucho más allá de su compatibilidad con el externismo. El autoconocimiento es un problema por sí mismo, con independencia de las dificultades adicionales que haya podido plantear el surgimiento del externismo semántico e intencional. Aunque no quisiera dedicar demasiado tiempo a esta cuestión, creo que es importante distinguir, en el marco del autoconocimiento, entre estados fenomenológicos y estados intencionales o con contenido intencional. Creo que el autoconocimiento presenta perfiles muy distintos en relación con unos y otros estados mentales.

Por lo que respecta a los estados fenomenológicos, tenemos, en primer lugar, órganos perceptivos que nos permiten sentir el sabor de los alimentos, el color de los objetos, el dolor de un pinchazo o el placer de la brisa en el rostro. Y lo que necesitamos para saber que tenemos esas sensaciones es, además de sentirlas, un conjunto de capacidades cognitivas y conceptuales que nos permiten identificarlas y atribuirnoslas. Esta relación cognitiva con nuestras propias sensaciones posee rasgos bastante próximos a la infalibilidad, incorregibilidad e inmediatez que Descartes atribuía a la conciencia.

¹ Utilizaré estos términos como sinónimos, aunque el primero connota sobre todo la relevancia del contexto social, destacada especialmente por Burge, (1979) en la determinación de los contenidos y conceptos de las mentes de los individuos, y el segundo sugiere más bien la importancia del entorno natural externo en dicha determinación, a partir de los trabajos de Putnam (1975) y del propio Burge (1982).

² Tal vez haya precedentes, pero los desconozco.

El caso de los estados intencionales es distinto. Una creencia no es sentida, no es equiparable a una sensación. Mientras que la posesión efectiva de una sensación se decide de modo definitivo por la conciencia inmediata y momentánea de la misma, la posesión efectiva de una creencia no se decide definitivamente de ese modo. La autoridad de la primera persona en relación con los estados intencionales se halla sometida a restricciones objetivas. Estas restricciones son, al menos, de dos tipos. Por un lado, de carácter lógico, relacionadas con la coherencia y la consistencia con otros estados intencionales y con la conducta. Así, la autoatribución sincera de una creencia puede verse cuestionada o desautorizada, por ejemplo, por la constatación de que dicha creencia entra en contradicción con otras que el sujeto profesa o que resulta razonable atribuirle. Hay también, por otro lado, restricciones de carácter constitutivo, derivadas del aspecto disposicional de estos estados, de su conexión con inferencias o acciones posibles. La ausencia de estas conexiones en circunstancias determinadas puede cuestionar justificadamente una autoatribución sincera de una creencia, un deseo o una intención. Y estas conexiones, de tipo lógico o constitutivo, de ningún modo son objeto de conciencia inmediata ni están sometidas a la autoridad irrestricta de la primera persona.

Podemos hablar así, en relación con los estados intencionales, de una especie de externismo, distinto del propugnado por Putnam y Burge. La mente individual se halla, en efecto, sometida a normas o principios objetivos de tipo lógico o constitutivo que la sobrepasan y la determinan, y el conocimiento y la autoatribución de nuestros estados intencionales se hallan asimismo limitados por dichas normas o principios. Estas normas y principios no tienen aplicación a las sensaciones y los estados fenomenológicos en general.

Así, pues, el autoconocimiento, en relación con los estados intencionales, constituye un problema en sí mismo, con independencia de su relación con el externismo y el antiindividualismo en las versiones de Putnam o Burge. Resulta, en efecto, enigmático cómo podemos atribuirnos sin esfuerzo aparente, de modo directo y no inferencial, independiente de la investigación empírica y con una especial autoridad y presunción de verdad (resumiré en lo sucesivo estos caracteres del autoconocimiento con la expresión “de modo privilegiado”), estados cuya posesión efectiva depende de condiciones objetivas, de tipo lógico o constitutivo, el acceso a las cuales requiere esfuerzo, inferencias e involucra proclividad al error. De ningún modo podemos considerar el autoconocimiento como una capacidad de naturaleza más o menos obvia, lamentablemente amenazada por insidiosas doctrinas externistas.

El autoconocimiento para el que el externismo supone presuntamente un obstáculo es el que tiene por objeto los estados intencionales, los estados con contenido intencional. Si consideramos tales estados sometidos a normas y principios objetivos que sobrepasan la conciencia individual, no tenderemos a ver en el externismo un problema completamente nuevo para la posibilidad del autoconocimiento. Veremos las doctrinas externistas como propuestas que destacan factores de tipo objetivo, natural o social, que determinan parcialmente el contenido de los estados intencionales y que se añaden a los factores objetivos que de todos modos son ya constitutivos de tales estados.

Sin embargo, cabe responder que, aun aceptando, por mor de la argumentación, que el externismo no representa un problema completamente nuevo para el autoconocimiento, parece no obstante agravarlo. En la medida en que el contenido de los estados intencionales depende también de factores objetivos de tipo natural o social que no podemos conocer de modo privilegiado, la cuestión de cómo podemos conocer de ese modo dichos estados se torna más difícil. Los factores naturales y sociales destacados por el externismo de Putnam o Burge como determinantes del significado de (algunas de) nuestras emisiones lingüísticas y del contenido de (algunos de) nuestros estados intencionales añaden dificultades adicionales a las suscitadas por los factores de carácter lógico o constitutivo (disposicional). Una diferencia que puede ser importante entre ambos grupos consiste en que el conocimiento de los factores del segundo grupo puede, en principio, lograrse por reflexión o análisis conceptual,³ y en este sentido a priori, mientras que el conocimiento de los factores del primer grupo parece claramente requerir el concurso de la experiencia y la investigación empírica. En principio, no parece aceptable que los factores naturales y sociales que determinan el significado y el contenido intencional puedan ser conocidos a priori. La relevancia de esta distinción aparecerá más adelante.

Entrando ya en el centro del debate, la compatibilidad o incompatibilidad entre autoconocimiento y antiindividualismo, es útil distinguir, siguiendo a Martin Davies (2000), dos problemas distintos involucrados en dicha cuestión, que él denomina el problema del logro y el problema de la consecuencia.

El primer problema ya nos resultará familiar por las observaciones anteriores. Puede plantearse así: ¿Cómo podemos atribuirnos y conocer de modo privilegiado

³ Pensemos, a este respecto, en el llamado “funcionalismo analítico” acerca de los estados intencionales, desarrollado por autores como David Armstrong (1990) y David Lewis (1991, 2003).

estados cuya naturaleza y contenido dependen de condiciones que no podemos conocer de ese modo? He abordado este problema con cierto detalle en otros trabajos (Moya 1998, 2003). La respuesta “no podemos” a la pregunta anterior, una respuesta que implica la incompatibilidad entre externismo y autoconocimiento, puede derivarse de un argumento que he denominado “el argumento simple”, cuya premisa principal es, en palabras de Bonjour (1992: 136), la siguiente: “Si el contenido depende de factores externos que pertenecen al entorno, entonces el conocimiento del contenido debería depender del conocimiento de tales factores”. Si añadimos a esta premisa esta otra: “Si el externismo es verdadero, el contenido depende de factores externos que pertenecen al entorno”, llegamos entonces a la conclusión según la cual si el externismo es verdadero, el conocimiento del contenido depende del conocimiento de factores externos que pertenecen al entorno. Ahora bien, puesto que no hay nada privilegiado en el conocimiento de tales factores externos, el autoconocimiento, entendido como un conocimiento privilegiado de los propios estados intencionales, no es posible si el externismo es verdadero. Externismo y autoconocimiento son incompatibles.

El argumento simple es válido, pero su primera premisa es problemática. Esta premisa constituye una aplicación del principio según el cual, si A depende de B, el conocimiento de A depende del conocimiento de B. Pero, como ya mostró McKinsey (1991: 182-3) no parece difícil hallar contraejemplos a este principio, tanto si la dependencia se entiende en sentido metafísico como en sentido conceptual o a priori. Así, en el primer sentido, mi existencia depende de la existencia de mis padres, pero el conocimiento de mi existencia no depende del conocimiento de la existencia de mis padres. Con respecto al segundo sentido de “dependencia”, podemos conocer una proposición sin conocer todas sus consecuencias lógicas, a pesar de que la proposición en cuestión depende conceptualmente de dichas consecuencias, de modo que, si la proposición tuviera otras consecuencias, ya no sería la misma proposición. De igual forma, podría saber que tengo una determinada creencia sin necesidad de conocer las condiciones externas de las que depende, metafísica o conceptualmente, que tenga dicha creencia.

Desde un punto de vista positivo, el problema del logro puede tal vez resolverse apelando a lo que se ha llamado el modelo de inclusión, defendido entre otros por Donald Davidson, Tyler Burge y John Heil. Según Burge, por ejemplo, “el conocimiento [de nuestros propios estados mentales] consiste en un juicio reflexivo que incluye pensar un pensamiento de primer orden sobre el que versa el juicio. El juicio reflexivo simplemente

hereda el contenido del pensamiento de primer orden” (1988: 656). Los juicios reflexivos que manifiestan el autoconocimiento son contextualmente autoverificantes. Así, si el entorno objetivo en la Tierra determina, para el pensamiento que Óscar 1 expresa diciendo “el agua calma la sed”, el contenido *el agua calma la sed*, el entorno objetivo de la Tierra Gemela determina, para el pensamiento correlativo que Óscar 2 expresa con las mismas palabras, el contenido *el agua gemela calma la sed*.⁴ Y cuando uno y otro se atribuyen reflexivamente ese pensamiento, el contenido del juicio reflexivo incluye en un caso el concepto *agua* y en el otro el concepto *agua gemela*, con lo que ambos juicios reflexivos, manifestación del autoconocimiento, son verdaderos. Así, externismo y autoconocimiento son compatibles.

El compatibilismo entre externismo y autoconocimiento fundado en el modelo de inclusión ha de afrontar algunas dificultades, como el argumento de las transiciones entre mundos, situaciones posibles en que un sujeto es trasladado, sin él saberlo, de la Tierra a la Tierra Gemela y viceversa, sufriendo, al cabo del tiempo, un cambio inadvertido en alguno de sus conceptos y en el contenido de sus pensamientos. El argumento trata de mostrar que los juicios reflexivos en el sentido de Burge, aun cuando sean autoverificantes, no constituyen realmente conocimiento (cf. Boghossian 1989, 1992). En un trabajo anterior (Moya 2003), traté de responder a este argumento concediendo a sus partidarios dos supuestos importantes, a saber, una concepción del conocimiento en términos de la noción de alternativas relevantes, según la cual para saber que P es necesario poder discriminar entre P y las alternativas relevantes a P, y el principio de transparencia según el cual “el significado es *transparente* en el sentido de que, si alguien asocia un significado a cada una de dos palabras, ha de saber si estos significados son el mismo” (Dummett, cit. por Boghossian 1992). Aun aceptando estos supuestos, que conceden al incompatibilista una notable ventaja inicial, conseguí defender, creo que con cierto éxito, el compatibilismo. Mi defensa descansaba en una concepción normativa del externismo, que privilegia, en la determinación del significado, las situaciones originales de aprendizaje del significado de ciertos términos y las muestras paradigmáticas empleadas en ellas para dicho aprendizaje. En su réplica a mi trabajo, Burge (2003) arguye, creo que con razón, que la epistemología de las alternativas relevantes, sea o no correcta en ciertos casos, no es aplicable a los juicios reflexivos característicos del

⁴ Las consecuencias del experimento de Putnam, referido al significado, para el contenido intencional fueron extraídas por Burge (1982).

autoconocimiento, y que el principio de transparencia de Dummett es falso.⁵ Sin estos dos supuestos, la posición del incompatibilista se debilita enormemente y la defensa del compatibilismo frente al argumento de las transiciones cobra mucha mayor fuerza. En cualquier caso, voy a aceptar, por mor de la argumentación, que el problema del logro puede ser resuelto sobre la base de alguna versión, quizá con ciertas modificaciones, del modelo de inclusión.

Más complicado para el compatibilismo resulta, en cambio, el otro problema al que me referí anteriormente y que, siguiendo a Davies, llamaremos “el problema de la consecuencia”, planteado originalmente por Michael McKinsey (1991) y reformulado con posterioridad por Paul Boghossian (1998) y Jessica Brown (1995), entre otros. En términos generales, el problema consiste en que el compatibilista, que acepta tanto el externismo como el autoconocimiento, se ve obligado a admitir que puede conocer a priori determinadas proposiciones sobre el entorno objetivo, lo que parece claramente inaceptable. El compatibilismo se ve así enfrentado así a una consecuencia inaceptable o absurda.

La primera formulación de este problema la debemos a McKinsey (1991). Según McKinsey, Burge (1988) “sostendría que las tres proposiciones siguientes son consistentes:

(1) Óscar sabe a priori que está pensando que el agua es húmeda.

(2) La proposición según la cual Óscar está pensando que el agua es húmeda depende necesariamente de E.

(3) La proposición E no puede ser conocida a priori, sino sólo mediante la investigación empírica” (McKinsey 1991: 178).

La proposición E expresaría las condiciones externas de las que depende el contenido de los pensamientos de Óscar.

Según McKinsey, las tres proposiciones anteriores son consistentes si la dependencia necesaria a la que se alude en (2) se entiende en términos de necesidad o implicación metafísica. Pero esa dependencia, sostiene McKinsey, debería entenderse como implicación conceptual. Creo que McKinsey tiene razón. Los experimentos mentales de Putnam o Burge sacan a la luz las consecuencias implícitas en nuestros *conceptos* de significado o de pensamiento; muestran lo que *entendemos por*

⁵ Falvey y Owens (1994) sostienen una posición cercana en espíritu a ésta de Burge cuando rechazan que el autoconocimiento se extienda a los juicios comparativos de un sujeto acerca de si dos de sus pensamientos o creencias tienen o no el mismo contenido.

“significado” o “pensamiento”. El externismo debe, pues, entenderse como una tesis de dependencia conceptual entre determinados estados intencionales y rasgos objetivos del entorno, no como una tesis de dependencia meramente metafísica. Pero, si entendemos así el externismo, (2) se convierte en

“(2b) La proposición según la cual Óscar está pensando que el agua es húmeda implica conceptualmente E” (McKinsey 1991: 182).

Y en ese caso, señala McKinsey, (1), (2b) y (3) son realmente inconsistentes. Sobre la base de (1) y (2b), que conoce a priori, Óscar puede deducir E, con lo que puede conocer E a priori. Pero esta consecuencia entra en contradicción con (3), que afirma que la proposición E no puede ser conocida a priori.

Planteado en términos de una reducción al absurdo, el argumento de McKinsey contra el compatibilismo puede formularse como sigue. Óscar, el compatibilista, puede razonar del modo siguiente:

- (1) Estoy pensando que el agua es húmeda.
- (2) Si estoy pensando que el agua es húmeda, entonces E.
- (3) Luego, E.

Las dos premisas, (1) y (2), pueden ser conocidas por Óscar a priori, sin necesidad de investigación empírica. (1) puede serlo en virtud de la capacidad de autoconocimiento y (2) en virtud de la reflexión y la argumentación filosófica a favor del externismo. Y la deducción de (3) a partir de (1) y (2) es claramente válida. Así, Óscar puede conocer E a priori. Pero es absurdo sostener que E, una proposición sobre el mundo externo, puede ser conocida a priori. Por lo tanto, el compatibilismo tiene una consecuencia absurda y deber ser rechazado.

Una cuestión importante para el éxito de esta crítica al compatibilismo es la interpretación de la proposición E. McKinsey la interpreta como “el mundo externo existe” o “hay objetos externos”. En efecto, según McKinsey, el argumento incompatibilista vendría a decir lo siguiente: “Mire, si usted pudiera conocer a priori que se halla en un determinado estado mental, y si hallarse en ese estado implica conceptual o lógicamente la existencia de objetos externos, entonces usted podría saber a priori que el mundo externo existe. Puesto que, obviamente, usted *no puede* saber a priori que el mundo externo existe, no puede saber a priori que se halla en el estado mental en cuestión. Es así de simple” (McKinsey 1991: 183). Presumiblemente, aceptando el autoconocimiento, la tesis que habría que rechazar es el externismo.

Un problema importante de la crítica de McKinsey es que no resulta tan obvio que no se pueda saber a priori que el mundo externo existe. Se me ocurre pensar, por ejemplo, en la refutación kantiana del idealismo en la *Crítica de la razón pura*. Kant trata de deducir la existencia de objetos externos, negada por el idealista, partiendo del conocimiento de sus propios pensamientos y representaciones, es decir, del autoconocimiento, procediendo, pues, a priori. Esta forma de argumentar puede tener fuerza. Cabe argüir, siguiendo la estela del fascinante estudio de Strawson sobre Kant, *Los límites del sentido*, que el curso de mis pensamientos sólo resulta inteligible para mí suponiendo que constituye un modo particular, entre otros posibles, de concebir un mundo objetivo independiente de tales pensamientos, un itinerario particular a través de ese mundo objetivo. El fracaso del imponente proyecto de Carnap, en *La construcción lógica del mundo*, de concebir el mundo objetivo como un constructo a partir de datos sensoriales habla en favor de la posición de Kant-Strawson. Independientemente del éxito final de esta argumentación, lo importante aquí, sin embargo, es que no resulta obviamente fallida y, por tanto, que, como hemos señalado, conocer a priori la existencia del mundo externo no es, frente a lo que McKinsey sostiene, obviamente imposible.

En posteriores planteamientos del problema de la consecuencia, los incompatibilistas han interpretado la proposición “externa” E de un modo mucho más concreto, haciendo así francamente implausible sostener que pueda ser conocida a priori. Boghossian (1998) interpreta esta proposición como *el agua existe*. La formulación de Boghossian del problema de la consecuencia es un alarde de sencillez y rotundidad. Quizá demasiado, como veremos. Según Boghossian, Óscar, un compatibilista, se halla en posición de argumentar, puramente a priori, como sigue:

1. Si tengo el concepto *agua*, el agua existe.
 2. Tengo el concepto *agua*.
- Por lo tanto,
3. El agua existe (Boghossian 1998: 202).

Un compatibilista ha de aceptar, según Boghossian, que tanto 1 como 2 son verdaderas y cognoscibles a priori. La premisa 2 no plantea dificultades; es una aplicación de la doctrina del autoconocimiento. La premisa 1 es más problemática. Aparentemente, razonando de modo paralelo, los compatibilistas estarían obligados a aceptar la existencia de unicornios, hadas y todo tipo de extrañas entidades, lo que sin duda sería una reducción

al absurdo de su posición. Boghossian no atribuye a los compatibilistas un error tan obvio. Un compatibilista sólo está obligado a aceptar la premisa 1 y razonar del modo problemático indicado en el caso de términos y conceptos aptos para construir experimentos de Tierras Gemelas. Y esta aptitud es cognoscible a priori, pues depende de condiciones que son ellas mismas cognoscibles a priori. En primer lugar, el término en cuestión (“agua”, en el ejemplo anterior) “ha de aspirar a nombrar una clase natural” (Boghossian 1998: 204), una condición que Boghossian parece considerar equivalente a que el usuario del término pretenda nombrar con él una clase natural (ibid.) En segundo lugar, el término en cuestión ha de expresar un concepto atómico, y no compuesto.⁶ En tercer lugar, el usuario del término ha de ser indiferente sobre la esencia (química) de dicha clase natural. Estas condiciones, por separado o conjuntamente, excluyen el caso de los unicornios, las hadas y muchos otros. Si la satisfacción de estas condiciones es cognoscible a priori, como sostiene Boghossian, un compatibilista que las satisfaga habrá de admitir que puede conocer la premisa 1 a priori. De este modo, y supuesta la premisa 2, un compatibilista habrá de aceptar que puede llegar a conocer, puramente a priori, que el agua existe, lo que resulta manifiestamente absurdo. Por lo tanto, el compatibilismo, que implica este absurdo, ha de ser falso. Al menos una de las dos doctrinas, el externismo o el autoconocimiento, de las que dependen las premisas, ha de ser abandonada.

Una respuesta natural por parte del compatibilista es insistir en que la primera condición es demasiado débil. Para aceptar la conclusión, lo que se requiere es que el término nombre de hecho una clase natural, y no meramente que aspire a hacerlo. Imaginemos, en efecto, a un compatibilista que posea el concepto *flogisto* y que satisfaga las tres condiciones indicadas. Podría entonces razonar así (cf. Moya 1998: 246):

A. Si tengo el concepto *flogisto*, el flogisto existe.

B. Tengo el concepto *flogisto*.

Por lo tanto,

C. El flogisto existe.

Nuestro compatibilista, una vez advertido de la inexistencia del flogisto, puede insistir en que la falsedad a la que llegó se debía a que el término “flogisto” no nombra

⁶ He de reconocer que mi idea de lo que significa que un concepto sea atómico es bastante vaga. Y no soy el único (cf. Brown 1999: 56). Pero concederé que hay una idea precisa de esta propiedad de los conceptos.

de hecho una clase natural, aun cuando él pretendía que lo hiciera. En realidad, tenía una extensión vacía. Y, claramente, que un término designe o no una clase natural no se puede saber a priori.

Frente a esta objeción, Boghossian responde, en primer lugar, que no es necesario, para construir experimentos putnamianos de Tierras Gemelas, que el término de que se trate designe de hecho una clase natural en la Tierra y, en segundo lugar, que un externista no puede admitir, de modo coherente con su posición, conceptos atómicos que aspiren a nombrar una clase natural y tengan una extensión vacía (el concepto *flogisto* podría ser un ejemplo). Y trata de defender esta idea mediante el experimento mental de la Tierra Seca, cuyos habitantes son réplicas psicológicas y funcionales nuestras, pero sufren una ilusión colectiva, pues de hecho no hay agua en su planeta. Según Boghossian, no hay una respuesta coherente que un externista pueda dar a la pregunta de qué concepto expresa el término “agua” en la Tierra Seca. Responder que se trata de un concepto compuesto, del tipo *el líquido que corre por los ríos y forma los mares, etc.*, conlleva aceptar que el mismo término, con el mismo papel funcional, expresa un concepto atómico en la Tierra y uno compuesto en la Tierra Seca. Pero el carácter compuesto o atómico de un concepto, insiste Boghossian, depende de su sintaxis interna, y no de circunstancias externas. Responder, en cambio, que expresa un concepto atómico no es tampoco una opción abierta. Si un externista acepta que los conceptos vienen (parcialmente) individuados por sus referentes externos, claramente el concepto que expresa el término “agua” en la Tierra Seca no es el mismo que el que expresa en la Tierra, donde su extensión no es vacía. ¿Cuál es, entonces, ese concepto? Dado que no hay “ninguna clase natural al final de la cadena causal relevante que lleva a los usos de ‘agua’ en la Tierra Seca” (Boghossian 1998: 209-10), no hay una respuesta que el externista pueda dar a esta cuestión. Así, Boghossian escribe: “Si, en un contexto dado, no hay un estado de cosas que establezca cuál es el referente de un concepto dado, en esa medida no hay tampoco un estado de cosas que establezca cuál es ese concepto” (1998: 210).

Así, supongo que, si Boghossian tiene razón, el externista no podría coherentemente afirmar que posee el concepto *flogisto*, de modo que su premisa B sería falsa. La falsedad de la conclusión queda así explicada y no es necesario aceptar que dicha falsedad se debe a que el término “flogisto” no designa una clase natural.

En un interesante artículo, Jessica Brown (1999) criticó la argumentación incompatibilista de Boghossian, en mi opinión de manera contundente. Según Brown, en su crítica al compatibilismo, Boghossian tiene en cuenta únicamente la versión del

externismo generada por los experimentos putnamianos de Tierras Gemelas, confiando en que su crítica se aplicará también a otras versiones del externismo, como la de Burge. Sin embargo, esta estrategia acaba mostrándose desastrosa, porque, según Brown, “si el externismo de Burge es verdadero, el argumento de Boghossian no consigue establecer que el externismo putnamiano es incompatible con el acceso privilegiado” ((1999: 53). De acuerdo con el externismo de Burge, un sujeto puede poseer un concepto aunque tenga una comprensión sólo parcial e imperfecta de sus condiciones de aplicación y, desde luego, aunque no conozca una definición correcta del mismo. La posesión del concepto en esas condiciones de dominio imperfecto se sustenta en la deferencia del sujeto hacia las reglas semánticas vigentes en su comunidad sociolingüística y, en su caso, en los expertos en la materia de que se trate. Como hemos visto, el argumento de Boghossian descansa crucialmente en que las condiciones para que un compatibilista acepte la premisa 1 han de ser cognoscibles a priori. Brown acepta que la tercera condición, la indiferencia sobre la esencia de la clase natural que el término aspira a nombrar, es cognoscible a priori, sobre la base del autoconocimiento. Sin embargo, dado el externismo de Burge, la primera condición, a saber, que un término aspire a nombrar una clase natural es algo sobre lo que un sujeto puede estar equivocado, aunque su intención, que sí puede conocer a priori, sea nombrar una clase natural. Pensemos, por ejemplo, en el socorrido ejemplo de “jade”. En nuestra comunidad sociolingüística, donde los expertos saben que el jade es tanto jadeíta como nefrita, el término “jade” no aspira a designar una clase natural, pero un sujeto, ignorante de este dato, podría usar el término con la intención de designar una tal clase. Finalmente, se desprende también del externismo de Burge que un sujeto puede creer que un concepto que posee y que domina imperfectamente es atómico (o compuesto) y estar equivocado, de modo que esto cuestiona que la segunda condición sea cognoscible a priori, y que el carácter compuesto o atómico de un concepto dependa únicamente de su “sintaxis” interna. Y lo importante es que, para corregir estos errores, se necesita información empírica acerca del uso lingüístico en la comunidad. Como Brown lo expresa:

Si el externismo de Burge es correcto, un sujeto podría tener el concepto expresado por un término y estar no obstante equivocado, de un modo sólo remediable por medios empíricos, acerca de si el término aspira a nombrar una clase natural y/o expresa un concepto atómico (1999: 57)

Por nuestra parte, añadiremos que el externismo de Burge puede mostrar también el error de la afirmación de Boghossian según la cual un externista no puede admitir conceptos atómicos que aspiran a nombrar clases naturales y tengan una extensión vacía, como “agua” en la Tierra Seca. Aun cuando el término no tenga un referente objetivo, las condiciones externas que, para un externista a lo Burge, sustentan la posesión del concepto *agua* en la Tierra Seca pueden venir dadas en este caso por la comunidad social y sus prácticas lingüísticas.

En suma, pues, a menos que Boghossian muestre concluyentemente la falsedad del externismo de Burge, su crítica al compatibilismo fracasa.

Finalmente, los experimentos mentales en que descansa el externismo de Burge, como el célebre ejemplo del término “artritis”, pueden llevarse a cabo con conceptos que no son atómicos y con términos que ni aspiran a nombrar una clase natural ni de hecho la nombran.

A pesar de su crítica al argumento incompatibilista de Boghossian, Jessica Brown es también defensora del incompatibilismo. La argumentación de Brown es más cauta que la de Boghossian y, por ello, más difícilmente rebatible. Según Brown (1995), la proposición “externa” E que un compatibilista que posea cierto concepto, como por ejemplo *agua*, ha de admitir que puede conocer a priori sería la siguiente:

O bien me encuentro en un entorno que contiene muestras de agua y *agua* es un concepto de clase natural, o bien soy parte de una comunidad que posee el concepto *agua*, sea o no *agua* un concepto de clase natural.

El compatibilista puede llegar a esta conclusión razonando a partir de dos premisas cognoscibles a priori. La primera, accesible al autoconocimiento, sería: Poseo el concepto *agua* y soy agnóstico sobre las condiciones de aplicación de ese concepto (p. ej., no tengo creencias sobre la composición química del agua). La segunda sería un condicional cuyo consecuente sería la proposición “externa” E y cuyo antecedente sería: Poseo el concepto *agua* y soy agnóstico sobre las condiciones de aplicación de ese concepto. Esta premisa sería cognoscible a priori mediante la argumentación y la reflexión filosófica a favor del externismo.

En el caso de Brown, la proposición E no contiene tanta información como en el caso de Boghossian. Aun así, contiene la suficiente para que siga siendo inaceptable

sostener que es cognoscible a priori, de modo que el problema de la consecuencia sigue en pie.

Vamos a centrarnos en la argumentación incompatibilista de Brown, puesto que es la más plausible de las que hemos examinado. Introduciremos, sin embargo, alguna modificación, debido a que, en nuestra opinión, es posible construir experimentos mentales putnamianos sin necesidad de que el término empleado designe una clase natural. Por ejemplo, se podría mostrar que el significado de un término como “agua” y el contenido de los pensamientos expresados con él es distinto en dos sujetos, Óscar 1 y Óscar 2, que son gemelos psicológicos y funcionales, suponiendo que, en la Tierra, el referente del término tiene de hecho la composición química A, B o C y en la Tierra Gemela la composición C, D o E. De todos modos, la modificación no es esencial para nuestra argumentación.

De acuerdo con esto, y eliminado de la formulación de Brown la referencia a clases naturales, el argumento problemático podría expresarse así, empleando el ejemplo del concepto *agua*:

1. Poseo el concepto *agua* y soy agnóstico sobre las condiciones de aplicación de este concepto.

2. Si poseo el concepto *agua* y soy agnóstico sobre las condiciones de aplicación de este concepto, entonces estoy, o he estado, en un entorno en el que hay agua, o bien pertenezco a una comunidad que posee el concepto *agua*, o tal vez ambas cosas.

Por lo tanto,

3. Estoy, o he estado, en un entorno en el que hay agua, o bien pertenezco a una comunidad que posee el concepto *agua*, o tal vez ambas cosas.

El problema, desde luego, no es la verdad de la conclusión, que parece bastante clara, sino el itinerario, puramente a priori, que conduce a ella.

El compatibilista no se halla en modo alguno indefenso ante este problema. Aun cuando, en la formulación anterior, las condiciones que ha de cumplir un concepto para construir experimentos mentales al estilo de Putnam o de Burge son bastante más relajadas que en el caso de Boghossian, cabe dudar también que sean cognoscibles a priori. En primer lugar, no es claro que un completo agnosticismo sobre las condiciones de aplicación de un concepto sea compatible con la posesión de dicho concepto. Y, en segundo lugar, no es claro que el grado de agnosticismo sobre las condiciones de

aplicación de un concepto compatible con la posesión de dicho concepto sea cognoscible a priori, y no requiera conocimiento empírico de las normas semánticas de la comunidad sociolingüística. Pensemos en el experimento de Burge con el concepto *arthritis*. En este caso, el desconocimiento del sujeto sobre las condiciones de aplicación del concepto parece compatible con la posesión de ese concepto. Pero el sujeto no sabe a priori que desconoce esas condiciones y remediar esa ignorancia parece requerir el conocimiento empírico del modo en que el término se usa en su comunidad. Por lo tanto, hay espacio, también aquí, para criticar la argumentación incompatibilista poniendo en duda la cognoscibilidad a priori de las premisas.

Por otra parte, cabe también la estrategia que los ingleses llaman “to bite the bullet”: admitir que, partiendo del conocimiento a priori de las premisas, es posible conocer a priori la conclusión, negando sin embargo que este conocimiento a priori sea misterioso o problemático (cf. Sawyer 1999, Phillips 2006). En este caso, no obstante, la estrategia parece menos prometedora que cuando la conclusión se limita a afirmar la existencia del mundo externo. En cualquier caso, dejaré también de lado esta posible respuesta del compatibilista.

Un modo más prometedor y novedoso de afrontar el problema, defendido por Martin Davies (2000, 2003), es el siguiente. Según Davies,⁷ la primera premisa del argumento problemático es verdadera y conocida de modo privilegiado en primera persona; la segunda premisa es verdadera y conocida por reflexión filosófica; ambas premisas, pues, son verdaderas y conocidas a priori; además, el argumento es válido. Sin embargo, el argumento pertenece a una clase de argumentos con un defecto de tipo epistemológico. En ellos, la justificación o fundamento que sirve de apoyo a las premisas no se transmite a la conclusión. El argumento es válido, pero no convincente (cogent), en el sentido de que alguien que, por alguna razón, pusiese en duda la conclusión, no vería, y con razón, las premisas como una justificación adecuada para aceptarla; a diferencia de los argumentos convincentes, no sería legítimo remitir a las premisas para convencer de la verdad de la conclusión a quien duda de ella. Hay una conexión entre esta deficiencia epistemológica y la falacia de la petición de principio, especialmente cuando ésta es sutil y no explícita. Un ejemplo de esta clase de argumentos es la famosa prueba de la existencia de un mundo externo ofrecida por G. H. Moore. Resulta bastante claro que un escéptico sobre el mundo externo no encontrará en las premisas del argumento de Moore

⁷ Una línea semejante la encontramos en Crispin Wright (2000).

una razón para aceptar la conclusión y abandonar su escepticismo. Hay un sentido en que la conclusión, la existencia de un mundo externo, es parte de lo requerido para aceptar las premisas, con lo que el argumento, como suele señalarse, comete una petición de principio contra el escéptico. Según Davies, el argumento que nos ocupa es también un caso de falta de transmisión de la justificación: la justificación a priori, “de sillón”, como la llama Davies, de la premisas no se transmite a la conclusión. Alguien que dudase de la existencia del agua en su entorno o de las convenciones lingüísticas de su comunidad no vería en las premisas del argumento razones que disiparan sus dudas.

La línea propuesta por Davies para responder al problema de la consecuencia es especialmente interesante porque sitúa este problema en un marco más general y abre un campo de investigación lógica y epistemológica novedoso e importante. Pero la cuestión es altamente compleja. A pesar de los esfuerzos de Davies, no es claro en qué consiste el fenómeno de la falta de transmisión de la justificación ni cuáles son los caracteres precisos de los argumentos que adolecen de este defecto. Ni siquiera es claro que existan tales caracteres comunes. Así, resulta prematuro formular un juicio sobre el éxito final de esta estrategia para resolver el problema de la consecuencia.

La línea de respuesta que voy a proponer frente al problema de la consecuencia tiene conexiones con la línea de Davies y estoy en deuda con ella. Trataré, sin embargo, de evadir sus dificultades suponiendo que la petición de principio es un fenómeno mejor conocido y más familiar que la falta de transmisión de la justificación en general. Y sostendré que el argumento que los incompatibilistas ponen en boca de los compatibilistas es un caso bastante claro de petición de principio. Si esto es así, el compatibilista no tendría por qué comprometerse con la posibilidad indeseada de conocer a priori el entorno objetivo. En lo que resta de mi exposición, trataré de exponer y justificar mi propuesta.

La premisa crucial del argumento que nos ocupa es la premisa 2, la premisa externista, podemos llamarla, ya que constituye una aplicación del externismo al caso concreto del concepto *agua*. La justificación para admitir esta premisa no va más allá que la justificación que puede llevarnos a abrazar el externismo como tal. Y esta justificación no tiene la forma de un argumento de cuyas premisas se seguiría lógicamente el externismo, o que aumentarían su plausibilidad. La principal justificación, como sabemos, son las intuiciones sobre el significado y el contenido intencional generadas por experimentos mentales como el de Putnam con el término “agua” o el de Burge con el término “artritis” y el concepto correspondiente.

El problema que plantea el argumento que nos ocupa es, como sabemos, el supuesto conocimiento a priori de hechos sustantivos sobre el mundo que se alcanza en la conclusión. Lo que trataré de sugerir es que ese pretendido conocimiento a priori es una ilusión derivada de los supuestos implícitos en los experimentos mentales indicados, en los que se presupone ya lo enunciado en la conclusión.

Pensemos que, si los experimentos de Putnam o Burge han de favorecer el externismo, la doctrina según la cual los contenidos intencionales no sobrevienen al estado interno, neurológico y psicológico, de los individuos, sino que se hallan también parcialmente determinados por la naturaleza objetiva del entorno, esos experimentos, digo, han de estar diseñados de tal modo que exista en ellos una separación nítida entre el estado interno de los individuos y el entorno objetivo. Ha de haber una diferencia en el entorno objetivo de ambos sujetos que no tenga reflejo alguno en su estado interno (individuado de modo no intencional), pues de otro modo, como en los experimentos científicos mal diseñados, no habríamos aislado la variable cuya contribución a cierto resultado queremos averiguar, distinguiéndola de otros factores presentes en la situación experimental y que tienen también sus propios efectos sobre el resultado final; no sería posible, en el caso de los experimentos mentales de Putnam o Burge, mostrar con claridad la contribución del entorno a la fijación del significado y el contenido intencional, distinguiéndola de la contribución del estado neurológico y psicológico interno del individuo. Ésta es la razón por la que Putnam sitúa a sus personajes en 1750, antes del desarrollo de la química moderna. En la actualidad, que el agua es H₂O es una creencia bastante general, como lo sería, en la fantaseada Tierra Gemela, el conocimiento de la composición química del agua gemela, XYZ. Y así, si el experimento tuviese lugar en nuestros días, a la diferencia objetiva en el entorno correspondería una diferencia en los estados psicológicos de ambos sujetos, de modo que no se podría establecer con claridad la tesis según la cual es la diferencia en el entorno la que explica la diferencia entre los estados intencionales de uno y otro sujeto.

Así, pues, los experimentos han de presuponer una noción robusta, fuertemente realista, del entorno objetivo, un conjunto de rasgos que este entorno posee realmente, con independencia de lo que los sujetos involucrados puedan creer o desear. Es el caso de la naturaleza real del agua (y de la naturaleza estipulada del agua gemela) en el experimento de Putnam, o de las convenciones semánticas efectivamente vigentes en la comunidad real y en la contrafáctica en el caso del experimento de Burge. Digamos que los experimentos mentales en cuestión han de dar por resuelta la cuestión del escepticismo

sobre el mundo externo, natural o social, para poder llevar a cabo su cometido, y presuponer que podemos conocer rasgos muy específicos del mismo. En efecto, este cometido, la generación de las intuiciones que respaldan el externismo, *no se podría alcanzar presuponiendo un mundo fenomenalista*, un mundo constituido por puras apariencias. Si presuponemos un mundo así, no habría diferencia alguna entre el significado y el contenido intencional de los dos sujetos involucrados en el experimento. Necesitamos asumir, por detrás y con independencia de la conciencia fenoménica, la existencia de entidades con una esencia propia, en el caso del experimento de Putnam, o de comunidades sociolingüísticas con convenciones semánticas efectivas de ciertos tipos, en el caso del experimento de Burge.

En este último, Burge (1979: 77) comienza presuponiendo la existencia de una comunidad sociolingüística con reglas semánticas determinadas que gobiernan el uso del término “arthritis”, y sobre esta base inicial, no cuestionada, comienza a introducir variaciones contrafácticas acerca de estas convenciones semánticas. Putnam (1975: 223), por su parte, con el fin de mostrar que el estado psicológico interno de un individuo no determina la extensión de un término, recurre a la ciencia-ficción y nos pide que supongamos la existencia de la Tierra Gemela, con sus peculiaridades en la naturaleza de lo que allí llaman “agua”. En ambos casos, los rasgos tanto de lo que se considera la situación real como la imaginada sólo podrían conocerse empíricamente, a posteriori. Y lo que se considera como la situación real no se pone en cuestión a los fines del experimento.

Estos supuestos fuertemente realistas, y sólo cognoscibles a posteriori, subyacen, pues, a los experimentos mentales que son la justificación principal para aceptar el externismo y se incorporan así implícitamente a esta doctrina. Y cuando ésta se aplica a determinados conceptos, estos supuestos se hacen explícitos en las premisas de argumentos como el que nos ocupa. Recordemos, en efecto, la premisa 2: “Si poseo el concepto *agua* y soy agnóstico sobre las condiciones de aplicación de este concepto, entonces estoy, o he estado, en un entorno en el que hay agua, o bien pertenezco a una comunidad que posee el concepto *agua*, o tal vez ambas cosas”. Y, sobre la base de la premisa 1, la conclusión descarga el consecuente del condicional y enuncia estos supuestos como hechos. Pero el pretendido conocimiento a priori de los hechos enunciados en la conclusión no es, en el fondo, más que el enunciado explícito de estados de cosas que no son a priori salvo en el sentido de que han sido presupuestos y estipulados sin pruebas en los experimentos mentales que conducen al externismo. No hay, así,

genuino conocimiento a priori de hechos sustantivos sobre el entorno objetivo en la conclusión del argumento que nos ocupa, y un externismo con conciencia crítica de sí mismo y de sus fundamentos no debería ni tendría por qué comprometerse con la posibilidad de semejante conocimiento a priori. Tampoco, pues, como ya insistió Burge, sería legítimo emplear el externismo para generar argumentos en contra del escepticismo.

REFERENCIAS

ARMSTRONG, D. M., 1990, "The Causal Theory of Mind", en W. G. Lycan (ed.), *Mind and Cognition*, Oxford, Blackwell.

BOGHOSSIAN, P. A., 1989, "Content and Self-Knowledge", *Philosophical Topics* 17, pp. 5-26.

———, 1992, "Externalism and Inference", en E. Villanueva (ed.), *Rationality in Epistemology. Philosophical Issues*, vol. 2, pp. 11-28.

———, 1998, "What the Externalist Can Know *A Priori*", en E. Villanueva (ed.), *Concepts. Philosophical Issues*, vol 9, pp. 197-211.

BONJOUR, L., 1992, Entrada "Externalism/internalism", en J. Dancy y E. Sosa (eds.), *A Companion to Epistemology*, Oxford, Blackwell, pp. 133-6.

BROWN, J., 1995, "The Incompatibility of Anti-Individualism and Privileged Access", *Analysis* 55, pp. 149-56. Reimpreso en P. Ludlow y N. Martin (eds.), *Externalism and Self-Knowledge*, Stanford, CSLI Publications. Citamos por esta edición.

———, 1999, "Boghossian on Externalism and Privileged Access", *Analysis* 59, pp. 52-9.

BURGE, T., 1979, "Individualism and the Mental", *Midwest Studies in Philosophy*, vol 4, pp. 73-121.

———, 1982, "Other Bodies", en A. Woodfield (ed.), *Thought and Object. Essays on Intentionality*, Oxford, Clarendon Press, pp. 87-120.

———, 1988, "Individualism and Self-Knowledge", *Journal of Philosophy* 85, pp. 649-65.

———, 2003, "Replies from Tyler Burge", en M. J. Frápolli y E. Romero (eds.), *Meaning, Basic Self-Knowledge, and Mind. Essays on Tyler Burge*, Stanford, CSLI Publications, pp. 243-96.

DAVIES, M., 2000, “Externalism and Armchair Knowledge”, en P. Boghossian y C. Peacocke (eds.), *New Essays on the A Priori*, Oxford, Oxford University Press, pp. 384-414.

———, 2003, “Externalism, Self-Knowledge and Transmission of Warrant”, en M. J. Frápolli y E. Romero (eds.), *Meaning, Basic Self-Knowledge, and Mind. Essays on Tyler Burge*, Stanford, CSLI Publications, pp. 105-130.

FALVEY, K. y OWENS, J., 1994, “Externalism, Self-Knowledge, and Skepticism”, *The Philosophical Review* 103, pp. 107-37.

LEWIS, D., 1991, “Psychophysical and Theoretical Identifications”, en D. M. Rosenthal (ed.), *The Nature of Mind*, Oxford, Oxford University Press.

———, 2003, “Reduction of Mind”, en T. O’Connor y D. Robb (eds.), *Philosophy of Mind. Contemporary Readings*, Londres, Routledge.

McKINSEY, M., 1991, “Anti-Individualism and Privileged Access”, *Analysis* 51, pp. 9-16. Reimpreso en P. Ludlow y N. Martin (eds.), *Externalism and Self-Knowledge*, Stanford, CSLI Publications, 1998, pp. 175-84. Las referencias se ajustan a esta última edición.

MOYA, C. J., 1998, “Boghossian’s *Reductio* of Compatibilism”, en E. Villanueva (ed.), *Concepts, Philosophical Issues*, vol. 9, pp. 243-51.

———, 2003, “Externalism, Inclusion, and Knowledge of Content”, en M. J. Frápolli y E. Romero (eds.), *Meaning, Basic Self-Knowledge, and Mind. Essays on Tyler Burge*, Stanford, CSLI Publications, pp. 165-86.

PHILLIPS, I., 2006, “Reflections on externalism and Self-Knowledge”, MindGrad Meeting 13/6/06.

PUTNAM, H., 1975, “The meaning of ‘meaning’”, en *Mind, Language and Reality*, Cambridge, Cambridge University Press, pp. 215-71.

SAWYER, S., 1999, “An Externalist Account of Introspective Self-Knowledge”, *Pacific Philosophical Quarterly* 80, pp. 358-78.

WRIGHT, C., 2000, “Cogency and Question-Begging: Some Reflections on McKinsey’s Paradox and Putnam’s Proof”, en E. Vaillanueva (ed.), *Skepticism, Philosophical Issues* 10, pp. 140-163.