

ESCEPTICISMO Y CONTENIDO MENTAL

Carlos Moya. Departamento de Metafísica
Universidad de Valencia

El presente trabajo se ocupa del escepticismo acerca del mundo externo y su conocimiento. En la primera sección se intenta mostrar que el internalismo acerca del contenido mental, es decir, la independencia constitutiva entre el contenido de los estados mentales y el mundo externo, es una premisa crucial en la argumentación escéptica. De este modo, la concepción opuesta del contenido mental, o externalismo, de ser correcta, podría cercenar un pilar central en la argumentación escéptica. En la segunda sección se expone brevemente la posición externalista y se formulan tres objeciones a ella desde el punto de vista del internalismo. Las tres secciones siguientes están dedicadas a las posibles respuestas del externalismo a dichas objeciones. En la sección final se exploran otros recursos a disposición del externalismo, se discute el problema del pensamiento demostrativo y se sugiere que, en conjunto, la balanza se inclina claramente del lado de la posición externalista.

1. Escepticismo e internalismo.

Un aspecto central del escepticismo sobre nuestro conocimiento del mundo externo consiste en la tesis de la independencia constitutiva entre nuestra vida mental y nuestro entorno objetivo. En una u otra forma, los argumentos escépticos defienden la posibilidad conceptual de una situación en la cual entornos objetivos completamente distintos puedan causar conjuntos recíprocamente indistinguibles de estados mentales. Recientemente se tiende a conocer esta concepción del contenido mental con el nombre de internalismo. La defensa de esta posibilidad conceptual constituye el núcleo de los argumentos escépticos de mayor alcance. En la primera de las *Meditaciones Metafísicas*, Descartes parte del hecho, reconocido por todos, de que los episodios que nos representamos en los sueños no suceden realmente en el mundo, para sostener a continuación que no disponemos de "indicios concluyentes ni señales que basten a distinguir con claridad el sueño de la vigilia."¹ El sueño nos proporciona un ejemplo de cómo nuestra representación de ciertas situaciones puede producirse con independencia de que tales situaciones se den en la realidad. Y si no disponemos de criterios infalibles de distinción entre el sueño y la vigilia, no tenemos justificación para suponer que las cosas son de cierta forma por el mero hecho de que nos las representemos de ese modo. El argumento del sueño trata de probar la posibilidad de una independencia general entre nuestra representación de situaciones, acontecimientos y objetos y la realidad de los mismos. Es posible que, cuando creemos que p , p sea realmente el caso, pero si p y la creencia de que p pueden variar de forma recíprocamente independiente, esto es algo que no podemos certificar.

Aun aceptando que no disponemos de criterios infalibles para distinguir entre el sueño y la vigilia y que, en consecuencia, el hecho de que p y nuestra creencia de que p pueden variar de forma recíprocamente independiente, cabría argüir que esa independencia no se extiende a los elementos que componen el contenido de la creencia, a las 'ideas' que intervienen en la representación del hecho. Supongamos que 'percibo' una casa marrón frente a mi ventana y que esa

¹ R. Descartes, *Meditaciones Metafísicas*, traducción de Vidal Peña, Alfaguara, Madrid, 1977, 18

'percepción' causa mi creencia de que hay una casa marrón frente a mi ventana. Puesto que podría estar soñando, es posible que mi creencia sea falsa. Sin embargo, ello no supone que las ideas que componen el contenido de la creencia, la idea de casa, del color marrón, de ventana, etc., sean asimismo independientes de casas, objetos de color marrón, ventanas, etc. Podría suceder que, durante la vida de vigilia, aun cuando no pudiéramos distinguirla concluyentemente del sueño, las cosas externas dejaran su impronta en nosotros en forma de ideas. Entre los espesos vapores del sueño, esas ideas, formadas a partir de cosas reales, se combinarían en nuestra mente en formas caprichosas, independientes del modo en que las cosas representadas se combinan en ese momento en el mundo real, induciendo en nosotros creencias falsas. Si ello fuese así, las consecuencias escépticas del argumento del sueño serían limitadas, puesto que cabría suponer cierto nivel de acuerdo entre los contenidos de nuestra mente y la realidad, cierto nivel de congruencia entre las relaciones causales y las relaciones epistemológicas entre mente y mundo. Descartes mismo admite la posibilidad de esta respuesta al argumento del sueño: "Con todo, hay que confesar que al menos las cosas que nos representamos en sueños son como cuadros y pinturas que deben formarse a semejanza de algo real y verdadero; de manera que por lo menos esas cosas generales —a saber: ojos, cabeza, manos, cuerpo entero— no son imaginarias, sino que en verdad existen. Pues los pintores, incluso cuando usan del mayor artificio para representar sirenas y sátiros mediante figuras caprichosas y fuera de lo común, no pueden, sin embargo, atribuirles formas y naturalezas del todo nuevas, y lo que hacen es sólo mezclar y componer partes de diversos animales"².

Cabría señalar que nuestras ideas de los objetos físicos son también el resultado de una combinación de otras ideas, de modo que no tenemos, tampoco en este caso, garantías de que tales combinaciones correspondan a algo real y verdadero. Estas combinaciones de ideas podrían ser el resultado de la misma actividad que en nuestros sueños da lugar a representaciones fingidas de acontecimientos y situaciones. Ahora bien, la extensión, la figura, la magnitud, el número, la duración y otras propiedades semejantes, con las que nos representamos objetos físicos, ya estemos despiertos o dormidos, deben al menos ser reales, de modo que nuestras ideas de tales propiedades no serían imaginarias, aunque puedan serlo las ideas de los objetos que nuestra mente forma por combinación de aquéllas: "... Aun pudiendo ser imaginarias esas cosas generales —a saber: ojos, cabeza, manos y otras semejantes— es preciso confesar, de todos modos, que hay cosas aún más simples y universales realmente existentes, por cuya mezcla, ni más ni menos que por la de algunos colores verdaderos, se forman todas las imágenes de las cosas que residen en nuestro pensamiento, ya sean verdaderas y reales, ya fingidas y fantásticas. De este género es la naturaleza corpórea en general, y su extensión, así como la figura de las cosas extensas, su cantidad o magnitud, su número, y también el lugar en que están, el tiempo que mide su duración y otras por el estilo"³.

La inexistencia de criterios concluyentes de distinción entre el sueño y la vigilia no nos permite saber, en un momento dado, si las cosas son como creemos. Ahora bien, si las ideas más simples que componen el contenido de nuestras creencias representan y están causadas por propiedades reales de las cosas, el argumento del sueño es compatible con cierto grado de dependencia epistemológica de nuestro pensamiento respecto del mundo y, así, con cierto nivel de acuerdo entre ambos. En mi opinión, el argumento del sueño es incluso compatible con la verdad de muchas de nuestras creencias sobre el mundo, a saber, las creencias que, sobre la base de la percepción o el razonamiento, y excluyendo ilusiones o errores, nos formamos estando despiertos,

²*Ibid.*, 18-19.

³*Ibid.*, 19.

ya que, en ese caso, es plausible pensar que no sólo las ideas que intervienen en el contenido de nuestras creencias, sino también la manera en que se combinan en él, dependen de la configuración real de las cosas. La verdad de tales creencias no equivale a conocimiento, ya que, no teniendo indicios concluyentes para distinguir el sueño de la vigilia, no tenemos tampoco medios infalibles de distinguir las creencias que son de hecho verdaderas, ni, por tanto, justificación subjetiva suficiente de las mismas. Ahora bien, si, en el espíritu del externalismo epistemológico (que no debería confundirse con el externalismo intencional del que se ocupa preferentemente este trabajo)⁴, se sostiene que es suficiente para el conocimiento que nuestras creencias verdaderas estén *de facto* justificadas (por ejemplo, por el hecho de estar causadas por la configuración de las cosas que las haría verdaderas), el argumento del sueño podría dejar abierta incluso la posibilidad de que tengamos conocimiento del mundo: creencias verdaderas objetivamente —aunque no subjetivamente— justificadas⁵.

Nuestro contraataque antiescético frente al argumento del sueño explota epistemológicamente un resquicio de dependencia de nuestras ideas respecto de la realidad, a saber, la admisión de que algunas de nuestras ideas más simples representan y están causadas por propiedades reales de las cosas. El escéptico necesita imperiosamente poner en cuestión ese último resquicio, elaborando una explicación alternativa de la presencia de tales ideas en nuestra mente que separe la relación causal de la relación representativa. El propio Descartes, empeñado en el proceso de la duda hiperbólica, se da perfecta cuenta de la necesidad de esta explicación alternativa a fin de mostrar la posibilidad conceptual de una independencia epistemológica *completa* de nuestros contenidos mentales respecto del mundo externo. Nuestros contenidos mentales, incluidas las ideas más simples, han de poder ser como son y variar como lo hacen aun cuando el mundo externo sea completamente distinto de como creemos que es e, incluso, aunque no exista. En este punto, Descartes introduce la posibilidad de un poder superior, más allá de nuestro control, que haga surgir en nuestra mente las ideas que de hecho tenemos: "Y, sin embargo, hace tiempo que tengo en mi espíritu cierta opinión, según la cual hay un Dios que todo lo puede, por quien he sido creado tal como soy. Pues bien: ¿quién me asegura que el tal Dios no haya procedido de manera que no exista tierra, ni cielo, ni cuerpos extensos, ni figura, ni magnitud, ni lugar, pero a la vez de modo que yo, no obstante, sí tenga la impresión de que todo eso existe tal y como lo veo?"⁶. Y si ese comportamiento ofende la suprema bondad y veracidad de Dios, atribuyámoslo entonces a un poderoso y maligno Genio que "ha usado de toda su industria para engañarme"⁷.

La función de este genio en la argumentación escéptica no es otra que asegurar la posibilidad de una independencia epistemológica completa entre nuestros contenidos mentales y

⁴ L. Bonjour define así el externalismo epistemológico: "Una teoría de la justificación es... *externalista* si acepta que al menos algunos de los factores de justificación [de las creencias de una persona] no necesitan ser... [cognitivamente] accesibles a esa persona" (Artículo «externalism/internalism», en J. Dancy y E. Sosa (eds.), *A Companion to Epistemology*, Blackwell, Oxford, 1992, 132. Cf. tb. D. M. Armstrong, *Belief, Truth and Knowledge*, Cambridge University Press, Cambridge, 1973, cap. 11.

⁵ No entraremos en la cuestión de si el conocimiento equivale o no a creencia verdadera justificada. En cualquier caso, disponer de creencias verdaderas (objetivamente) justificadas es hallarse en una situación epistémica bastante satisfactoria.

⁶ R. Descartes, *op.cit.*, 19.

⁷ *Ibid.*, 21.

los objetos, propiedades y situaciones externas. La posición escéptica requiere que la sucesión de los cambios en nuestras percepciones, creencias, imágenes y representaciones en general *pueda* tener lugar con independencia de los cambios que puedan darse en el mundo externo. Así, la posición escéptica se torna más fuerte y persuasiva si se logra mostrar que dicha posibilidad es, no sólo pensable, sino perfectamente congruente con lo que creemos saber sobre el funcionamiento de los sistemas perceptivos y sobre la dependencia de los procesos mentales respecto del cerebro. Así, por ejemplo, de acuerdo con nuestras creencias de raigambre científica, la experiencia visual tiene como causa próxima los impulsos transmitidos al cerebro por el nervio óptico. Parafraseando a Descartes, podemos preguntarnos: ¿Quién nos asegura que el patrón de tales impulsos responde fielmente a la luz que reflejan los objetos en nuestro entorno? Si ese patrón de impulsos nerviosos pudiera mantenerse en su estado actual, variando al mismo tiempo el entorno objetivo tanto como se desee, nuestras experiencias visuales serían exactamente las mismas que ahora tenemos: en mi caso, seguiría 'viendo' la superficie negra de la mesa donde escribo, mi mano sujetando el bolígrafo y recorriendo el papel de izquierda a derecha, el teléfono situado un poco más lejos y a la izquierda, etc., y teniendo las creencias correspondientes aun cuando no hubiera nada de todo eso. ¿Quién nos asegura entonces, utilizando los términos de la pesadilla ideada por Putnam⁸, que no somos cerebros en una cubeta de nutrientes, con las terminaciones nerviosas conectadas a un potentísimo ordenador programado por un genial y no menos maligno neurólogo? En el caso de Descartes, el entorno objetivo causalmente responsable de los pensamientos está constituido, en parte, por el Genio Maligno. En el caso de Putnam, por el ordenador y el Malvado Neurólogo. Estos respectivos entornos causan los contenidos mentales, pero no hay en éstos clave epistemológica alguna sobre la naturaleza de dicho entorno. Suponiendo ahora que nuestros contenidos mentales sean indistinguibles de los que tendríamos si fuésemos una mente cartesiana bajo el poder de un Genio Maligno o un cerebro en una cubeta, suponiendo, pues, que en este último caso no nos parecería que las cosas son distintas de como ahora nos parece que son, la conclusión es que no podemos saber cómo son realmente las cosas y que, *aun cuando no seamos mentes cartesianas o cerebros en cubetas*, nuestras creencias sobre el mundo no sólo carecen de justificación, sino que podrían ser sistemáticamente falsas.

Así, pues, limitándonos a las creencias, es crucial para la argumentación escéptica que, si fuésemos cerebros en una cubeta o mentes cartesianas sometidas a un Genio Maligno, creeríamos exactamente lo mismo que ahora creemos, es decir, nuestras creencias tendrían exactamente el mismo contenido intencional. La diferencia estaría, obviamente, en su valor de verdad, pero a éste no tendríamos acceso, puesto que, desde nuestro punto de vista, las mismas experiencias que ahora nos parece que confirman nuestras creencias parecerían confirmarlas igualmente si nos halláramos en la situación descrita por Descartes o Putnam.

Dada la conexión entre intencionalidad y significado, el internalismo intencional conlleva un internalismo semántico: suponiendo que 'el gato está sobre la estera' exprese el contenido de una creencia, el significado de estos signos (o de las representaciones mentales correspondientes) ha de poder ser el mismo tanto si la creencia expresada en ellos pertenece a un cerebro en una cubeta como si es la creencia de un sujeto situado en el entorno en el que suponemos hallarnos. De otro modo, el contenido intencional en ambos casos no podría ser el mismo. Si el contenido

⁸ H. Putnam, *Razón, verdad e historia*, traducción de J. M. Esteban, Tecnos, Madrid, 1988, cap. 1.

intencional es constitutivamente independiente del entorno objetivo, el significado ha de serlo también.

El internalismo dista mucho de ser una tesis obvia o evidentemente verdadera. En los experimentos mentales escépticos se nos pide que aceptemos, sobre bases intuitivas, que nuestra vida mental sería exactamente como es ahora tanto si somos cerebros en una cubeta como si no. Esas bases intuitivas, sin embargo, no son inocentes ni prefilosóficas, sino que requieren y suponen un compromiso previo con el internalismo⁹. Lo cierto es que, a menos que nos resulte ya atractiva por otras razones la concepción cartesiana de la mente, la pregunta relativa a cómo sería la vida mental de un cerebro en una cubeta puede dejarnos sencillamente perplejos. Desde luego, la respuesta: "Exactamente igual a la que ahora tenemos" no tiene especiales títulos de plausibilidad si no estamos previamente inclinados a aceptarla. La apelación a la intuición no puede ser decisiva en la defensa del internalismo si la intuición que se pretende inducir está ya alimentada por prejuicios internalistas.

Lo que he tratado de mostrar en esta sección es el carácter central de las premisas internalistas en la argumentación escéptica. Entre las diversas estrategias antiescépticas que podrían ser o han sido efectivamente desarrolladas, la defensa del externalismo con respecto al contenido intencional y semántico (en adelante, simplemente 'externalismo') representa una de las más recientes y, al mismo tiempo, más prometedoras. Dado el carácter central del internalismo en los argumentos escépticos, la estrategia externalista no es periférica, sino que atenta contra una de las piedras angulares del escepticismo. Trataremos de caracterizar ahora los fundamentos del externalismo.

2. El externalismo. Putnam y la Tierra Gemela.

Podemos caracterizar el externalismo, en términos generales, como la doctrina según la cual (en palabras de Bilgrami) "los contenidos de las creencias de un agente no son independientes del mundo externo al agente en cuestión". El externalismo, continúa diciendo Bilgrami, "constituye un rechazo de la concepción según la cual la intencionalidad puede caracterizarse plenamente con independencia del mundo externo, o, por expresarlo en los términos de la Primera Meditación Cartesiana, niega la idea de que los contenidos intencionales de un agente serían exactamente los que son aun cuando resultara que no hay un mundo externo"¹⁰.

En éste, como en muchos otros campos, la obra de Hilary Putnam ha sentado las bases de desarrollos posteriores. En un artículo justamente célebre¹¹, Putnam introduce la cuestión del externalismo como una contribución al análisis del concepto de significado, pero es natural, como el propio Putnam hace, extender sus consideraciones al problema del contenido intencional de los estados mentales, especialmente las creencias.

El principal argumento de Putnam en favor del externalismo descansa en el experimento mental de la Tierra Gemela. Aun a riesgo de aburrir a algunos lectores, permítasenos exponer brevemente este experimento mental. Putnam nos pide que imaginemos que, en un lugar de cierta galaxia existe un planeta al que llamaremos la 'Tierra Gemela'. La Tierra Gemela es muy semejante a la Tierra. Sus habitantes hablan lenguas prácticamente indistinguibles de las lenguas que se

⁹ Cf. A. Bilgrami, *Belief and Meaning*, Blackwell, Oxford, 1992, 214-216.

¹⁰ *Ibid.*, 2.

¹¹ H. Putnam, «The Meaning of 'Meaning'», en Id., *Mind, Language and Reality* (Philosophical Papers, vol. 2), Cambridge University Press, Cambridge, 1975, 215-271.

hablan en la Tierra. Podemos suponer que la Tierra Gemela es un duplicado, una réplica exacta de la Tierra, y que cada uno de nosotros tiene su doble en la Tierra Gemela, de modo que cada sujeto terrestre y su doble son físicamente iguales, molécula a molécula.

La única diferencia entre la Tierra y la Tierra Gemela es la siguiente: el líquido que en castellano-gemelo se denomina 'agua' no es en realidad H₂O, como sucede en la Tierra, sino una substancia de composición química distinta, cuya fórmula podemos abreviar como XYZ. (Cabría objetar que este supuesto es incompatible con la identidad física de cada sujeto terrestre y su gemelo, ya que el cuerpo humano está formado en gran parte por H₂O. Pero esta objeción es irrelevante, ya que es sencillo modificar el ejemplo para sortearla). Aparte de esta diferencia en la composición química, XYZ es indistinguible macroscópicamente del agua terrestre, sabe como agua, calma la sed y los habitantes de la Tierra Gemela la usan del mismo modo que en la Tierra se utiliza el agua. Para distinguirla del agua terrestre, llamaremos a XYZ 'gagua', aunque, como hemos indicado, los hablantes del castellano gemelo la llaman 'agua'.

Retrocedamos ahora hasta 1750, una época en la que la química moderna no se había desarrollado todavía, ni en la Tierra ni en la Tierra Gemela. El hablante terrestre de esta época no sabía que lo que él llamaba 'agua' era H₂O, como tampoco sabía el hablante gemelo que lo que él llamaba asimismo 'agua' era XYZ.

Comencemos ya a elaborar algunas consecuencias del experimento mental. Sea Oscar1 un habitante castellanohablante de la Tierra en la época indicada y Oscar2 su doble en la Tierra Gemela. Preguntémosnos ahora si el término 'agua' tiene el mismo significado en boca de Oscar1 y en boca de Oscar2. Parece que la respuesta ha de ser negativa. En el primer caso, 'agua' significa agua, mientras que, en el segundo caso, 'agua' significa gagua. En boca de Oscar1, la emisión 'el agua calma la sed' significa que el agua calma la sed, mientras que, en boca de Oscar2, significa que el gagua calma la sed. Las condiciones de verdad de ambas emisiones son distintas. En boca de Oscar1, 'el agua calma la sed' es verdadera si, y sólo si, el agua calma la sed, mientras que, en boca de Oscar2, dicha emisión es verdadera si, y sólo si, el gagua calma la sed. La aplicación de estas consecuencias al contenido intencional es clara. La creencia que Oscar1 expresa al emitir sinceramente las palabras 'el agua calma la sed' es la creencia de que el agua calma la sed, mientras que, en el caso de Oscar2, es la creencia de que el gagua calma la sed. Las condiciones de verdad de ambas creencias son distintas. Así, tratándose de creencias distintas, ambas pueden ser verdaderas a la vez. Es también claro cómo estas consideraciones podrían extenderse a los deseos, intenciones y otros estados mentales.

Si estas consecuencias son correctas, la premisa internalista según la cual entornos distintos pueden causar conjuntos de estados mentales absolutamente indistinguibles, tanto en su cualidad como en su carácter intencional, es falsa. La diferencia entre los estados mentales de Oscar1 y Oscar2 se debe *exclusivamente* a la diferencia entre los entornos objetivos de ambos sujetos, pues el ejemplo de Putnam está construido de tal modo que no tiene por qué haber ninguna otra diferencia entre ellos. Por lo tanto, frente a la concepción internalista, el entorno objetivo es constitutivo del contenido intencional. Y si esto es así, la argumentación escéptica descansa en un supuesto falso. En otro lugar Putnam explota de manera explícita el externalismo para rechazar la posibilidad de que fuésemos cerebros en una cubeta. *In nuce*, su argumentación es la siguiente: si realmente fuésemos cerebros en una cubeta, no podríamos pensar que lo somos; pero podemos pensarlo; luego no lo somos¹². El argumento es sorprendente, brillante en su (aparente) simplicidad.

¹² H. Putnam, *Razón, verdad e historia*, cit. *supra*, cap. 1.

La forma lógica del argumento es impecable: un *modus tollens*. Dado que la segunda premisa es obviamente verdadera, la cuestión decisiva reside en la verdad de la primera. El argumento en favor de ésta podría exponerse del modo siguiente: del mismo modo que Oscar2 no puede tener pensamientos sobre el agua, debido a la naturaleza del entorno causalmente responsable de sus pensamientos, ya que en ese entorno no hay agua, tampoco un cerebro en una cubeta puede tener pensamientos sobre cerebros y cubetas debido a que no hay tales cosas en el entorno causalmente responsable de sus pensamientos, que consta solamente de impulsos electroquímicos transmitidos por el ordenador. Parece claro que la verdad de la premisa depende por completo de la verdad del externalismo. Volvamos, pues, a la discusión de éste.

El experimento mental de Putnam es lo que Dennett denomina una "bomba de extraer intuiciones"¹³. Y no todas las intuiciones apuntan en la misma dirección. El escéptico tiene a su disposición diversas alternativas para resistirse a abandonar su posición. Podría incluso, como veremos, explotar el experimento mental de Putnam para *favorecer* su causa.

El escéptico podría contraatacar diciendo, en primer lugar, que el experimento de Putnam descansa sobre un concepto de creencia artificioso, construido expresamente para favorecer el externalismo. Sólo si se *parte* de que el entorno objetivo es constitutivo del contenido intencional de una creencia se puede llegar a la conclusión de que Oscar1 y Oscar2 creen cosas distintas. La conclusión descansaría en una *petitio principii*. La objeción, hasta aquí, no me parece correcta. El punto de partida de Putnam no es un compromiso previo con el externalismo. Lo que Putnam pretende sacar a la luz a través del experimento son los compromisos implícitos en nuestro concepto *cotidiano* de creencia, cualesquiera que éstos sean. Su punto de partida es este concepto cotidiano de creencia, no un concepto artificial elaborado sobre la base del externalismo como doctrina filosófica. La intuición que nos lleva a afirmar que Oscar1 y Oscar2 no creen lo mismo se limita a sacar a la luz el compromiso fuertemente externalista de dicho concepto, no lo introduce en él. El escéptico no puede rechazar sin más ese concepto cotidiano, sustituyéndolo por un concepto distinto, ya que sostiene que son nuestras creencias actuales, y no un sucedáneo de las mismas inventado *ad hoc*, las que podrían ser sistemáticamente erróneas. Si el escéptico elaborase un nuevo concepto de creencia, puramente internalista, para sostener a continuación que todas nuestras creencias, entendidas según ese concepto, podrían ser falsas, su posición perdería toda su fuerza e interés¹⁴. Ha de reconocer, pues, que el aspecto externalista está efectivamente presente en nuestro concepto cotidiano de creencia, retirando así su objeción inicial. Este reconocimiento, sin embargo, no obliga al escéptico a aceptar la verdad del externalismo y con él la falsedad de una premisa esencial para su posición. Puede simplemente considerar el aspecto externalista del concepto cotidiano de creencia como una excrecencia injustificada, fruto del instinto natural y no de la reflexión racional. El escéptico es consciente de que su posición no coincide con la actitud natural del ser humano, fuertemente antiescéptica. Pero esta actitud, al igual que el aspecto externalista del concepto de creencia, estrechamente relacionado con ella, puede ser un resultado de la "desidia" que, en palabras de Descartes, "me arrastra insensiblemente hacia mi manera ordinaria de vivir"¹⁵,

¹³ D. Dennett, *Elbow Room. The Varieties of Free Will Worth Wanting*, Clarendon Press, Oxford, 1984, 17-18.

¹⁴ Cf., para una defensa de los argumentos escépticos frente a una objeción semejante aplicada al concepto de conocimiento, B. Stroud, *El escepticismo filosófico y su significación*, trad. de Leticia García, Fondo de Cultura Económica, México, 1991, cap. 2.

¹⁵ R. Descartes, *op. cit.*, 21.

desidia que nos aparta instintivamente de la duda escéptica y que, por lo mismo, nos lleva a considerar el mundo externo como constitutivo del contenido de nuestras creencias.

Esta respuesta proporciona al escéptico un momentáneo punto de reposo desde el que puede iniciar una contraofensiva más ambiciosa, explotando paradójicamente el experimento mental de Putnam en favor del internalismo. Lo que nos lleva a decir que Oscar1 y Oscar2 no comparten sus creencias es, no sólo el aspecto externalista de nuestro concepto de creencia, sino la posición externa que adoptamos ante la historia. Puesto que nosotros sabemos que las sustancias que ambos sujetos llaman 'agua' son distintas, les *atribuimos* creencias distintas. Pero ellos no lo saben. Si Oscar1 y Oscar2 llegaran a encontrarse en uno de los dos planetas, descubrirían asombrados no sólo su completo parecido físico, sino también su plena concordancia en todas sus creencias. Descubrirían que ambos llaman 'agua' a (lo que aparentemente es) la misma sustancia y que coinciden en todas sus opiniones sobre dicha sustancia. Ambos creen, por ejemplo, que la sustancia que ambos llaman 'agua' calma la sed. Supongamos que el externalismo es, a pesar de todo, correcto. En ese caso, todas las creencias relacionadas con la sustancia en cuestión de uno de los dos sujetos serían sistemáticamente falsas (recordemos que los dos se hallan ahora en el mismo planeta), sin que el sujeto en cuestión tuviera posibilidad alguna de descubrir tal cosa. Si el encuentro tiene lugar en la Tierra, serían falsas las creencias de Oscar2. Si el lugar de encuentro es, en cambio, la Tierra Gemela, sería Oscar1 el infortunado.

Así, pues, si el externalismo es correcto, no sólo nuestros sujetos podrían tener creencias sistemáticamente falsas sobre el mundo externo, sino que ni siquiera serían capaces de saber exactamente qué es lo que creen (pues cada uno de ellos piensa que cree lo mismo que el otro cuando no es así). El externalismo no es sólo incapaz de atajar el escepticismo sobre el mundo externo, sino que lo extiende al contenido de la propia mente. Este es un argumento que nos debería llevar a rechazar el externalismo y a abrazar el internalismo. Este no excluye el escepticismo sobre el mundo externo, pero al menos mantiene la autoridad cognoscitiva del sujeto sobre sus propios estados mentales, la certeza indudable del *Cogito*. El sujeto sabe al menos cuáles son sus creencias, aunque no pueda saber si son verdaderas.

El propio Putnam hace una importante concesión al internalismo al admitir que el terrestre y su gemelo comparten sus estados mentales *estrechos*, es decir, caracterizados de modo no relacional sino sólo en términos de lo que hay en sus respectivas cabezas, por más que no compartan sus estados mentales *amplios* o externalistas. La distinción de Putnam parece tener dos objetivos principales. En primer lugar, pretende salvaguardar un ámbito mental en el que la autoridad del sujeto quede respetada, a saber, los estados mentales caracterizados de forma puramente internalista, con lo que Putnam parece aceptar la idea de que el externalismo constituye una amenaza a la autoridad de la primera persona. En esta aceptación, Putnam coincide no sólo con los internalistas, sino también con algunos partidarios del externalismo¹⁶. En segundo lugar, la distinción trata de preservar cierta concepción de la explicación psicológica basada en el "solipsismo metodológico" (Putnam), según la cual las propiedades relacionales (externalistas) de los estados mentales son irrelevantes para dar cuenta de la conducta y sólo lo que hay en la mente o en el cerebro puede ser legítimamente usado con fines explicativos. Los estados mentales

¹⁶ Cf., por ejemplo, A. Woodfield, «Foreword», en Id. (ed.), *Thought and Object. Essays on Intentionality*, Clarendon Press, Oxford, 1982, reimpr. 1984, v-xi; J. McDowell, «Singular Thought and the Extent of Inner Space», en Ph. Pettit y J. McDowell (eds.), *Subject, Thought and Context*, Clarendon Press, Oxford, 1986, 137-168.

amplios, individuados en términos externalistas, carecen de valor explicativo. Sólo los estados mentales estrechos pueden tener dicho valor. Así, el externalismo, que atribuye a Oscar1 y a Oscar2 creencias *distintas*, es incapaz de explicar por qué actúan del mismo modo, como *ex hypothesi* lo hacen. Necesitamos el internalismo para dar cuenta de esta similitud en la conducta de ambos 'gemelos'. Así, pues, los internalistas puede utilizar la supuesta inanidad explicativa de los estados mentales individuados según criterios externalistas como un argumento adicional en favor de su posición y en contra del externalismo como concepción de la mente.

Tras este contraataque escéptico, el externalismo, y con él sus presuntas consecuencias antiescéticas, se hallan en una difícil situación. El balance, tras la confrontación, es bastante negativo para él. A su favor, el externalismo sólo tiene su capacidad de explicar, al menos parcialmente, las intuiciones que nos llevan a atribuir creencias distintas a Oscar1 y a su gemelo. El internalismo ha reconocido este punto (o debería hacerlo), como vimos, aceptando que el concepto cotidiano de creencia contiene un componente externalista. En su contra, sin embargo, el externalismo tiene las objeciones siguientes. En primer lugar, no impide que un sujeto pueda tener creencias sistemáticamente falsas sobre el mundo, de tal modo que, además, le sea imposible descubrir que las tiene. No supone, pues, ninguna ventaja frente al escepticismo sobre el mundo externo. En segundo lugar, conlleva la negación de la autoridad cognoscitiva del sujeto sobre sus propios estados mentales, ya que individúa éstos en términos de objetos externos cuya naturaleza real puede ser desconocida para el sujeto. Con ello, extiende el terreno sometido a la duda escéptica, incluyendo en él, no sólo el mundo externo, sino también el mundo interior. En tercer lugar, los estados mentales, individuados en términos externalistas, resultan ociosos desde el punto de vista explicativo. Oscar1 y Oscar2, siendo internamente idénticos, actúan también del mismo modo en lo que se refiere a lo que ambos llaman 'agua', siendo irrelevante que la sustancia así llamada sea distinta en uno y otro caso. La individuación internalista de los estados mentales, que atribuye las mismas creencias a ambos sujetos, da cuenta de la similaridad de su conducta, mientras que la individuación externalista hace de esta similaridad un completo misterio.

El externalismo ha de hacer frente a estas objeciones o abandonar la escena. En lo que resta del presente trabajo, examinaremos si tiene recursos para defender sus posiciones y, con ello, para sentar las bases de una sólida estrategia antiescética.

3. Externalismo y escepticismo.

El internalista sostenía, como vimos, que el externalismo no excluía la hipótesis escéptica según la cual nuestras creencias sobre el mundo externo podrían ser sistemáticamente falsas sin que tuviésemos posibilidad alguna de descubrir que lo son. Así, si suponemos que Oscar2 es trasladado a la Tierra, sus creencias sobre lo que llama 'agua' serían todas falsas, ya que su término 'agua' significa gagua, y no agua, y esto último es lo que debería significar para que tales creencias fuesen verdaderas. Lo contrario sucedería si Oscar1 es trasladado a la Tierra Gemela. Y, esto es lo importante, ni el uno ni el otro tienen posibilidad alguna de descubrir su error.

Considero que debe admitirse que, en el supuesto considerado, las creencias del sujeto trasladado serían falsas. Imaginemos que, mientras estoy ausente de mi habitación, el teléfono que hay sobre mi mesa es sustituido por una imitación externamente perfecta, pero carente de todo mecanismo interior. Si admitimos que un objeto así no es un teléfono, mis creencias relativas al teléfono que hay sobre mi mesa serían, a partir de ese momento, falsas, puesto que no hay ya un teléfono sobre mi mesa. Este eventualidad, ciertamente concebible, no constituye, sin embargo, una seria amenaza escéptica, puesto que me sería posible descubrir que mis creencias son falsas:

por ejemplo, cuando intentase efectuar una llamada desde el nuevo 'teléfono' o cuando los técnicos de la compañía telefónica viniesen a revisarlo.

Para mostrar que nuestras consideraciones no dependen esencialmente de supuestos de ficción científica, relativas a hipotéticos mundos gemelos en los que se hablan lenguas que logramos entender, cabe recurrir a un ejemplo posible que involucra únicamente nuestro mundo y las lenguas habladas en él. Supongamos que en Chile hay animales a los que los chilenos llaman 'gatos', muy similares externamente a los gatos en España, pero que de hecho constituyen una especie diferente, de modo que los gatos chilenos y los gatos españoles no pueden tener descendencia viable. Si me traslado a Chile, mis creencias acerca de los 'gatos' que allí veo, en la medida en que dependen del supuesto de que pertenecen a la misma especie que los gatos españoles, serían falsas. Cabría incluso la posibilidad de que nunca saliese de mi error. Ahora bien, *podría* salir de él consultando a los expertos correspondientes. Y si un español se hubiese trasladado a Chile hace un siglo, antes del desarrollo de la genética moderna, la imposibilidad de descubrir su error habría sido sólo contingente.

A nuestro parecer, el caso de los gemelos es esencialmente del mismo tipo que los dos ejemplos anteriores. La imposibilidad en que se halla nuestro sujeto (digamos, Oscar2) de descubrir la falsedad de sus creencias no es una imposibilidad metafísica o necesaria, sino sólo contingente. Si la química moderna se hubiera desarrollado en ambos planetas antes de su traslado a la Tierra, de modo que Oscar2 supiera que el líquido al que en su planeta llaman 'agua' es XYZ, sería perfectamente capaz de descubrir que en la Tierra no había XYZ y que sus creencias hasta ese momento habían sido erróneas, en la medida en que descansaban sobre el supuesto falso de la identidad entre las dos sustancias llamadas 'agua' en la Tierra y en la Tierra Gemela. Pero lo que el escepticismo requiere, para sacar adelante sus tesis, es una imposibilidad metafísica, de principio, no una imposibilidad meramente práctica, dependiente sólo de una contingencia histórica. El escepticismo no sostiene que nosotros nos hallemos en una situación epistémica más favorable que los antiguos mesopotámicos, por ejemplo. Para él, la posibilidad conceptual del error sistemático depende de la condición epistémica del ser humano, una condición insuperable mediante el desarrollo de la ciencia.

4. El externalismo y la autoridad de la primera persona.

La segunda objeción contra el externalismo, recordémoslo, era que éste se veía obligado a negar la especial autoridad que un sujeto tiene sobre los contenidos de su propia mente, en la medida en que hacía depender éstos del entorno objetivo, sobre cuya naturaleza el sujeto no tiene autoridad especial alguna.

Es importante distinguir la cuestión relativa a la autoridad de la primera persona (una cuestión tal vez mejor conocida bajo el rótulo 'autoconocimiento') de la disputa entre el internalismo y el externalismo como teorías del contenido mental. Desde la obra de Descartes, el internalismo ha solido recurrir a la autoridad de la primera persona (en el caso de Descartes, a la verdad indudable de juicios del tipo 'yo pienso') en apoyo de su posición, como si la autoridad de la primera persona sólo pudiese ser explicada sobre la base del internalismo. Algunos externalistas han aceptado también este último supuesto, viéndose así llevados a negar la autoridad de la primera persona¹⁷. Así, Andrew Woodfield escribe: "Una tercera persona podría muy bien estar en mejor

¹⁷ Este no es ciertamente el caso de Davidson, quien defiende tanto el externalismo como la autoridad de la primera persona, y trata de mostrar que no existe incompatibilidad entre ambos.

posición que el sujeto para saber en qué objeto está pensando éste, y por lo tanto en mejor situación, en este respecto, para saber cuál era ese pensamiento"¹⁸. Algunos externalistas parecen considerar la crítica a la autoridad de la primera persona como parte de su ataque al internalismo de raíz cartesiana. Desde nuestro punto de vista, se hallan en un grave error. La llamada 'autoridad de la primera persona' no es, creemos, obra de una teoría filosófica sobre la mente, sino la expresión de un fenómeno prefilosófico que ocupa un lugar de central importancia en nuestra vida y en la concepción de nosotros mismos como personas. Descartes no crea esa autoridad, sino que la *descubre*.

Es justo reconocer, sin embargo, que otros partidarios del externalismo han reconocido la independencia entre ambas cuestiones¹⁹, señalando que el externalismo es compatible con la autoridad de la primera persona.

Por nuestra parte, quisiéramos sostener una tesis más fuerte, a saber, que el externalismo no es sólo compatible con la autoridad de la primera persona, sino que puede dar cuenta positivamente de ella, e incluso hacerlo de forma más plausible que el internalismo²⁰.

Cuando el internalismo acusa al externalismo de destruir la autoridad de la primera persona lo hace desde el supuesto, implícito o explícito, de que sólo es posible dar cuenta de dicha autoridad sobre la base de un acceso privilegiado a los propios contenidos mentales que en último término ha de tener carácter *introspectivo*. Así, como la introspección no puede practicarse, por definición, sobre objetos externos, hacer de éstos un aspecto constitutivo del contenido mental imposibilita la introspección y, con ello, la autoridad de la primera persona.

El supuesto es, creemos, falso. Para verlo, construiremos un ejemplo en el cual el contenido del deseo de un sujeto involucra claramente un objeto externo. Con ello no estamos cometiendo una petición de principio contra el escepticismo, al suponer que hay objetos externos a los que tenemos acceso. Lo único que pretendemos mostrar es que la autoridad del sujeto sobre sus contenidos mentales no resultaría amenazada *si* hubiera objetos externos y éstos fuesen constitutivos del contenido de nuestros estados mentales. Vayamos al ejemplo.

Imaginemos que un coleccionista de arte pasea entre los puestos de un mercado de viejo, en el que ha encontrado, en ocasiones, piezas de interés. Ese día no ha descubierto, hasta el momento, nada digno de mención. De repente, su expresión, algo ausente, se transforma. Sus ojos se iluminan. En uno de los puestos, no lejos de él, un objeto ha llamado poderosamente su atención. Se trata de un grabado en el que le ha parecido reconocer ciertos rasgos del estilo de Dürero. Con

Cf. D. Davidson, «First Person Authority», *Dialectica* 38 (1984), 101-111 y «El conocimiento de la propia mente», en Id., *Mente, mundo y acción*, trad. de Carlos Moya, Paidós/UAB, Barcelona, 1992, 119-152.

¹⁸ A. Woodfield, *op. cit.*, viii.

¹⁹ Además de Davidson, muestran también una clara conciencia de ello T. Burge y C. Wright. Cf., del primero, «Individualism and Self-Knowledge», *Journal of Philosophy* 85 (1988), 649-663 y del segundo su crítica al libro de C. McGinn *Wittgenstein on Meaning*, *Mind* 98 (1989), 289-305.

²⁰ Los argumentos desarrollados a continuación se basan en dos ponencias no publicadas, «Mind and the First Person» y «Intentional Attitudes: First and Third Person», presentadas respectivamente al Primer Congreso Europeo de Filosofía Analítica (Aix-en-Provence, 1993) y al Tercer Simposio Internacional de Ciencia Cognitiva (San Sebastián-Donostia, 1993).

paso rápido se dirige hacia el puesto en cuestión, toma el grabado y empieza a examinarlo atentamente. Dejaremos al lector en suspenso acerca de si se trataba realmente de un Durero.

Nuestro sujeto desea o pretende examinar el grabado. No creemos que sea controvertido decir de él también que sabe que tiene ese deseo o esa intención. Si le preguntamos, mientras se dirige hacia el puesto callejero, por qué camina hacia allí, nos dirá sin vacilar (a menos que nos considere posibles competidores) que desea examinar el grabado. Ahora bien, ¿cómo sabe que es eso lo que desea? Esta pregunta es confundente. Preguntar *cómo* lo sabe puede sugerir que el sujeto ha estado desarrollando algún tipo de investigación para llegar a saberlo, lo que no parece ser el caso. En una primera aproximación, cabe responder negativamente a la pregunta diciendo, plausiblemente, que en cualquier caso *no lo sabe por introspección*. Precisamente, el interés que el objeto ha despertado en él ha hecho que su atención se concentre *en el objeto mismo*, no en sus procesos mentales. Difícilmente puede practicarse la introspección cuando la atención del sujeto se ve absorbida por un objeto o acontecimiento externo. La respuesta podría tal vez discurrir en los siguientes términos: el sujeto sabe que desea (examinar) el objeto sencillamente porque *ese objeto* ha absorbido todo su interés y atención, porque no ha dejado de ocupar su mente mientras se dirigía hacia él.

Comparemos ahora esta posición del sujeto con respecto al objeto de su deseo con la posición de un observador externo y tendremos las bases de la autoridad de la primera persona en un marco externalista. Mientras que la pregunta "¿Cómo sabe el sujeto lo que desea?" resulta extraña o no pertinente (debido en parte a que su interés está precisamente absorbido por el objeto de su deseo), la pregunta "¿Cómo sabemos nosotros lo que el sujeto desea?" no resulta extraña ni impertinente. Supongamos que estamos observando a nuestro coleccionista de arte sin que éste lo advierta. En un momento dado, percibimos un cambio en su expresión y en su conducta. Su cansino deambular se convierte en un caminar resuelto y rápido. ¿A qué se debe ese cambio? Tratamos de seguir la dirección de su mirada y de sus pasos. Tal vez se trata de algo que hay en aquel puesto de venta. En efecto, hacia allí se encamina. Cuando vemos que nuestro sujeto toma en sus manos el grabado, fijamos ya el objeto de su deseo o de su intención.

Así, mientras que nosotros empleamos la conducta del sujeto como evidencia de lo que desea, el sujeto mismo no emplea su propia conducta como evidencia de lo que desea. En realidad, no podría hacerlo. Si el sujeto observase su propia conducta a fin de descubrir lo que desea, esto supondría precisamente que no hay nada que descubrir. Que desea el objeto se manifiesta precisamente en que atiende a él, no a su propia conducta. Al hacerlo así, el sujeto crea la evidencia conductual en que nosotros nos basamos para descubrir lo que desea, pero él mismo no utiliza esa conducta como evidencia de lo que desea. Lo que desea se halla frente a él, centrando su atención y guiando su comportamiento. Nuestro juicio sobre el contenido de su deseo, formulado *a posteriori* sobre bases conductuales, presupone que ese contenido ya está fijado para el sujeto.

Si el internalista insiste en que no comprende cómo el sujeto puede saber lo que desea si no es a través de la introspección, tendremos que responderle de este modo: el sujeto no puede saber lo que desea por introspección porque, si comienza a practicarla, el fenómeno que trata de captar, su deseo *hacia* el objeto, queda con ello destruido²¹. La insistencia del internalista en la introspección pasa por alto un fenómeno muy importante acerca del deseo en los seres humanos, a saber, que el deseo de un objeto dirige la atención y el interés del sujeto hacia el objeto deseado, no hacia sus propios estados o procesos internos. Y esto es así aun cuando el objeto ya no esté

²¹ Debo a Harold Noonan esta forma de expresar mi posición.

presente. Si nuestro coleccionista, de vuelta a casa, sigue pensando en el grabado y deseando poseerlo, lo que desea y aquello en lo que piensa es el grabado mismo, no su representación interna de éste. Tampoco ahora sabe lo que desea por introspección, pues lo que desea no está en su interior y no puede hallarlo allí. He aquí el núcleo del externalismo.

No es claro que las consideraciones aquí desarrolladas puedan extenderse a todos los contenidos mentales. Tal vez sea posible. Sin embargo, nuestra preocupación no era exactamente ésta. Lo único que pretendíamos era poner de manifiesto, mediante un ejemplo restringido al deseo, pero representativo de muchos otros casos, que el externalismo, el supuesto de que el objeto externo es constitutivo del contenido mental, no sólo es compatible con la autoridad de la primera persona sobre sus propios contenidos mentales, sino que puede ofrecer una explicación plausible de la misma. Más aún: si los supuestos de esta explicación son correctos, es más bien el internalismo el que tiene problemas para dar cuenta de la autoridad de la primera persona.

5. Contenido externalista y relevancia causal.

La tercera objeción a la que el externalismo debe hacer frente puede resumirse así: Oscar1 y Oscar2 han de tener las mismas creencias, puesto que actúan del mismo modo en relación con lo que ambos llaman 'agua'. Puesto que el externalismo les atribuye creencias distintas, el externalismo debe ser erróneo o, en cualquier caso, esas 'creencias' carecen de realidad psicológica y de relevancia causal. Las creencias, o al menos las creencias psicológicamente reales y explicativamente relevantes, son, pues, constitutivamente independientes del entorno. El externalismo, al atribuir a ambos sujetos creencias distintas, hace de la similaridad de su conducta un misterio. En cambio, el internalismo, puesto que les atribuye las mismas creencias, da cuenta perfectamente de dicha similaridad.

Hay conexiones importantes entre esta objeción al externalismo y determinados intentos de sentar las bases de una psicología científica. Así, por ejemplo, de acuerdo con lo que Stephen Stich ha denominado el "principio de autonomía psicológica", dos seres humanos (o cualesquiera otros organismos) que sean réplicas exactas, molécula a molécula, el uno del otro (como sucede con nuestros Oscar1 y Oscar2) habrán de ser considerados como psicológicamente idénticos, de modo que "cualquier propiedad psicológica ejemplificada por uno de estos sujetos será también ejemplificada por el otro"²². El principio de autonomía psicológica tiene una clara afinidad con el internalismo. De hecho, el tipo de psicología inspirado en él constituiría una especie de paradójica alianza entre el internalismo de raíz cartesiana y el materialismo. Así, otra forma de exponer este principio consistiría en sostener "que las propiedades y relaciones que una teoría psicológica explicativa ha de invocar han de sobrevenir a [esto es, han de estar determinadas o fijadas por, C. M.] las propiedades y relaciones *físicas actuales, internas* de los organismos (esto es, precisamente aquellas propiedades que un organismo comparte con todas sus réplicas)"²³.

La objeción tiene como punto de partida la función causal y explicativa de las creencias con respecto a la conducta de un sujeto. La objeción insiste en separar esta función de las creencias de su función como portadoras de verdad o falsedad, como representaciones de cómo son las cosas. Las propiedades causales de las creencias tienden a aislarse de sus propiedades semánticas. La

²² S. Stich, «Autonomous Psychology and the Belief-Desire Thesis», en W. G. Lycan (ed.), *Mind and Cognition*, Blackwell, Oxford, 1990, 346.

²³ *Ibid.*, 347. (Énfasis en el original).

radicalización de esta separación ha llevado a algunos filósofos²⁴ a la tesis de que el concepto ordinario de creencia, que incluye ambos aspectos, es inconsistente y debe ser reemplazado por un concepto distinto. En una línea afín se hallan las propuestas de escindir la noción de contenido y de asignar las propiedades causales al contenido estrecho, reducido a lo que hay en la cabeza del sujeto, y las propiedades semánticas al contenido amplio, individuado por las relaciones con el entorno. Las creencias comunes que el internalismo atribuye a Oscar1 y a Oscar2 serían creencias cuyo contenido es estrecho.

Es difícil ver, sin embargo, cómo esta separación puede dar cuenta de la especificidad que caracteriza la eficacia causal de la creencia. Las creencias causan nuestra conducta en la medida en que las consideramos como representaciones verdaderas de cómo son las cosas y, por tanto, en la medida en que les atribuimos capacidad semántica. En la concepción puramente internalista del contenido, la capacidad representativa de las creencias se convierte en un dato bruto, carente de explicación²⁵.

Es importante distinguir la objeción que estamos examinando de una consideración plenamente correcta con la que podría confundirse. Oscar1 y Oscar2 (como el resto de nosotros) actúan según *creen* que son las cosas, sean éstas así o no. Una creencia falsa causa una conducta tan eficazmente como una verdadera, siempre que el sujeto la tenga por verdadera. Esta consideración, sin embargo, no sustenta la idea de que las propiedades semánticas de las creencias son irrelevantes para su relación causal con la conducta. Pensar de otro modo supone confundir el valor de verdad que una creencia tiene con su capacidad de tener un valor de verdad (verdadero o falso), con su capacidad de representar las cosas. Por otra parte, la consideración de que hablamos no representa una amenaza para la tesis externalista según la cual el entorno es constitutivo del contenido de las creencias. 'Agua' significa gagua en el caso de Oscar2 y la creencia que él expresa diciendo 'el agua calma la sed' es la creencia de que el gagua calma la sed. Esto es así en virtud del entorno en el que Oscar2 se encuentra. Pero si Oscar2 cree (falsamente) que tiene ante sí gagua y tiene sed, beberá. Un contenido externalista es perfectamente compatible, pues, con la consideración de que actuamos según creemos que son las cosas, sean éstas así o no.

Internalistas y externalistas deben aceptar la consideración anterior. A partir de aquí, sin embargo, el internalista puede construir su objeción del siguiente modo: tanto Oscar1 como Oscar2 actúan según creen que son las cosas, sean éstas así o no; ahora bien, uno y otro forman estas creencias sobre la base de cómo les *parece* que son las cosas; en sus respectivos entornos, a ambos les parece que las cosas son del mismo modo; por lo tanto, las creencias que formarán serán también las mismas; la naturaleza real de su entorno es irrelevante para las creencias que poseen; las creencias (distintas) atribuidas sobre una base externalista no son psicológicamente reales y, por ello, no son causalmente relevantes.

Mediante esta reformulación, el internalista ha conseguido fortalecer su objeción al incorporar a ella la capacidad representativa de las creencias mediante la alusión a cómo le *parece* a un sujeto que son *las cosas*, al tiempo que sostiene que el modo en que las cosas *son* es irrelevante para la apariencia que presentan a un sujeto y para las creencias que forma sobre la base de esa apariencia. El escenario de la Primera Meditación cartesiana se perfila de nuevo ante nosotros.

²⁴ Cf., por ejemplo, S. Stich, *From Folk Psychology to Cognitive Science*, MIT Press, Cambridge Mass., 1983.

²⁵ Cf. J. McDowell, *op. cit.*, 152, 158.

La cuestión es si la apelación a las apariencias ha de forzarnos a retroceder al internalismo. La concepción internalista de la apariencia ha de hacer de la referencia a las cosas un rasgo prescindible de la apariencia. Si su argumento ha de funcionar, la apariencia ha de ser un estado interno, especificable sin referencia alguna al entorno externo²⁶. Pero es dudoso que esto pueda hacerse. La apariencia que una cosa presenta a un sujeto es una propiedad *de* la cosa misma. Si me parece que un perro es un lobo, es el *perro* el que me parece ser tal. La apariencia de lobo es una propiedad del perro. Esta dependencia del modo como las cosas le parecen ser a un sujeto con respecto a las cosas que le parecen ser de ese modo hace que la especificación de la apariencia no pueda llevarse a cabo sin referencia a la naturaleza real de las cosas. El internalista supone que el agua y el gagua, aunque son realmente distintas, presentan la misma apariencia para Oscar1 y Oscar2. Y en algún sentido ha de tener razón, dado el modo en que ha sido construido el experimento mental. Pero consideremos la cuestión más de cerca. Supongamos que Oscar1 se halla ante un vaso de agua y Oscar2 ante un vaso de gagua. ¿Cómo podemos especificar, desde el punto de vista de los propios sujetos, cómo les parece que son las cosas? No podemos decir que a ambos les parece que el líquido es agua, pues esto sólo es cierto de Oscar1. A Oscar1 le parece, en efecto, que *ese* líquido que hay en el vaso es del *mismo* tipo que *el que* bebió ayer (es decir, agua). Sobre esa base, pasa a creer que ese líquido es agua. A Oscar2 le parece asimismo que *ese* líquido que hay en el vaso es del *mismo* tipo que *el que* bebió ayer (es decir, gagua). Sobre esta base, pasa a creer que ese líquido es gagua. Es decir, tanto Oscar1 como Oscar2 han de especificar cómo les parece que son las cosas sobre la base de cómo son las cosas. No necesitan, sin embargo, *saber cómo son realmente*. La referencia demostrativa contenida en ambas especificaciones introduce en ellas la naturaleza real de uno y otro líquido, de modo que ambas especificaciones son distintas ya que el líquido es distinto en uno y otro caso.

La noción de apariencia que el internalista necesita para su propósito de mostrar que ambos sujetos comparten las mismas apariencias y, *con ellas*, las mismas creencias no ha de contener esencialmente una referencia a objetos. Lo que el internalista quiere decir es que ambos sujetos comparten entidades puramente internas, tal vez datos sensoriales. El fracaso de la reducción fenomenista hace pensar que no es posible especificar las apariencias concebidas de ese modo. Pero, aunque fuese posible, una apariencia así entendida es bastante menos de lo que se necesita para dar lugar a una creencia y, *a fortiori*, a las mismas creencias en ambos. Imaginemos que contemplo una fotografía de Franco saludando a Millán Astray. Podemos suponer que, si existen datos sensoriales, un ser dotado de un aparato perceptivo similar al mío comparte conmigo tales datos. Pero eso no garantiza en modo alguno que comparta conmigo la creencia de que, en cierta ocasión, Franco saludó a Millán Astray. Tampoco mi doble en la Tierra Gemela compartiría conmigo esa creencia.

El internalismo insiste, sin embargo, en que, si Oscar1 y Oscar2 no comparten las mismas creencias, resulta inexplicable que desarrollen la misma conducta. Una primera respuesta consistiría en poner en cuestión precisamente que desarrollen la misma conducta. Si Oscar1 tiene sed y tiene ante sí un vaso de agua, beberá de él. Si Oscar2 tiene sed y tiene ante sí un vaso de gagua, beberá de él. Pero beber agua y beber gagua son dos conductas distintas, tan distintas como vender (o comprar) oro verdadero y vender (o comprar) oro falso. Y ambas conductas se explican sobre la base de creencias distintas.

²⁶ Me inspiró aquí, modificándola libremente, en una sugerencia de J. McDowell (*Ibid.*, 151, 157).

El internalista insistirá, con razón, en que no hacemos justicia a lo que quiere decirnos: que, en algún sentido, ambos hacen lo mismo y tienen las mismas creencias. Concedamos, en efecto, que, para ciertos propósitos explicativos, no llevamos a cabo la clasificación de la conducta y de las creencias mediante una malla tan fina²⁷. Las conductas de Oscar1 y Oscar2 pueden, por ejemplo, subsumirse bajo el tipo: 'beber del vaso que uno tiene ante sí'. Y sus respectivas creencias pueden subsumirse bajo el tipo: 'creer que el líquido contenido en el vaso que uno tiene ante sí calma la sed'. El internalista insistirá en que los agentes pueden tener creencias explicativamente relevantes subsumibles bajo ese tipo, con independencia de la naturaleza real de sus respectivos entornos.

¿Cuáles son las creencias particulares respectivas de Oscar1 y Oscar2 que subsumimos bajo ese tipo? En el caso de Oscar1, sería la creencia que él podría expresar diciendo: "el líquido contenido en este vaso calma la sed". En el caso de Oscar 2, sería la creencia que él expresaría con las mismas palabras. La cuestión, sin embargo, es que el contenido de las creencias que uno y otro expresan con las mismas palabras es distinto, prescindiendo incluso de la distinta naturaleza del líquido en cada caso. La creencia de Oscar1 es verdadera si, y sólo si, el líquido contenido en *el vaso que hay delante de él* calma la sed y la creencia de Oscar2 es verdadera si, y sólo si, el líquido contenido en *el vaso que hay delante de él* calma la sed. Puesto que los vasos son distintos, las condiciones de verdad son distintas. La diferencia se debe al elemento demostrativo incluido en las creencias ('este vaso'), que determina referencias distintas en uno y otro caso. La cuestión importante, sin embargo, es que no parece posible retroceder hacia contenidos de creencias que sean causalmente eficaces y se hallen *a la vez* libres de elementos demostrativos. Oscar1 puede estar sediento, puede creer, frente a un vaso de agua, que el agua calma la sed y no hacer nada a menos que crea, además, que *ese* líquido, o el líquido que hay en *ese* vaso (o sencillamente *eso*) es agua²⁸. El elemento demostrativo estaba también incluido en la especificación del tipo de creencia: 'creer que el líquido contenido en *el vaso que uno tiene ante sí* calma la sed'.

Pero si las creencias de Oscar1 y Oscar2 no tienen el mismo contenido (individuado por sus condiciones de verdad), ¿qué hay entonces en común entre sus creencias que les lleva a beber del vaso que cada uno de ellos tiene ante sí? Si el externalismo no puede responder a esta pregunta, su posición se ve debilitada. La respuesta podría ser la siguiente. Al tener sus respectivas creencias, ambos comparten ciertos componentes doxásticos, ciertas sub-creencias que incluyen predicados comunes, a saber, que algo es líquido y calma la sed. El externalista puede aceptar esto sin dificultad, dado que en el entorno de ambos sujetos hay sustancias líquidas y ambos han aprendido, en sus respectivos entornos sociales, el significado de 'calmar la sed'. El internalista, aunque por razones distintas, ha de aceptar también esta comunidad de componentes de ambas creencias. Estas sub-creencias son causalmente inertes por sí mismas. Pero ambos comparten además un modo de pensamiento demostrativo o singular acerca de un objeto (en cada caso distinto) al que atribuyen los predicados anteriores. Es la combinación de ambos aspectos, el componente general o descriptivo y el modo demostrativo de referencia al objeto lo que ambos comparten y lo que lleva a ambos a beber de los vasos respectivos.

Si lo dicho hasta aquí es correcto, el externalismo ha respondido con éxito a la objeción considerada en esta sección. Ha conseguido explicar razonablemente la similitud en la conducta de

²⁷ Pero cf., para una malla aún más fina, H. Noonan, «Fregean Thoughts», en C. Wright (ed.), *Frege. Tradition and Influence*, Blackwell, Oxford, 1984, 23; y G. Evans, *The Varieties of Reference*, ed. por J. McDowell, Clarendon Press, Oxford, 1982, 203.

²⁸ Cf. J. Perry, «The Problem of the Essential Indexical», *Noûs* 13 (1979), 3-21.

ambos sujetos, en un determinado nivel de clasificación, sin necesidad de atribuirles las mismas creencias. El externalismo ha atribuido a ambos sujetos ciertos componentes comunes a las creencias (distintas) que ambos poseen, a saber, determinados predicados o conceptos (líquido, calmar la sed, etc.) y un modo demostrativo de pensar en un objeto al que atribuyen dichos predicados. Si la combinación de ambos componentes es causalmente eficaz, las creencias mismas que incluyen esos componentes lo son. El objeto externo interviene constitutivamente en estos componentes a través del modo demostrativo de referencia al mismo y como causa que desencadena ese modo de pensar, indispensable en la eficacia causal de la creencia. Veremos finalmente, en la última sección, si el externalismo, una vez que ha respondido satisfactoriamente a las objeciones presentadas, puede llegar algo más lejos.

6. Hacia un balance.

Los principales logros que el internalismo puede invocar en su favor, en especial su capacidad para explicar la autoridad de la primera persona y para dar cuenta de la eficacia causal de las creencias, son logros que el externalismo puede también invocar en favor de su propia causa. Con respecto al último, cabría señalar que el externalismo puede hacer también justicia a la intuición de que los estados mentales que explican la conducta han de ser internos al sujeto que desarrolla esa conducta. Inspirándonos en una penetrante indicación de Davidson²⁹, podemos decir que la inclusión constitutiva del entorno en el contenido de un estado mental o de un componente de éste no implica en modo alguno que el estado mental o el componente en cuestión no sean estados del propio sujeto. Oscar1 y Oscar2 piensan demostrativamente *en* el objeto que tienen ante sí, de modo que el objeto mismo es constitutivo de su pensamiento, pero este pensamiento demostrativo *acerca del* objeto es un estado o una actividad *de* nuestros sujetos.

En segundo lugar, alguien podría insistir en que considera contraintuitiva la atribución de creencias distintas a Oscar1 y Oscar2, dadas las condiciones del experimento mental de Putnam. En la respuesta a la tercera objeción, hemos seguido la estrategia consistente en mantener, sobre la base del externalismo, que, dado que los objetos en que piensan Oscar1 y Oscar2 son distintos (agua y gagua, respectivamente), sus creencias son también distintas. Sin embargo, el externalismo no está necesariamente obligado a seguir esta estrategia. Puede aceptar, incluso, que Oscar1 y Oscar2 comparten las mismas creencias³⁰ sin renunciar a su posición. Para ello, no necesita sino negar que el agua y el gagua sean sustancias distintas. A fin de cuentas, lo que nos lleva a considerarlas distintas son propuestas metafísicas sobre los tipos naturales y su individuación. Estas propuestas, aunque plausibles, no son, sin embargo, incontestablemente evidentes. Si las dejamos a un lado, podemos simplemente considerar H₂O y XYZ como dos clases de agua, dada la completa similitud entre sus funciones y apariencia externa.

En realidad, al aceptar que el agua y el gagua son sustancias distintas, aunque externamente similares, nos hemos situado en el terreno más desfavorable para el externalista y más propicio para el internalista. Este no puede acusarnos de trabajar sobre una hipótesis prefabricada en contra suya. Pensemos que los entornos respectivos de Oscar1 y Oscar2 son externamente indistinguibles para ambos y, en general, para un perceptor normal. Y si el externalismo ha podido mantener su posición bajo el supuesto indicado, cabe pensar que podrá hacerlo tanto más fácilmente en supuestos más favorables para él, como sucede en el ejemplo de

²⁹ D. Davidson, «El conocimiento de la propia mente», cit. *supra*, 143.

³⁰ Seguimos de nuevo una observación de Davidson, *Ibid.*, 139.

los cerebros en una cubeta. Pues si un perceptor normal entrase en la ominosa sala que contiene los cerebros conectados al ordenador, encontraría diferencias muy notables entre su propio entorno habitual y el entorno de los cerebros en cuestión. El escéptico, dado su compromiso con el internalismo, ha de sostener que un sujeto que habite fuera de la sala (llamémosle F) y uno de los cerebros del interior (al que llamaremos D) podrían ser psicológicamente indistinguibles.

El internalista debe aceptar que F y D pueden compartir no sólo conceptos generales, sino también un modo demostrativo de pensar, sistemáticamente ejercido, acerca de determinados objetos particulares que ambos sujetos consideran externos a ellos, puesto que F posee ese modo de pensar y lo ejerce de manera sistemática y D, según el internalista, sería psicológicamente igual a F. El internalista ha de mostrar, pues, que los pensamientos demostrativos son sistemáticamente posibles sin que exista ninguno de los objetos particulares sobre los que pretenden versar y, además, sin necesidad de que existan, en el entorno del sujeto de esos pensamientos, objetos del tipo al que, desde el punto de vista del sujeto, pertenece el supuesto objeto particular en el que piensa. Es decir, el internalista ha de mostrar que un sujeto puede tener pensamientos del tipo 'ese pájaro es precioso' sin que haya tal pájaro ni objeto particular alguno ante el sujeto y, además, sin que haya pájaros en el entorno del sujeto.

En este punto, el argumento obvio del escéptico es la apelación a las alucinaciones (los espejismos o las ilusiones perceptivas no son suficientes). Un sujeto que sufre una alucinación puede pensar 'ese pájaro es precioso' mirando hacia un lugar donde no hay tal pájaro ni otro objeto. Lo más que casos de este tipo podrían probar (no decimos 'prueban', por la razón que se verá) es que hay, de forma esporádica, pensamiento demostrativo sin objeto, lo cual es bastante menos de lo que necesita el internalista. Las alucinaciones no prueban que el pensamiento demostrativo pueda darse *sistemáticamente* en ausencia del objeto. El sujeto de la alucinación muestra una capacidad para pensar demostrativamente y para atribuir al supuesto objeto predicados generales (pájaro, precioso) y el internalista necesita probar que dicho sujeto puede tener esas capacidades sin haber tenido jamás percepciones verídicas de objetos particulares y sin que haya pájaros en su entorno. Y esto es algo que no ha probado.

De acuerdo con algunos autores, en especial Gareth Evans y John McDowell, un caso como el anterior ni siquiera prueba que el sujeto tenga un pensamiento demostrativo. Según estos autores, los pensamientos demostrativos son "russellianos" o dependientes del objeto, en el sentido de que no pueden ser pensados si el objeto particular en cuestión no existe³¹. El principal argumento de Evans en favor de esta tesis parece ser el siguiente: si no existe el objeto particular, no es posible expresar el contenido del supuesto pensamiento demostrativo; ahora bien, un pensamiento cuyo contenido no es posible expresar no tiene realmente contenido y un pensamiento sin contenido no es un pensamiento en absoluto. Pensemos que, en el caso anterior, no se puede decir: 'El sujeto piensa que *ese pájaro* es precioso', puesto que no hay tal pájaro³². Tampoco el propio sujeto podría expresar así lo que pensaba, por la misma razón. Por otra parte, el contenido de un pensamiento demostrativo no puede ser expresado sin pérdidas en términos puramente descriptivos.

La tesis de Evans y McDowell es digna de un examen detallado cuyo desarrollo excedería del espacio disponible aquí. Su carácter contraintuitivo ha alejado de ella incluso a algunos

³¹ Cf. J. McDowell, *op. cit.*, 137-146 y G. Evans, *op. cit.*, 80-85, 201-204.

³² Cf. G. Evans, *op. cit.*, 179 y P. Carruthers, «Russellian Thoughts», *Mind* 96 (1987), 21-22 para esta forma de exponer el argumento de Evans.

partidarios del externalismo³³. Una forma de oponerse a la misma es hallar algún modo de expresar el contenido del pensamiento demostrativo. Carruthers, por ejemplo, propone una forma de expresarlo que, aplicada a nuestro ejemplo, daría como resultado lo siguiente: 'El contenido del pensamiento del sujeto es el mismo que habría tenido si hubiera habido un pájaro real que causara sus experiencias y él hubiera pensado que ese pájaro era precioso'³⁴. Si esta descripción es correcta, el argumento de Evans no resulta concluyente. Más allá de esto, sin embargo, es importante que la expresión indirecta propuesta por Carruthers se lleve a cabo tomando como término de comparación el contenido de un pensamiento demostrativo basado en la percepción verídica de un objeto particular. Como Carruthers señala, sólo alguien que cree percibir un objeto real puede pensar seriamente en términos demostrativos³⁵. Esta tesis guarda alguna relación con la sugerencia de McDowell según la cual la idea de cómo son realmente las cosas forma parte de la comprensión de cómo nos parece que son³⁶. No parece, en efecto, que se pueda describir el contenido de un pensamiento demostrativo sin objeto real sin recurrir a la idea de percepción verídica de objetos reales.

El escéptico insistirá en que, en todo caso, esto prueba que necesitamos las *ideas* de percepción verídica, existencia de objetos particulares y otras relacionadas para expresar nuestros pensamientos demostrativos, pero no prueba que tengamos necesidad de la *realidad* de lo expresado por esas ideas. En este punto, sin embargo, el escéptico comienza a deslizarse por una peligrosa pendiente, ya que, si ha de hacer plausible su posición, necesita ofrecer *alguna* explicación de la existencia de esos pensamientos demostrativos. Ese modo demostrativo de pensar es literalmente ubicuo en nuestra economía psicológica e indispensable en la eficacia causal de nuestras creencias. Y la cuestión es que la existencia y ubicuidad de los pensamientos demostrativos resultan un obstáculo muy serio para alguien que ha de sostener que un cerebro en una cubeta puede ser psicológicamente indistinguible de un sujeto corpóreo normal. La superioridad de la explicación externalista frente a las posibles alternativas internalistas es aquí abrumadora. Esta superioridad no es ciertamente un argumento *decisivo* contra el internalismo, siempre que éste disponga de alguna explicación, por poco plausible que sea. Pero sospechamos que, por lo que se refiere al pensamiento demostrativo, la explicación internalista ha de resultar *especialmente* implausible. Un cerebro en una cubeta debería estar pensando con gran frecuencia demostrativamente, como todos lo hacemos, sin que hubiera objeto particular alguno ni objetos del tipo relevante en su entorno. Para explicar esto no basta con apelar a la condición cognitiva natural del ser humano. Para este tipo de engaño sistemático consistente en un constante apuntar demostrativamente a la nada necesitamos *literalmente* un poderosísimo agente engañador, un Genio Maligno o un Malvado Neurocirujano, que nos oculte arteramente cómo es el mundo y nos haga pensar constantemente en objetos particulares allí donde no los hay. Pues el pensamiento demostrativo no tiene función ni explicación natural alguna en el supuesto contexto de un cerebro en una cubeta. Ciertamente, no es la intención de Descartes o de Putnam sugerir que haya realmente un Genio Maligno o un Malvado Neurocirujano. Sin embargo, en la explicación de la existencia y

³³ Cf. H. Noonan, «Russellian Thoughts and Methodological Solipsism», en J. Butterfield (ed.), *Language, Mind and Logic*, Cambridge University Press, Cambridge, 1986, 67-90; P. Carruthers, *op. cit.* y A. Bilgrami, *op. cit.*, 10.

³⁴ P. Carruthers, *op. cit.*, 24.

³⁵ *Ibid.*, 25.

³⁶ Cf. J. McDowell, *op. cit.*, 151 y la sección 5 del presente trabajo.

ubicuidad del modo demostrativo de pensar nada inferior a la existencia de estos Seres parece suficiente. La alternativa a ello es avanzar resueltamente hacia el milagro y, en realidad, hacia el milagro perpetuo de que nos hallemos constantemente ocupados en un modo demostrativo de pensar carente de cualquier función inteligible. La posibilidad de la huida hacia el milagro hace que la posición escéptica no pueda ser definitivamente refutada³⁷. Sin embargo, esta explicación coloca al escéptico en una posición que debe calificarse, si no como estrictamente contradictoria, sí ciertamente como paradójica: pues, al tiempo que nos pide que desconfiemos de nuestras creencias relativas al mundo exterior, está solicitando de nosotros que creamos en el milagro³⁸. Así, pues, tal vez no sea casual que una de las motivaciones del escepticismo moderno fuese subrayar la debilidad de la razón humana y la necesidad del recurso a la revelación y a la fe³⁹.

Este trabajo ha estado concebido en su mayor parte como una *defensa* del externalismo frente a los ataques y objeciones del internalismo. Tal vez no esté de más, para concluir, que el externalismo invierta su estrategia, pasando a la ofensiva y solicitando del internalismo una explicación del carácter público del significado y del hecho de la comunicación entre los seres humanos. Si el entorno externo no es constitutivo del contenido de nuestros estados mentales, presumimos que tal explicación ha de resultar una tarea harto ardua⁴⁰. Aunque siempre cabe recurrir al milagro y, por qué no, a la armonía preestablecida⁴¹.

³⁷ Cf. A. Bilgrami, *op. cit.*, 219.

³⁸ Esta referencia al milagro se inspira en Bilgrami (*op. cit.*, 219, 225). El argumento aquí desarrollado sería, en la clasificación de este autor, un argumento de "ambición intermedia", como el argumento que en su libro desarrolla el propio Bilgrami.

³⁹ Cf. R. H. Popkin, *La historia del escepticismo desde Erasmo hasta Spinoza*, trad. de J. J. Utrilla, F.C.E., México, 1983, cap. 3.

⁴⁰ Donald Davidson ha subrayado especialmente el papel del externalismo como condición necesaria de la comunicación y de la interpretación intersubjetiva. Cf., por ejemplo, D. Davidson, «Verdad y conocimiento: una teoría de la coherencia» y «Las condiciones del pensamiento», en *Id.*, *op. cit.*, 73-97 y 152-161, respectivamente. Cf. también mi «Introducción al pensamiento de Davidson: mente, mundo y acción», *ibid.*, 9-45, especialmente 33-38.

⁴¹ Agradezco a T. Grimaltos, J. Marrades y N. Sánchez sus valiosas observaciones y sugerencias.