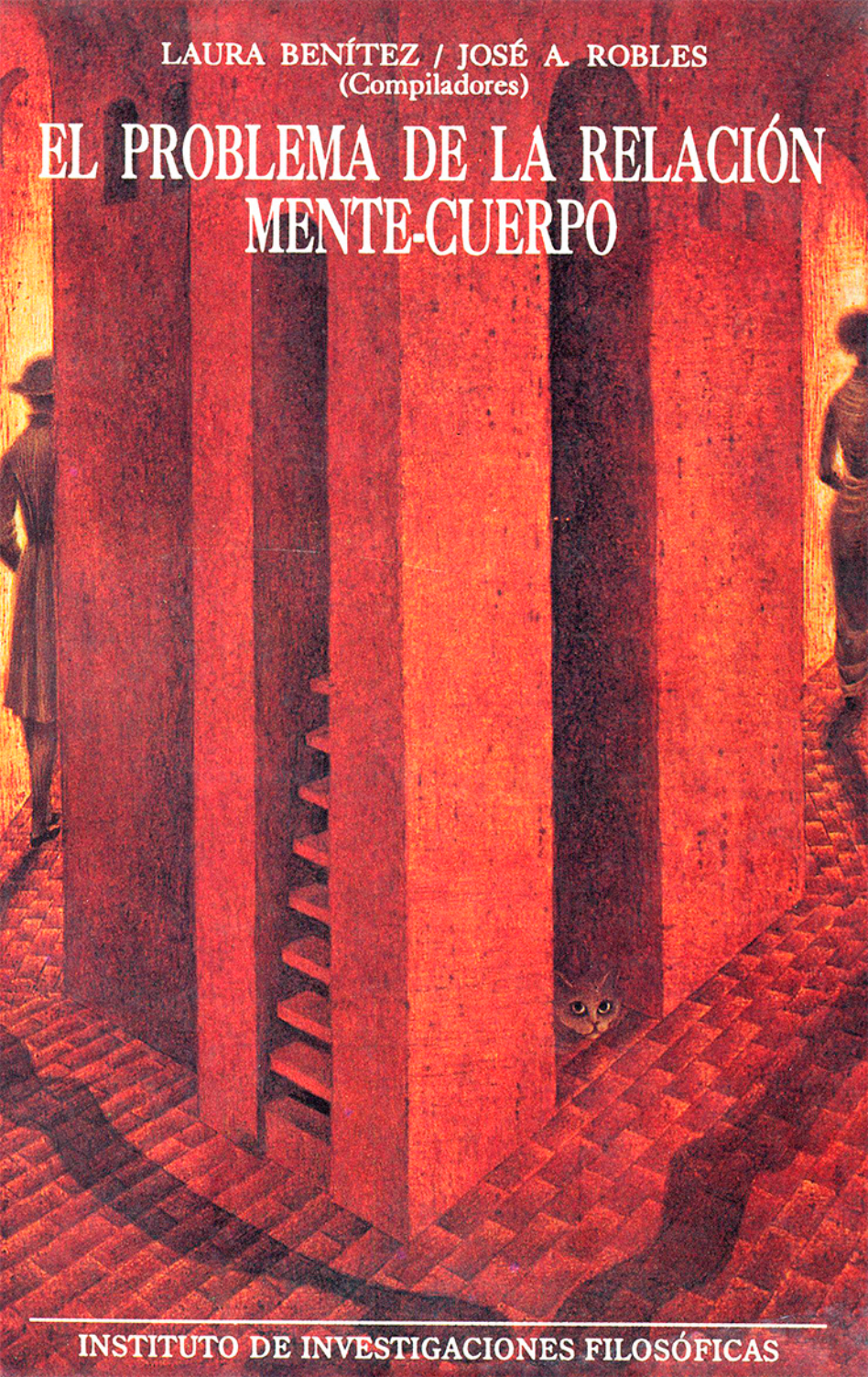
LAURA BENÍTEZ / JOSÉ A. ROBLES
(Compiladores)

EL PROBLEMA DE LA RELACIÓN MENTE-CUERPO



INSTITUTO DE INVESTIGACIONES FILOSÓFICAS

El problema de las relaciones mente-cuerpo, que aquí nos ocupa, se remonta a los orígenes del pensamiento filosófico y, aun hoy en día, siguen debatiéndolo no sólo los filósofos, sino también los científicos preocupados por las ciencias cognitivas, esto es, psicólogos, neurofisiólogos, computólogos interesados en programas de Inteligencia Artificial, etc. En este libro se estudian temas como la relación entre el cuerpo y lo mental, la base fisiológica de la memoria, la explicación de las creencias, las voliciones y el pensamiento, desde las perspectivas fisicalista y mentalista. Además se presentan teorías clásicas y contemporáneas sobre lo mental, pasando revista a la historia de la filosofía medieval, moderna y contemporánea. De acuerdo con esto, entonces, el volumen que le presentamos al lector ofrece no sólo la perspectiva clásica sobre la problemática aludida, sino también la contemporánea vinculada a las neurociencias. Así, consideramos que el material que aquí hemos recogido podrá ser de utilidad tanto para los que se asoman por primera vez al tema, como para quienes ya han lidiado con él por algún tiempo.

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

LA FILOSOFÍA Y SUS PROBLEMAS

1

Instituto de Investigaciones Filosóficas
Director: Dr. León Olivé
Secretario: Dr. Ambrosio Velasco

El problema de la relación mente-cuerpo

Laura Benítez / José A. Robles
(compiladores)

Universidad Nacional Autónoma de México

En la portada:
Despedida
Remedios Varo

1ª edición, 1993
© 1993 UNAM
Circuito Mario de la Cueva
Ciudad de Investigación en Humanidades
Ciudad Universitaria, 04510, México, D.F.
Instituto de Investigaciones Filosóficas

Impreso en México / *Printed in Mexico*

ISBN 968-36-3121-5

Presentación

La Colección *La filosofía y sus problemas*, que se inicia con este volumen sobre *El problema de la relación mente-cuerpo*, recogerá los resultados de estudios vivos y actuales provenientes, de manera central, de los proyectos de investigación que se desarrollan en el Instituto de Investigaciones Filosóficas de la UNAM. De este modo se dará a conocer, a un público amplio, al igual que al estudiante y al experto, un grupo de escritos que les mostrarán diversas cuestiones filosóficas importantes que hoy se debaten con gran interés, así como las formas mediante las cuales se las aborda y discute.

Los trabajos que aquí aparecerán, como sucede con este libro, serán el resultado de la investigación y el análisis de especialistas del propio Instituto en conjunción con otros colegas, tanto de México como del extranjero. El lector encontrará, entre los autores de los textos de esta Colección, a muchos de los escritores de mayor renombre internacional en su campo y hallará múltiples propuestas en torno a una amplia gama de enigmas filosóficos que, en estos momentos, se estudian en todo el mundo.

En la variada temática que se cubrirá –de lógica, epistemología, filosofía del lenguaje y de la mente, de filosofía de la ciencia, ética, filosofía política e historia de la filosofía, para sólo mencionar algunos campos–, la constante será la de presentar escritos inéditos y novedosos, siempre analizados previamente en seminarios y *simposia* promovidos por nuestro Instituto y realizados con la ya mencionada participación de expertos.

Por las razones anteriores, creemos que esta Colección será útil no sólo en el salón de clase sino que servirá, también, para guiar a los lectores interesados en profundizar sobre los varios temas que se publicarán en este y los próximos volúmenes y que

incluso el investigador avanzado encontrará nuevos enfoques y conjeturas dentro de su propio campo de estudio.

En suma, consideramos que *La filosofía y sus problemas* ofrecerá una muestra de los estudios más actualizados sobre las cuestiones filosóficas de mayor interés hoy en día, ya sea abordadas desde una perspectiva histórica o bien desde un enfoque problemático plenamente contemporáneo. Así pues, sólo nos resta esperar que, mediante esta Colección, el lector se acerque a las formas presentes, vivas, de la investigación filosófica, de sus problemas y de sus propuestas.

León Olivé

Director

Instituto de Investigaciones Filosóficas

UNAM

ÍNDICE

Laura Benítez / José A. Robles:
El problema de la relación mente-cuerpo 9

Mauricio Beuchot (IIF, UNAM)
Reflexiones en torno a las relaciones alma-cuerpo en Santo Tomás de Aquino 13

Laura Benítez (IIFilos, UNAM)
Reflexiones en torno al interaccionismo cartesiano 25

Leiser Madanes (Universidad de Buenos Aires)
¿Abandonamos la partida? Consideraciones sobre problemas mente-cuerpo en Descartes 43

José A. Robles (IIFilos, UNAM)
Berkeley: mente e ideas 73

Martha B. Bolton (Rutgers University)
La explicación de la conciencia en los Nouveaux essais *de Leibniz* 101

Alejandro Herrera (IIFilos, UNAM)
¿Fue Leibniz un paralelista sicofísico? 119

Margarita Costa (Universidad de Buenos Aires)
El problema mente-cuerpo en Descartes y Hume 135

Carmen Silva (FFyL, UNAM)
El naturalismo humeano y el problema mente-cuerpo 163

Quassim Cassam (Wadham College, Oxford)
La unidad trascendental de la apercepción 179

Alejandro Tomasini (IIFilos, UNAM)
Materialismo, interaccionismo y análisis gramatical 201

William Ramsey (University of Notre Dame):
Prototipos y análisis conceptual 243

Nydia Lara / Francisco Cervantes (CI, UNAM)
Mente-cuerpo y ciencias cognitivas 275

EL PROBLEMA DE LA RELACIÓN MENTE-CUERPO

El presente libro recoge los ensayos en torno al problema de la relación mente-cuerpo, que se presentaron en la segunda semana de agosto de 1991, como resultado de la investigación de los profesores que forman el área de Historia de la Filosofía del Instituto de Investigaciones Filosóficas (UNAM). Este simposio representó la 7ª reunión anual del grupo y la 2ª como parte del Proyecto 'Historia de la Filosofía y Pensamiento Científico Contemporáneo' (IN 600789), auspiciado por la DGAPA. El simposio sobre el Problema de la relación mente-cuerpo se organizó después de un año de búsquedas y análisis sobre la temática, con la presencia de los propios investigadores y distinguidos invitados extranjeros. El problema que nos ocupa se remonta a los orígenes del pensamiento filosófico y aún hoy en día siguen debatiéndolo no sólo los filósofos, sino también los científicos preocupados por las ciencias cognitivas, esto es, psicólogos, neurofisiólogos, computólogos interesados en programas de Inteligencia Artificial, etc. En esta reunión se discutieron temas como la relación entre el cuerpo y lo mental, la base fisiológica de la memoria, la explicación de las creencias, las voliciones y el pensamiento, desde las perspectivas fisicalista y mentalista. Además se estudiaron las teorías clásicas y contemporáneas sobre lo mental, pasando revista a la historia de la filosofía medieval, moderna y contemporánea. Así, el volumen que presentamos al lector, ofrece no sólo la perspectiva clásica sobre la problemática aludida, sino también la contemporánea vinculada a las neurociencias. Tales problemas surgen al analizar las manifestaciones psíquicas del ser humano (pensamientos, creencias, voliciones, etc.) y al preguntarse cómo éstas pueden relacionarse con las manifestaciones somáticas (movimientos conscientes diversos del cuerpo). Aún se tienen que considerar posibles explicacio-

nes dualistas (la existencia de un elemento anímico diferente del somático corpóreo), pues hasta ahora no se ha dado una explicación definitiva de esta cuestión en términos puramente físicos (y no es claro que se pueda dar una explicación semejante). En la actualidad, los debates sobre la relación mente-cuerpo no sólo se dan en los ámbitos científico y filosófico, sino que representan, para cada individuo en nuestra cultura, una fuente de constante reflexión por sus importantes implicaciones. En efecto, ¿puede hablarse del alma al margen del cuerpo? ¿Cuáles, si las hay, son las características específicas de lo mental? ¿Qué problemas suscita la negación del ámbito mental? ¿Puede sostenerse consistentemente un fisicalismo? Éstos y muchos otros interrogantes han dado origen a las reflexiones filosóficas y científicas sobre el tema; por ello, este volumen resultará de interés no sólo para los especialistas, sino para el público en general y podrá servir como texto complementario de diversos cursos universitarios, por lo que esperamos que el mismo sea una adición valiosa a la bibliografía académica en las áreas de epistemología, historia de la filosofía y ciencias cognitivas en general. Los problemas señalados, en la reunión a la que nos referimos, se trataron desde muy diversos puntos de vista y teniendo en cuenta las propuestas de varios autores, por lo que creemos que un posible lector de este volumen se encontrará con una temática variada que podrá atraer su atención. Por otra parte, el material que aquí hemos recogido será de utilidad tanto a personas que se asoman por primera vez al tema, como a quienes ya han lidiado con él por algún tiempo.

En este libro se recogen tanto textos interpretativos de la posición que acerca del problema plantearon filósofos del pasado como consideraciones muy recientes de trabajos en proceso, los que representan la labor de profesores de la UNAM, así como de profesores visitantes de universidades como Rutgers, Notre Dame, Oxford y Buenos Aires. Creemos que el éxito de la reunión se refleja de manera objetiva en los contenidos del libro, que

esperamos aliente la preocupación filosófica de nuestros lectores. Aquí nos corresponde, finalmente, expresar nuestro agradecimiento tanto a la Dirección General de Asuntos del Personal Académico (DGAPA), a través de su programa de Proyectos de Investigación, como al Instituto de Investigaciones Filosóficas por el apoyo que le ha prestado a nuestro grupo de trabajo para llevar a cabo las labores de investigación y de discusión de los resultados de la misma.

Laura A. Benítez, José A. Robles

REFLEXIONES EN TORNO A LAS RELACIONES ALMA-CUERPO EN SANTO TOMÁS DE AQUINO

Mauricio Beuchot*

1. Planteamiento

En este trabajo me planteo el problema de cómo deben conectarse entre sí, en el sistema de Santo Tomás, la mente y el cuerpo (o, si hacemos caso a Anthony Kenny, y ampliamos la noción de *mente* a operaciones no sólo cognoscitivas, podemos simplemente decir "el alma y el cuerpo"). Hay dos interpretaciones entre los estudiosos de Santo Tomás al respecto, y quiero optar por una y defenderla. Una interpretación exegética dice que tal unión debe ser la de dos entidades no-substanciales que constituyen una substancia al unirse como materia y forma respectivamente, y la otra interpretación establece que tal unión se da entre dos substancias de tipo muy especial, que se llaman substancias incompletas, que juntas constituyen una substancia completa. En la primera interpretación es claro que se salvaguarda muy bien el hilemorfismo aristotélico que sirve de base a Santo Tomás; pero creo que la segunda, a saber, la de las dos substancias incompletas formando una completa, es la correcta. Mi tesis será, pues, que esta interpretación no va contra el esquema metafísico aristotélico, porque conserva un sentido hilemórfico, y además, como trataré de hacer ver, la primera interpretación, tal como se ha manejado en el tomismo, tiene que arrostrar las mismas consecuencias que achaca a la segunda; pero esta últi-

* Centro de Estudios Clásicos del Instituto de Investigaciones Filológicas-UNAM

ma, la segunda, tiene la ventaja de ser más coherente con el marco conceptual general filosófico y teológico de Santo Tomás.

2. *Los correlatos: alma y cuerpo*

Así pues, en cuanto al *status* ontológico del cuerpo y el alma, de las dos opiniones interpretativas en el ámbito del tomismo que he referido, la primera es la que parece más unitaria y, por lo mismo, más en la línea del hilemorfismo aristotélico. Pero comenzaré recalcando que el carácter hilemórfico no se pierde en la segunda interpretación exegética, la del cuerpo y el alma como dos substancias incompletas. Ciertamente la otra interpretación, la primera, suena más monista y por ello más atractiva; pero el argumento para probar que la segunda interpretación, que defiendo, no se sale del hilemorfismo aristotélico, surge de que para Santo Tomás el cuerpo humano es una substancia incompleta que se comporta y funge como materia prima, esto es, no constituye una substancia totalmente independiente o autónoma, sino que hace las veces de la materia en la composición, y el alma del hombre, a su turno, es también una substancia incompleta, pero que funge como forma primera; y, dado que el lazo de unión entre ambas es substancial, de ellas resulta una substancia completa, que es el hombre. No se olvide que para Aristóteles y Santo Tomás lo principal y distintivo de la substancia no es recibir y sustentar accidentes, sino tener unidad *per se*, y aquí lo que tiene unidad *per se*, en la composición, no es el alma por su lado ni el cuerpo por el suyo, sino el ser humano. En esta línea de los que interpretan cuerpo y al alma como substancias incompletas, Robert Brennan lo expresa así:

> Aplicada al hombre, la doctrina hilemórfica afirma que entre la materia prima y la forma primera de todo ser humano existe un lazo de unión substancial perfecto. Esto quiere decir, más concretamente,

que el cuerpo del hombre, como materia prima, es una sustancia incompleta; que el alma del hombre, como forma primera, es también una sustancia incompleta; que el cuerpo y el alma juntamente, unidos como complemento mutuo, constituyen una substancia completa que es el hombre.[1]

Y en otro lugar añade que el alma es una forma substancial distinta de las demás:

Por una parte, no es tan independiente de la materia que pueda conseguir su perfección en el orden operativo sin el cuerpo. Y, por otra, no depende de la materia en tal grado que deje de existir actualmente si se separa del cuerpo. Por consiguiente, no puede ser sustancia completa en cuanto ésta equivale a esencia o especie completa, pues necesita unirse con la materia para adquirir la perfección total de su naturaleza y para, unida con la materia, constituir una esencia completa que es el hombre. No obstante, en cierto sentido podemos llamarla completa, a saber, desde el punto de vista de su substancialidad, puesto que puede continuar existiendo después de su separación de la materia.[2]

El alma es forma del cuerpo en el sentido de que es el principio estructurante y organizador de la materia corpórea humana; precisamente aquello que la hace ser humana. Es lo que explica el que a ese contenido le advenga una estructura; también explica que a ese conjunto de elementos, funciones y operaciones le corresponda una unidad ontológica; más aún, el alma es una entidad inferida a partir del comportamiento del hombre para explicar su unidad substancial. A nadie se le ocurriría argumentar

1 R.E. Brennan, *Psicología tomista. Análisis filosófico de la naturaleza del hombre*, Madrid, Javier Morata Ed., 1960, p. 97.
2 *Ibid*, p. 109.

para probar la existencia del cuerpo; lo que se tiene que probar es la existencia del alma. El argumento de Santo Tomás para probar que el alma es la forma del cuerpo corre así:

> Aquello por lo que cualquier organismo vive es su alma. Ahora bien, la vida se manifiesta en las diversas clases de substancias según sus distintas clases de operaciones. Síguese, por consiguiente, que la raíz de los actos vitales del hombre, en el último análisis, es el alma. Porque el alma es el principio primero por el cual vivimos, sentimos y nos movemos; e igualmente el principio primero por el cual entendemos. Y así este principio radical de nuestras operaciones intelectuales, llámese mente o alma racional, ha de ser la forma del cuerpo.[3]

De hecho, de que tenemos operaciones no meramente corpóreas, sino que Sto. Tomás llama espirituales, como el entendimiento y la voluntad, cuya operación no se reduce a la del organismo, es como infiere que tenemos un alma espiritual.

Viendo al alma como forma del cuerpo, i.e. la *psique* con relación al *soma*, se da una explicación hilemorfista de la interacción entre el cuerpo y el alma en los fenómenos que podemos llamar *psicosomáticos*. Santo Tomás alude a esas operaciones del conjunto; pues, aunque el alma tiene su operación propia, que es la de entender, tiene otras que sólo puede realizar por el compuesto, como funciones comunes al alma y al cuerpo, p. ej. la sensación y la emoción.[4] Y también se explica el que las facultades superiores puedan actuar sobre las inferiores, o a la inversa; p. ej., por afecciones del alma, el cuerpo puede cambiar de color o de calor, produciéndose la enfermedad, e incluso la muerte, pues se puede morir de tristeza, de amor o de gozo; y las afecciones del cuerpo pueden repercutir sobre el intelecto; o

3 *Sum. Theol.*, I, q. 76, a. 1.
4 Cfr. *Cont. Gent.*, l. II, c. 57; *Sum. Theol.*, I, q. 75, a.3.

las facultades superiores pueden actuar sobre las inferiores, como los actos de contemplación espiritual pueden retardar o paralizar movimientos de la naturaleza animal; y, a la inversa, p. ej. la intensidad de las pasiones puede obnubilar la inteligencia.[5] Además, se explica la unidad e identidad del yo, y su permanencia; pues, si es una substancia, eso da cuenta de por qué las operaciones tienen unidad y por qué la identidad personal se conserva a pesar de los cambios que sufre la persona en su cuerpo y en su mente a través de su historia. Igualmente Sto. Tomás aduce el argumento de que la unión substancial cuerpo-alma explica la aversión al sufrimiento y a la muerte, pues todo sufrimiento viene de la desarmonía o desequilibrio en la relación entre alma y cuerpo, siendo la muerte la descomposición o ruptura completa entre ambos. Cuando el hombre está bien, el cuerpo y el alma tienen perfectamente su unión substancial; de lo cual se concluye una vez más la composición substancial de cuerpo y alma en el hombre.[6]

Sin embargo, hay que insistir en que el alma es forma del cuerpo de una manera especial, ya que supera con mucho en dignidad ontológica al cuerpo. Sto. Tomás previene que hay casos en que la forma es muy superior a la materia, a pesar de estar condicionada a serle inherente, y por esa superioridad puede separarse de ella. Dice:

> Aunque el ser de la materia y de la forma sea uno, no es necesario, sin embargo, que la materia se acomode al ser de la forma. Antes bien, cuanto más noble es la forma, tanto más sobrepasa su ser a la materia. Esto puede comprobarse fijándose en las operaciones de las formas, por cuya consideración conocemos la naturaleza de las mismas, pues cada una obra en conformidad con su ser. Luego, la forma

5 Cfr. *De Veritate*, q. 26, a. 10.
6 Cfr. *Sum Theol.*, I-II, q. 5, a. 3; II-II, q. 164, a. 1.

cuya operación sobrepasa la condición de la materia, superará también por la dignidad de su propio ser a la materia.[7]

En suma, que alma y cuerpo sean vistos como substancias incompletas no rompe con el hilemorfismo, pues en la unión quedan perfectamente compenetradas, y no se dan como entidades inconexas de modo irreconciliable. Por el contrario, se unen con carácter de coprincipios, el cuerpo con las funciones normales de la materia receptora y el alma con las de forma substancial que da el ser a la materia en el compuesto. De este modo, la denominación de "substancias incompletas" que después configuran una substancia completa es de una substancialidad *sui generis*; no se trataría de un agregado de substancias, sino de un compuesto hilemórfico, pues preserva esa característica fundamental de constituir una unidad *per se*, o de suyo. Es una unión substancial, una unión de materia y forma en un compuesto hilemórfico que es una substancia, ella sí con pleno derecho substancia, y no como sus componentes, que lo son de manera incompleta.

3. El modo de unión entre cuerpo y alma

Además, la unión entre el cuerpo y el alma es la más perfecta que puede pensarse:

> Entre todas las cosas, el ser es aquello que más inmediata e íntimamente conviene a las cosas... De donde conviene que, ya que la materia tiene el ser por la forma, al dar la forma el ser a la materia, ante todo se entienda que le adviene a la materia, y que le es inherente de manera más inmediata que las demás cosas. Pero esto es propio

7 *Cont. Gent.*, l. II, c. 67.

de la forma substancial, que dé a la materia el ser de modo simple; pues ella misma es por la que la cosa es lo mismo que es. Pero no tiene el ser de modo simple por las formas accidentales, sino un ser según algún respecto; por ejemplo el ser grande, o coloreado, o algo así. Luego si alguna forma es aquella que no da a la materia el ser de modo simple, sino que adviene a la materia ya que existe en acto por alguna forma, no será la forma substancial.[8]

Pero quedan por conocer algunos aspectos de este modo de unirse el alma y el cuerpo, y eso nos lo puede declarar otro texto de Sto. Tomás:

La más perfecta de las formas, esto es, el alma humana, la cual es el fin de todas las formas naturales, tiene una operación que excede completamente a la materia, pues no se hace por un órgano corporal, a saber, el entender. Y, ya que el ser de la cosa es proporcional a su operación, según se dijo, pues cada cosa opera en cuanto es ente, conviene que el ser del alma humana sobreexceda a la materia corpórea, y no sea totalmente comprendido por ella, pero que de algún modo sea tocado por ella. Por consiguiente, en cuanto supera al ser de la materia corpórea, pudiendo subsistir y obrar por sí, el alma humana es una substancia espiritual; pero en cuanto es tocada por la materia, y le comunica a ella su ser, es la forma del cuerpo.[9]

Aquí nos dice Santo Tomás que, según un respecto o consideración, el alma es cierta substancia, a saber, una substancia espiritual (además de ser, según otra consideración, la forma del cuerpo). ¿Qué ha de entenderse por eso de que el alma es una substancia? ¿Qué tipo de substancia? ¿Cómo hay que interpretarla? Esto llevó a una polémica exegética dentro del tomismo.

8 *QQ. DD. de Anima,* a. 9, c; cfr. *Sum. Theol.*, I, q. 76, aa. 4-7.
9 *De spirit. creat.*, a. 2.

Algunos tomistas dicen, por un lado, que el alma y el cuerpo son entendidos por Sto. Tomás como dos substancias incompletas que, juntas, constituyen una substancia completa; y otros, por otro lado, dicen que alma y cuerpo no pueden ser substancias de ningún tipo, ni siquiera incompletas, y que incluso hablar de substancias incompletas es un despropósito. Oigamos las razones de unos y otros.

Por una parte, podemos ver en Verneaux la siguiente interpretación de Sto. Tomás. El Aquinate ha sostenido que el alma humana no sólo es inmaterial, cosa que tiene toda alma, sino que a diferencia de las otras es espiritual. Es decir, no sólo no es un cuerpo, sino que no depende del cuerpo en cuanto a su existencia, *in esse*, aunque dependa de él en cuanto a su esencia. Tiene, pues, el alma una suficiencia o independencia que bien podría llamarse *subsistencia*.

> Pero –dice– este término se presta a muchos equívocos: en particular, nos lleva a pensar que el alma es una substancia completa, lo que es falso..., y produce nefastas consecuencuas en la antropología. Santo Tomás es muy prudente en la formulación de su tesis: el alma humana, dice, es *aliquid subsistens, quod per se existit* [algo subsistente, que existe por sí]. Lo que equivale a atribuirle un ser de tipo substancial sin decir, no obstante, que sea una substancia.[10]

Pero reconoce que la unión del cuerpo y el alma es una unión substancial, pues de ella resulta una sola substancia, lo cual sólo se explica si se dice que ese resultado es una substancia completa y el cuerpo y el alma son substancias incompletas. Y no encuentra problema en el hecho de que el alma sea subsistente para que sea forma del cuerpo; pues, desde el punto de vista del ser, nada impide que le dé al cuerpo su acto de existir por tener ella una

10 R. Verneaux, *Filosofía del hombre*, Barcelona, Herder, 1967, pp. 215-216.

existencia de orden superior, sino que eso ayuda más. Cita al Aquinate, quien dice: "Pero nada impide a la substancia intelectual, por ser subsistente, que sea el principio formal del ser de la materia, como comunicando su ser a la materia".[11] Y tampoco, desde el punto de vista del obrar, se impide que una forma tenga alguna actividad propia en la que el cuerpo no participe. De nuevo deja hablar a Sto. Tomás, quien agrega:

> Hay que considerar que, cuanto una forma es más noble, tanto más domina a la materia corporal, y tanto menos está inmersa en ella, y más la excede con su operación o virtud... Y cuanto más se avanza en la nobleza de las formas, tanto más se encuentra que la virtud de la forma excede a la materia elemental: como el alma vegetal más que la forma del metal; y el alma sensible más que el alma vegetal; pero el alma humana es la última de las formas en nobleza; de donde excede tanto con su virtud a la materia corpórea, que tiene una operación y virtud en la cual de ningún modo comunica con la materia corporal. Y esta virtud se llama intelecto.[12]

Pero siempre insiste Verneaux en que el alma, para Sto. Tomás, aun cuando es algo subsistente o substancial, no es substancia completa. Le es esencial la relación al cuerpo, pues es para informarlo. Y, aunque el Aquinate llama al alma "substancia intelectual" (sobre todo en la *Suma Cont. Gent.*, l. II, c. 68), no deja de insistir en que no es un *hoc aliquid*, no es un ente completo o substancia completa. El alma no es el hombre, la persona, sino una parte suya. Además, el alma está presente a la vez en todo el cuerpo y en cada una de las partes del cuerpo: "Ya que el alma se une al cuerpo como forma, es necesario que esté en todo el cuerpo y en cualquier parte suya".[13]

11 *Cont. Gent.*, l. II, c. 68.
12 *Sum. Theol.*, I, q. 76, a. 1.
13 *Sum. Theol.*, I, q. 76, a. 8.

Como ejemplo de los que no quieren ver el alma y el cuerpo como dos substancias incompletas formando una completa (el hombre, la persona), sobre todo que no quieren ver el alma como substancia, queremos proponer a Enrique Dussel. A él le parece que, si Sto. Tomás hubiera dicho que el alma es una substancia incompleta, habría en él un dualismo velado, y no habría aportado nada nuevo frente al dualismo platónico de la cristiandad. Pero Sto. Tomás no es tal dualista. Para él el alma es una forma. "Una forma *sui generis*. La única forma que puede subsistir, porque es la forma substancial de un ente espiritual: el hombre".[14] (Pero si el alma es una forma que puede subsistir, ya es algo subsistente, un cierto tipo de substancia, una cosa substancial.) Dussel afirma que Sto. Tomás explica el que esa forma substancial es subsistente porque recibe de Dios y posee ella misma el ser del ente, el *esse*. Para ello aprovecha la expresión de la doctrina tomista de la participación elaborada por Cornelio Fabro. Dice Dussel:

> En esto nuestro pensador se distingue ya radicalmente de Aristóteles. El Estagirita admitía que el alma era forma, pero forma mortal, en la que participaba una realidad separada y divina, el *noûs*. Para Tomás en cambio por la doctrina de la creación, por la metafísica del *esse* (que no es ya la de la *ousía*), por la defensa de la substancia del alma como forma substancial, constituye óntica y ontológicamente al hombre, como siendo el que ejerce las funciones intelectuales y el mismo intelecto agente.[15]

Pero dice que Sto. Tomás nunca habla de "substancia incompleta", más aún, no cabe en su metafísica del *esse*; algo sólo puede ser substancia o accidente, no algo intermedio. Sin em-

14 E. D. Dussel, *El dualismo en la antropología de la cristiandad*, Buenos Aires, Ed. Guadalupe, 1974, p. 244.
15 *Ibid.*, p. 245.

bargo, reconoce que Sto. Tomás habla del alma como no-completa: "El alma no tiene la perfección de su naturaleza fuera del cuerpo, ya que no es por sí misma especie completa de alguna naturaleza, sino que es parte de la naturaleza humana".[16] No es completa en cuanto a la especie o a la naturaleza, o esencia, sino a la existencia. El alma no es substancia, pues se predica de un sujeto y existe en un sujeto, y el alma no es la persona, esto es, la substancia misma.

Mas, aunque es parte, subsiste. (Aun cuando es substancia no completa en cuanto a la esencia, es substancia completa en cuanto a la existencia). ¿Por qué subsiste? Porque así fue creada, es un caso especial; pues, en relación con el cuerpo, "ella misma es la que tiene el ser (*esse*), no sólo porque está compuesta por el ser como las otras formas, sino, aún más, porque está compuesta por su ser".[17] Pero la subsistencia que tiene es *sui generis*; ¿cómo es ella? Es una subsistencia fuera de la naturaleza del alma, que necesita del cuerpo para formar la persona; mas no es una subsistencia anti-natural.

> Pero, metafísicamente, la subsistencia del alma como forma no es absoluta y en sentido estricto, sino "de algún modo" (*secundum quid*). En efecto, sólo se dice "subsistente" en sentido estricto la persona humana... La persona es el sujeto metafísico, la que subsiste en sentido estricto y plenario, en sentido propio. El alma, si subsiste, subsiste o puede subsistir *en* la persona, que es el sujeto propio de la inmortalidad o resurrección. A nuestro conocimiento, Tomás nunca vinculó el problema de la subsistencia del alma con la cuestión de la persona.[18]

16 *De Pot.*, q. 3, a. 10.

17 *De unitate intellectus*, cap. 1, n. 38.

18 E. Dussel, *op. cit.*, p. 248. La individualidad del alma después de la muerte y antes de la resurrección de la carne no será abordado aquí, ya que requeriría un tratamiento muy extenso y que rebasaría los límites que nos hemos fijado.

Con ello se ve que, al ser subsistente el alma dentro del compuesto humano, queda el mismo problema de explicar esa subsistencia; en todo caso es igual que el problema de explicar el alma como substancia incompleta.

4. Conclusión

Para concluir, creo que se puede aceptar la interpretación del alma y del cuerpo en el sistema de Santo Tomás como lo ha hecho la línea tradicional, a saber, la que se remonta al Cardenal Cayetano, a Capreolo, Soncinas y otros connotados exégetas y comentaristas que han entendido la unión substancial de una y otro como la de dos substancias incompletas que forman una substancia completa, a saber, el hombre individual y concreto. Interpretándolos así no se pierde la estructura hilemórfica, pues son componentes que actúan como materia y forma; y, además, al conferir así cierta substancialidad al alma incompleta pero superior a la del cuerpo, se explica mejor la subsistencia que conserva a pesar de la descomposición o corrupción del cuerpo. En efecto, la otra interpretación tiene que adjudicar de cualquier manera cierta autonomía al alma con relación al cuerpo, y esa autonomía queda mejor explicada con la noción de substancia incompleta pero más perfecta que la otra que es el cuerpo, de modo que pueda subsistir después de la descomposición o desunión o ruptura del compuesto.

REFLEXIONES EN TORNO AL INTERACCIONISMO CARTESIANO

Laura Benítez G.*

Introducción

La problemática mente-cuerpo vista a la altura de nuestro siglo es no sólo cada vez más amplia y compleja sino que se ha recortado como objeto de estudio, de diversas areas de conocimiento, con enfoques y métodos distintos y aun contrarios.

Este sugerente panorama me ha llevado a someter a análisis la vieja teoría cartesiana del dualismo sustancial mente-cuerpo en relación con las nuevas perspectivas filosóficas. Dedico este trabajo a la consideración del problema de la interacción mente-cuerpo como aparece en la filosofía cartesiana a la luz de algunas críticas contemporáneas que, a mi juicio iluminan, de manera interesante dicha problemática.

La estrategia que pretendo seguir es la siguiente:

1. Presentar la teoría cartesiana haciendo énfasis en el análisis del interaccionismo.
2. Retomar los argumentos del interaccionismo a partir de dos modos de interpretación: el de quienes consideran que se trata de una propuesta infundada (Kenny, Williams, Wilson) y el de quienes piensan que pueden encontrarse interesantes fundamentos a esta tesis cartesiana (Richardson, Garber, Ariew).
3. Expondré las ventajas y límites de la lectura ontológica de Ri-

* Instituto de Investigaciones Filosóficas, UNAM.

chardson y de la física de Garber y propondré que, en último análisis, ambas lecturas se sustentan en una consideración peculiar de la noción de unión alma-cuerpo que aunque se acerca al dualismo atributivo es el peculiar dualismo de la unión como la entiende Descartes más que de la unidad sustancial.

Sobre el interaccionismo en René Descartes

En su respuesta a la segunda carta de un corresponsal no identificado,[1] Descartes escribió que era su propósito explicar *Las pasiones del alma* al margen de la oratoria y la filosofía moral para abordar este tema como médico. En efecto, toda la primera sección, del párrafo 1 al 45, puede leerse como una explicación fisiológica de las pasiones con una fuerte base mecanicista. Así, entiende que la función del cuerpo es moverse; que los nervios y el cerebro contienen un aire sutil al que llama, siguendo la tradición, espíritus animales; que nervios y espíritus juegan un papel fundamental en el movimiento del cuerpo.

Descartes introduce la explicación del funcionamiento de las partes del cuerpo diciendo:

Para hacer esto más inteligible explicaré aquí en pocas palabras el modo en que la máquina de nuestro cuerpo está compuesta.[2]

Según Descartes el principio corporal de todos nuestros movimientos es el calor que hay en el corazón que calienta la sangre cuyas partes más sutiles son espíritus animales, cuerpos pequeños que se mueven con enorme rapidez y que pueden penetrar

1 Los editores han hecho diversas conjeturas pero realmente no se sabe a quién escribió Descartes una carta fechada el 14 de agosto de 1649en Egmont y que aparece en la edición de Adam & Tannery como prefacio al *Tratado de las pasiones*.

2 Para citar las obras de Descartes me refiero a la edición canónica de Adam & Tannery. Todas las traducciones son mías. A/T XII VII.

por los poros del cerebro hasta su parte interna o salir de allí, por los nervios hasta los músculos para mover el cuerpo. Sin embargo, los espíritus animales tienen diversas formas pues no todos están constituidos por los mismos elementos, ni provienen de los mismos órganos del cuerpo, ni de la misma cantidad de sangre en el corazón. La diversidad de espíritus y los objetos de nuestros sentidos explican todo movimiento corporal sin recurrir a ningún principio anímico.[3]

Sin embargo, Descartes no se queda en físico pues aunque el propósito de *Las Pasiones* no es describir ontológicamente las sutancias, si lo es explicar sus funciones y ello implica no sólo las funciones del cuerpo sino también las del alma. De las funciones del alma la primera, naturalmente, es pensar. Nuestros pensamientos se dividen en acciones (voliciones) y pasiones (percepciones y operaciones cognitivas). Descartes ubica las pasiones como un tipo especial de percepciones referidas al espíritu; como sentir en la propia alma los efectos de la percepción; aunque también hay pasiones de la imaginación que son mucho menos vivaces. Descartes define las pasiones como: "Las percepciones, sensaciones o emociones del alma que se remiten particularmente a ella y que son causadas, mantenidas y reforzadas por algunos movimientos de los espíritus animales".[4]

Una vez delimitadas las funciones tanto del alma como del cuerpo, Descartes pasa al desarrollo de la teoría de la interacción entre ambos.

Parte de las nociones ya propuestas en las *Meditaciones* de que el alma es simple y el cuerpo un conjunto susceptible de descomponerse. Naturalemnte no ha perdido de vista la diferencia ontológica básica que ha establecido. Que extensión y pensamiento no admiten características comunes y que el cuerpo y

3 Pas A/T XI XVI.
4 Pas. A/T XI XXVII.

la mente tienen diversas funciones. Sin embargo, la necesidad de la unión parece imponer a Descartes la tarea de bucar la manera de conjuntar lo diverso por definición. Un primer paso para hablar de la unión entre el alma y el cuerpo es la propuesta cartesiana de que la aproximación no puede darse por el lado de la extensión y su división en partes, absolutamente incompatibles con la sustancia pensante, sino por el de la unidad y aun, "indivisibilidad" del cuerpo cuando se le considera desde el punto de vista de su organicidad o integración orgánica. Aquí, Descartes cambia, del enfoque físico del cuerpo como divisible en partes, al biológico de organicidad al afirmar que el cuerpo es uno en vista de que la separción de cualquiera de sus partes lo hace defectuoso. El alma está unida al cuerpo en tanto unidad orgánica y no en tanto extensión divisible.

La explicación de la unión del alma con el todo del cuerpo representa una propuesta límite, que no encaja dentro del marco de la ontología dualista cartesiana. En efecto, aunque Descartes alegue que es más fácil en el orden ontológico la unión entre sustancias que la unión entre la sustancia y los accidentes,[5] el carácter excluyente de la dos sustancias creadas (extensión y pensamiento) que propone las hace, ontológicamente, independientes, esto es concebibles la una al margen de la otra, y distintas en el orden de lo real, esto es efectivamente separadas.

Descartes no ignora la dificultad que acarrea el pensar en la unión alma-cuerpo incluso considera que no es una noción transparente al entendimiento sino de experiencia cotidiana y sentido común. Y si tal expediente no bastara siempre se puede recurrir a la omnipotencia divina capaz de salvar cualquier escollo dualista que se plantee a nuestra corta razón.

5 Res. a Obj. A/T pend. H.R. II p. 132. Con esto Descartes quería contestar a la objeción propuesta por Gassendi de la imposibilidad de la unión entre la extensión y el pensamiento.

A pesar de todas las dificultades, la propuesta cartesiana va más lejos que el planteamiento general de la unión alma-cuerpo. Se propone explicar detalladamente el mecanismo de la interacción. Refiere en *Las Pasiones*, que el alma ejerce sus funciones más inmediata y particularmente en la glándula pineal desde donde irradia la acción a todo el cuerpo y, a la vez, en la glándula se recoge todo cuanto ha de pasar del cuerpo al alma.

La elección de la glándula como sitio privilegiado de la comunicación entre el alma y el cuerpo se debe a un criterio en parte topológico y en parte fisiológico y, se constituye, para Descartes, en un aspecto importante de la explicación de la percepción sensible. Así nos dice:

> La razón que me persuade de que el alma no puede tener en todo el cuerpo ningún otro lugar que esta glándula [...] es que considero que las otras partes de nuestro cerebro son todas dobles [...] e incluso que todos los órganos de nuestros sentidos externos son dobles; y puesto que no tenemos sino un solo y simple pensamiento de una misma cosa en un mismo tiempo, es necesario que exista un lugar donde las dos imágenes [...] puedan juntarse antes de llegar al alma a fin de que no le representen dos objetos sino uno.[6]

En el nivel meramente fisiológico Descartes describe la glándula pineal como suspendida por encima del conducto por el cual los espíritus de las cavidades anteriores del cerebro se comunican con los de las posteriores. La dota, además, de la peculiar capacidad de agitarse con el menor movimiento de los espíritus, a la vez que su movimiento agita los espiritus animales.

La hipersensibilidad de la glándula no se limita al mecanicismo fisiológico de mover *a* y ser movida *por* los espíritus animales, sino que puede ser movida por el alma de diversos modos que

6 Pas. A/T XI XXXII.

repercuten en el cuerpo y también puede ser movida por el cuerpo (concretamente por los espíritus animales) de diversas maneras que repercuten en el alma. En suma, la glándula pineal es según Descartes el punto nodal de la interacción alma-cuerpo.

Naturalemnte se trata de una propuesta que lejos de aclarar el problema de la interacción lo hace aún más confuso y, por otro lado, tampoco es pertinente para la explicación fisiológica de la actividad cerebral.

En suma, si la unión alma-cuerpo resulta ser una noción oscura para el entendimiento, la interacción fincada en ella resulta, a primera vista, ininteligible.

La unión alma-cuerpo como interacción. ¿Propuesta ininteligible?

Al seguir cuidadosamente *Las Meditaciones* lo que se obtiene es el cuadro de dos sustancias completamente distintas con series de atribuciones tales que podemos concebir a cada una, de manera completa, por separado. Por ello, desde la ontología cartesiana es muy difícil explicar la unión e interacción entre una sustancia material y una espiritual que no tienen, en tanto sustancias, ninguna propiedad común.

Elisabeth de Bohemia, Gassendi, More y otros contemporáneos exigieron a Descartes una respuesta al problema concreto de *cómo el alma del hombre puede determinar a los espíritus del cuerpo para realizar las acciones voluntarias (no siendo sino una sustancia pensante).*[7]

Descartes dio una respuesta. Este intento ha sido calificado por muchos como infructuoso pero tal vez convenga revalorarlo

7 A Descartes, 16 de mayo de 1643, en *Descartes Lettres*. Presses Universitaires de France, París, 1964, p. 96.

pues se trata de proponer la inteligibilidad de la interacción entre la mente y el cuerpo.

Así, existen dos posiciones en relación con este problema:

1. Quienes encuentran la propuesta ininteligible porque, como se desprende de las objeciones de Elisabeth y Gassendi, la falta de una homogeneidad sustancial impide la aplicación de un concepto de causalidad tal que la mente pueda ser causa de efectos en el cuerpo o el cuerpo causa de efectos en la mente; objeción que es reiterada en nuestros días por Ryle quien encuentra semejantes relaciones imposibles, por Williams que las ve como un escándalo, por A. Kenny y por M. Wilson que atacan de ella aspectos concretos.
2. Quienes proponen que de la lectura atenta de las respuestas de Descartes a sus objetores pueden obtenerse interesantes ideas con respecto a la inteligibilidad de la interacción. En este caso se hallan, Garber, Richardson y Ariew, entre otros.

En 1643 Descartes escribió a Elisabeth un par de cartas donde la cuestión central es hacer inteligible la interacción. En esas cartas al decir de Garber: "Descartes ofrece una versión filosóficamente interesante y sofisticada de por qué piensa que la noción de interacción mente-cuerpo es perfectamente inteligible, en sus propios términos, y por qué no requiere ni admite clarificación".[8]

En su respuesta a Elisabeth, del 21 de mayo de 1643, Descartes refiere que hay dos modos de entender al alma, uno que subraya la *distinción* entre alma y cuerpo, como lo hizo en las *Meditaciones Metafísicas*, mostrando que el modo esencial de esta sus-

8 Garber, Daniel, "Understanding Interraccion: What Descartes should have told Elisabeth", en *Southern Journal of Philosophy*, núm. 21 (4): 15-32, 1983.

tancia es el pensamiento; otro, que hace hincapié en su *unión* con el cuerpo y su capacidad para actuar sobre él y que era uno de los objetivos de *Las Pasiones*.[9] Esta precisión hace ver que no existe en Descartes, una sola vía de explicación con respecto a los problemas de la interacción. Cabría en efecto, además de la explicación metafísica, otra de carácter "moral" o de "filosofía de la acción", y ésta es una cuestión que hay que conservar en mente para entender el desarrollo de su propuesta.

La objeción de Elisabeth a Descartes descansa en la concepción mecanicista del movimiento de los cuerpos. Si son condiciones del movimiento el contacto y la extensión ¿cómo puede producir movimiento algo sin semejantes características como el alma?

La respuesta de Descartes es que realmente esta explicación mecánica de la acción entre cuerpos no es importante para la comprensión de la interacción. No lo es porque poseemos una noción especial en términos de la cual podemos entender la interacción mente-cuerpo, que es completamente distinta a las nociones por medio de la cuales comprendemos por separado a la mente y al cuerpo. Si estas nociones fundan la metafísica y la física respectivamente, aquella funda la filosofía moral o práctica.

Existen, para Descartes varias nociones primitivas como las de ser, número y duración a las que se agregan las más importantes para esta discusión: la de extensión, que se refiere a los cuerpos y la de pensamiento que se refiere a la mente. Con todo,

9 "Porque habiendo dos cosas en el alma humana de las cuales depende todo el conocimiento que podemos tener sobre su naturaleza, una de las cuales es que piensa, y la otra, que estando unida al cuerpo puede actuar y padecer con él, no he dicho casi nada de esta última y solamente me he aplicado a hacer comprender bien la primera, a causa de que mi principal objetivo era probar la distinción que hay entre el alma y el cuerpo..." A Elisabeth, 21 de mayo de 1643, en *op. cit.*, p. 97.

el cuadro de nociones primitivas sería incompleto si no contáramos con una noción que hiciera explícita la unión de mente y cuerpo. A este respecto Descartes dice:

> ...finalmente, en relación con la mente y el cuerpo juntos tenemos únicamente la noción de su unión de la cual depende nuestra noción del poder de la mente para mover al cuerpo y del poder del cuerpo para actuar sobre la mente y causar sensaciones y pasiones.[10]

Nociones como la del alma, el cuerpo y su unión son primitivas en vista de que no pueden ser explicadas en términos unas de otras y sólo pueden ser entendidas por sí mismas; pero más importante aún, cada una tiene su propio dominio de aplicación. Desde esta perspectiva, Descartes quiere hacer ver a Elisabeth que confunde fuerza mental con la capacidad que un cuerpo tiene para actuar sobre otro. En suma, para Descartes el error de Elisabeth consiste en trastocar los dominios aplicando la explicación mecánica de la acción de un cuerpo sobre otro al dominio de la interacción mente-cuerpo.[11]

Si ha de tomarse en cuenta la propuesta cartesiana de las nociones primitivas es importante subrayar que no sólo implica dificultades ontológicas sino algunas epistemológicas. En efecto, al paso que la extensión y el pensamiento son ideas claras y distintas que se conocen por el entendimiento, la noción de la unión mente-cuerpo no puede conocerse claramente pues se obtiene por mediación de los sentidos.[12] Descartes dice a Elisabeth

10 A Elisabeth, 21 de mayo de 1643, en *Ibid.*, p. 97.

11 A este respecto ver el análisis de Richardson, R. C. "The 'Scandal' of cartesian interactionism", en *Mind* 91 (361): 20-37, enero de 1982.

12 "...y finalmente, las cosas que pertenecen a la unión del alma y del cuerpo no se conocen sino oscuramente mediante el entendimiento, e incluso mediante el entendimiento auxiliado por la imaginación, pero se conocen muy claramente por los sentidos." A Elisabeth, 28 de junio de 1643, en *op. cit.*, p. 98.

que la noción de la unión mente-cuerpo no puede ser objeto de la reflexión filosófica sino que debe extraerse de la vida cotidiana tal como lo hace el hombre de la calle. Esto implica que, aunque primitiva o fundante, la noción de unión entre alma y cuerpo es más bien una noción adquirida y, consecuentemente confusa para el entendimiento. No obstante, lo más importante es que se señala a esta noción como perteneciente a un dominio distinto de conocimiento en el que cuenta mucho más la experiencia vivida que la especulación.

¿No es ininteligible hablar de interacción dada la heterogeneidad sustancial?

En su análisis de la noción cartesiana de la unión alma-cuerpo Richardson sostiene que las respuestas a Elisabeth no sólo intentan resolver el problema de la heterogeneidad sustancial sino que muestran, hasta cierto punto, que la interacción *per se* no es incoherente. La unión de los diversos se hace ininteligible sólo cuando se considera desde la perspectiva de la causación mecánica, es decir, cuando se trata de imponer el esquema de la acción de un cuerpo sobre otro a la acción entre mente y cuerpo.

La sugerencia de Richardson es que la interacción mente-cuerpo no cae bajo el esquema de la causación mecánica sino bajo un esquema de causalidad más amplio que puede admitir causas agentes, esto es, libres o no determinadas.

En todo caso, desde mi punto de vista, lo importante es considerar que Descartes no aplicó el modelo mecánico irrestrictamente a otros dominios que no fueran los físicos y con ello dejó abierta la puerta a otras posibilidades de explicación en relación con la acción de la mente.

Por otra parte, es muy importante observar que no todas las nociones primitivas son simultáneamente fundantes de los di-

versos dominios de conocimiento. Descartes dice, en la segunda carta a Elisabeth, que no podemos concebir al mismo tiempo la distinción y la unión entre cuerpo y alma porque para ello es necesario concebirlos como dos cosas y al mismo tiempo como una sola cosa, lo cual es absurdo. De otra manera, las nociones primitivas de la metafísica no pueden ser fundamento de la filosofía práctica y en este dominio la prioridad debe recaer sobre la noción de la unión entre alma y cuerpo antes que sobre su distinción.

En esta misma línea Richardson advierte que si bien no podemos pensar en la mente y el cuerpo como separados y a la vez como unidos, sí podemos pensar en ellos como separables y como unidos. (Tal vez en el sentido tomista de sustancias incompletas)

En suma, el que se trate de naturalezas distintas e incluso el que no puedan concebirse simultáneamente como unidas y separadas no impide que interactuen y se conciban como unidas y sólo en este sentido como unidad sustancial.[13] Desde esta perspectiva el error sería concebirlos como una entidad compuesta y no como un todo sustancial.

Richardson considera que el problema de la interacción mente-cuerpo conduce a Descartes a plantear la unidad sustancial. Esta propuesta se sustentaría en la consideración atenta de una inmensa cantidad de procesos tales como sensaciones, pasiones, imaginación, etc., que se vuelven ininteligibles sin la consideración de la unidad como base de la interacción. Sin embargo, Descartes no propone sino la "cercana e íntima unión que existe entre la mente y el cuerpo" para explicar los procesos de inte-

13 "La naturaleza me enseña mediante estas sensaciones de dolor, hambre, sed, etc., que no estoy únicamente alojado en mi cuerpo como un piloto en su nave, sino que estoy íntimamente unido a él y, por así decirlo, tan entremezclado con él que parece que compongo con él un todo".

racción lo cual no supone cancelar la dualidad, ni postular una unidad sustancial en sentido escolástico.

Para concluir este apartado podemos decir que según Richardson, Descartes, en sus respuestas a Elisabeth, no está preocupado por la heterogeneidad sustancial o la distinción numérica de las sustancias, cuanto por el ser de la unidad sustancial, y yo me permito agregar que ello no es así. En todo caso Descartes ha dejado de lado los problemas ontológicos de la heterogeneidad sustancial para concentrarse, nó en el de una unidad sustancial que gestaría una contradicción en el ámbito de la ontología, sino en el de la unión mente-cuerpo, esto es en el de la persona como sustento de la acción moral. En suma, lo importante es que Descartes está construyendo la filosofía moral sobre una noción primitiva muy distinta a las de la metafísica, la de unión alma-cuerpo.

La inteligibilidad de la interacción en la perspectiva de un esquema causal ampliado

Descartes rechaza el esquema mecánico como fundamento de la interacción alma-cuerpo no sólo porque se transgreden los límites de su aplicabilidad, reduciendo la fuerza anímica a fuerza mecánica; sino porque la noción primitiva de unión, que está a la base de las relaciones entre el alma y el cuerpo, las cuales requieren una explicación causal, se inscribe en un ámbito de causalidad que rebasa el marco mecanicista.

La primera afirmación se halla explícita en las respuestas de Descartes a Elisabeth; la segunda, aunque puede inferirse de la primera, no se encuentra en las mencionadas cartas pero sí desarrollada fragmentariamente en algunos escritos de Descartes. Esta segunda afirmación, piensa Daniel Garber, es la respuesta que Descartes, siendo consistente con su física, debió dar a Elisabeth.

La lectura que Garber hace de la física cartesiana, por cierto muy peculiar, remite a sus fundamentos metafísicos. Las leyes del movimiento están fundadas en la inmutabilidad y perfección de Dios. O de manera más estricta: las leyes del movimiento y del choque derivan de las causas del movimiento. Descartes afirma que la causa general del movimiento es Dios, sin embargo, en la perspectiva de Garber, los movimientos particulares también tienen como causa a Dios ya que los cuerpos en movimiento no son la causa real del cambio en el choque sino que tal causa es Dios quien empuja a los cuerpos como parte del proceso de conservación de la misma cantidad de movimiento en la materia desde el momento en que la creó. "La imagen que uno tiene de la física es la de la materia inerte empujada momento a momento por un Dios activo y, de tiempo en tiempo, por mentes incorpóreas activas".[14]

Me parece que se trata de una lectura extrema. Que Dios sea el fundamento último de las leyes o causa general del movimiento es una opinión común; sin embargo, que la conservación de la misma cantidad de movimiento en el universo deba interpretarse como constante y contínua acción de Dios sobre la materia inerte, de modo que toda causa particular de movimiento deba remitirse a Dios, no sólo es una exageración sino que está en contra del modelo desarrollista del mundo natural que Descartes propone tanto en *El Mundo* como en *Los Principios*. No obstante, lo verdaderamente importante para el tema que nos ocupa es la acción de Dios en el mundo (sea general o particular).

En efecto, al preguntarnos cómo entender la acción de Dios sobre el mundo nos preguntamos por la acción de la sustancia inextensa sobre la extensa. A esto Descartes responde en una carta a More:

14 Garber, Daniel, "Understanding...", en *op. cit.*, p. 26.

> No es una desgracia para un filósofo creer que Dios puede mover un cuerpo, sin considerar a Dios como corpóreo; no es mucha más desgracia el que piense lo mismo con respecto a otras sustancias incorpóreas. Por supuesto no pienso que ningún modo de acción pertenezca unívocamente a ambas, Dios y las creaturas, pero debo confesar que la única idea que puedo encontrar en mi mente para representar el modo en que Dios o un ángel pueden mover la materia es el que me muestra el modo en el que soy consciente de que puedo mover mi propio cuerpo mediante mi propio pensamiento.[15]

Para Daniel Garber, este texto muestra que la idea que tengo del modo en que Dios actúa sobre el mundo se deriva de la noción que tengo de cómo puedo actuar yo sobre mi cuerpo; es decir, se deriva de la noción primitiva de la unión mente-cuerpo.

Para Garber, la respuesta de Descartes no debió ser tanto que Elisabeth mezcla las nociones primitivas pues, según él, la noción primitiva de la unión alma-cuerpo, está a la base de la comprensión de la acción causal en el mundo físico. Aún más, la interacción mente-cuerpo es el paradigma de todas las explicaciones causales y si no fuera inteligible por sí misma ninguna otra lo sería. "Cuestionar la intelegibilidad de la interacción mente-cuerpo es cuestionar la empresa entera de la explicación causal."[16]

En suma, no es la interacción mente-cuerpo la que debe someterse a la explicación causal de la mecánica, sino que, para comprender la causalidad entre los cuerpos físicos, hasta sus primeras causas, debemos entender cómo actúa la mente sobre el cuerpo.

15 *Ibid.*, p. 27. A More, 15 de abril de 1649. Roger Ariew sugiere que la noción de Dios actuando sobre el mundo debería considerarse otra noción primitiva y no derivada de la unión alma-cuerpo, en "Mind-Body interaction in cartesian philosophy: a replay to Garber", en *Southern Journal of Philosophy* 21 (4): 32-37, 1983.

16 *Ibid.*, p. 29.

Me parece que, según Descartes, la noción de unión mente-cuerpo es primitiva y fundante del ámbito de la reflexión sobre la acción humana pero no puede leerse como el paradigma de toda otra forma de causalidad cuando no es de suyo una noción teóricamente explícita, ni existe ningún texto que avale semejante propuesta, pues, la carta a More que se invoca es confusa y se presta a muy diversas interpretaciones, v. gr. que Dios y el mundo están unidos igual que el cuerpo y el alma. Así, parece lo más atinado interpretar la respuesta a More como una analogía y no como que Descartes intente fundamentar la noción de la acción de Dios en el mundo, en la noción primitiva de la unión alma-cuerpo, lo cual implicaría, desde la propuesta que intento sostener, que Descartes fundaría la física y la metafísica en la filosofía moral.

Si la noción de la unión mente-cuerpo es tan inteligible como las nociones separadas de mente y de cuerpo, entonces hay dos conceptos de "persona" en la filosofía cartesiana

A partir del análisis del problema de la interacción en la filosofía cartesiana, encontramos que hay dos formas de abordar sus propuestas. La primera que llamaré la posición A, aborda el problema desde el ámbito de la metafísica; la segunda, posición B, lo hace desde la filosofía moral.

En la posición A se subraya la *separación* ontológica entre las dos sustancias (extensa y pensante) lo cual da por resultado lo que Williams llama la versión cartesiana de persona, esto es, "cosas no físicas asociadas con un cuerpo". Si se busca una imagen más familiar, la dualidad ontológica, llevada a sus últimas consecuencias se plasma cabalmente en la del *piloto en la nave*. Sin embargo, Descartes rechaza esta idea en diversas ocasiones, precisamente, cuando se propone hablar ya no, ontológicamente,

de la separación de las sustancias sino justo de lo contrario, de su unión, para dar cuenta de la acción moral del hombre.

En la posición B, se hace énfasis en la *unión* alma-cuerpo como una noción primitiva y se concluye en la unidad de la persona que no apela al expediente ontológico de la unidad sustancial. Se trata de una sola persona en que se mezclan tan íntimamente el alma y el cuerpo que la persona puede entenderse como un sujeto con dos series de atribuciones distintas: estados físicos y estados mentales. En este caso las personas no se reducen ni a sus cuerpos ni a sus mentes sino que son sujetos de dos tipos o clases de propiedades (físicas y mentales), nos hallamos, pues, me parece, frente a una propuesta cercana al dualismo atributivo.

En suma, las nociones primitivas con que Descartes explica la interacción mente-cuerpo muestran, en su imposibilidad de concebirse conjuntamente, no sólo que son fundantes de diversas disciplinas, a saber, la física, la metafísica y la filosofía moral, dentro del sistema, sino que estas disciplinas albergan distintas perspectivas teóricas que, en el caso de la definición de la noción de 'persona' no son reducibles: el dualismo sustancial y una especie de dualismo de propidades.

Averiguar que tan cerca está Descartes de un dualismo atributivo, o de un monismo neutral, o de cualquier otra posición que intente dar cuenta del problema de la interacción, siempre en relación estrica con la noción primitiva de la unión mente-cuerpo es lo que se ha pretendido hacer ene este trabajo del cual pueden extraerse tres conclusiones:

1. En la noción primitiva de la unión si bien se da un dualismo atenuado no se da la cancelación de una sustancia en favor de otra, esto es Descartes no es un reduccionista, le interesa mantener, aunque íntimamente unidos al alma y al cuerpo como separables.

2. Aunque cercano al planteamiento tomista de la unidad sustancial Descartes no concibe al alma y al cuerpo como sustancias incompletas de ahí que no pase de la unión a la unidad sustancial. La unión es entremezcla pero en principio separable.
3. La perspectiva cartesiana de la "persona" como la unión alma-cuerpo aunque no cancela el dualismo pone el acento en las "propiedades" físicas y anímicas del individuo, este único sujeto cuyos procesos sólo pueden explicarse por la interacción de sus sustancias constitutivas. En suma, Descartes se aproxima al dualismo atributivo, y a la unidad sustancial al proponer su concepto de persona como unión alma-cuerpo, esto es intenta cerrar, aunque no del todo, el hueco que abrió con su dualismo sustancial.

¿ABANDONAMOS LA PARTIDA? CONSIDERACIONES SOBRE PROBLEMAS MENTE-CUERPO EN DESCARTES

Leiser Madanes*

Los aniversarios suelen ser pretextos inmejorables para hablar acerca de algunos autores y algunos libros en determinados momentos; libros y autores que –dicho sea de paso– nos encontrábamos leyendo y teníamos la intención de hablar acerca de ellos antes aún de percatarnos de que podíamos recordar un aniversario como excusa para hacerlo. Pues –tal es el espíritu de la época– parecería ser que para hablar acerca de Descartes hoy en día necesitamos por lo menos la excusa de un aniversario.

Agosto de 1641-agosto de 1991; 350 años han pasado desde la publicación de las *Meditationes de prima philosophia, in qua Dei existentia et animæ inmortalitas demonstratur*, con un primer conjunto de seis objeciones y respuestas. No se trata tan sólo del aniversario de la primera edición de un libro; es principalmente el 350 aniversario del principal problema que el libro propuso a la filosofía en su forma moderna: la distinción cuerpo-alma y, como su consecuencia, la relación cuerpo-alma.

Si se necesita una excusa para hablar acerca de Descartes, se debe a la diversidad de razones que se han dado durante los últimos 350 años mostrando por qué haríamos mejor en abandonar el rumbo trazado por sus *Meditaciones*. Spinoza comenzó expresando su asombro acerca de las opiniones ingeniosas de este "preclarísimo varón" –así llamaba a Descartes–, y confesando que le resultaba imposible formarse una idea clara y dis-

* Universidad de Buenos Aires.

tinta de la íntima unión de un pensamiento y una pequeña porción de cantidad [*Ética*, 1677, Prefacio a la parte Va.]. Hoy en día nos aconsejan que ignoremos simplemente este y otros problemas heredados de la filosofía moderna y que nos dediquemos a otro tipo de tarea intelectual. Es tal la cantidad de críticas que se han ido acumulando contra Descartes desde hace tres siglos y medio que hemos adquirido la tendencia a leer las *Meditaciones* por el placer de su destrucción, así como en el juego del boliche esperamos pacientemente que la máquina vuelva a colocar los pinos bien parados para ver si podemos arrasar con ellos de una vez. Leibniz –para tomar otro ejemplo de este deporte de disparar contra Descartes– recuerda cómo se fue alejando del cartesianismo cuando comenzó a considerar el problema de la unión del alma y del cuerpo. Nos cuenta que no encontraba la manera de explicar cómo el cuerpo podía pasarle algo al alma y viceversa. Ni cómo una sustancia creada podía comunicarse con otra sustancia creada. "Hasta donde lo podemos saber por sus escritos –dice Leibniz– Descartes había abandonado aquí la partida".[1]

¿Pero es esto verdad? ¿Descartes abandonó realmente la partida? ¿Debemos nosotros abandonarla, tal como lo aconsejan algunos pensadores posmodernos? Yo no estoy en posesión de las respuestas a estos problemas. Incluso tengo serias dudas acerca de cuál fue la respuesta que el propio Descartes dio a estos problemas. Les pido, por tanto, que consideren este trabajo como la presentación de un problema o de una serie de problemas que nos vienen preocupando desde la publicación de las *Meditaciones*.

Quizás convendría que resuma mis principales puntos de vista, así ustedes podrán luego seguir el hilo de mi exposición con mayor

1 "Nuevo sistema de la naturaleza y de la comunicación de las substancias, así como de la unión que hay entre el alma y el cuerpo", en *Journal des Savants*, 27 de junio de 1695; Gerhardt IV, 477 y ss. Trad. de Ezequiel de Olaso, en *G. W. Leibniz. Escritos filosóficos*, Buenos Aires, 1982, pp. 459 y ss.

facilidad reconociendo en qué están de acuerdo y en qué en desacuerdo. Considero que el estado de la cuestión es el siguiente.

Por un lado, están quienes discuten si el dualismo sustancialista de Descartes, sumado a su concepción de la causalidad, permite o no permite explicar la unión y la interacción cuerpo-alma. Los temas básicos en discusión son, qué entiende Descartes por sustancia, atributo y modo, y en qué consiste su principio de la causalidad.[2] Por otro lado están quienes acusan a Descartes de haber legado a la posteridad un problema mal planteado. La relación del conocimiento –afirman– no es la relación fundamental entre el hombre y el mundo. No hay por qué reducir al hombre a un sujeto epistemológico que busca certeza. La filosofía no tiene por qué limitarse al proyecto moderno que consiste en examinar ideas en la mente. No entremos en discusiones –nos dicen–, simplemente abandonemos el camino de Descartes. Hay que emprender otro tipo de tarea intelectual, aun cuando no siempre quede claro qué clase de tarea es.[3]

Ambas posiciones comparten un común denominador. Unos para ahondar su estudio, otros para abandonarlo, consideran que Descartes mismo planteó el problema de la distinción cuerpo-mente elaborada a lo largo de las *Meditaciones*. Esta interpretación de Descartes es tan antigua como los propios textos de Descartes. Consta en las objeciones; la formulan Spinoza, Leibniz, etcétera.

Ahora bien, es sabido que Descartes advirtió que concebir la distinción entre pensamiento y extensión es una tarea propia de la reflexión filosófica, pero que concebir la unión cuerpo-mente

2 Por ejemplo, la discusión entre Daisie Radner: "Is There a Problem of Cartesian Interaction?" –en el que critica trabajos de Robert C. Richardson y Louis E. Loeb– y la respuesta de ambos, en *Journal of the History of Philosophy*, vol. XXXIII, núm. 1 y núm. 2.

3 En el mundo de habla inglesa, Richard Rorty ha expresado este punto de vista, reconociendo a quienes lo antecedieron en esta crítica.

es una tarea no filosófica. Esta indicación o advertencia de Descartes a menudo es citada como una curiosidad, pero no suele ser explorada. Quienes se dedican a la historia de la filosofía y al examen de la consistencia o inconsistencia del sistema cartesiano, quizás ignoren esta advertencia de Descartes dado que no es un tema que haya sido retomado por los postcartesianos (Malebranche, Spinoza, Leibniz) y, por lo tanto, no ha tenido en la historia de la filosofía moderna las consecuencias que tuvo el problema de la posibilidad de la unión e interacción entre sustancias. Quienes, por otro lado, buscan abandonar la herencia cartesiana muestran tal urgencia por enterrar a Descartes que apenas la mencionan como una curiosidad en su obituario. Considero que en una época como ésta, en la que lo que en general se conoce como la modernidad está sufriendo constantes embates, vale la pena explorar esta vía que Descartes señala. El problema que propongo ante ustedes es el siguiente: ¿qué sentido tiene afirmar que la distinción entre pensamiento y extensión o entre mente y cuerpo es un problema filosófico, pero que la unión mente-cuerpo deberá ser abordado abandonando la filosofía? ¿Es posible separar el problema de la distinción mente-cuerpo y el de la unión mente-cuerpo, al punto tal de proponer que sean temas de tareas tan radicalmente diferentes como la filosofía y lo que no es filosofía?

El planteo sustancialista

Aun cuando mi intención es concentrarme en la manera no sustancialista de plantear el problema de la relación mente-cuerpo, quisiera comenzar con algunas observaciones sobre el examen tradicional sustancialista. El problema básico que quiero presentar se encuentra en un párrafo de la carta que Descartes le envía a Clerselier el 12 de enero de 1646 y que se publica en la

ediciónfrancesa de las *Meditaciones* de 1647 en lugar de las quintas objeciones de Gassendi y las respectivas respuestas [Adam y Tannery IX, 202-207].[4] Descartes –seis años después de la primera edición de las *Meditaciones*– le dice a Clerselier lo siguiente:

> Por lo que toca a las dos cuestiones que añaden al final, a saber, *cómo el alma mueve al cuerpo si ella no es material, y cómo puede recibir las especies de los objetos corpóreos*, me dan sólo ocasión para advertir que nuestro autor no ha tenido razón cuando, so pretexto de dirigirme objeciones, me ha propuesto muchas cuestiones cuya solución no era necesaria para probar las cosas que yo he escrito, acerca de cuyas cuestiones pueden los más ignorantes hacer más en un cuarto de hora de lo que todos los mayores sabios podrían resolver en toda su vida; y ésta es la causa de que no me haya tomado el trabajo de responder algunas. Y entre otras, éstas que acabo de mencionar presuponen la explicación de la unión que hay entre el alma y el cuerpo, de la cual aún no he tratado. Mas os diré que toda la dificultad que contienen procede de una suposición que es falsa, y que no puede ser probada en modo alguno, a saber: que si el alma y el cuerpo son dos sustancias de diferente naturaleza, ello les impide obrar una sobre la otra; al contrario, quienes admiten accidentes reales, como el calor, el peso y otros semejantes, no dudan de que esos accidentes obren sobre el cuerpo y, sin embargo, hay más diferencia entre ellos y él, es decir, entre unos accidentes y una sustancia, de la que hay entre dos sustancias.

Hasta aquí Descartes. Este texto es, ante todo, un voto de confianza en mi exposición, pues Descartes les está advirtiendo

4 En adelante: AT, número de volumen y de página. Las traducciones han sido tomadas de la edición de Vidal Peña, *René Descartes. Meditaciones metafísicas con objeciones y respuestas*, Madrid, 1977; de la de Ezequiel de Olaso, *Descartes. Obras escogidas*, Buenos Aires, 1967; y en ocasiones fueron realizadas por mí directamente de los textos cartesianos.

a ustedes que si yo resuelvo alguna cuestión en un cuarto de hora, soy un ignorante, y si no resuelvo ninguna, podemos suponer que soy un sabio. Pero además es un texto sorprendente por varias razones:

1. Es el año 1646 y Descartes advierte que aún no trató el tema de la unión del alma y el cuerpo (*"de laquelle je n'ai point encore traité"*). Ahora bien, si le tomamos la palabra a Descartes no deberíamos considerar que textos tales como la *Sexta meditación* o las respuestas a las objeciones planteadas por Arnauld explican la unión cuerpo-alma. Es decir, ningún texto publicado con anterioridad a esta carta debería considerarse como una explicación de la relación cuerpo-alma.
2. Según le escribe Descartes a Clerselier, Gassendi supone que, si el alma y el cuerpo son dos sustancias de diferente naturaleza, ello les impide obrar una sobre la otra. Según Descartes, esta suposición es falsa y no puede ser probada en modo alguno. Es decir, Descartes invierte el peso de la prueba: antes de explicar cómo se comunican dos sustancias diferentes, advierte de todas maneras que quienes deben explicaciones son quienes afirman que dos sustancias diferentes no se pueden comunicar. Según Descartes, Gassendi partió (i) de una suposición (ii) que es falsa y (iii) que no puede ser probada en modo alguno, es decir, que Descartes sabe que nunca se podrá demostrar. Nuevamente, si le tomamos la palabra a Descartes en ninguno de sus escritos anteriores a 1646 se estaría afirmando o implicando que dos sustancias diferentes no se pueden comunicar. Esto es, sin embargo, lo que con mayor frecuencia se le objetará a Descartes: según sus propias definiciones de sustancia y atributo, dos sustancias distintas no pueden comunicarse entre sí.
3. En este párrafo de la carta a Clerselier también nos sorprende el razonamiento de Descartes. Es así: Hay más diferencia entre

los accidentes y la sustancia que entre dos sustancias. Hay quienes admiten que los accidentes obran sobre la sustancia. [Premisa tácita: Si se puede pensar lo más difícil, se puede pensar lo más fácil]. Por lo tanto, es concebible que una sustancia obre sobre otra. Este razonamiento nos sorprende pues Descartes mismo no cree en los accidentes reales, ni tampoco parece clara cuál es la lógica o doctrina sustancialista que le da sustento a la argumentación.

Hemos identificado un conjunto de problemas a los cuales Descartes se está refiriendo en su carta a Clerselier. Ante todo, se queja de que le están presentando objeciones a temas acerca de los cuales él aún no trató. Además, considera que no es él quien debe una explicación --es decir, no reconoce que a partir de lo escrito hasta 1646 surja un problema sobre la relación entre mente y cuerpo.

Ahora bien, una manera de interpretar este texto consiste en suponer que Descartes no está siendo sincero con Clerselier. Podríamos suponer que Descartes ha caído en la cuenta que tras la distinción mente-cuerpo establecida en las *Meditaciones* resulta sumamente dificultoso o sencillamente imposible concebir una relación entre ambos. Más aún, podríamos suponer que Descartes ha decidido ignorar ante Clerselier una serie de textos escritos y publicados con anterioridad en los cuales ha tratado de explicar esta relación pero que no han sido todo lo convincentes que hubiera sido dable esperar. Se nos ocurre no sólo la última meditación, sino además diversas partes de las objeciones y respuestas. Sin embargo, antes de inclinarnos hacia esta interpretación tan poco caritativa hacia Descartes, tratemos de examinar hasta dónde es posible aceptar lo que Descartes escribe a Clerselier. Con este fin propongo las siguientes preguntas:

(i) Descartes advierte que aún no ha explicado la unión entre cuerpo y alma; que esta explicación no era necesaria para probar las cosas ("*la preuve des choses*") acerca de las cuales sí ha escrito y que la suposición de que dos sustancias diferentes no pueden interactuar no puede ser probada ("*ne peut aucunement être prouvée*"). La pregunta que surge aquí es: ¿Qué quiere explicar Descartes y qué quiere probar, y cuál es la diferencia entre una explicación y una prueba?

(ii) ¿En qué sentido –o en qué medida– es Descartes un sustancialista?

(iii) ¿Qué quiere hacer Descartes con la unión del alma y el cuerpo? ¿Pretende probarla? ¿Cree que es necesario explicarla? ¿Qué tipo de explicación o de prueba considera que es pertinente dar cuando se trata el problema de la unión mente-cuerpo?

Creo conveniente adelantarles hacia dónde apuntarán las respuestas que daré a estas preguntas, así les resultará más fácil seguirme. Ya he advertido que no pretendo ofrecer una interpretación a mi juicio concluyente sino más bien volver a plantear ciertos problemas. Mi posición es la siguiente: considero que Descartes no es un pensador sustancialista propiamente dicho sino que recurre a un lenguaje sustancialista cuando su interlocutor así lo requiere o para abordar algunos problemas muy específicos. Es decir, no tiene un compromiso sustancialista de la misma índole que Spinoza. Además, considero que Descartes piensa que diversos problemas requieren aproximaciones también diversas y sugiere que hay otra manera de abordar el problema de la relación mente-cuerpo que nada tiene que ver con la posibilidad de la interacción entre sustancias. En definitiva, desearía aportar algunos elementos gracias a los cuales podríamos considerar que Descartes estaba siendo –relativamente– sincero con Clerselier. Veamos hasta qué punto puedo aportar pruebas para estas interpretaciones.

El tímido sustancialismo cartesiano

Antes de entrar de lleno en la exposición de la vía alternativa para encarar el problema de la relación mente-cuerpo, deseo aclarar qué quiero decir cuando afirmo que presentar este problema en términos sustancialistas no es del todo adecuado. Descartes, obviamente, utilizó el término 'sustancia' en diversos textos. Pero creo que Descartes acomodaba en cierta medida su vocabulario a su interlocutor. El lenguaje sustancialista abunda en las respuestas a las objeciones, pero no es frecuente en las propias *Meditaciones*, en las que se recurre a estos términos básicamente al presentar el principio de causalidad en la *Tercera meditación*. Nuestra información más sistemática acerca del sustancialismo en Descartes proviene de la exposición *more geometrico* que con visible desgano redacta a pedido de Mersenne al final de la respuesta a las segundas objeciones y de *Los principios de la filosofía* [1645], obra destinada a la enseñanza en las escuelas.

No hay mención de sustancia en la *Segunda meditación* donde Descartes nos ofrece la primicia de su pensamiento, aun cuando utiliza este término en un pasaje análogo del *Discurso* [AT VI 33] y en la "Síntesis de la Segunda meditación" escrita con posterioridad. Es notorio, a lo largo de todas las *Meditaciones*, el deseo de Descartes de apartarse, en la medida de lo posible, del vocabulario de las escuelas. También en el *Discurso* escribe "aquí emplearé libremente, si les parece, palabras de la escuela" [*idem*, p. 34].

Descartes no parte de una definida doctrina sustancialista –como es el caso de Spinoza–, y con ese marco conceptual piensa lo que hay. Descartes emprende la búsqueda de certeza y recurre a una concepción sustancialista si ese vocabulario le sirve para expresar lo que piensa. Incluso en sus últimas obras utiliza el término 'sustancia' de manera totalmente libre. Por ejemplo, en el *Tratado de las pasiones del alma* [1649] se refiere a la masa encefálica como la sustancia del cerebro (art. 10 y art. 31) y también escribe

que la médula interior de los nervios es su "sustancia interior" (art. 12). Según Burman, Descartes había manifestado que "todos los atributos tomados en conjunto son en verdad lo mismo que la sustancia" [AT V, 154-155: *Omina attributa, collective sumta, sunt quidem idem cum substantia...* (16 de abril de 1648)]. En la más sustancialista de sus *Meditaciones*, la tercera, se refiere de manera desprolija: "Cuando pienso que la piedra es una sustancia o bien una cosa que es capaz de existir por sí misma" y luego se referirá a Dios como sustancia infinita. En la *Quinta meditación* dice: "Cuando imagino un triángulo... no deja de haber cierta naturaleza, forma o esencia determinada de esa figura". Descartes quiere que lo sigamos en su itinerario y recurre a cualquier término para hacernos comprender lo que está pensando. Por ejemplo: sirve a su propósito afirmar que cada naturaleza tiene un atributo principal, pero la relación entre atributo y sustancia no queda bien definida. En la *Sexta meditación* encontramos más ejemplos: en la edición original en latín leemos "Y por consiguiente, puesto que sé con certeza que existo y que, sin embargo, no advierto que ninguna otra cosa pertenece necesariamente a mi naturaleza o a mi esencia, sino que soy una cosa que piensa" [AT VII 78] y la versión francesa agrega "o una sustancia cuya esencia o naturaleza íntegra no consiste más que en pensar" [AT IX 62]. Un párrafo más abajo aclara, al escribir 'concepto formal', que se está sirviendo de términos de la escuela. En la síntesis de la *Sexta meditación* se lee: "muestro allí que el alma del hombre es realmente distinta del cuerpo y, sin embargo, que está tan estrechamente fundida y unida con él que compone como una misma cosa con él". El latín no contiene nada que equivalga al "como" y tampoco emplea el término 'res'. Lo que es francés se traduce por "una misma cosa", solamente es *unum quid* [*qu'elle ne compose que comme une meme chose avec lui - ut unum quid cum ipsa componat*].

En definitiva, no es mi intención argumentar que Descartes no era sustancialista porque contaba con una ontología definida y diferente. Más bien deseo mostrar que limitar el examen de la relación mente-cuerpo al problema de la posible o imposible interacción entre sustancias implica forzar una conceptualización sustancialista que Descartes mismo intentó eludir.

Paso ahora a examinar la forma alternativa de plantear el problema de la unión mente-cuerpo que sugiere Descartes. Comencemos recordando cuál era el proyecto de Descartes.

El Discurso del método [1637]

(i) Probar y explicar

En la sexta y última parte del *Discurso del método* Descartes advierte la diferencia que existe entre probar algo y explicar algo. El texto dice así:

> Si alguna de las cosas de que hablo al comienzo de la *Dióptrica* y de los *Meteoros* en principio son chocantes, porque las llamo suposiciones y no parezco dispuesto a probarlas, téngase la paciencia de leer todo atentamente y espero que se sentirá satisfacción. Pues me parece que las razones se siguen unas a otras de tal manera que, como las últimas son demostradas por las primeras, que son sus causas, estas primeras lo son recíprocamente por las últimas que son sus efectos. Y no se debe imaginar que yo cometa en esto la falta que los lógicos llaman un círculo; pues como la experiencia muestra que la mayoría de estos efectos son muy ciertos, las causas de las cuales lo deduzco no sirven tanto para probarlos como para explicarlos; mas, muy por el contrario, son las causas las que son probadas por ellos. Y las he llamado suposiciones para que se sepa que pienso poderlas deducir de esas primeras verdades que he explicado antes. [AT VI, 76].

En la posterior versión latina del *Discurso* el término 'probar' es remplazado por 'demostrar', atendiendo una crítica formulada por Morin a Descartes: la apariencia de los movimientos celestes se desprende (*se tire*) con igual certeza de la suposición de la estabilidad de la Tierra como de la suposición de su movilidad [véanse las cartas de Morin a Descartes, 22 de febrero de 1638, y la respuesta de Descartes del 13 de julio del mismo año, en la que Descartes subraya que "hay una gran diferencia entre probar y explicar"].

Desde el punto de vista de la operación lógica involucrada, no parece haber diferencia entre explicar y probar, pues en ambos casos se infiere una conclusión a partir de determinadas premisas. La distinción entre probar y explicar se refiere a lo que una persona ya sabe o ignora. Si la persona sabe cuál es la conclusión, las premisas cumplen la función de explicar dicha conclusión; si, por el contrario, la persona ignora si la conclusión es verdadera o falsa, las premisas cumplen la función de probar la conclusión.

En la "Síntesis de las seis meditaciones siguientes" Descartes afirma que en la *Sexta meditación* aduce "todas las razones de las que se puede concluir la existencia de las cosas materiales"; y aclara: "no se trata de que las juzgue muy útiles para probar lo que prueban, a saber, que hay un mundo, que los hombres tienen cuerpos, y otras cosas semejantes que jamás han sido puestas en duda por ningún hombre de buen sentido, sino porque considerándolas de cerca se llega a conocer que no son tan firmes ni tan evidentes como las que nos conducen al conocimiento de Dios y de nuestra alma, de modo que éstas son las más ciertas y las más evidentes que pueden entrar en el conocimiento del espíritu humano". En este texto 'prueba' parece tener otro matiz. Sabemos algo, es decir, no ponemos en duda algo y, sin embargo, no encontramos razones que justifiquen plenamente nuestra creencia. Pero adviértase que no por eso le restamos credibilidad a nuestra creencia: ningún hombre de buen sentido jamás ha

puesto en duda que tiene un cuerpo. Simplemente, no encontramos una justificación para esta creencia que sea tan firme y evidente como la justificación de la creencia en la existencia de nuestra alma y en la existencia de Dios. El hecho de que no dispongamos de una ciencia de la unión cuerpo-alma no implica que debamos abandonar nuestra creencia en dicha unión.

Tomando como ejemplo el tema que nos ocupa, podríamos en principio suponer que Descartes se propuso en las *Meditaciones* probar la distinción entre cuerpo y alma, y que en otros textos habríase visto ante la necesidad de explicar su interacción. Así como no dudo de que ahora estoy leyendo, tampoco dudo de que mi alma y mi cuerpo forman una unión. Quizás necesite explicar esta unión, pero no necesito probarla. Por el contrario, como el punto de partida es que mi alma y mi cuerpo están unidos, necesito probar que son distintos y separables, para así explicar (porque lo sé por verdad revelada) que mi alma es inmortal. Cada problema –unión cuerpo-alma, distinción cuerpo-alma, inmortalidad del alma, etc.– requiere una estrategia adecuada: o explicación, o prueba.

(ii) La filosofía práctica

La prueba parecería no tener límites: todo lo que aún no sabemos si es verdadero o falso merece el intento de ser probado. La explicación, en cambio, tiene límites más estrechos: si sabemos que algo es falso, por lo general requeriremos que se explique por qué es falso. Dentro del ámbito de lo que sabemos que es verdadero, quizás no estemos urgidos por explicarlo todo. Pongamos nuevamente como ejemplo el llamado por nosotros problema cuerpo-alma. En la última parte del *Discurso* Descartes contrapone la filosofía especulativa de las Escuelas y su propia

filosofía que caracteriza de "filosofía práctica". Descartes describe aquí su proyecto intelectual. Este proyecto consiste principalmente en la conservación de la salud, que es el primer bien y el fundamento de todos los demás bienes de esta vida. Descartes se ve a sí mismo ante todo como médico. Y no sólo desea que nos hagamos "como dueños y poseedores de la naturaleza" gracias al conocimiento de la naturaleza, sino que agrega que "el espíritu depende tan fuertemente del temperamento y de la disposición de los órganos del cuerpo que, si es posible encontrar algún medio de que los hombres sean comúnmente más sabios y más hábiles de lo que han sido hasta aquí, creo que es en la medicina donde hay que buscarlo" [*Idem*, p. 61-63].

Vemos que en el *Discurso* de 1637 Descartes presenta su proyecto de investigación sobre las vinculaciones del espíritu y el cuerpo como parte de un problema de filosofía práctica. A nosotros, hoy en día, el problema puede interesarnos en sí mismo. Descartes, en cambio, lo plantea como un tema de medicina, destinado a mejorar la salud de los hombres. Con este mismo tenor lo retomará diez años después en el *Tratado de las pasiones del alma*, que lo escribe *"en physicien"* [carta de Descartes a xx, 14 de agosto de 1649]. Al final del *Discurso* Descartes enfatiza aun más su resolución: "diré solamente que he resuelto no emplear el tiempo que me queda de vida en otra cosa que en tratar de adquirir algún conocimiento de la naturaleza que sea tal que se puedan sacar de él, para la medicina, reglas más seguras que las que ha habido hasta ahora." [*Idem*, p. 78]. Esto significa que se desatiendan determinados problemas teóricos o inconsistencias que aparecen a lo largo de la tarea intelectual que se está llevando a cabo. Pero deseo destacar que el por nosotros identificado como problema mente-cuerpo Descartes no comienza planteándoselo como problema teórico –tal como lo encontramos en la carta a Clerselier– sino como tema de investigación práctica de un médico.

El médico buscará explicar aquellas cosas que le resulten útiles para la terapia, tal como lo vemos en el *Discurso.*

La Sexta meditación [1641]

Al promediar la *Sexta* –y última– *meditación* Descartes ya cumplió con el plan trazado para su obra. La *Sexta meditación* lleva por título "De la existencia de las cosas materiales y de la distinción real entre el alma y el cuerpo del hombre". Ambos temas ya han sido considerados cuando Descartes parece dar un extraño giro a su pensamiento y se refiere a lo que nos enseña la naturaleza. ¿Qué es la naturaleza entendida como fuente de conocimiento? ¿Cómo llegamos a conocer lo que la naturaleza nos enseña? Ante todo, Descartes advierte que todo lo que la naturaleza nos enseña encierra alguna verdad (*omnia quæ doceor a natura aliquid habeant veritatis*, AT VII 80). La justificación de este principio parece preanunciar a Spinoza. En efecto, dice Descartes que "por la naturaleza, considerada en general, entiendo ahora a Dios mismo, o bien el orden y la disposición que Dios ha establecido en las cosas creadas" (*per naturam enim, generalites spectatam, nihil nunc aliud quam vel Deum, ipsum, vel rerum creaturum coordinationem a Deo institutam intelligo; idem*). La naturaleza tomada en general se diferencia de "mi naturaleza en particular", que es únicamente la complexión (o la reunión, agrega la versión francesa) de todas las cosas que Dios me ha dado. Descartes no aclara qué tipo de relación se da entre la naturaleza en general (*generaliter spectatam*) y mi naturaleza en particular (*in particulari*). Pero cualquiera sea la forma en que mi naturaleza en particular se inscribe en la naturaleza en general, esta última funciona como garante de verdad de lo que sabemos gracias a nuestra naturaleza particular. En alguna medida, lo natural particular sabe algo acerca de la naturaleza en general.

Ahora bien, ¿qué nos enseña la naturaleza? Ante todo, que tenemos un cuerpo. Esto es lo que la naturaleza nos enseña de manera más expresa (la versión francesa agrega a expresamente: y sensiblemente). También la naturaleza nos enseña, dice Descartes, que mi cuerpo está mal dispuesto cuando siento dolor y que tiene necesidad de comer o de beber cuando tengo los sentimientos de hambre o de sed. "Y por consiguiente –concluye Descartes– no debo de ningún modo dudar que hay en esto algo de verdad".

¿Qué es lo que hay de verdad en estas enseñanzas de la naturaleza? No será, ciertamente, que cada vez que siento sed es verdad que mi cuerpo necesita reponer líquidos, ni que cada vez que siento hambre mi cuerpo necesita reponer energías. Hay contraejemplos que indican que puedo equivocarme en determinados casos. El "algo de verdad" se refiere a que tengo un cuerpo y que no estoy alojado en él como un capitán en su barco, sino que "le estoy muy estrechamente unido y confundido y mezclado de tal modo que formo como un único con todo él" (*sed illi arctissime esse conjunctum & quasi permixtum, adeo ut unim quid cum illo componam; que je lui suis conjoint tres étroitement et tellement confondu et melé, que je compose comme un seul tout avec lui*).

Hacia la mitad de la *Sexta meditación* la reflexión filosófica propia de las *Meditaciones* parece haber cumplido su propósito habiéndose demostrado la distinción entre el alma y el cuerpo y la existencia de las cosas corpóreas. Lo que queda de la última *Meditación* parece ser más bien una descripción de nuestra experiencia de la vida (por ejemplo, el alma no está alojada en el cuerpo como un piloto en su nave). Esta experiencia encierra algo de verdad, pero no todo lo que nos ofrece es verdadero. En cierto sentido, Descartes ha abandonado aquí la partida –recordando la crítica afirmación de Leibniz– pero este abandono no obedece a la incapacidad de resolver inconsistencias surgidas

del propio sistema filosófico, sino que se debe al fin de un recorrido y a la indicación de un nuevo camino a seguir. El itinerario de las *Meditaciones* en su búsqueda de creencias ciertas y con su examen de ideas claras y distintas nos llevó a concluir que es cierta la creencia de que yo existo, de que Dios existe, de que las cosas materiales existen. Descartes parece estar indicando que hay otro camino que debe seguirse si se desean investigar otro tipo de problemas. La tarea consistiría –según la manera en que el propio Descartes comienza a ejercerla– no en probar o explicar algo, sino en describir cierta clase de experiencias que reciben el nombre de enseñanzas de la naturaleza. Algo de verdad hay en estas enseñanzas (Dios también es garante de esta verdad), pero difícilmente el criterio de verdad sea la evidencia entendida como claridad y distinción de nuestras ideas.

Habiendo ejercitado al lector en la búsqueda de certeza durante seis meditaciones, Descartes obviamente no quiere que estas enseñanzas de la naturaleza vuelvan a confundirlo todo. "Pero para que no haya en esto nada que no conciba distintamente debo definir con precisión lo que entiendo propiamente cuando digo que la naturaleza me enseña algo" (AT VII 82). "Tomo aquí la naturaleza con significado más restringido que cuando la llamo reunión o complejo de todas las cosas que Dios me ha dado, puesto que esta reunión o complejo comprende muchas cosas que pertenecen sólo al espíritu, de las que no entiendo hablar aquí, al hablar de naturaleza." Descartes desea que quede en claro que todo lo que se conoce por la sola luz natural del espíritu, sin la ayuda del cuerpo, no está incluido en la categoría de enseñanza de la naturaleza. Además, tampoco están incluidas como enseñanzas de la naturaleza las propiedades de los cuerpos. Por naturaleza no se entiende en este contexto ni lo que pertenece sólo al espíritu ni lo que pertenece sólo a los cuerpos. Por naturaleza se entiende "las cosas que Dios me ha dado en cuanto estoy compuesto de espíritu y de cuerpo". Es decir, to-

do aquello que hace a la unión cuerpo-alma –y no a su distinción– se conoce gracias a estas enseñanzas de la naturaleza. Estas enseñanzas son fundamentalmente prácticas: me enseñan a evitar cosas que provocan en mí el sentimiento de dolor y a dirigirme hacia aquellas cosas que me comunican algún sentimiento de placer. Descartes concluye con una advertencia: "pero no veo que además de esto me enseñe que de estas diversas percepciones de los sentidos debamos sacar conclusiones respecto de las cosas que están fuera de nosotros, sin que el espíritu las haya examinado cuidadosa y maduramente. Pues sólo al espíritu y no al compuesto de espíritu y de cuerpo corresponde, me parece, conocer la verdad de esas cosas". No debemos perder de vista esta aclaración: solamente por el espíritu conocemos qué es el espíritu y qué son los cuerpos. Las enseñanzas de la naturaleza se refieren a un grupo reducido de fenómenos, vinculados con la descripción de la unión cuerpo-alma y a cierta guía para nuestra vida práctica.

Al final de la *Sexta meditación* Descartes se propone mostrar que el hecho de que a veces nos equivoquemos en lo que consideramos que es una enseñanza de la naturaleza, no contradice la bondad de Dios. Interesa destacar que su argumentación da por verdadera la unión del cuerpo y del alma y se limita a examinar, como si fuese un médico, cómo funciona esa relación. La naturaleza nos enseña que el alma está unida al cuerpo. El médico examina cómo trabaja de hecho esa unión, cómo funcionan los músculos y los nervios. El filósofo no cuestiona cómo es posible esa unión. La enseñanza de la naturaleza es una prueba de la unión cuerpo-alma. Quizás no podamos explicar cómo es posible esa unión. O quizás el pedido de explicación sea un pedido mal formulado. Pero el no estar capacitados para explicar la unión no puede hacernos renunciar a la enseñanza de la naturaleza que nos muestra que esa unión es verdadera. Tal como advertía Descartes al final del *Discurso* es preferible el cono-

cimiento de algunas pocas verdades a la vanidad de aparentar que no se ignora nada.

La correspondencia con Elisabeth [1643]

(i) La primera explicación a Elisabeth

La correspondencia de Descartes con la princesa Elisabeth resulta especialmente relevante para nuestros problemas, pues en el transcurso del intercambio epistolar el problema de la relación mente-cuerpo comienza planteándose como un problema de relación entre sustancias y concluye con la propuesta del abordaje desde la no filosofía. Veamos cómo se desarrolla.

En 1643 Elisabeth le había planteado a Descartes un problema similar al que Descartes discute en la carta a Clerselier de 1646. La correspondencia entre Descartes y Elisabeth no forma parte de los textos editados o conocidos por Clerselier en el momento de escribirse con Descartes. La princesa quería saber cómo es que el alma, que no es sino una sustancia pensante, podía determinar los espíritus del cuerpo para llevar a cabo acciones voluntarias. Descartes no reacciona ante la pregunta planteada por Elisabeth de la misma manera en que posteriormente reaccionará en la carta a Clerselier. Comienza admitiendo que "la cuestión que Vuestra Alteza me propone es la que se me puede hacer con mayor razón después de los escritos que he publicado" [carta a Elisabeth, Egmond, 21 de mayo de 1643; AT III, 663]. Para nuestro caso, los escritos relevantes publicados son las *Meditaciones* con objeciones y respuestas. La respuesta de Descartes comienza con la siguiente aclaración:

1. Hay dos cosas en el alma humana de las que depende todo el conocimiento que podamos tener de su naturaleza: (a) que piensa;

(b) que está unida al cuerpo y, por lo tanto, puede obrar y padecer con el cuerpo.

2. Descartes indica luego que su principal propósito había sido probar la distinción entre el alma y el cuerpo, y agrega que saber que el alma piensa servía a su propósito, mientras que saber que está unida al cuerpo perjudicaba dicho propósito. Por lo tanto, casi había dicho nada acerca de la unión del alma y del cuerpo.
3. Habiendo aclarado que en sus publicaciones casi nada ha dicho acerca de este problema, se propone a continuación explicarle a Elisabeth de qué modo concibe la unión del alma y del cuerpo y cómo es que el alma tiene fuerza para mover el cuerpo. La explicación propiamente dicha procede de la siguiente manera:

Hay en nosotros ciertas nociones primitivas que son como unos originales sobre el patrón de los cuales formamos todos nuestros restantes conocimientos. Hay muy pocas nociones como éstas:

(i) están las nociones más generales –ser, número, duración– que convienen a todo lo que podemos concebir;
(ii) para el cuerpo en particular tenemos la noción de extensión, de la que se siguen las de figura y movimiento;
(iii) para el alma sola no tenemos sino la noción del pensamiento, en la que están comprendidas las percepciones del entendimiento y las inclinaciones de la voluntad;
(iv) para el alma y el cuerpo juntos sólo tenemos la noción de su unión, de la que depende la de la fuerza que tiene el alma para mover el cuerpo y el cuerpo para actuar sobre el alma causando sus sentimientos y sus pasiones.

Cuando queremos explicar alguna dificultad por medio de una noción que no le pertenece, nos equivocamos. Cuando queremos explicar una de estas nociones por otra, también nos equivoca-

mos, pues siendo primitivas, cada una de ellas no puede ser entendida sino por ella misma. De lo que se desprende que la unión cuerpo-alma no puede ser explicada por el agregado de cuerpo más alma.

Ahora bien, ocurre que el empleo de los sentidos nos ha hecho mucho más familiares las nociones de la extensión, de las figuras y de los movimientos, que las demás nociones. Por lo tanto, la principal causa de los errores consiste en que queremos servirnos de estas nociones para explicar cosas a las que no pertenecen, como cuando se quiere emplear la imaginación para concebir la naturaleza del alma o cuando se quiere concebir el modo como el alma mueve al cuerpo por la manera como un cuerpo mueve a otro. Según Descartes, toda la ciencia de los hombres consiste en distinguir estas nociones y en no atribuir cada una de ellas sino a las cosas a que pertenecen. Esta tarea se realizó en las *Meditaciones.*

Lo primero que ahora hay que explicar es la manera de concebir las nociones que pertenecen a la unión del alma con el cuerpo, sin las que pertenecen al alma sola o al cuerpo solo. Descartes señala que para esto puede servir lo que escribió al final de las Respuestas a las Sextas Objeciones. Solamente en nuestra alma podemos buscar estas nociones simples. El alma, por su naturaleza, contiene en sí todas las nociones simples. Pero no siempre las distingue de otras. O no las atribuye a los objetos a los cuales se las debe atribuir. Por ejemplo, antes de llevar a cabo el ejercicio de las *Meditaciones* solía confundir la noción de la fuerza con que el alma actúa en el cuerpo con la fuerza con que un cuerpo actúa sobre otro. Este error ocurría porque todavía no conocíamos el alma y atribuíamos la noción de la fuerza con que el alma actúa sobre el cuerpo y la de la fuerza con la cual un cuerpo actúa sobre otro a diversas cualidades de los cuerpos (como serían, el calor, la gravedad...). Además, imaginábamos que estas cualidades eran reales, es decir, que tenían

una existencia distinta de la del cuerpo y, por consiguiente, eran sustancias aunque solían llamarse cualidades. ¿Cómo concebíamos –equivocadamente– estas "sustancias"? Nos servíamos tanto de las nociones que están en nosotros para conocer el cuerpo (*i.e.* extensión y sus derivados) como de las que están para conocer el alma (pensamiento y derivados). Si lo que le atribuíamos a esas "sustancias" o "cualidades reales" era material, recurríamos a las nociones comunes referidas a la extensión; si era inmaterial, a las propias del pensamiento. Así, equivocadamente suponíamos que la gravedad es una cualidad real o sustancia y le atribuíamos fuerza para mover el cuerpo en el cual está hacia el centro de la tierra. No pensábamos que esto ocurría por un contacto real de una superficie contra otra (ya sea porque se trataba de una relación entre dos cuerpos distantes uno del otro o porque se concebía la gravedad como inmaterial). Pero en nosotros mismos experimentamos que tenemos una noción particular –la de la unión del cuerpo y del alma– y equivocadamente la aplicamos a la gravedad. El error consiste en no advertir que la gravedad no es algo realmente distinto del cuerpo (*i.e.* no es una cualidad real o sustancia). La noción primitiva de la unión nos ha sido dada para concebir "el modo como el alma mueve el cuerpo", pero no para concebir cómo un cuerpo mueve a otro.

(ii) La segunda explicación a Elisabeth

La princesa no queda satisfecha con la explicación de Descartes. La analogía cuerpo-alma cuerpo-gravedad no le resulta convincente. Así no se alcanza a comprender por qué un cuerpo puede ser empujado por algo inmaterial. Más aún, si un ser inmaterial –como la gravedad– pudiera mover un cuerpo, no habría por qué su-

5 Véase, en la edición de Alquié, vol III, p. 43, nota.

poner que en el mundo se conserva la cantidad de movimiento, tal como sostiene Descartes. [Elisabeth a Descartes, 10-12 de junio de 1643].[5]

Habiendo leído la respuesta de Elisabeth, Descartes le envía una segunda carta [a Elisabeth, Egmond de Hoef, 2 de junio de 1643], en la cual busca aclarar su anterior explicación. Nuevamente Descartes advierte que:

1. Hay que distinguir tres géneros de ideas o de nociones primitivas: la noción que tenemos del alma, la del cuerpo y la de la unión que hay entre el alma y el cuerpo.
2. Cada una de estas nociones se conoce de una manera particular y no por comparación de una con la otra.
3. A partir de esta distinción fundamental, hay que:

   (i) explicar la diferencia que existe entre estas tres clases de nociones,
   (ii) explicar la diferencia entre las operaciones del alma mediante las cuales las poseemos,
   (iii) formular los medios de hacer que cada una de estas nociones nos sea familiar y fácil.

Descartes reconoce que un punto oscuro en su anterior explicación es la comparación con la gravedad. Después de aclarar por qué se había servido de la comparación con la gravedad, considera que hubiera correspondido advertir que aunque se quiera concebir el alma como material (lo que propiamente es concebir su unión con el cuerpo), no se deja de conocer más adelante que ella es separable del cuerpo.

Ahora bien, habiendo enumerado el orden en el cual desea responder, comienza la respuesta propiamente dicha. Con respecto a la diferencia entre las tres clases de nociones, advierte que:

(i) el alma se concibe por el entendimiento puro;
(ii) el cuerpo (la extensión, las figuras y los movimientos) también se pueden entender por el entendimiento solo, pero mucho mejor por el entendimiento con el auxilio de la imaginación;
(iii) las cosas que pertenecen a la unión del alma con el cuerpo no se conocen sino oscuramente por el entendimiento solo, ni siquiera por el entendimiento ayudado por la imaginación, pero se conocen muy claramente por medio de los sentidos.

Y luego agrega una aclaración particularmente significativa:

> De donde resulta que los que no filosofan nunca y sólo se sirven de sus sentidos no dudan de que el alma mueve al cuerpo y de que el cuerpo actúa sobre el alma; más aún, consideran ambos como una sola cosa, es decir, conciben su unión, pues concebir la unión que hay entre dos cosas es concebirlas como una sola.

¿Por qué medios se obtienen las diversas nociones primitivas? Los pensamientos metafísicos que ejercitan el entendimiento puro sirven para hacernos familiar la noción del alma; el estudio de las matemáticas, que ejercita principalmente la imaginación en las consideraciones de las figuras y de los movimientos nos acostumbra a formar nociones del cuerpo distintas; por fin, sirviéndonos solamente de la vida y de las conversaciones habituales, y absteniéndonos de meditar y de estudiar estas cosas que ejercitan la imaginación, se aprende a concebir la unión del alma y del cuerpo.

Descartes advierte que habla en serio, que su consejo de apartarse de la filosofía para concebir la unión del alma y del cuerpo no es algo dicho al pasar, ni una manera elegante de comunicarle a una princesa que la complejidad del problema supera su capacidad de resolverlo. Absteniéndose de meditar se aprende a con-

cebir la unión del alma y del cuerpo o, mejor dicho, se llega a concebir el alma y el cuerpo como unidad. Descartes ensaya una explicación teórica de por qué, para ello, es imprescindible dejar de filosofar: ocurre que el espíritu humano no puede concebir muy distintamente la distinción que hay entre el alma y el cuerpo y al mismo tiempo su unión, porque para esto hay que concebirlos como una sola cosa y simultáneamente concebirlos como dos, lo que es contradictorio.

Descartes describe sucintamente esta noción de la unión que sin filosofar experimenta siempre cada uno en sí mismo: que es una sola persona que posee conjuntamente un cuerpo y un pensamiento, los que son de tal naturaleza que este pensamiento puede mover el cuerpo y sentir los accidentes que le ocurren. Ahora bien, alguien que tiene presentes las *Meditaciones* o cuyo espíritu está cautivado aún por el ejercicio de las *Meditaciones*, no puede concebir esta unión. Para estas personas que todavía tienen presentes en su espíritu las razones que prueban la distinción del alma y del cuerpo, y no queriendo que las abandonen para representarse la noción de la unión que cada uno experimenta en sí mismo sin filosofar, dice Descartes que se ha

> ...servido antes de la comparación de la gravedad y de las demás cualidades que comúnmente nos imaginamos que están unidas a algunos cuerpos, así como el pensamiento está unido al nuestro y no me ha preocupado que esta composición ni fuera exacta en el sentido de que estas cualidades no son reales como se las suele imaginar, porque he creído que Vuestra Alteza estaba ya enteramente convencida de que el alma es una sustancia distinta del cuerpo.

Sabía que la comparación con la gravedad y las cualidades reales era inexacta y equivocada. Pero su estrategia de discusión consiste en utilizar –dentro de lo posible– los marcos conceptuales de sus interlocutores. Recurre al ejemplo porque supone

que así le hará comprender el problema a Elisabeth. Esta estrategia la observamos con frecuencia en las respuestas a las objeciones de las Meditaciones. Es tal su flexibilidad conceptual que, ante las nuevas objeciones de Elisabeth, abandona rápidamente estos ejemplos y le recomienda que examine el problema con otro marco teórico:

> Pero ya que Vuestra Alteza ha observado que es más fácil atribuir materia y extensión al alma que atribuirle la capacidad de mover un cuerpo y ser movida por él, sin tener materia, os ruego que queráis atribuir libremente esta materia y esta extensión al alma, pues esto no es más que concebirla unida al cuerpo. Y después de haber concebido bien esto y haberlo experimentado en vos misma, os será fácil considerar que la materia que hayáis atribuido a este pensamiento no es el pensamiento mismo y que la extensión de esta materia es de otra naturaleza que la extensión de este pensamiento, porque la primera está determinada en cierto lugar, del que excluye cualquiera otra extensión de cuerpo, lo que no hace la segunda. Y así Vuestra Alteza no dejará de volver fácilmente al conocimiento de la distinción del alma y del cuerpo, pese a haber concebido su unión.

Tal es el remedio que Descartes encuentra para que, quienes lo han acompañado a lo largo de las Meditaciones y han finalmente aprendido a concebir la distinción entre el cuerpo y el alma, puedan ahora concebir la unión del cuerpo y del alma, que es una noción tan primitiva como la del alma o la del cuerpo. Nuevamente, Descartes cambia el marco conceptual. Nuevamente, lo que le sugiere a la princesa es muy difícil de comprender: ¿en qué consiste atribuirle materia o extensión al alma? ¿Cómo se le atribuye materia a un pensamiento? Desde las *Meditaciones*, esto no es comprensible. Quizás convenga entonces abandonar la filosofía y simplemente vivir, experimentar, la unión del cuerpo y del alma.

De Arnauld, del 29 de julio de 1648

Cinco años más tarde Descartes mantiene la misma posición que sostuvo en sus cartas a Elisabeth. En correspondencia con Arnauld le advierte que ningún razonamiento ni ninguna comparación tomada de otras cosas puede mostrar que la mente, que es incorpórea, pueda impulsar al cuerpo [*Quod autem mens, quæ incorporea est, corpus possit, sed certissima & euidentissima experientia quotidie nobis ostendit*; AT V 222]. Recuérdese que en la carta a Clerselier, escrita el año anterior, Descartes afirmaba que ningún razonamiento puede probar que si el cuerpo y el alma son sustancias de diferente naturaleza, no pueden obrar una sobre la otra.

Ahora le afirma a Arnauld que ningún razonamiento puede probar lo contrario, a saber, que la mente incorpórea pueda mover el cuerpo. ¿Estamos ante una manifestación de escepticismo por parte de Descartes, en la cual muestra una equipolencia de razones, frente a lo cual no podemos tomar partido? De ninguna manera. Descartes inmediatamente agrega a Arnauld que el hecho de que el alma mueva el cuerpo es algo que diariamente nos muestra "una experiencia muy cierta y muy evidente, pues ésta es una de las cosas que conocemos por ahí (*per se notis*), que oscurecemos cuando queremos explicar otras" [AT V, 222]. La mente es consciente de su unión con el cuerpo. Esta conciencia se manifiesta en la inclinación de nuestra voluntad a mover los miembros. Descartes le propone a Arnauld una comparación... Obviamente, la comparación es: "La mayoría de los filósofos que consideran que la gravitación de la piedra es una cualidad real... etcétera".

Observaciones finales

Según Descartes, el conocimiento humano consiste en distinguir claramente nuestras nociones y atribuir correctamente cada una

de ellas a la cosa a la que se aplica. De alguna manera, esta concepción del conocimiento humano se asemeja a nuestra concepción común de justicia, entendida como darle a cada uno lo que le corresponde. Traté de mostrar que el propio Descartes divide el problema mente-cuerpo en dos problemas diferentes: el de la distinción de la mente y el cuerpo y el de la unión e interacción mente-cuerpo. La distinción de mente y cuerpo depende de nuestra comprensión de dos nociones primitivas diferentes: la noción de una mente que piensa y de la materia extensa. La unión e interacción mente-cuerpo depende de nuestra concepción de una tercera noción primitiva: la noción de la unidad cuerpo-mente. No debemos intentar comprender la unión mente-cuerpo –que es una noción primitiva en sí misma–como un compuesto de las otras dos nociones primitivas. Llegamos a la distinción mente-cuerpo a través de los ejercicios mentales propuestos en las *Meditaciones*. Descartes llama filosofía a estos ejercicios. La unión cuerpo-alma no puede concebirse desde el punto de vista de las *Meditaciones*. Descartes consideró que debíamos aproximarnos a ella desde la no-filosofía. La tarea propia de la filosofía consiste en dar pruebas y explicaciones. Parecería que Descartes considera que la tarea propia de la no-filosofía consiste en proporcionar descripciones. ¿Cómo sabemos, en cada caso, cuál es la actitud adecuada? Intenté mostrar que el proyecto de Descartes fue concebido como un proyecto de filosofía práctica, en el sentido de que su propósito principal consistía en el avance de nuestros conocimientos de física y medicina para ayudar al prójimo. Una manera de justificar la decisión de elegir entre filosofía y no-filosofía podría ser, por lo tanto, que cada problema ante el cual nos enfrentamos tenga su propia manera de ser abordado. Por su parte, esto se justificaría argumentando que, de hecho, tenemos tres nociones primitivas distintas: la de la mente, la del cuerpo y la de la unión mente-cuerpo. Retomando palabras de Leibniz, Descartes abandona la partida de la filosofía, que consistía en distinguir mente y cuerpo, para jugar otra partida

–la de la no-filosofía–, que consiste en la unión mente-cuerpo. El conocimiento humano es la tarea abarcadora que consiste en decidir a cuál de estos juegos jugaremos.

BERKELEY: MENTE E IDEAS*

José A. Robles**

0. Introducción

En su lucha en contra del dualismo cartesiano (que para él daba origen al escepticismo y a la irreligión), Berkeley aparentemente logra deshacerse de uno de los polos del mismo: la sustancia material (= extensa para Descartes). Ciertamente retiene la sustancia espiritual y considera que la misma, junto con las ideas, es lo único que necesita para dar cuenta y razón del mundo físico, tanto en el aspecto ontológico, como en el de nuestra relación epistémica con el mismo. Conforme a su tesis *esse es percipi*, Berkeley cree que ha anulado la duda escéptica ya que, de acuerdo con tal tesis, no hay que dar ningún imposible salto mortal epistémico, de las ideas presentes al entendimiento, al sustrato del que las mismas provienen (conforme a los autores más representativos de su época): la sustancia material. Considero que todo lo anterior es correcto: Berkeley emplea su versión de la navaja de Ockham y se deshace de entidades que sólo pueden producir moles-

* Presenté una versión anterior de este escrito como ponencia en el VII Simposio del Área de Historia de la Filosofía del Instituto de Investigaciones Filosóficas, el martes 13 de agosto de 1991. En la discusión participaron Laura E. Benítez, Martha B. Bolton, Alejandro J. Herrera, Guillermo M. Hurtado y Quassim Cassam; las observaciones y sugerencias que formularon me motivaron a redactar un texto más extenso en el que intento dar una fundamentación más amplia y mejor de la tesis que aquí sostengo. Aquí también quiero agradecerle a Carmen Silva su cuidadosa lectura de este escrito y sus sugerencias.

** Instituto de Investigaciones Filosóficas, UNAM.

tias y confusión. Sin embargo, cuando intentamos estudiar con algún detalle la relación entre espíritu e ideas –la propuesta berkeleyana de la relación mente-cuerpo–, se nos presenta, una vez más, una fuerte versión dualista, de la que Berkeley creía haberse desembarazado en su rechazo de la sustancia material. Lo que me interesa sostener en este escrito es que, conforme a la tesis berkeleyana de la relación mente-ideas, surgen serios problemas, similares a los que tiene Descartes para salvar el abismo que él creara con su tajante distinción entre sustancia (material) extensa y sustancia pensante. Un problema similar le surge a Berkeley cuando éste caracteriza espíritus e ideas como entidades que son fuertemente opuestas en sus propiedades ontológicas y se plantea, nuevamente, la dificultad de encontrar la manera de explicar la relación entre ellas. Berkeley sostiene que hay dos tipos de ideas, uno es el de las que no dependen de nosotros, son producto de la voluntad de Dios, a saber, las ideas de percepción; el otro corresponde a las ideas que dependen de nuestra voluntad para estar presentes ante nuestra conciencia y éstas son las ideas acerca del pasado (recuerdos) y las de la imaginación. Lo que sostengo en este ensayo es que en todos los casos, tanto en el de las ideas de percepción como en el de las de memoria e imaginación, habrá que apelar a Dios de manera esencial, esto es, en ninguno de estos casos los espíritus finitos se bastan a sí mismos para manejar tales ideas, e insisto en señalar que esto es así incluso con respecto a las que Berkeley supone que dependen meramente de nuestra voluntad, esto es, las de la memoria y las de la imaginación.[1]

1 A.A. Luce, de los pocos autores berkeleyanos que ha tratado el tema de la relación mente-ideas de la imaginación, nos recuerda (en [8], p. 90, n. 1) que siempre que se tratan estos u otros temas afines, en los que se pueden apuntar problemas para la posición de Berkeley, "sería muy temerario inferir que Berkeley no vió la dificultad y que no tenía una respuesta a la mano. El *C.P.B.* [*Common Place Book*; nombre dado por Frazer a los cuadernos de notas de Berkeley que Luce luego reordenó y rebautizó con el nombre de *Philosophical Commentaries* (*Comentarios filosóficos*), que es el nombre con el que

La conclusión que considero que se puede extraer de mi alegato es que, en contra de lo que señalan los siguientes pasajes optimistas de Berkeley –uno de los *Comentarios filosóficos* (redacción, 1707-8) y el otro de los *Principios* (1710)–, habrá que ver el espíritu, en su sistema, como imposibilitado de ejercer el aspecto imaginativo activo que él estaba tentado a concederle:

> G. Por qué no podemos concebir que sea posible para Dios crear cosas a partir de la Nada. ciertamente nosotros de alguna manera creamos siempre que imaginamos. (CF 830)

> Encuentro que puedo despertar a placer ideas en mi mente y variar y cambiar el escenario con tanta frecuencia como lo crea adecuado. No hay más que tener voliciones y directamente esta o aquella idea surge en mi fantasía y con este mismo poder se elimina y da paso a otra. Este hacer y deshacer ideas debe, con propiedad, denominar a la mente activa. Todo esto es cierto y se funda en la experiencia; pero cuando hablamos de agentes no pensantes o de despertar ideas con exclusión de la voluntad, sólo nos engañamos con palabras. (PCH I, 28)

A.A. Luce, expresa esta versión optimista de la propuesta de Berkeley, de la siguiente forma:[2]

> Berkeley, por así decir, sigue un curso medio entre Malebranche y Locke. Él le concedió a Malebranche que algunas (la mayoría) de nuestras ideas son también ideas de Dios. Él le concedió a Locke que

---

ahora se les conoce] contiene diversas indicaciones de que las cuestiones psicológicas se reservaban para la segunda parte de los *Principios*, que Berkeley perdió en Italia luego que 'ya había avanzado bastante'". La noticia se la da Berkeley a Johnson, su corresponsal norteamericano, en una carta del 25 de noviembre de 1729 (*cf.*, acerca de esto, [3], p. 238).

2 El pasaje se encuentra en [8], pp. 90-1.

algunas ideas son hechas por el hombre. La conexión entre idea y voluntad se toma en cuenta repetidas veces en el *Commonplace Book* (801-45). 'Toda Idea tiene una Causa esto es, la produce una Voluntad' (843). Esto no es, como en la apologética cristiana ordinaria, una deducción a partir de los atributos divinos, sino que es una cuestión de observación factual de la que correctamente se infiere el ser de Dios. 'Cuando en plena luz del día abro mis ojos, no está en mi poder elegir si he de ver o no' (*Princ.* § 29). Las ideas sensoriales 'no son criaturas de mi voluntad. Hay, por tanto, otra voluntad o espíritu que las produce'. Así pues, las ideas reales que son el mundo real, son las ideas de Dios descubiertas a nosotros y que, por tanto, se hacen nuestras ideas. Pero, para Berkeley, también está la clara cuestión de hecho que cualquiera puede verificar en su experiencia, de que nosotros mandamos, controlamos, hacemos y deshacemos algunas de nuestras ideas. 'Este hacer y deshacer ideas debe, con propiedad, denominar a la mente activa. Todo esto es cierto y se funda en la experiencia'. Así, la imaginación con sus ideas tiene un lugar definido en el sistema de Berkeley. Para él, la sensación propiamente dicha es una esfera de la operación manifiesta de Dios. Yo creo que él considera la imaginación como una esfera de la libre actividad del hombre, en la que Dios determina (*wills*) no entrar. En teoría, las dos esferas son distintas; pero no es probable que Berkeley las considerase como realmente separadas ... Difícilmente podría haber hecho una escisión tajante entre ellas. De hecho, la imagen se combina con el percepto y éste con la imagen. Así, conforme a los principios de Berkeley, no hay necesidad de tomar nuestras fantasías como totalmente sin Dios, ni nuestras imaginaciones como algo fuera de la Voluntad general de Dios.

Ésta es la propuesta optimistas de Berkeley y lo que aquí sostengo es que, *dentro de su sistema*, incluso esta facultad imaginativa dependerá, para poder ejercerse, de la buena voluntad de Dios. De lo anterior me parece que se puede concluir que los espíritus finitos berkeleyanos se ven reducido a ser voluntades

sin que haya materia alguna donde aquélla pueda ejercerse. Por otra parte, también me interesa poner de manifiesto que, sobre estas cuestiones, en el contexto de su época, fue Malebranche quien tuvo una mayor sensibilidad de los problemas y, en consecuencia, ofrece algunos intentos mejores de solución. De manera concreta, el atribuir la posesión de *todas* las ideas (como elementos *inteligibles* del mundo perceptual) a Dios y ninguna a las almas o espíritus finitos. Antes de concluir esta sección me interesa subrayar el hecho de que el mismo Luce, en el pasaje que cité líneas arriba, señala que (conforme a los principios de Berkeley) "no hay necesidad de tomar nuestras fantasías como totalmente sin Dios, ni nuestras imaginaciones como algo fuera de la Voluntad general de Dios". Mi propuesta elimina el 'totalmente' y es la aseveración tajante de que (lógicamente) Berkeley no puede sostener que haya ideas de la imaginación, fantasías, recuerdos, etc., sin el auxilio total de Dios.

1. Propuestas de Berkeley sobre los espíritus y las ideas

En esta sección presento diversos pasajes de la obra de Berkeley en los que éste formula algunas de las propuestas a las que he aludido líneas atrás: el dualismo con respecto a espíritus e ideas, mostrando la fuerte diferencia tanto ontológica como funcional entre ambos tipos de entidades (en esto, como veremos, la posición de Berkeley es claramente cartesiana) y su deseo de mantener una fuerte relación de proximidad entre ellas, en base a su tesis central, *esse es percipi*, para evitar la postulación de una sustancia material. A partir de esto, creo que se puede formular un serio cargo en contra de Berkeley en el sentido de que su posición preserva los arduos problemas que surgieron de la formulación dualista cartesiana de sustancia extensa *vs.* sustancia pensante. Paso ahora a presentar mi alegato.

Que hay nuevamente un dualismo, lo muestra claramente, entre otros de los pasajes que aquí presentaré, el siguiente tomado del *De Motu* (1721; traduzco de la versión inglesa de Luce):

> Para aclarar la naturaleza es ocioso aducir cosas que no son ni evidentes a los sentidos ni inteligibles a la razón. Veamos, entonces, qué es lo que nos dicen los sentidos y la experiencia y la razón que los acompaña. Hay dos clases supremas de cosas, el cuerpo y el alma. Con ayuda de los sentidos conocemos la cosa extensa, sólida, móvil, con figura y dotada de otras cualidades que se muestran a los sentidos; pero la cosa que siente, que percibe, que piensa, la conocemos mediante cierta conciencia interior. Además, vemos que esas cosas son totalmente distintas y muy heterogéneas la una de la otra. Hablo de las cosas conocidas, pues de las desconocidas no tiene caso hablar. (*De Motu*, 21)

La formulación que aquí expresa Berkeley no podía ser más cartesiana: la extensión (aunque no sólo ella) le pertenece al cuerpo (pero el mismo ya no será una sustancia material) y enfrentada a éste, se presenta la cosa que piensa (con respecto a la que Berkeley retendrá la caracterización de sustancia espiritual). Las siguientes consideraciones son pertinentes para subrayar la diferencia: la extensión, así como los colores y toda la gama de (ideas de) cualidades primarias y secundarias lockeanas, no son atributos del espíritu, sino que son propias de las ideas perceptuales berkeleyanas como tales (son "cualidades que se muestran a los sentidos").[3] Berkeley, en los Principios, subrayó claramente lo anterior:

3 Malebranche, en la *Investigación*, III, ii, 7, S iv, donde trata de 'Cómo se conoce el alma', tiene el siguiente pasaje (añadido en la 2ª edición; en [9] 1, pp. 257-8) que Berkeley pudo haber tomado en cuenta:

Cela peut aussi servir à prouver que les idées qui nous représentent quelque chose hors de nous ne sont point des modifications de notre âme; car si l'âme

...podría objetarse, quizás, que si la extensión y la figura existen sólo en la mente, se sigue que la mente es extensa y tiene figura, puesto que la extensión es un modo o atributo que (para decirlo como en la Escuela) se predica del sujeto en el que existe. Respondo [que] esas cualidades están en la mente sólo como ella las percibe, esto es, no a manera de *modo* o *atributo*, sino sólo a manera de *idea* y de ninguna forma se sigue que el alma o la mente sea extensa porque la extensión sólo existe en ella, así como tampoco es roja o azul porque esos colores existen en ella y en ningún otro lugar, como todo el mundo lo reconoce... (*PCH*, 1, 49)

voyait toutes choses en considérant ses propres modifications, elle devrait connaître plus clairement son essence ou sa nature que celle des corps, et toutes les sensations ou modifications dont elle est capable que les figures ou modifications dont les corps sont capables. Cependant elle ne connaît point qu'elle soit capable d'une telle sensation par la vue qu'elle a d'elle-même en consultant son idée, mais seulement par expérience; au lieu qu'elle connaît que l'étendue est capable d'un nombre infini de figures par l'idée qu'elle a de l'étendue. Il y a même de certaines sensations, comme les couleurs et les sons, que la plupart des hommes ne peuvent reconnaître si elles sont ou ne sont pas des modifications de l'âme, et il n'y a point de figures que tous les hommes, par l'idée qu'ils ont de l'étendue, ne reconnaissent être des modifications des corps.

[Esto también puede servir para probar que las ideas que nos representan alguna cosa fuera de nosotros no son, en manera alguna, modificaciones de nuestra alma; pues si el alma viese todas las cosas al considerar sus propias modificaciones, ella debería conocer más claramente su esencia o su naturaleza que la de los cuerpos y todas las sensaciones o modificaciones de las que ella es capaz, que las figuras o las modificaciones de las que son capaces los cuerpos. Sin embargo, ella para nada sabe ser capaz de cierta sensación por la visión que tiene de sí misma al consultar su idea, sino solamente por experiencia; en lugar de esto, ella sabe que la extensión es capaz de recibir un número infinito de figuras por la idea que tiene de la extensión. Incluso hay ciertas sensaciones, como los colores y los sonidos, que la mayoría de los hombres no puede reconocer si son o no son modificaciones del alma y no hay figura alguna que todos los hombres, por la idea que de ella tienen, no reconozcan como modificación de los cuerpos.]

En los *Diálogos* (1713), Berkeley subraya que:

> *Filonús.* ... me es evidente que no hay sustancia en la que las ideas puedan existir que no sea el espíritu; todos están de acuerdo en que los objetos que se perciben inmediatamente son ideas y nadie puede negar que las cualidades sensibles son objetos que se perciben inmediatamente. Por tanto, es evidente que no puede haber ningún *substratum* de esas cualidades sino el espíritu, en el que existen, no a la manera de modo o propiedad, sino como una cosa percibida en aquello que la percibe ... (3D, p. 237)

El espíritu, en tanto que actividad, estará caracterizado por tener voliciones, tanto para alcanzar ciertos resultados no perceptuales como para percibir las ideas, *eminentemente pasivas*, que Dios pone ante él. De esta manera, las ideas *no* son algo en el espíritu, sino que son algo *para* el espíritu, al menos para los espíritus finitos que somos nosotros. Nuevamente en los *Diálogos*, Berkeley nos dice:

> *Filonús.* ... Concedo que propiamente no tengo ninguna idea sea de Dios o de cualquier otro espíritu pues, al ser éstos activos, no pueden representarlos cosas perfectamente inertes como lo son nuestras ideas. Sin embargo, sé que yo, que soy un espíritu o sustancia pensante, existo, tan ciertamente como sé que existen mis ideas. Además, sé lo que significan los términos *yo* y *yo mismo* y sé esto de manera inmediata o intuitiva, aun cuando no lo perciba, como percibo un triángulo, un color o un sonido. La mente, el espíritu o el alma es esa cosa indivisible, inextensa que piensa, actúa y percibe. Digo *indivisible* porque es inextensa e *inextensa* porque las cosas extensas, con figura y móviles son ideas y claramente no es una idea, ni semejante a una idea, aquello que percibe las ideas, que piensa y que tiene voliciones. Las ideas son cosas inactivas y percibidas y los espíritus son una especie de ser totalmente distinto de aquéllas. (3D, p. 231)

Aquí, además de afirmar la distinción extenso-pensante (que transforma mejor en la dicotomía pasivo-activo), Berkeley presenta su tesis del espíritu como unidad: es esa unidad que soy yo; por otra parte, disocia claramente la manera que tenemos de conocer (percibir) las ideas, de la manera como entramos en contacto cognitivo con nuestro espíritu. Finalmente, subraya también una tesis semántica no ideísta, en la que el significado de las expresiones no está dado por ideas, ya que ni de Dios ni de nuestros espíritus tenemos ideas, aun cuando entendemos perfectamente el significado de estos términos.[4] Lo que ciertamente no figura en el pasaje citado es la identificación cartesiana materia = extensión. Según bien se sabe, lo extenso dejó de tener en Berkeley las características de ser material y sustancial, sus ideas son algo dependiente y están más relacionadas con lo espiritual. El siguiente pasaje de los *Principios* muestra lo que ahora he dicho:

4 Finalmente, para cerrar esta sección, deseo presentar un último pasaje de Berkeley en donde se matizan ideas ya expresadas y se añaden algunos detalles; el pasaje es de su diálogo en siete secciones *Alcifrón* (1732):

Eufranor: Te ruego me digas, Alcifrón, ¿no es totalmente inactiva una idea?
Alcifrón: Lo es.
Eufranor: Por tanto, un agente, una mente o espíritu activo no puede ser una idea o similar a una idea. De aquí parece seguirse que las palabras que denotan un principio activo, un alma o un espíritu, en un sentido estricto y propio, no pueden representar ideas. Y, sin embargo, las mismas tampoco carecen de significado, puesto que entiendo lo que significa el término Yo o yo mismo o sé a qué se refiere, aunque no sea ninguna idea o cosa parecida, sino lo que piensa, tiene voliciones, aprende ideas y opera sobre ellas. Ciertamente debe concederse que tenemos alguna noción, que entendemos o sabemos lo que significan los términos yo mismo, voluntad, memoria, amor, odio y demás aun cuando, para hablar con precisión, estas palabras no sugieren otras tantas ideas distintas. (*Alc.*, VII, 5)

Cosa o *ser* es el nombre más general de todos; bajo él caen dos clases totalmente distintas y heterogéneas que nada tienen en común sino el nombre, a saber, *espíritus* e *ideas*. Los primeros son *sustancias activas, indivisibles*; las últimas son *seres inertes, fugaces, dependientes*, que no subsisten por sí mismos, sino que se apoyan o existen en las mentes o en las sustancias espirituales. (*PCH*, I, 89)

2. *Espíritu e ideas*

Algo que surge con claridad de las anotaciones citadas es la distinción fuerte que Berkeley establece entre el espíritu (la mente) y las ideas. Estas últimas son inertes, el primero es activo y actúa sobre aquéllas en términos de un perceptor sobre el objeto percibido. Además de querer establecer una fuerte diferencia entre ideas y espíritu, Berkeley desea mantener una gran relación de cercanía entre estas dos entidades: éste es su principio central, *esse es percipi* o sea que la existencia de las ideas depende de que sean percibidas. Pero entonces, teniendo en cuenta estas dos propuestas, es claro que las ideas *no* pueden ser aspectos o modificaciones del espíritu. De las dos propuestas se sigue que las ideas no son partes o atributos del espíritu, sino que son algo ajeno al mismo sobre las que éste puede ejercer la acción de percibirlas. Esto el mismo Berkeley lo expresó de manera clara en uno de los pasajes de los *PCH* que cité páginas atrás (*cf. supra PCH* I, 49 en p. 75). Lo anterior significa que la existencia de ciertas ideas no depende de la percepción de (mentes) espíritus finitos, ya que éstos las perciben como algo ajeno a sí mismos. Hay, es verdad, según Berkeley lo señala (*cf. supra*, en p. 71, las citas de *CF* y de *PCH*, así como la cita de A.A. Luce en pp. 71-2), ideas sobre las que, *aparentemente*, los espíritus tienen un dominio autónomo e inmediato, como lo son las de la fantasía o imaginación, pues éstas, claramente, pueden traerse, a voluntad, al primer plano de la conciencia. Pero, ¿no es

extraño, dadas las tesis generales de Berkeley aquí reseñadas, que pueda el espíritu conjurar, a voluntad, entidades de un carácter tan totalmente ajeno a sí mismo? Dado que las ideas no son parte del espíritu, es difícil explicar cómo puede la mente *producirlas* a voluntad. Éste es un problema general de la propuesta cartesiana de la mente como inmaterial e inextensa. Descartes apelaría, aunque sea dudosa (o nula) su fuerza explicativa, a rastros dejados por los espíritus animales en un cuerpo para explicar los recuerdos del espíritu ligado al mismo; pero, ni siquiera esta salida falsa la tiene Berkeley. Ciertamente, Descartes mismo concede que las ideas dejan su rastro en el espíritu, por lo que de allí mismo sería posible recuperarlas de alguna manera.[5]

5 Le agradezco a Laura Benítez el haberme señalado esta última propuesta cartesiana que me parece muy objetable en Descartes, quizás por las razones que da Berkeley. Con respecto a la primera propuesta, Berkeley mismo se oponea ella –y ciertamente no tiene ninguna posibilidad de aceptarla–, según lo muestra la discusión de este asunto en los *Diálogos*, donde Filonús rechaza la "explicación moderna" del origen de las ideas que le presenta Hylas –siguiendo, quizás, las propuestas a este respecto de Malebranche en la *Investigación*–:

Filonús. Y, ¿llamas a esto una explicación de la manera como nos vemos afectados por ideas?

Hylas. ¿Porqué no, Filonús; tienes algo que objetar acerca de esto?

Filonús. Primero me gustaría saber si entiendo correctamente tu hipótesis. Tú haces que ciertos trazos en el cerebro sean las causas u ocasiones de tus ideas. Por favor dime, ¿con cerebro te refieres a alguna cosa sensible?

Hylas. ¿Qué otra cosa crees que podría querer decir?

Filonús. Todas las cosas sensibles son percibibles inmediatamente, las cosas que son percibibles inmediatamente son ideas y éstas existen sólo en la mente. Si no me equivoco, hace tiempo que, hasta esto, habíamos quedado de acuerdo.

Hylas. No lo niego.

Filonús. Por tanto, el cerebro del que hablas, al ser una cosa sensible, existe sólo en la mente. Ahora me gustaría saber si piensas que es razonable suponer que

2.1. Berkeley, Malebranche y Dios

Lo que parece seguirse de lo anterior es que incluso las ideas de la imaginación dependen de la buena voluntad de Dios para que las mismas se presenten ante la mente. Ésta no puede conjurar, por sus propios medios, la presencia de estas entidades extensas, inertes, etc., sino que tiene que contar con el auxilio de Dios. Aquí aparece una variante ocasionalista en la filosofía de Berkeley: Dios se vale de nuestras voliciones (de las voliciones de los espíritus finitos), como una ocasión para poner ante nosotros las ideas, sean de imaginación o de percepción, relacionadas, estas últimas, con nuestros movimientos corpóreos. Ciertamente el ocasionalismo está presente en este último caso, ya que para que las voliciones de

---

una idea o cosa que existe sólo en la mente ocasiona todas las otras ideas. Y, si así lo piensas, dime, ¿cómo das cuenta del origen de esa misma idea primaria o cerebro? (3D, p. 209)

La tesis *moderna*, cartesiana, también se expresa en una reseña de los *3D* de Berkeley (en el *Journal Litteraire* de mayo y junio de 1713, en pp. 147-60; en [4] 1, pp. 113-26), en las siguientes palabras (el pasaje está en p. 117 de [4]):

Quand on dit qu'un Corps a quelqu'une de ces dernieres qualitez [secondes], on ne veut pas dire que ce Corps ressente quelque chose de semblable à ce qu'il nous fait sentir: on veut simplement dire que ce Corps, par la configuration de ses parties, ou par le mouvement qu'il comunique à la matiere qui l'environne, ou qui se réfléchit sur ce Corps, a la propriété de mettre nos fibres dans un certain mouvement, qui par le cerveau se communique à l'ame, en y faisant naître quelqu'une de ces idées qu'on nomme Sensations.

[Cuando se dice que un cuerpo tiene alguna de estas últimas cualidades [secundarias], no se quiere decir que este cuerpo resienta alguna cosa similar a la que él mismo nos hace sentir; simplemente se quiere decir que este cuerpo, por la configuración de sus partes o por el movimiento que él comunica a la materia que lo rodea o que se refleja sobre este cuerpo, tiene la propiedad de poner nuestras fibras en cierto movimiento que se comunica al alma por el cerebro y allí hace nacer alguna de las ideas que denominamos sensaciones.]

los espíritus finitos se muestren en términos de movimientos corpóreos es preciso que éstos se manifiesten como cambios en las ideas de percepción: tales movimientos corpóreos no son sino la manifestación de la voluntad divina en términos de cambios de tales ideas perceptuales presentes a los perceptores finitos. En esto Berkeley no tiene sino que estar de acuerdo, pues las ideas del mundo perceptual Dios las presenta a nuestros espíritus, sin que dependan, para nada, de nuestra voluntad. En *PCH* I, 147, donde Berkeley se propone poner de manifiesto la enorme variedad de formas como se muestra la presencia de Dios, nos dice:

> Podemos incluso aseverar que la existencia de Dios se percibe de manera mucho más evidente que la existencia de los hombres, porque los efectos de la naturaleza son infinitamente más numerosos y considerables que los que se ascriben a agentes humanos. No hay ninguna marca peculiar que denote a un ser humano o algún efecto que éste produzca que no evidencie con mayor fuerza el ser de ese espíritu que es el *Autor de la Naturaleza.* Pues es evidente que, al afectar a otras personas, la voluntad de un hombre no tiene más objeto que tan sólo el movimiento de los miembros de su cuerpo; pero que a tal movimiento lo acompañe, o que despierte, cualquier idea en la mente de otros, esto depende totalmente de la voluntad del Creador. Él es el único que *por mantener todas las cosas por la Palabra de su Poder,* mantiene esa comunicación [*intercourse*] entre espíritus mediante la cual somos capaces de percibir la existencia los unos de los otros ...

En este sentido, la siguiente propuesta de Malebranche describe perfectamente este aspecto de la filosofía de Berkeley (*malgré lui*):

> Ahora bien, me parece muy cierto que la voluntad de los espíritus no es capaz de mover el cuerpo más pequeño que haya en el mundo,

pues es evidente que no hay ninguna liga necesaria entre la voluntad que tenemos, por ejemplo, de mover nuestro brazo y el movimiento de nuestro brazo. Es verdad que se mueve cuando nosotros lo queremos y que, así, somos nosotros la causa natural del movimiento de nuestro brazo. Pero las causas *naturales* de ninguna manera son causas verdaderas, no son sino causas *ocasionales* que no actúan sino por la fuerza y la eficacia de la voluntad de Dios ... (*Investigación*, VI, 2.ª pte., iii)

Y, aunque Berkeley haya dicho en sus *Comentarios filosóficos* que,

S Movemos nuestras Piernas nosotros mismos. somos nosotros quienes inducimos su movimiento. En esto difiero de Malebranche. (*CF* 548)

no es claro, en manera alguna, que pueda haber una diferencia en el aspecto clave de sostener que las voliciones de un espíritu finito son *la causa* de que se modifiquen las ideas que no dependen de su voluntad. Claramente, en este caso, Berkeley tiene que aceptar la propuesta ocasionalista de Malebranche.

Si el ocasionalismo, en el caso de mis voliciones de mover mi cuerpo, por ejemplo, parece que puede explicarse plausiblemente en términos de *la presencia perceptual ante mi espíritu de mi cuerpo*, ¿cómo se puede explicar mi volición de tener un recuerdo de ...? Y aquí es difícil saber cómo eliminar los puntos suspensivos, ya que para precisar aquello acerca de lo que quiero tener un recuerdo, ¿no sería necesario que en mi mente estuviera *ya* eso que pretendo recordar? Pero si la mente es diáfana, además de inextensa, etc.[6] *¿dónde* se puede decir que esté tal cosa? Berkeley puso de manifiesto estos problemas en sus *CF*:

6 *Cf.*, con respecto a la mente diáfana y otros temas relacionados, [10], pp. 21 y ss.

S Mem: inquirir diligentemente por ese extraño Misterio a saber Cómo es que puedo buscar, pensar en este o en ese Hombre, lugar, acción cuando nada parece Introducirlos en mis pensamientos. cuando no tienen conexión perceptible alguna con las Ideas que ahora me sugieren mis sentidos. (*CF*, 599)

Y, en los mismos *Comentarios*, ahora teniendo en mente la tesis de la inquietud (*uneasiness*) de Locke, vuelve a formular el problema, aun cuando sin ser consciente de la seriedad que el mismo tiene para su posición general:

S Yo puedo inducir el traer a la mente algo que es pasado, aun cuando al mismo tiempo aquello que traigo a la mente no estuviera en mis pensamientos antes de esa Volición mía y consecuentemente no podría haber tenido inquietud por su causa (*CF*, 707)

El problema que le surge a Berkeley quiero formularlo de esta otra manera: dentro del esquema berkeleyano de explicación ni siquiera podemos actuar directamente en el caso de aquellas ideas que supuestamente tenemos más a la mano, las propias de los recuerdos y la imaginación, pues la mente, el espíritu, no tiene manera de crear *ex nihilo* (desde su perspectiva) ninguna de las ideas que, *supuestamente* desea imaginar, ya que las mismas, por definición, no son modos o atributos del espíritu, por lo tanto, son ajenas al mismo, por lo tanto, el espíritu no tiene manera de hacerlas presentes ante sí.[7] En esta coyuntura es per-

7 Cuando Malebranche, en el Prefacio al tomo II de la *Investigación* rechaza que él hubiese afirmado, como se lo achaca Simon Foucher, que "todas nuestras ideas no son sino maneras de ser de nuestra alma" el abate Rabbe en su libro *Simon Foucher* (París, 1867) señala, en contra de Malebranche, que si nuestras ideas no son formas de ser de nuestra alma, ellas nos son tan extrañas como los objetos mismos; nosotros no las conocemos inmediatamente. *cf.* [5] p. 250 y n. 7 y [12] p. 96.

tinente recordar la crítica de Malebranche a quienes '...creen que nuestras almas tienen el poder de producir las ideas de las cosas en las que quieren pensar...' y luego más adelante sigue diciendo que

> ...si los hombres no se precipitasen en sus juicios ... deberían tan sólo concluir que, según el orden de la naturaleza, su voluntad es ordinariamente necesaria a fin de que tengan esas ideas, pero no que la voluntad es la causa verdadera y principal que las hace presentes a su espíritu y aún menos que la voluntad las produce de la nada ... (*Investigación* III, 2ª pte., iii)

y considero que esto, que la voluntd se vea ('ordinariamente'), a lo más, como razón necesaria *pero nunca* sufiente de la producción de ideas se aplica perfectamente a la filosofía de Berkeley en el caso de las ideas de imaginación, aun cuando, en última instancia, no es claro (ni en el caso de Malebranche, ni en el de Berkeley) cómo pueda intervenir la voluntad para producir algo *de lo que no se tiene idea* alguna.

El único que puede presentarnos las ideas de recuerdos o de la imaginación, es Dios; pero, entonces, es también Dios quien *tiene que* producir en nosotros el deseo de imaginar φ, para que podamos llegar a imaginar φ, lo que, nuevamente, será un producto de Dios en nuestro espíritu. La visión que parece surgir de la propuesta anterior es la de Dios susurrando a nuestro espíritu sugerencias de deseos e inclinaciones y nuestro espíritu, nuestra voluntad, tendrá, como única finalidad, la de asentir o rechazar las insinuaciones divinas.

3. *¿Está Berkeley en un callejón sin salida?*

Tal como he formulado el problema berkeleyano, la conclusión que parece seguirse es que, a pesar de Berkeley, Dios está presente

de manera aún más inmediata a nuestras vidas que lo que aquél estaba dispuesto a conceder. Ciertamente, la escasa libertad espiritual parece conducir a serios problemas de tipo teológico y moral: si es verdad que Dios es quien susurra en nosotros sugerencias de deseos, ¿no se seguiría de esto que puede susurrarnos deseos profundamente abyectos y revestirlos con atuendos por demás llamativos y atrayentes que nos hagan aceptarlos, ponerlos en práctica y, así, pecar?[8] ¿Sería justo que por esto nos condenásemos? Por otra parte, ¿qué sentido tendría atribuirle recuerdos a un espíritu puramente volitivo? En *PCH* I, 138, Berkeley ofrece la siguiente caracterización del espíritu, con la cual podemos estar de acuerdo,

8 Acerca de esto es posible citar el siguiente pasaje ejemplar de Malebranche, en el que el filósofo francés saca una conclusión similar a la mía, sin darle los visos heréticos que yo encuentro en la misma:

No hay, pues, sino un solo Dios verdadero y una sola causa que sea verdaderamente causa y no se debe imaginar que lo que precede a un efecto sea su verdadera causa. Dios ni siquiera puede comunicar su potencia a las criaturas, si seguimos las luces de la razón, él no puede hacerlas causas verdaderas, no puede hacerlas dioses. Pero, aun cuando pudiese, no podemos concebir por qué lo desearía. Los cuerpos, los espíritus, las inteligencias puras, todo esto nada puede. Es aquel que ha hecho los espíritus quien los esclarece y quien los mueve. Es aquel que ha creado el cielo y la tierra quien rige sus movimientos. En fin, es el autor de nuestro ser quien ejecuta nuestras voluntades. *Semel jussit, semper paret.* Incluso él mueve nuestro brazo cuando nos servimos de él en contra de sus órdenes; pues él se queja, a través de su profeta, que nosotros hacemos que sirva a nuestros deseos injustos y criminales. ([9] VI, II, iii)

El profeta al que alude Malebranche es Isaías, quien en 43:24 dice:

Antes bien, te has servido de mí en tus pecados y me has causado gran pena con tus iniquidades.

La lectura que hace Malebranche de este pasaje es que el hombre peca por desobediencia; mi lectura es que el hombre cae (pero no peca) por ignorancia. Quizás, si la lectura primera fuera correcta, el castigo podría ser justo. En el segundo caso, con él se cometería una injusticia total.

siempre y cuando por "pensar" Berkeley se refiera a razonar y a imaginar en el sentido restringido que de inmediato propondré:

> ... Pues por la palabra *espíritu* sólo nos referimos a aquello que piensa, tiene voliciones y percibe; esto y esto sólo constituye el significado de ese término.

Como lo he señalado, Berkeley tiene un serio problema para dar cuenta de la imaginación y de los recuerdos dentro de su sistema; sin embargo, el problema parece no ser tan grave para la imaginación o la fantasía que actúe sobre lo presente en la percepción, en tanto que Dios nos provee de los perceptos que conforman el mundo real; de esta manera, el espíritu puede variar y recomponer esas presentaciones perceptuales: la voluntad elige ciertos aspectos separados del mundo real y desea unirlos; Dios se los presenta unidos. Los siguientes pasajes muestran la dependencia, con respecto a la percepción, que Berkeley les atribuye a las ideas de la imaginación:

> S Tener Ideas no es lo mismo que la Percepción. un Hombre puede tener Ideas cuando él sólo Imagina. Pero entonces esta Imaginación presupone Percepción. (*CF*, 582)

> M hablando con propiedad Idea es el cuadro que hace la Imaginación éste es el parecido y se refiere a la Idea o (si lo deseas) a la cosa. (*CF*, 657a)

> M Las Ideas del Sentido son las cosas Reales o Arquetipos. Las Ideas de la Imaginación, los Sueños etc son copias, imágenes de éstas. (*CF*, 823)

En los dos primeros pasajes, Berkeley parece aún titubear acerca del uso del término "idea"; en ellos, esta expresión parece

referirse a algo secundario, no real –en un sentido fuerte de este término– y dependiente de lo real que se da en la percepción. En el tercer pasaje ya acepta llamar "idea" tanto a las presentaciones perceptuales como a las presentaciones de la imaginación, pero en los tres pasajes Berkeley mantiene la dependencia señalada de las ideas de la imaginación con respecto a la percepción. No obstante, al menos en estos pasajes, nada se dice acerca de cómo es la relación de dependencia entre percepción e imaginación. Lo que en este escrito he sostenido es que la misma es tal que la segunda requiere de la presencia *en acto* de la primera para poderse efectuar, esto es, dentro del esquema berkeleyano sólo si hay datos (perceptuales, generados por la bondad divina) presentes ante la mente, es que ésta puede actuar sobre ellos y recombinarlos (con auxilio divino mínimo, quizás).

Por otra parte, según lo señalé páginas atrás, hay pasajes en los que Berkeley parece tomar conciencia de los problemas que encierra su posición pero, lamentablemente, no tenemos ninguna formulación suya en la que proponga alguna salida. Presento aquí nuevamente, para facilidad de referencia, los textos a los que me refiero:

> S Mem: inquirir diligentemente por ese extraño Misterio a saber Cómo es que puedo buscar, pensar en este o en ese Hombre, lugar, acción cuando nada parece Introducirlos en mis pensamientos. cuando no tienen conexión perceptible alguna con las Ideas que ahora me sugieren mis sentidos. (*CF*, 599)

> S Yo puedo inducir el traer a la mente algo que es pasado, aun cuando al mismo tiempo aquello que traigo a la mente no estuviera en mis pensamientos antes de esa Volición mía y consecuentemente no podría haber tenido inquietud por su causa (*CF*, 707) -

G. Por qué no podemos concebir que sea posible para Dios crear cosas a partir de la Nada. ciertamente nosotros de alguna manera creamos siempre que imaginamos. (*CF* 830)

Encuentro que puedo despertar a placer ideas en mi mente y variar y cambiar el escenario con tanta frecuencia como lo crea adecuado. No hay más que tener voliciones y directamente esta o aquella idea surge en mi fantasía y con este mismo poder se elimina y da paso a otra. Este hacer y deshacer ideas debe, con propiedad, denominar a la mente activa. Todo esto es cierto y se funda en la experiencia; pero cuando hablamos de agentes no pensantes o de despertar ideas con exclusión de la voluntad, sólo nos engañamos con palabras. (*PCH* I, 28)

Estos pasajes son centrales para el estudio de la cuestión que aquí me ocupa. Anteriormente sugerí cómo, *a partir de los datos perceptuales presentes*, podría el espíritu modificar (imaginativamente y con el auxilio de Dios) algunas relaciones entre tales datos sin tener que esforzarse mucho. Empero, algo diferente sucede con los recuerdos, ya que éste es el caso en el que la imaginación actúa sin tener los datos perceptuales presentes. En *CF* 599 Berkeley claramente tiene en cuenta el problema, al caracterizar correctamente como "Misterio" el hecho de recordar, de tener recuerdos, "cuando no tienen conexión perceptible alguna con las Ideas que ahora me sugieren mis sentidos"; en *CF* 707 Berkeley toma como un hecho que los recuerdos suceden (como creación *ex nihilo*, desde la perspectiva de la mente finita) y esto le permite ir en contra de la tesis de la inquietud de Locke.[9]

9 Para un análisis de la propuesta de Berkeley sobre esta tesis de Locke, véanse mis *Estudios berkeleyanos*, [11], cap. 6, pp. 131-42.

El tercer pasaje de los *CF*, 830, presenta el aspecto importante y básico en los presupuestos de la época, que sería el único, quizás, que apoyaría la propuesta de Berkeley: según la *Biblia* (Gen. 5: 1) el hombre fue creado a imagen y semejanza de Dios[10] y, en este sentido, se le podría atribuir cierta limitada capacidad creativa. El pasaje en cuestión, ciertamente toma en cuenta las dificultades de la tesis de la imaginación para un espíritu como el berkeleyano: si éste no tiene capacidad de crear, entonces no puede imaginar.

Aun cuando la propuesta sea problemática, aceptémosla sin conceder su verdad. A pesar de todo, queda el problema de cómo escoger los recuerdos, de cómo saber qué recordar, y para esto no puedo sino remitirme nuevamente a la buena voluntad de Dios.[11] ¿Cómo podríamos crear *ex nihilo* (con respecto a los *contenidos del espíritu*) un recuerdo si no sabemos qué queremos recordar? Y, ¿cómo podemos saber qué recordar si el espíritu

10 Mucho insistió Laura Benítez en subrayar que el alma debería ser una especie de ser creativo, aun cuando en un sentido restringido, no comparable con el de la Divinidad. Esto, según su sugerencia, de alguna manera explicaría, dentro del sistema de Berkeley, la posibilidad de que los espíritus finitos produjesen ideas de la memoria y de la imaginación.

Lo que señalo acerca de esto es que no puedo ver la fuerza explicativa de decir algo como 'no sé cómo sucede, pero sucede de alguna manera'. Lo que creo que se puede decir acerca de la posición general de Berkeley, es que él no tiene forma de enfrentarse al problema que le crean las ideas de la memoria y de la imaginación sin que Dios le dé una mano para producir tales ideas. Lo que, en un sentido, es una explicación tan buena o tan mala como la de dotar al alma con poderes creativos.

11 Martha Bolton, en observaciones propuestas al final de la lectura de mi escrito, sugería que quizás podríamos suponer el espíritu berkeleyano como dotado por Dios de capacidades para llevar a cabo la proeza de recordar o de imaginar (sin apoyos perceptuales presentes). Mi impresión entonces y ahora sigue siendo la de que, si no se tienen a la mano elementos mejores para explicar el ser, la existencia, de los recuerdos de α cuando α no los recuerda, que el que Dios los conserve en el ser, entonces no hay mucha o ninguna diferencia con mi propuesta de Dios susurrando sugerencias al espíritu.

no tiene contenidos? Si tiene contenidos de algún tipo, *no* pueden ser ideas; lo que caracteriza los espíritus berkeleyanos serán, más bien, capacidades; ciertamente la de percibir, la de razonar y la de ejercer su voluntad sobre contenidos presentes. Pero ¿cómo puede allegarse contenidos pasados que *estuvieron* presentes ante él a la manera de ideas (de percepción)?

Aquí me parece pertinente señalar que la propuesta de Malebranche viene a resolver algunas de las dificultades que se le plantearon a Berkeley. El padre del Oratorio mantuvo que Dios nos muestra las ideas y nos impone las sensaciones, los deseos, las inclinaciones ('sentiments' en la expresión de Malebranche), etc. Conforme al filósofo francés, en Dios están las ideas, en tanto que inteligibles para los hombres y las sensaciones, sentimientos, etc., Dios, aun cuando no tenga ni pueda tener estas afecciones, las impone a los espíritus en tanto que Él sabe (como su Creador omnisciente) cuáles les convienen a éstos. Así, el espíritu finito, a la manera como lo he señalado, puede llevar a cabo (y son las únicas que puede realizar) funciones de selección y acción, siempre con el auxilio directo de Dios. Sin embargo, si lo que se desea mantener es una posición dualista, con algún tipo de interacción entre los dos elementos ontológicos, Locke tiene un interesante y perceptivo pasaje, en su *Examen de la opinión del padre Malebranche de ver todas las cosas en Dios*, con respecto al "misterio" (palabra que empleó Berkeley) de la conversión de impresiones sensoriales en ideas sensibles:

> Pienso que entiendo las impresiones que hacen en la retina los rayos de luz y pueden concebirse los movimientos que de allí se continúan al cerebro y estoy persuadido que éstos producen ideas en nuestras mentes, pero de una manera que me es incomprensible. Esto lo puedo resolver sólo por la buena disposición de Dios, cuyas vías nos rebasan.[12]

12 En [7], § 10, p. 217.

En el mismo escrito Locke ofrece buenas razones para dudar de que la propuesta de Malebranche, de ver todas las cosas en Dios, sea algo que pueda tomarse como una explicación de nuestra percepción del mundo.[13] Creo que esta propuesta agnóstica es el único camino que le queda a Berkeley, dadas las premisas que figuran en su propuesta, o bien tendrá que modificar profundamente tesis centrales a su sistema.

13 Para Malebranche, cualquier cosa que perciba el espíritu debe estar realmente presente ante y unida a él; y si hay ciertas experiencias que son *propias* de los espíritus finitos, hay otras que les son externas y para las cuales es preciso que sea Dios quien las presente o las imponga a los espíritus y así, de alguna manera, es en Dios que obtenemos realmente tales experiencias. O, dicho de otra manera, conforme a Malebranche, Dios presenta las ideas (en tanto que manifestaciones *inteligibles*, no sensoriales) al espíritu y produce en él las sensaciones y sentimientos que nos permitan tener una captación *sensorial* del mundo físico. Acerca de esto, Locke ha señalado serias dificultades en su escrito sobre Malebranche, subrayando que, en la doctrina de éste,

... hay muchas expresiones que, a mi entender, al no contener ninguna idea clara, es probable que, con su sonido, eliminen muy poco de mi ignorancia. P.ej. ¿qué es "estar íntimamente unido al alma"? ? qué es que dos almas o espíritus estén íntimamente unidos, pues la unión íntima es una idea que se toma de los cuerpos cuando las partes de uno penetran la superficie del otro y tocan sus partes internas; cuál es la idea que debo tener de unión íntima entre dos seres tales que ninguno de ellos tiene extensión o superficie alguna? Y si esto no se explica de tal manera que me dé una idea clara de su unión, me hará entender muy poco más de la naturaleza de las ideas en mi mente, cuando se dice que las veo en Dios, quien por estar "íntimamente unido al alma" se las exhibe, que cuando sólo se dice que por designio de Dios se producen en la mente por ciertos movimnientos de nuestros cuerpos, a los que están unidas nuestras mentes. Forma que, por imperfecta que sea para explicar este asunto, será tan buena como cualquier otra que no elimine, mediante ideas claras, mi ignorancia de la manera de mi percepción. [En [8] § 4, pp. 212-3]

Con respecto a la última parte del escrito de Locke podemos estar plenamente de acuerdo y aquí aparece una crítica muy pertinente a la posición de Berkeley. La única ventaja que pueda tener el prescindir de la materia, de los cuer-

4. Recapitulación y conclusión

El camino que he recorrido en este escrito consta de los siguientes pasos:

i) mostrar que Berkeley establece una tajante separación entre espíritu e ideas que, ontológicamente, los hacen ser por completo heterogéneos;
ii) recordar que Berkeley hace que el espíritu sea una sustancia activa que percibe ideas y que éstas no son (ni pueden serlo ontológicamente) modos o atributos de los espíritus;
iii) indicar que Berkeley hace una separación importante entre ideas ajenas a cualquier espíritu finito e ideas de la memoria y de la imaginación, propias de los espíritus particulares, en el sentido de que se pueden manejar a voluntad y placer.

Pero, dadas las condiciones que señalo en i)-iii), concluyo que

iv) es imposible, sin la ayuda de Dios, que un espíritu finito pueda conjurar, a voluntad, ideas (sean de percepción o de imaginación) que no son ni pueden ser parte de ese espíritu.

Además,

v) si las ideas de imaginación (así como las perceptuales) no son ni pueden ser parte del espíritu, entonces es difícil (si no

---

pos, es una de economía ontológica, pero no de eficacia explicativa. Tanto la versión malebranchista como la cartesiana, con respecto a la percepción, están a la par acerca de poder explicativo: ninguna de las dos *explica* cómo es que un espíritu inextenso, para el que no valen las caracterizaciones propias de los objetos de percepción, llega a tener las ideas que sus autores dicen que tiene. Esto mismo, según los argumentos que he presentado en este escrito, creo que vale para la propuesta de Berkeley.

es que imposible) explicar cómo pueda éste escoger un tema para imaginar, si no es mediante *la presentación que* Dios haga ante él de diversas ideas que el espíritu podrá (ahora sí) *escoger* (pero *no conjurar*) a voluntad.

Por lo tanto, desde esta perspectiva,

vi) el espíritu, dentro del esquema filosófico berkeleyano, se reduce a ser puramente perceptor (no conjurador) de ideas, sin que pueda, en manera alguna, actuar sobre ellas (conjurarlas, modificarlas, eliminarlas, etc.), a no ser por la voluntad de Dios; de ser esto así, la única libertad que le resta es la de escoger (determinar volitivamente) alguna(s) de las sugerencias que Dios le presenta.

En alguno de los pasajes de los *CF* Berkeley nos preguntó,

S Pr: ¿cómo se distingue el alma de sus ideas? no podría haber alma ciertamente si no hubiese ideas sensibles, ninguna percepción, recuerdos, amor, temor, etc. ninguna facultad podría ejercerse. (*CF* 478)

Y él mismo, en su comentario a la nota anterior, nos dice:

S El alma es la voluntad hablando con propiedad y como tal distinta de las Ideas. (CF 478a)

Ésta debería ser la propuesta básica, última, de Berkeley acerca de la mente o espíritu: lo único que éste puede ser (*per se*), dentro de esta filosofía, es una voluntad y un entendimiento discursivo; esto es, dada la forma que tiene Berkeley de caracterizar los términos de su propuesta, mentes o espíritus e ideas, la tesis de Malebranche puede ofrecerse para mantener, así, la

consistencia. Berkeley, sin embargo, optó por algo diferente, según lo muestra el siguiente pasaje tomado de los *Diálogos* (así como los pasajes citados con anterioridad, en pp. (71) 87-88 *CF* 830 y *PCH* I, 28):

> *Filonús.* Las ideas formadas por la imaginación son débiles y confusas; además, *tienen una dependencia total de la voluntad* ...
>
> (3D, p. 235) [El subrayado me pertenece.]

Pero, por las razones que he dado en este escrito, esto no puede sostenerlo de manera consistente con otras propuestas básicas de su pesamiento.[14]

14 Aquí deseo señalar la existencia de un artículo de Nicholas Jolley, 'Berkeley and Malebranche on Causality and Volition' (en *Central Themes in Early Modern Philosophy*, Hackett Publishing Company, Inc.; Indianapolis y Cambridge, 1990; pp. 227-44) que yo desconocía cuando redacte este ensayo y con el cual tengo profundas afinidades. En otra ocasión espero escribir un ensayo más amplio sobre este tema en el que se le dé un lugar especial al escrito de Jolley. [Nota añadida en febrero de 1993.]

Bibliografía

[1]. Bennett, Jonathan: *Locke, Berkeley, Hume: temas centrales.* Traducción al español de *Locke, Berkeley, Hume: Central Themes* (Oxford University Press, 1971), por José Antonio Robles. Instituto de Investigaciones Filosóficas; UNAM, México, 1988.

[2]. Berkeley, George: *The Works of George Berkeley Bishop of Cloyne* (1948-57); vols. 1-9. Thomas Nelson and Sons Ltd., 1964.

[2a]. _____: *Three Dialogues between Hylas and Philonus* (1713). En [2] 2, pp. 147-263. Hay diversas traducciones al español.

[2b]. _____: *Alciphron*, or the Minute Philosopher (1732). En [2] 3. Hay traducción al español: *Alcifrón o el filósofo minucioso.* Trad. Pablo García Castillo; Introd. y notas, Cirilo Flórez Miguel. Ediciones Paulinas; Madrid, 1978.

[3]. _____: *Comentarios filosóficos; Introducción manuscrita a Los principios del conocimiento humano; Correspondencia con Johnson.* Introducción, traducción y notas de José Antonio Robles. Instituto de Investigaciones Filosóficas, UNAM; México, 1989.

[4]. Berman, D. (ed.) *George Berkeley*, Eighteenth Century Responses Vols. 1-2. Garland Publishing, Inc.; New York & London, 1989.

[5]. Gouhier, H.: *La philosophie de Malebranche* et son expérience religieuse. Librairie Philosophique J. Vrin; Paris, 1948. (2eme ed.).

[6]. Grayling, A.C.: *Berkeley*, the Central Arguments. Gerald Duckworth & Co. Ltd.; London, 1986.

[7]. Locke, John: *An Examination of P. Malebranche's Opinion of seeing all Things in God.* En el vol. 9, pp. 211-55 de *The Works of John Locke.* A new edition, corrected. In ten volumes. London: 1823. Reprinted by Scientia Verlag Aalen, Germany, 1963. [Hay traducción al español en José A. Robles y Carmen Silva (comps.): *Escritos varios y correspondencia de (y sobre) John Locke*, colección Σ , UAM-I; México, 1993: pp. 35-81.

[8]. Luce, A.A.: *Berkeley and Malebranche* (1934). Oxford University Press, Oxford, 1967.

[9]. Malebranche, Nicolas: *Recherche de la Verité.* Introduction et texte établi par Geneviève Rodis-Lewis. Vols. 1-3. Librairie Philosophique J. Vrin; Paris, 1962 (vols. 1, 2), 1946 (vol. 3).

[10]. Pitcher, George: *Berkeley.* Traducción al español de *Berkeley* (Routledge & Kegan Paul Ltd., 1977), por José A. Robles García. Fondo de Cultura Económica; México, 1983.

[11]. Robles, José Antonio: *Estudios berkeleyanos*. Instituto de Investigaciones Filosóficas (UNAM); México, 1990.
_____ (comp): Véase [3].
[12]. Rodis-Lewis, Geneviève: *Nicolas Malebranche*. Presses Universitaires de France; Paris, 1963.
[13]. Winkler, Kenneth P.: *Berkeley, an Interpretation*. Clarendon Press; Oxford, 1989.

La explicación de la conciencia en los *Nouveaux essais* de Leibniz

Martha Brandt Bolton*

1. Una tesis que subyace a gran parte de la argumentación de Leibniz en los *Nouveaux essai*, sostiene que todo lo que sucede en el mundo de manera natural, es explicable o inteligible. Puesto que Leibniz mantiene que las sustancias básicas que constituyen el mundo son o bien almas o bien entelequias animiformes, esta tesis prominente implica que es explicable cualquier cosa que suceda naturalmente en las almas.[1] Creo que esta doctrina es uno de los núcleos de sus desacuerdos de amplio espectro con su antagonista Locke. En la investigación del conocimiento de Locke, el agnosticismo acerca de "cómo piensa un hombre" es tanto una presuposición metodológica como una conclusión positiva. Locke duda de que seamos jamás capaces de *entender* cómo las partículas en movimiento causan las ideas de color, etc., cómo puedan las partículas causar jamás cualquier otro tipo de ideas, cómo opera la mente sobre el cuerpo o cómo pensamos.[2] En este escrito, me interesa estudiar cómo pretende Leibniz establecer que está mal

* Rutgers University

1 Véanse las páginas finales del prefacio (*NE* 60-68) en donde Leibniz arguye que cualquier cosa que pase naturalmente en una sustancia es explicable y, así, que el poder de pensar no le pertenece naturalmente a la materia. Véanse, también, *NE* 179, 380-1, 390. Las referencias y las citas de los *Nouveaus essais* son de la traducción de Peter Remnant y Jonathan Bennett (Cambridge University Press; Cambridge, 1981; lo abrevio como *NE*).

2 Véase el *Ensayo* I, i,2; II, viii, 22; IV, iii, 6, 13, 17, 28; IV, vi, 7, 14; II, xxiii, 23, 25, 28-29; II, xxvii, 12-3. Las referencias y las citas del *Ensayo* son del *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch (Clarendon Press; Oxford, 1975).

concebido el escepticismo explicativo de Locke; en particular, ¿cuál es la estrategia de Leibnizpara explicar la conciencia de los estados sensoriales y de los objetos externos? Hoy en día, está muy viva la intuición lockeana de que somos simplemente incapaces de entender cómo es que surgen los estados de conciencia a partir de sucesos en el cuerpo o, incluso, de cualquier otra cosa.[3] Esto le añade algún interés a la cuestión acerca de cómo interpretó Leibniz el problema presentado por Locke y cómo es que él se propone resolverlo.

El problema es serio, pues abarca varias aseveraciones centrales de la metafísica de Leibniz. Yo pasaré muy por encima de algunos problemas y propuestas importantes. Tan sólo voy a considerar la posición de Leibniz en los *NE*, aun cuando en las notas haga referencias a cartas del mismo periodo (1703-1705). El examen que hizo Leibniz de Locke le dio cierta forma al problema, lo tornó en un problema acuciante y yo estoy interesada en la respuesta de Leibniz a ese reto particular.

2. En parte, la solución de Leibniz al problema de explicar lo que sucede en un alma o sustancia, descansa en dos de sus famosas doctrinas, a saber, las mónadas sin ventanas y la armonía preestablecida. Conforme a la primera, no hay ninguna interacción causal, sea ésta la que sea, entre un alma (sustancia o mónada) y otras cosas creadas, incluyendo los cuerpos. Sea lo que sea que sucede en una sustancia:

> ... sucede espontáneamente en ella y surge de sus propias profundidades; pues ninguna sustancia creada puede tener una influencia sobre cualquier otra, de tal manera que todo le llega a una sustancia de sí misma (aun cuando, en última instancia, de Dios). *NE* 210

3 Véanse, p.ej., el 'Panpsychism' de Thomas Nagel en *Mortal Questions* (Cambridge University Press; Cambridge, 1979), especialmente pp. 189-90; Colin McGinn, *The Problem of Consciousness* (Blackwell; Oxford, 1991), especialmente cap. 1.

Así pues, todos los cambios que acontecen en un alma, los causan naturalmente esfuerzos o tendencias ya presentes en el alma.[4] Sin embargo, conforme a la doctrina de la armonía preestablecida, la serie autogenerada de sucesos dentro de un alma, corresponde exacta y completamente a sucesos tal como éstos acaecen en el resto del mundo creado. Hay una perfecta correspondencia entre lo que sucede en un espíritu y lo que sucede en su cuerpo; hay también una completa correspondencia entre lo que sucede en su cuerpo y en todos los otros cuerpos, porque los cuerpos están infinitamente divididos y forman un *plenum* (según lo mantiene Leibniz); como un resultado de esto, hay una correspondencia entre la serie de los sucesos en un alma y la serie de los sucesos en cualquier otra alma, *vía* sus cuerpos respectivos (p.ej., *NE* 55, 72, 116, 155, 177, 212, 242, 381). Ésta es la armonía preestablecida.

Leibniz sostiene que la naturaleza de una sustancia es la de *percibir* (p.ej., *NE* 378). Y un alma percibe sólo si se satisfacen estas condiciones: (a) el alma tiene estados con contenido representativo y (b) el contenido de estos estados es verídico, esto es, realmente acaecen las cosas y los sucesos representados en los estados. El alma causalmente aislada puede, a lo más, generar estados con contenido representativo; no puede generar los sucesos fuera de ella especificados por esos contenidos. Pero lo que Leibniz sostiene es que la condición de veracidad se satisface debido a la armonía preestablecida. Leibniz mantiene que, porque la misma vale, cada alma *percibe* (o expresa) sucesos en su propio cuerpo y en los cuerpos circundantes y *actúa* para producir cambios en su cuerpo conforme a sus apetitos o deseos.[5] Pero, a pesar del hecho

4 Especialmente en *NE*, Leibniz tiende a decir que cualquier cosa que pase naturalmente en un alma, le viene de sus propias profundidades, cualquier cosa podría suceder milagrosamente *NE* 66, 379.

5 En *NE* hay sólo unos cuantos enunciados explícitos en el sentido de que el alma percibe sucesos en su propio cuerpo (véanse *NE* 55 y 132). Con más fre-

de que una mente percibe sucesos corpóreos y responde en términos de su propio cuerpo, aquélla no tiene comercio causal con su cuerpo. Un alma actúa con la intención de mover su cuerpo y, como resultado de esto, las percepciones del alma representan apropiadamente los cambios corpóreos (conforme a leyes mecanicistas), pero el alma no afecta más que las representaciones que tiene. Eliminando la interacción mente-cuerpo, Leibniz mantuvo haber resuelto un problema central para comprender lo que sucede en la mente (p.ej. *NE* 220, 224, 380-1).[6]

Es verdad que uno de los argumentos de Locke a favor del agnosticismo explicativo fue que aun cuando observemos que los sucesos corpóreos causan percepciones sensibles en las mentes y que las voliciones en las mentes causan movimientos en los cuerpos, nosotros no podemos concebir cómo se realiza esto (p.ej., *Ensayo* IV, iii, 6). Leibniz respondió que la teoría de la armonía preestablecida hacía plenamente inteligible la conexión entre la mente y el cuerpo. Pero, incluso si esto fuese así, no se llegaría a la fuente real del agnosticismo explicativo de Locke.

---

cuencia dice que las percepciones del alma corresponden a o expresan sucesos en su cuerpo (p.ej. *NE* 54, 55, 116, 117). Pero la expresión es una forma de representación y creo que este uso restringido de 'percepción' es tan sólo una concesión al uso ordinario del término.

6 Evidentemente Leibniz concede que es insostenible la interacción mente-cuerpo (véase, p.ej. *NE* 55), aun cuando no es totalmente claro por qué lo hizo. La prohibición en contra de la interacción causal entre *sustancias* no se aplica a este caso, porque en los *NE* Leibniz niega rotundamente que las cosas materiales sean sustancias (p.ej. *NE* 211). Esto lo señaló Robert McRae en *Leibniz: Perception, Apperception & Thought* (Toronto University Press; Toronto, 1976), p. 47. En cartas a Damaris Masham que, sin duda, deseaba que se compartiesen con Locke, Leibniz subrayó las dificultades que presentaba la física mecanicista a la interacción mente-cuerpo. Si las almas actúan sobre los cuerpos, entonces "... se perturban las leyes corpóreas por la acción del alma". (*Die Philosophischen Schriften von Gottfried Wilhelm Leibniz*, ed. C.I. Gerhardt, v. III; Berlín, 1887; p. 341).

Pues a Locke lo impresionó la intuición cartesiana de que el fenómeno de la conciencia es *sui generis*, aun cuando no estuviese de acuerdo con Descartes en que este hecho acerca de la conciencia probaba que una mente es una sustancia cuya única esencia es pensar. Conforme a esto, Locke dijo, no sólo que no podemos concebir cómo es que los sucesos físicos producen ideas en nosotros sino también, y más importante, que no podemos concebir cómo es que *lo que sea* produce ideas o pensamientos (especialmente *Ensayo* IV, vi, 14). Para enfrentarse a esta posición, Leibniz necesita algo más que la doctrina de la armonía preestablecida. Él necesita poner de manifiesto aspectos no cartesianos de la mente en términos de los cuales el acaecer de estados conscientes pueda recibir una explicación de algún tipo. De hecho, él tiene este tipo de estructura. Leibniz definió el alma de manera diferente a la de Locke y, con ello, propuso un método distinto para investigar sus operaciones.

Locke definió un alma como una sustancia que realiza las operaciones conscientes de pensar, querer, saber y demás (p.ej. en *Ensayo* II, xxiii, 5). Así, cualquier hipótesis con respecto a las causas de las operaciones mentales se evaluará exclusivamente por su capacidad para explicar los fenómenos del pensamiento consciente. En contraste con esto, Leibniz se alía a un movimiento anticartesiano, "vitalista", en el que el alma se definió como la fuente de la actividad orientada hacia metas, que va desde las funciones básicas de sustento vital, a través de los hábitos inconscientes, hasta los actos autoconscientes deliberados.[7] Según esto, los informes de primera persona acerca de lo

7 Al comienzo de los *NE*, Leibniz se alía, de manera general, con las filosofías anticartesianas de "... quienes en todo ponen vida y percepción p.ej. de Cardano, Campanella y (mejor que ellos) de la difunta platonista, la Condesa de Conway y nuestro amigo el difunto Sr. Franciscus Mercurius van Helmont ... junto con su amigo el difunto Sr. Henry More". (*NE* 72)

que uno capta dentro de uno mismo, no son los únicos, quizás tampoco los mejores tipos de evidencia con respecto a lo que sucede dentro de un alma.[8] En principio, son pertinentes los datos biológicos, fisiológicos y conductuales y es apropiado formular hipótesis acerca de operaciones del alma, ocultas, no conscientes, a la par con los corpúsculos insensibles de la teoría mecanicista. Conforme a esto, Leibniz avizora una ciencia de las almas que él denomina 'pneumatología' (*NE* 56). Esta forma de concebir un alma (o mente) proporciona un enfoque a la explicación de la conciencia que no está disponible desde la perspectiva de Locke. Esto es, explicar las operaciones conscientes de un alma como funciones de una sustancia definida más ampliamente como la fuente de una gama de procesos y actividades dirigidos hacia una meta. Pasamos a considerar ahora la manera como Leibniz elaboró su enfoque alternativo con respecto a las operaciones del alma.

3. La teoría de Leibniz es que un alma es inmaterial y esencial (o naturalmente) activa (*NE* 65, 53). Considerada de manera abstracta, es una "fuerza actuante primitva" constante que toma la forma concreta de una multitud de fuerzas derivadas o de tareas específicas coexistentes y sucesivas (*NE* 216, 169). La naturaleza de un alma es la "percepción y sus consecuencias", esto es, la acción (*NE* 378, también 210). La propuesta de que las percepciones comprenden deseos y generan acciones, se indica en pasajes tales como los siguientes:[9] "Lo que usualmente nos mueve son esas percepciones diminutas que podrían denominarse sufrimientos que no llegamos a captar ..." (*NE* 188).

8 Para Locke, los informes de primera persona son nuestra única fuente de información acerca de la naturaleza y de las operaciones de la mente; véase, especialmente, *Ensayo* IV, iii, 17.

9 Jacques Jalabert, en "La psychologie de Leibniz: ses caractères principaux", *Revue philosophique de la France et de l'etranger*, 136 (1946), pp. 453-72, subraya tanto que Leibniz identifica percepción y apetitos como que les adscribe contenido proposicional a todas las operaciones del alma.

"Toda impresión tiene un efecto, pero no nos damos siempre cuenta de los efectos" (*NE* 115). "Creo que no hay percepciones que nos sean indiferentes por completo ..."(*NE* 162) y "... todo sentimiento es la percepción de una verdad ..." (*NE* 94).

Así pues, las percepciones en un alma desembocan naturalmente en intentos de cambio; los intentos desembocan en acciones a menos que se les impida (*NE* 173) y el alma actúa de una manera determinada por la multitud de intentos que están copresentes en ella (*NE* 193). El alma se impulsa a sí misma de un estado a otro percibiendo sucesos en su cuerpo y en sus alrededores, actuando en respuesta conforme a estos apetitos y percibiendo los sucesos corpóreos que se siguen.

Parte integral de esta concepción anticartesiana de las almas es que éstas están percibiendo y actuando constantemente, sea que sean o no conscientes de estar haciéndolo. Por virtud de la armonía preestablecida todas las impresiones corpóreas se perciben en el alma y "... algo sucede en el alma que corresponde a la circulación de la sangre y a todo movimiento interno de las vísceras ..." (*NE* 116). Ésta es la famosa doctrina de Leibniz de las percepciones diminutas. En los *NE*, él presenta una diversidad de argumentos en apoyo de su tesis de que el alma con frecuencia percibe cosas sin darse cuenta de ello (específicamente *NE* 112-18). Pero en el contexto de su disputa con Locke, hay algunas cuestiones técnicas con respecto a la noción de percepción no consciente.

Las cuestiones centrales se refieren tanto al contenido representativo de las percepciones diminutas como a su veracidad. ¿Qué es lo que determina el contenido de una percepción de la que no se da cuenta un alma?[10] Conforme a Leibniz, ¿qué es que

10 Esta pregunta se ha tratado con diversos grados de seriedad en la literatura; p.ej., compárense McRae, *op. cit.* pp.23-4 y Robert Brandom, 'Leibniz and Degrees of Perception', *Journal of the History of Philosophy* 19 (1981), esp. pp. 459 y ss.

realmente existan los cuerpos representados por las percepciones? Estas cuestiones adquieren un apremio especial en el contexto de la disputa de Leibniz con Locke,[11] pues Locke mantenía que el contenido de las ideas lo fijan dos consideraciones: lo que la mente que tiene la idea capta que es lo que ésta representa y, en el caso de las ideas sensoriales, las causas físicas de las ideas, independientes de la mente.[12] Es claro que Leibniz no dispone de ninguna de estas maneras de fijar los contenidos de las percepciones. Es tentador suponer que Leibniz piensa que el contenido de las percepciones de un alma está determinado por el papel que desempeña en conexión con los cambios adaptativos en su cuerpo; esto es, cuando el cuerpo de un alma está en ciertas circunstancias, el alma tiene percepciones correspondientes que generan otros estados en el alma que corresponden (por la armonía preestablecida) a cambios adaptativos en su cuerpo. Parece que Leibniz podría decir que el contenido de las percepciones está determinado por los estados corpóreos entre las que éstos median y el principio de los apetitos del alma.[13] Pero esta solución puede no estar disponible para Leibniz pues, como lo han argüido muchos comentadores, él parece haber vacilado sobre la cuestión del *status* metafísico de los cuerpos.[14] En ocasiones parece haber sostenido que los cuerpos no son sino objetos intencionales, "bien fun-

11 No pretendo sugerir que Leibniz o Locke se propuso explicar el contenido intensional de las percepciones o de las ideas en términos de condiciones no intensionales. A ellos, sin embargo, les preocupaba considerar qué tipos de factores determinan el contenido representacional de una idea o de una operación de una mente dada.

12 He analizado esto en "The Epistemological Status of Ideas: Locke compared to Arnauld", *History of Philosophy Quarterly* 9 (1992), pp. 409-24.

13 Algo como esto se sugiere como una explicación del contenido de las ideasen el artículo de Michael Losonsky, "Leibniz's Adamic Language of Thought", *Journal of the History of Philosophy* 30 (1992), esp. pp. 538-9.

14 Entre los comentadores recientes véanse Catherine Wilson, *Leibniz's Metaphysics* (Princeton University Press; Princeton 1989), esp. seccs. 17-9;

dados" en el acuerdo entre representaciones de todas las sustancias reales.[15] En tal caso, Leibniz no puede apelar a los cambios corpóreos para determinar los contenidos intencionales de las percepciones de las almas. En otras ocasiones, Leibniz sugirió que los cuerpos se identifican con agregados de sustancias inmateriales; en tal caso, los cuerpos no son extensos, no tienen posición espacial ni están en movimiento unos con respecto a otros como en la percepción se representa que están. Entonces, no es claro cómo los sucesos en los cuerpos, que realmente *no* son extensos, etc., determinan los contenidos de las percepciones en las que están representados como *siendo* extensos, etc. Rebasa las limitaciones de escrito, analizar estas cuestiones con respecto a la tesis de Leibniz en los *NE* acerca de la naturaleza de los cuerpos y el contenido de las percepciones diminutas. Estos asuntos son importantes para tener una explicación plena de lo que sucede dentro de las almas, pero aquí supondré que Leibniz puede dar una explicación satisfactoria de ellos.

5. Nos interesa la explicación de Leibniz de cómo las percepciones diminutas que corresponden con lo que sucede en el cuerpo dan origen a las percepciones sensibles de las que nos damos cuenta. Leibniz dice, "Nunca estamos sin percepciones pero, necesariamente, con frecuencia no tenemos conciencia ('*apperception*'), a saber, cuando ninguna de nuestras percepciones sobresale" (*NE* 162). (Remnant y Bennett traducen '*apperception*' ['apercepción'] como 'captación' ['*awareness*']; en este escrito uso indistintamente estas palabras.)

Glen A. Hartz estudia dos explicaciones de los cuerpos en "Leibniz's Phenomenalisms", *Philosophical Review* 101 (1992), pp. 511-50

15 Esta línea de pensamiento la identificó Montgomery Furth en "Monadología", reimpreso en *Leibniz: A Collection of Critical Essays*, Harry Frankfurth (comp.), Doubleday; Nueva York, 1972.

La apercepción requiere de percepciones que se distingan o que resalten con respecto a las demás percepciones presentes en el alma. Un conjunto de percepciones se distingue si sus miembros se refuerzan unos a los otros de tal manera que es mayor o más fuerte que otro conjunto de percepciones en el alma o si es novedoso con respecto a otras percepciones en el alma:

> ... siempre hay objetos que golpean nuestros ojos y oídos y, por tanto, que también tocan nuestras almas, sin que les prestemos atención. Pues nuestra atención la retienen otros objetos hasta que un objeto dado se hace lo suficientemente poderoso como para atraerla, sea por actuar con más fuerza sobre nosotros o de alguna otra manera. (*NE* 115; véanse, también, 53, 118, 134)

El que un alma tenga percepciones agudizadas es una condición necesaria, pero no suficiente, para que el alma aperciba (p.ej. *NE* 118). También el alma debe darse cuenta o prestarles atención a sus percepciones agudizadas:[16]

> ...La memoria se necesita para la atención; cuando, por así decir, no estamos alertas para prestar atención a algunas de nuestras propias percepciones, las dejamos pasar sin considerarlas e incluso sin darnos cuenta de ellas. Pero si alguien ... nos hace notar, por ejemplo, algún ruido que oímos en ese momento, entonces lo recordamos y somos conscientes de haber acabado de tener alguna sensación del mismo ... (*NE* 54; también 113, 115, 238)

Esto nos da el diseño básico de la explicación de Leibniz de la apercepción, esto es, que si el alma tiene un conjunto de

16 Esta propuesta la enfatizan tanto McRae, en *op. cit*, como Mark Kulstad en "Leibniz on Consciousness and Reflection", *Southern Journal of Philosophy*, vol. supl. 21 (1983). pp. 39-66.

percepciones diminutas que es suficientemente fuerte o novedoso como para ser notado y el alma recuerda inmediatamente ese conjunto de percepciones y le presta atención, entonces el alma capta ese conjunto de percepciones. ¿Cómo es que esto proporciona lo que Leibniz prometió, a saber, una *explicación inteligible* de cómo surge la apercepción de estados de los que uno no es consciente?

Al buscar la respuesta, hemos de tener en mente que Leibniz mantiene que cualquier cosa que suceda naturalmente en el alma es una modificación de su "naturaleza inherente constante" que puede entender una mente finita (*NE* 66). La naturaleza de un alma (o sustancia) es la "percepción y sus consecuencias" y, según he argüido, esto significa que la naturaleza del alma es la de representar los estados de su cuerpo (y de los cuerpos circundantes) y determinarse a sí misma a cambiar de maneras apropiadas al contenido de esas percepciones y de sus metas. ¿Cómo es que surge la captación en un alma con esta naturaleza? Un tipo de respuesta es que la apercepción tiene un papel funcional específico en las operaciones del alma.[17]

Si esto especifica correctamente el proyecto de Leibniz, él no necesita explicar la naturaleza de la apercepción. Ni necesita negar que ésta es una propiedad emergente que pertenece a algunos estados del alma, pero no a los estados que la causan. Pero es instructivo mirar rápidamente una versión en la que Leibniz proporciona algo más que una mera explicación funcional de la apercepción. Como veremos, no es una estrategia prometedora y no se adecúa a lo que Leibniz dice realmente.

17 Sobre la noción de explicación funcional véase el artículo "Functional Analysis", de Robert Cummings, reimpreso en Ned Block, *Readings in the Philosophy of Psychology*, vol. 1 (Harvard University Press; Cambridge, 1980). *No* pretendo decir que Leibniz sostiene lo que hoy se denomina una teoría "funcionalista" de la mente; véase la introducción de Block a "Functionalism", pp. 171 y ss.

(a) Algunos comentadores han pensado que la diferencia entre percepción y apercepción consiste en el hecho de que la apercepción es reflexiva, esto es, es un acto por el cual un agente se percibe a sí mismo.[18] Conforme a esta versión, las percepciones y las apercepciones tienen un contenido esencialmente diferente. Toda captación es, *inter alia*, captación de uno mismo, en tanto que el yo nunca está presente en el contenido de las percepciones que no capta el alma. Se ha alegado que esta diferencia sea constitutiva de lo que es la apercepción. Esto se podría aplicar al proyecto explicativo de Leibniz, pues si él puede explicar cómo un sujeto puede, él mismo, representarse a sí mismo, eso constituiría una explicación de la apercepción.

Pero, ¿sostiene Leibniz que la captación comprende esencialmente autoconciencia? Leibniz usa el término 'reflexión' para una *especie* de conocimiento de sí. Y la tesis de que el conocimiento reflexivo es constitutivo de lo que en francés se denominaba '*conscience*' ('*conscientia*' en latín) estaba bien establecida en el siglo diecisiete.[19] Leibniz la apoyó en algunas obras.[20] Pero esto no ayuda a responder nuestra pregunta sobre la apercepción, pues en los *NE*, Leibniz claramente distingue la reflexión, por una parte, y la apercepción, por la otra.[21] Leibniz dice que la reflexión es el poder de atender a uno mismo y a los

18 P.ej., Jalabert, pp. 471-2; McRae, p. 26. La propuesta la examina cuidadosamente Mark Kulstad en *loc. cit.*

19 Véase el artículo de Emily Michael y Fred S. Michael, "Early Modern Concepts of Mind: Reflecting Substance vs. Thinking Substance", *Journal of the History of Philosophy* v. 27 (1989), pp. 29-48. La palabra francesa '*conscience*' *significa* 'conocimiento de sí'. Así, aun cuando había una tradición que identificaba el conocimiento reflexivo y la *conscience* (o *conscientia*), no se sigue que la captación se identificase con el conocimiento reflexivo. Los filósofos dentro de esta tradición, tendían a mantener que los animales inferiores no tienen un conocimiento reflexivo pero sienten dolores y tienen percepción sensorial.

20 Véase Kulstad, *op. cit.*

estados propios (p.ej. *NE* 51). Este poder está ligado a la posesión de distintos conceptos del yo, de la sustancia, de la acción, etc.; a la habilidad de conocer las verdades necesarias, la identidad personal y pertenecer a la comunidad moral (p.ej. *NE* 236). De acuerdo con esto, las bestias no humanas *no* tienen el poder de la reflexión, pero sí tienen apercepción:

> ...las bestias no tienen entendimiento [poder de reflexionar y de derivar verdades necesarias], aun cuando tienen la facultad de la apercepción de las impresiones más conspicuas y distintivas –como cuando un jabalí apercibe a alguien que le está gritando y va directo hacia esa persona, habiendo tenido previamente sólo una mera percepción de ella, confusa, como sus percepciones de todos los objetos que están ante sus ojos y reflejan rayos de luz dentro de las lentes. (*NE* 173; véase también, 134)

Así, en la doctrina de los *NE*, la apercepción no comprende conocimiento *reflexivo* del yo. Pero, ¿comprende esencialmente la apercepción *algún* tipo de captación del yo? El pasaje que acabamos de citar dice que el jabalí capta al hombre y deja sin especificar si el jabalí se capta a sí mismo.

Acerca de esto, nada nos aclara la explicación básica de Leibniz, de que uno apercibe prestando atención a las propias percepciones, pues 'percepción' puede significar el objeto de percepción, así como un acto perceptual. Quizás podamos plausiblemente asumir que la apercepción presupone una captación mínima de los actos perceptuales previos de uno; al menos parece comprender captación de la modalidad sensorial de los actos. Por ejemplo, Leibniz dice que captamos un *ruido* (no meramente captamos, por ejemplo, un

21 En *NE*, el uso que Leibniz hace de '*conscience*' se confina estrictamente al asunto de la captación de sí para determinar la identidad personal (véase *NE* 238).

mosquito cerca de la oreja) y es plausible suponer que el jabalí capta las sensaciones visuales y auditivas cuando capta al hombre gritando. Pero ésta es una conjetura. El texto no da ninguna indicación clara de cuál es la tesis de Leibniz. Esto pesa mucho en contra de la sugerencia de que Leibniz pretendía analizar la apercepción en términos de autorrepresentación y de que esto subyace a su explicación de la apercepción.

(b) Veamos la sugerencia de que él explica la apercepción en términos de su papel funcional dentro del alma. Cuando el alma capta un conjunto de percepciones realzadas, al menos parte de lo que capta son los contenidos de esas percepciones. Por ejemplo, cuando el jabalí presta atención a (un subconjunto de) las percepciones diminutas del arreglo de objetos que reflejan la luz a sus ojos, el jabalí capta a un hombre que es uno de esos objetos. En general, porque una apercepción representa los contenidos de un conjunto de percepciones diminutas, habrá una relación sistemática entre el contenido apercibido y los contenidos de las percepciones diminutas en ese conjunto. Esto es algo que Robert Brandom ha desarrollado en otro contexto.[22] Mi sugerencia es que ésa es la base del papel funcional distintivo que desempeña la captación sensorial en las operaciones del alma. Aproximadamente, su papel es representar la totalidad o parte de lo que hace un conjunto de percepciones diminutas y representarlo de manera que tienda a causar acciones más adecuadas.

¿Cómo *están* relacionados los contenidos de las apercepciones a los contenidos de las percepciones diminutas que representan? Una apercepción combina en una representación contenidos previamente distribuidos entre muchas percepciones. Para usar uno de los conocidos ejemplos de Leibniz, cuando una persona capta el rugido del mar, ella capta una totalidad compuesta de ruidos hechos por olas singulares chocando contra la costa; si las per-

22Brandom, *op. cit.*

cepciones diminutas pasasen inadvertidas, no habría tenido sino percepciones de olas singulares. Es como la diferencia entre percibir los árboles individuales en el bosque y percibir el bosque.

Las apercepciones del tipo que Leibniz denomina 'percepciones confusas' tienen contenidos, con una relación sistemática adicional, con los contenidos de las percepciones diminutas. Los paradigmas de las percepciones confusas son imágenes sensoriales de colores, sabores, olores, calor, placeres y dolores.[23] En ocasiones se nos dice que una percepción de verde contiene o está compuesta de percepciones de amarillo y azul (*NE* 56, 134, 219, 255, 403). Pienso que esto significa que la imagen del verde es la captación de un conjunto de percepciones diminutas de amarillo y azul. (Como una explicación de la captación del verde, esto genera algunas preguntas que aquí no consideraré.[24]) El conjunto contiene percepciones de amarillo y azul como miembros y el contenido representativo de la imagen del verde se deriva de esas percepciones. La sugerencia es que la imagen representa una masa de partículas amarillas y azules sin representar colores, figuras o tamaños individuales. Leibniz dice, "... no podemos ordenar todo lo que tiene una parte en nuestras percepciones confusas; ellas son expresiones de los detalles de lo que sucede en los cuerpos e incluso tienen en ellas algo de infinito" (véase *NE* 381 y también el famoso pasaje de la rueda dentada giratoria, *NE* 403).

El siguiente pasaje da alguna indicación de cómo las imágenes que resumen, al omitir detalles, aumentan la capacidad del alma para la acción apropiada:

23 Pero también hay imágenes confusas de figura (*NE* 120).

24 Algunas de estas cuestiones son, i) si la imagen de verde diminuto representa (se deriva de) percepciones de amarillo y azul, ¿qué es lo que representan las imágenes de amarillo y de azul? ii) ¿Cómo es de entender lo que es una percepción diminuta de amarillo o de azul (dada la tendencia a pensar que las percepciones de amarillo o azul comprenden imágenes sensoriales?).

> Con respecto a las ideas o, más bien, imágenes confusas ... tales como colores, sabores y demás, que son el resultado de diversas ideas diminutas que son distintas en sí mismas aun cuando no las captemos de manera distinta, ... el papel de estas impresiones es el de proporcionarnos inclinaciones naturales y aportarnos un fundamento para observaciones de la experiencia. (*NE* 487)

La falta de resolución en las imágenes que captamos es ventajosa porque, "en ocasiones, nos confunde la multiplicidad de las cosas que hay que tomar en cuenta". Más aún, la captación distinta de imágenes como totalidades nos da la oportunidad de pensar en ellas en términos de diversos conceptos, de leyes generales y de principios de explicación.

De las sensaciones confusas de incomodidad o de dolor también se dice que son útiles:

> ...si se amplificase lo suficiente lo que sucede en nosotros cuando tenemos apetito o deseo, causaría sufrimiento. Esto es por lo que el infinitamente sabio Autor de nuestro ser actuaba en nuestro interés cuando hizo que, con frecuencia, seamos ignorantes y estemos sujetos a percepciones confusas –de tal manera que pudiésemos actuar más rápidamente por instinto y que no nos preocupasen sensaciones excesivamente distintas de múltiples objetos que ... no nos son del todo gratos. (*NE* 166)

Estos pasajes muestran que Leibniz piensa que las imágenes sensoriales confusas son especialmente útiles a criaturas con una capacidad limitada para lidiar con una multitud de detalles (en ocasiones desagradables). Lo que he estado sugiriendo es que ésta es la base de la prometida explicación de Leibniz de la apercepción sensorial. Resumo la explicación que propongo: la apercepción sensorial es un artificio mediante el cual se combinan y (en ocasiones) se resumen en una percepción los contenidos

de muchas percepciones separadas; que la percepción hace una aportación distintiva a las operaciones del alma, porque tiende a causar acciones más apropiadas que las que causaría la masa de percepciones que aquélla representa.

Si ésta es la versión correcta de la estrategia de Leibniz, entonces su explicación no responde directamente la exigencia de los filósofos actuales quienes sostienen que la conciencia es esencialmente subjetiva y, por esa razón, reacia a que se la explique en términos objetivos. Ni dará respuesta a quienes sostienen que los estados mentales no son plenamente objetivos en base a que hay tipos de *qualia* (p.ej. dolor o la imagen de verde) que no pueden *definirse* en términos de un papel funcional causal.[25] Leibniz mismo propuso que las imágenes de color habrían de explicarse en términos de imágenes de color menores y que no podemos esperar tener un conocimiento completo de ellas (*NE* 120, 376). Pero Leibniz no estaba comprometido con explicar las operaciones de la mente en términos objetivos, no intencionales y ciertamente no en términos físicos. Después de todo, el blanco de Leibniz era Locke y su agnosticismo explicativo. La ventaja de Leibniz estaba en promulgar una noción del alma que permitiese una explicación funcional o teleológica de sus operaciones, de la que no se disponía dentro de la estructura de la noción cartesiana de Locke de la mente. De esta manera, Leibniz montó un importante desafío en contra de dos supuestos centrales de la teoría lockeana de la mente y de la filosofía general, esto es, que todos los estados cognitivos son conscientes y que la introspección es la autoridad final acerca de los contenidos de los estados conscientes (ideas).

Trad. de José A. Robles

25. P.ej., Ned Block, "Troubles with Functionalism" y Sidney Shoemaker, "Functionalism and Qualia", en Block, *Readings*.

¿FUE LEIBNIZ UN PARALELISTA SICOFÍSICO?

Alejandro Herrera Ibáñez*

Cuando se habla del problema mente-cuerpo, es común catalogar a Leibniz como defensor del paralelismo sicofísico (*v gr.*: Mates, Bunge), es decir, de la teoría que sostiene que no hay interacción entre mente y cuerpo, o entre fenómenos mentales y fenómenos corporales, sino que éstos se dan paralelamente. En Leibniz esta teoría toma forma en su famosa doctrina de la armonía preestablecida.

En este trabajo me propongo examinar, por un lado, en qué sentido se puede decir que Leibniz fue paralelista. Por otro lado, propondré una lectura de Leibniz que nos lo revela como un monista y no como un dualista. Ello me llevará a preguntarme si es quizá más correcto ver a Leibniz como un monista neutral, o sea, como defensor de la teoría según la cual tanto lo mental como lo físico son producto de una entidad neutra que no es ni mental ni física. En el curso de esta investigación también me veré llevado a preguntarme si es posible atribuir a Leibniz un tipo de epifenomenismo, específicamente del tipo que llamaré "epifenomenismo invertido".

Mi estrategia será la siguiente. En la primera parte haré una somera exposición de la ontología dinamista de Leibniz, para de esta manera mostrar su visión monista del universo. En la segunda parte mostraré cómo plantea Leibniz el paralelismo que normalmente se le atribuye. Finalmente, examinaré las tensiones que se dan entre el paralelismo y el epifenomenismo de Leibniz y su esquema monista del universo.

* Instituto de Investigaciones Filosóficas, UNAM.

I

El concepto de fuerza en la filosofía de Leibniz es de capital importancia. En un trabajo anterior mostré la importancia de este concepto para entender la noción leibniziana de materia. A Leibniz no le pareció suficiente la caracterización cartesiana de la sustancia física en términos de extensión, y las noción newtoniana de inercia le pareció muy pasiva. Ello lo llevó a desarrollar una dinámica que –como la etimología del nombre lo sugiere– explicara el movimiento en términos de la noción de fuerza. No me detendré aquí a examinar en detalle los distintos tipos de fuerza, o las distintas manifestaciones de ésta, que aparecen en la dinámica de Leibniz. Baste aquí decir que Leibniz adoptó una noción dinámica de sustancia, definiéndola como "un ser capaz de acción" (*GP* VI, 598/*Q* 597). En esta sucinta fórmula se encierra perfectamente la visión dinámica de Leibniz. Como veremos, la capacidad de actuar ya es en sí un elemento dinámico. No se trata, por tanto, de una potencia pura.

Lo que hay que recalcar es que la definición que da Leibniz de sustancia se aplica no sólo a los objetos físicos, sino también, y sobre todo, a los objetos espirituales, es decir, a las mónadas de su sistema, en el cual ellas son las únicas y auténticas sustancias. Por tanto, a ellas, más que a nada, se aplica la definición dinámica que Leibniz proporciona. Hay, por consiguiente, fuerza en las mónadas, y hay fuerza en los cuerpos y, más aún, es posible decir que hay fuerza antes de la existencia de las mónadas creadas. A la noción de fuerza está íntimamente ligada la de conato, como veremos más adelante. El conato es una tendencia, un impulso presente en todo ente, aun si tal ente es sólo posible.

Tenemos, pues, que la ontología dinamista de Leibniz está permeada por la noción de fuerza en tres niveles importantes. En el primer nivel, el del mundo de las posibilidades, sólo unos cuantos posibles o esencias son actualizados por Dios, mientras

que otros se quedan sin pasar al mundo de lo existente, es decir, al mejor de los mundos posibles. Sin embargo, el proceso de selección que hace Dios requiere de toda su infinita sabiduría, porque para Leibniz todos los posibles tienden a, o pugnan por, existir. En uno de sus escritos sin título, conocido como "Todo posible exige existir", Leibniz asienta que "si en la naturaleza misma de la existencia no hubiera cierta *inclinación* a existir, nada existiría" (*GP* VII, 194/*O* 152. Las cursivas son mías). En otros lugares Leibniz utiliza, aún más claramente, la noción de *tendencia* a existir (*GP* VII, 303/*O* 473) y la de *conato* de existir (*C* 534/*O* 501),[1] y el conato, como veremos, es una tendencia. Así pues, la posesión de esta tendencia, inclinación o conato, o sea la posesión de fuerza, es una condición necesaria aunque no suficiente, de todos los posibles, para existir.

En el segundo nivel, ya instalados en el mundo de los existentes, las mónadas, únicas y verdaderas sustancias, son por naturaleza activas. Nunca son real y totalmente pasivas, aunque a veces Leibniz divide las fuerzas –especialmente en la dinámica– en activas y pasivas. Esta fuerza se manifiesta en las mónadas de dos maneras: como apetito (fuerza activa) y como percepción (fuerza pasiva). El apetito es la tendencia a cambiar de percepciones, y la percepción es la fuerza pasiva de las mónadas, análoga a la fuerza pasiva de la resistencia de los cuerpos. Si a las mónadas les quitamos estas fuerzas, las privamos de su naturaleza.

Al tercer nivel pertenecen los objetos materiales. Ahora bien, la materia es siempre compuesta. Los cuerpos no pueden, por tanto, ser auténticas sustancias en la filosofía de Leibniz, puesto que las auténticas sustancias para él deben ser simples. De aquí que –como lo mostré en otro trabajo– para Leibniz la materia es sólo un fenómeno, aunque bien fundado. La materia es un

1 El traductor interpretó "*conatus*" como "tendencia".

producto fenoménico de las mónadas, pero no es una ilusión, como tampoco lo es, nos dirá, el arcoiris. Este mundo fenoménico, objeto de estudios de la ciencia de la dinámica, también está gobernado por fuerzas, y sin ellas no se puede entender. Pero si la materia es un fenómeno bien fundado, y la materia está gobernada por fuerzas, ¿cuál es la naturaleza de éstas? Desde luego que no puede ser más que fenoménica. Leibniz divide, en el *Bosquejo dinámico*, las fuerzas en primitivas y derivadas. Estas últimas son el objeto de la dinámica. Se sigue que las fuerzas derivadas son fenoménicas.

Hemos visto cómo la noción de fuerza opera en la filosofía de Leibniz tanto en el mundo de los posibles como en el de los existentes, y en este último tanto entre las mónadas o auténticas sustancias como entre los cuerpos o productos fenoménicos de las mónadas. Veamos ahora en más detalle cómo es posible sostener el carácter univerasal de la noción de fuerza en la filosofía leibniziana.

En "Un nuevo método para aprender y enseñar jurisprudencia" (*PA* VI, i, 266-89/*L* 85-92) Leibniz caracteriza la *actividad o fuerza* como "el estado de una cosa, del que se sigue el cambio". Esta definición es perfecta para cubrir tanto el caso de los cuerpos –pues es por la fuerza que se explica el movimiento– como el caso de las mónadas, ya que el apetito que éstas poseen es concebido como aquello que hace posible el cambio de una percepción a otra. Leibniz agrega que la fuerza es doble, a saber, la de actuar y la de resistir. De la primera dice que es inmaterial, y de la segunda que es material. Las fuerzas inmateriales son las inteligencias separadas (¿Dios y los ángeles?) y las almas primarias o entelequias que están en los cuerpos (*i. e.*, las mónadas). Para Leibniz la forma más simple de la fuerza es el conato de acción. Este puede darse en entes simples y en entes compuestos. Los primeros son entes percipientes y su conato en general se llama apetito (y cuando éstos son además pensantes, su conato

se llama voluntad). Los segundos son cuerpos y su conato es la fuerza motriz. Este esquema se enriquece y se complica en el *Bosquejo dinámico*, pero aquí es suficiente para mostrar, por un lado, que la noción de fuerza, ligada a la de actividad, es utilizada para dar cuenta tanto del mundo monádico como del mundo corpóreo y, por otro lado, para mostrar que el conato es concebido como la fuerza en su forma más simple.

Otra definición iluminadora se encuentra en una carta dirigida por Leibniz a Eckhard. En ella le dice que "la *fuerza* es grado de acción" (*GP* I, 266/*L* 177). Esta definición es sumamente importante, pues nos permite comprender, por un lado, que la fuerza "pasiva" de la resistencia tiene algún grado de acción y, por otro, que el conato es también activo en algún grado. Luego, todas las fuerzas son activas (a este respecto, véase *GP* IV, 395/*O* 437. También dice ahí que la fuerza "envuelve un conato o tendencia a la acción de modo tal que la acción se sigue si algo no lo impide").

Además, en un escrito en que examina la física de Descartes, Leibniz caracteriza la materia como resistencia (*i.e.*, fuerza) difundida ("una difusión de la resistencia constituye la materia". *GP* IV, 394/*O* 436). Esto, en el plano de la simultaneidad. En el plano de la sucesión, "la acción necesita el transcurso del tiempo" (*GP* IV, 396/ *O* 437). En una carta a De Volder explicita que la acción es fuerza conducida en el tiempo (*GP* II, 270/*L* 537). De manera que cuando la fuerza se extiende espacialmente, tenemos la materia; y cuando se extiende temporalmente, tenemos la acción. Luego, todo el mundo fenoménico puede entenderse en términos de fuerza extendida.

Es importante señalar que el conato es una fuerza, pues éste aparece en todos los niveles de la realidad: el conato de existencia es tendencia, pugna, inclinación, exigencia de existir; el conato de las mónadas o percipientes es apetito; el conato de los cuerpos es "fuerza muerta" o motriz. Está presente en todos los objetos

y se manifiesta cuando los obstáculos (fuerzas activas de otros objetos) cesan. Lo esencial del conato es la tendencia. Leibniz lo define como "fuerza que tiende al futuro" (*GP* IV, 399/*O* 443). No es posible exagerar la importancia de esta noción, pues en un universo cuyas partes careciesen de conato, es decir, en un universo absolutamente en reposo, éstas serían intercambiables entre sí sin que hubiese variación perceptible, y si esto es así, entonces "el estado de las cosas en un instante no puede distinguirse de un estado en cualquier otro instante" (*ibid*).

Esto es perfectamente claro para las mónadas, pues si no tuviesen un apetito y una percepción propios no podrían diferenciarse las unas de las otras. No hay, por tanto, que concluir que el conato es el grado cero de acción. Es un grado sumamente pequeño, más pequeño "que cualquier magnitud dada" (*GP* IV, 229/*L* 140). "El conato es el inicio del movimiento" (*GP* VII, 573/*L* 107). El par conato-movimiento es explicado por Leibniz mediante la comparación con otros pares de nociones análogas. En un sitio lo compara con los pares de nociones punto-espacio y uno-infinito (*GP* IV, 229/*L* 140), y en otro compara el par conato-ímpetu con los pares fuerza muerta-fuerza viva, punto-línea y línea-plano (*GM* VI, 117-9/*L* 299).

La capacidad de acción es –veíamos– lo que define a la sustancia. Esta capacidad no es otra cosa que fuerza, pues para Leibniz "la esencia de las sustancias consiste en la *fuerza primitiva* de la acción" (*L* 155. Las cursivas son mías).[2] En su primera formulación de un nuevo concepto de sustancia –en marzo de 1694– dice: "La noción de *fuerzas* o virtud (que los alemanes llaman *Kraft* y los franceses *la force*) ... arroja muchísima luz para entender la verdadera *noción de sustancia*" (*GP* IV, 469/*O* 457). Para dejar bien claro lo que propone, Leibniz

2 Loemker toma este texto de F. Rabbe, *L'Abbé Simon Foucher*, París, 1867, Apéndice, p. xlii (*L* 156, n. 6)

muestra en seguida la diferencia entre su noción y la noción escolástica de potencia:

> La fuerza activa difiere de la potencia desnuda ... en que la potencia activa, o facultad, de los escolásticos, no es sino una posibilidad próxima de actuar, la cual, sin embargo, para pasar al acto necesita de una excitación ajena o como de un estímulo. Pero la fuerza activa contiene cierto acto o *entelécheia*, y se encuentra entre la facultad de actuar y la acción misma, e involucra un conato. Así, se pone en operación por sí misma y no necesita de auxilios, excepto la sola eliminación del obstáculo (*ibid.*)

En este texto proporciona dos ejemplos: el de un cuerpo pesado que pende de una cuerda, y el de un arco tenso, y concluye que "esta virtud de actuar es inherente a toda sustancia, y de ella siempre nace alguna acción; de modo que la propia sustancia corpórea (lo mismo que la espiritual) jamás cesa de actuar" (*ibid.*). Luego, no hay ni cuerpos ni mónadas perfectamente quietos. Leibniz insiste en que los cuerpos nunca están en absoluto reposo para señalar que, aunque la materia es por sí misma inerte, está sin embargo siempre en acción en virtud de una "*entelequia primera*..., de cierto *próton dektikón* [capacidad primera] de actividad, a saber, la fuerza motriz primitiva sobreañadida a la extensión... y a la masa, que actúa siempre" (*GP* IV, 511/*O* 493). Esta fuerza –habíamos sugerido– juega un papel importante también en el mundo de los posibles.

Si bien la noción de fuerza tiene una gran importancia en la teoría física de Leibniz, expresada en su dinámica, no debe pensarse que su alcance es más limitado de lo que aquí hemos señalado. Para Leibniz la noción de fuerza es totalmente filosófica. En el *Nuevo sistema* dice: "Al tratar... de profundizar en los principios mismos de la mecánica... advertí... que era preciso emplear... la noción de *fuerza*, que es muy inteligible, aunque

pertenezca al dominio de la metafísica" (*GP* IV, 478/*O* 460). Y en otro sitio recalca que ya ha dicho muchas veces "que el origen del propio mecanismo no fluye del mero principio material y de razones matemáticas, sino, por así decirlo, de una fuente metafísica más alta" (*GP* IV, 505/*O* 486). En una afirmación todavía más atrevida, asevera que "los principios generales de la naturaleza corpórea y de la mecánica misma son más bien metafísicos que geométricos" (*Discurso de metafísica*, par. 18. *GP* IV, 444/*O* 302). En todo caso, para Leibniz la dinámica se subordina a la metafísica (*GP* IV, 398/*O* 440; *GP* VI, 588/*L* 624). No debe olvidarse que Leibniz llega a la noción de fuerza y la desarrolla a partir de su rehabilitación de las formas sustanciales escolásticas o de las entelequias primeras aristotélicas para encontrar los principios de la verdadera unidad de los objetos, principios que para Leibniz no pueden residir en la mera pasividad (*GP* IV, 479/*O* 461).

Con lo anterior espero haber mostrado que para Leibniz el mundo está gobernado por fuerzas. Más aun, es tentador pensar que lo único que para él hay en el mundo son fuerzas. En una carta a De Volder señala que "nada hay en las cosas sino sustancias simples y, en éstas, percepción y apetito. La materia y el movimiento, sin embargo, no son tanto sustancias o cosas como fenómenos de los percipientes" (*GP* II, 270/*L* 537). Más claramente, dice en otro sitio que "en cuanto al *movimiento*, lo que hay de real en él es la *fuerza* o poder [puissance] ... El resto son sólo fenómenos y relaciones [rapports]" (*GP* IV, 523/*L* 496).

II

Había dicho que para Leibniz el conato es el inicio del movimiento. Puesto que hay conato en todos los entes, hay una tendencia universal al movimiento. Ello crea un conflicto de fuerzas entre

objetos. Unos ejercen presión sobre otros. Esta presión es definida por Leibniz como "conato de penetración" (*GP* VII, 573/*L* 107), y para él la penetración es unión (*GP* IV, 230/*L* 141). Pero puesto que los cuerpos poseen también la fuerza de la resistencia o impenetrabilidad, esto impide su unión.

Leibniz se enfrenta a dos problemas. Por un lado, tiene que dar cuenta de la interacción de sustancias del mismo tipo, es decir, de la interacción de las mónadas y de la interacción de los cuerpos. La segunda es explicada mediante la dinámica. Por otro lado, tiene que dar cuenta de la interacción de "sustancias" de tipo diferente, es decir, tiene que resolver –dentro de su sistema– el problema de la relación entre cuerpo y alma. Puse "sustancias" entre comillas, porque –como ya vimos– los cuerpos no son realmente sustancias, por más que Leibniz a menudo habla de ellos como si lo fueran, indudablemente para poder dialogar con otros. De esta manera, Leibniz aborda el problema mente-cuerpo refiriéndose a las soluciones de su época (*GP* IV, 477-87/*O* 459-71; *GP* IV, 498-500/*L* 459-60; *GP* IV, 517-24/*L* 492-7). Este problema se plantea para él como el de la unión entre cuerpo y alma.

Para Leibniz, hay en su época dos soluciones insatisfactorias. La primera es la solución escolástica, a la que llama "vía de la influencia" o "sistema de la mutua influencia". La segunda es la de los cartesianos, y la llama "vía de la asistencia" o "sistema del supervisor perpetuo". En referencia más directa a Malebranche, también la llama "sistema de las causas ocasionales".

La primera, que es propiamente la solución interaccionista, es rápidamente rechazada por Leibniz, pues le parece obvio –y en esto dice estar de acuerdo con los cartesianos– que, dada la naturaleza material del cuerpo y la naturaleza inmaterial del alma, no es posible explicar cómo uno puede influir sobre el otro. La segunda solución ya no es interaccionista, sino paralelista. A Leibniz, sin embargo, le parece inútil, y hasta ofensivo para

la sabiduría divina, el requerir de la continua asistencia o supervisión de Dios para mantener siempre sincronizados el reloj del alma y el reloj del cuerpo. Ello equivaldría, piensa, a postular un *deus ex machina* en cada acción, lo cual llevaría a sostener la existencia continua de milagros. Para Leibniz es suficiente el milagro de la creación y de la fijación, de una vez para siempre, de sus leyes. La solución leibniziana es, pues, muy similar a la malebranchiana, con la diferencia de que, al no requerir de la acción continua de Dios, es más económica. A ésta la llama Leibniz la "vía de la armonía preestablecida" o "sistema del acuerdo natural". También se refiere a ella como la "hipótesis de los acuerdos". Esta famosa teoría paralelista de Leibniz también le sirve para resolver el problema que él llama de la "comunicación entre las sustancias", es decir, entre las mónadas, que, como se recordará, están encerradas en sí mismas, "no tienen ventanas". La causalidad es, pues, sólo una apariencia o fenómeno. Otra manera en que Leibniz plantea la pertinencia de su solución es señalando que el cuerpo pertenece al ámbito de las causas eficientes en tanto que el alma pertenece al de las causas finales, causas que no pueden entrar en contacto mutuo.

Con la teoría de la armonía preestablecida y con su teoría de la fuerza, Leibniz creyó haber dado un golpe maestro con el que satisfacía las inquietudes de diversas escuelas a lo largo de la historia de la filosofía. En una respuesta a Pierre Bayle concluía:

> La consideración de este sistema hace ver... que cuando uno penetra al fondo de las cosas, observa más razón de la que creía en la mayor parte de las sectas de filósofos. La falta de realidad sustancial de las cosas sensibles de los escépticos; la reducción de todo a las armonías o números, ideas y percepciones de los pitagóricos y los platónicos; lo uno y también el todo de Parménides y de Plotino, sin ningún spi-

nozismo; la conexión estoica, compatible con la espontaneidad de otros; la filosofía vital de los cabalistas y herméticos, que introducen el sentimiento en todo; las formas y entelequias de Aristóteles y los escolásticos; y hasta la explicación mecánica de todos los fenómenos particulares según Demócrito y los modernos, etc., se encuentran reunidas como en un centro de perspectiva (*GP* IV, 523-4/*L* 496).

Hay, sin embargo, elementos de fuerte tensión en la teoría leibniziana. Por un lado, he mostrado que hay un monismo en su visión del mundo, basado en la fructífera noción de fuerza. Por otro lado, la solución dualista paralelista de la teoría de la armonía preestablecida parece ir en contra de la visión monista sugerida. Y, finalmente, la concepción de la materia como algo meramente fenoménico no parece ser compatible ni con la solución dualista ni con la tesis monista. ¿Es posible resolver esta triple tensión? Veamos.

III

Preguntémonos primero si Leibniz es realmente dualista. Indudablemente lo es, pero su dualismo es distinto del cartesiano. Mientras el dualismo de Descartes y Malebranche es sustancial, y el paralelismo de este último es del tipo sustancia-sustancia, el parelelismo de Leibniz es del tipo sustancia-fenómeno, es decir, en el sistema leibniziano hay dos series de acontecimientos paralelos: una es la serie de sucesos sustanciales o monádicos, que consisten en una serie de apeticiones y percepciones, y otra es la serie de fenómenos materiales, fundados en combinaciones monádicas. El paralelismo leibniziano no es, por tanto, tan fuerte como el cartesiano. En este último la separación es tajante, pues la naturaleza del cuerpo es totalmente distinta de la del alma. Para Leibniz, en

cambio, es posible explicar en términos de una misma noción los sucesos monádicos (sustanciales) y los sucesos corporales (fenoménicos). Esta noción es la de fuerza. Luego, el dualismo leibniziano del tipo sustancia-fenómeno es compatible con su monadismo dinamista. Las mónadas son fuerzas primitivas y los fenómenos materiales son fuerzas derivadas o fenoménicas, estudiadas por la dinámica.

En segundo lugar, tenemos la tesis leibniziana de que la materia es un fenómeno producido por las mónadas. De una combinación de sustancias simples, que son las mónadas, se obtiene un conjunto de fenómenos de naturaleza compuesta, que son los cuerpos. A diferencia del epifenomenismo típico, en el que se hace de los sucesos mentales un añadido fenoménico de los sucesos materiales, en Leibniz hay una posición original que puede calificarse como "epifenomenismo invertido". Dicha posición se encuentra muy cerca de una posición emergentista, pero si se requiere de esta última que la nueva entidad sea una sustancia o una propiedad, la teoría leibniziana no es emergentista, pues los cuerpos no son ni sustancias surgidas de las mónadas (éstas son las únicas sustancias que hay), ni propiedades de éstas. Las únicas propiedades esenciales de las mónadas son el ser simples y el tener apetito o percepción, o sea, fuerza. De estas propiedades se derivan otras, como la de expresar cada una el universo entero desde su perspectiva, pero nunca la de ser materiales. Ahora bien, ¿cómo entra este epifenomenismo invertido en el esquema del monismo dinamista? La respuesta es sencilla. Hay fuerzas sustanciales (llamadas fuerzas primitivas), y las fuerzas físicas (llamadas fuerzas derivadas) son epifenómenos. Recordemos que lo fenoménico no es para Leibniz ilusorio. La materia es un fenómeno real, bien fundado. No es una alucinación. Podríamos decir que para Leibniz la esencia de lo mental radica en sus impulsos y percepciones –entendidas éstas

como actos– y que lo material es una serie de perceptos –entendidos éstos como contenidos de las percepciones.[3]

Este tipo de epifenomenismo genera, sin embargo, una grave dificultad en el sistema de Leibniz. Decir que la materia es un producto de las mónadas no es sino decir que éstas son causa de la materia. Pero Leibniz no puede aceptar esta afirmación, pues las mónadas son de una naturaleza radicalmente distinta a la de los cuerpos. Al menos, ése es el supuesto que lo lleva a adherirse a la solución paralelista. Las causas son, por tanto, sólo aparentes o fenoménicas. Pero si esto es así, se viene abajo la interpretación monista que sugiero. Ésta podría defenderse diciendo que puesto que los cuerpos también son fuerzas, no hay entonces un abismo que haga imposible la causalidad entre las mónadas y ellos. Pero si aceptamos esta tentadora interpretación, la solución paralelista pierde su razón de ser. Si queremos mantener que en Leibniz hay un paralelismo y una ontología de fuerzas, tendríamos que sostener también que esta ontología no es monista, pues el paralelismo sostendría que las fuerzas monádicas y las fuerzas corpóreas son de dos tipos radicalmente distintos e incomunicables. Pero cuando Leibniz sostiene que su dinámica se funda en su metafísica y a unas fuerzas las llama *derivadas* de las otras, indudablemente está pensando en un esquema monista en el que hay un solo tipo de fuerza que se manifiesta de manera distinta debido al ámbito en que opera. El monismo que Leibniz parece tener en mente no podría ser desarrollado por él sin resolver tensiones como ésta.

3 Si esta posición le parece a alguien anacrónica, lea esta reflexión de un connotado físico de nuestro tiempo: "¿Qué es la materia? ¿Cuál es nuestro esquema *mental* de la *materia*? La primera pregunta es ridícula. (¿Cómo vamos a decir *qué es* la materia –o, por precisar, *qué es la electricidad*– si se trata de fenómenos observables una sola vez?) La segunda trasluce ya un cambio radical de actitud: la materia es una imagen de nuestra mente –por lo tanto la mente es anterior a la materia" (Schrödinger, 21-2).

El monismo subyacente en la filosofía de Leibniz difiere del monismo spinozista en que mientras en éste hay una sola sustancia y lo demás son modos de ésta, en el leibniziano hay cabida para una pluralidad de sustancias, las cuales se diferencian entre sí por los grados de claridad de sus percepciones, pero entendidas todas como fuerzas. Aquí cabe preguntarse si sería correcto caracterizar el monismo leibniziano como un monismo neutral. Si éste se entiende como la tesis de que tanto los sucesos mentales como los materiales están constituidos por un componente neutro que no es ni mental ni material, no es fácil responder. ¿Hasta qué punto las percepciones de las mónadas no dominantes son mentales? La filosofía leibniziana anda rozando el pansiquismo y el panvitalismo sin trazar una línea divisoria clara. Pero hay mónadas –las dominantes– con actividades típicamente mentales: como sensación, memoria, conciencia, autoconciencia y raciocinio. Éstas no son producto de fuerzas neutrales, sino que son las fuerzas primitivas llamadas por Leibniz almas, en unos casos, y espíritus, en otros. Además, los fenómenos corporales no son producto directo de fuerzas neutrales. Las mónadas son fuerzas, y de la organización de éstas provienen las otras fuerzas, las materiales. Por momentos Leibniz parece estar cerca de un monismo neutral, pero no lo formula explícitamente.

He sugerido en mi interpretación que las mónadas son fuerzas y que, por tanto, el mundo leibniziano es un mundo de fuerzas. Podría objetarse que si bien el tener fuerza es una propiedad esencial de las mónadas, esto no es lo mismo que afirmar que las fuerzas tienen el estatus ontológico de sustancias. La objeción es fuerte, porque para Leibniz el apetito y la percepción son propiedades, no sustancias. Las sustancias son las mónadas, y éstas no son apetitos ni percepciones. ¿Qué son entonces las mónadas? Quitémosles el apetito y la percepción y es como si no nos quedara nada, pues perdemos la posibilidad de diferenciarlas. ¿Nos queda acaso su simpleza? Pero el ser simples con-

siste en no tener extensión. No pueden ser potencia pura, pues Leibniz rechaza tajantemente esta noción. Tampoco pueden ser materia prima en el sentido aristotélico, pues ésta es indeterminación absoluta, y Leibniz equipara su noción de materia prima con la fuerza pasiva y primitiva de resistencia, en el nivel metafísico (hay otra fuerza pasiva y derivada, a la que también llama resistencia, al nivel físico o dinámico), y habíamos dicho que ninguna fuerza es de grado cero, por más pasiva que ésta sea.

Ciertamente Leibniz no formuló una teoría monista explícita de la fuerza, pero todo indica que tuvo los elementos para hacerlo, y que estuvo muy cerca de ello. Esta posibilidad hace a la teoría dinamista leibniziana muy atractiva para la visión científica contemporánea del universo, incluyendo en ésta el problema de la relación entre sucesos mentales y sucesos cerebrales

Bibliografía

BUNGE, Mario, *The Mind-Body Problem*, New York, Pergamon Press, 1980. Trad. al esp. *El problema mente-cuerpo*, Madrid, Tunos, 1988.

GM Die Mathematische Schriften von Gottfried Wilhelm Leibniz 7 vols., ed por C. J. Gerhardt, Berlín, Halle, 1849-55.

GP Die Philosophische Schriften von Gottfried Wilhelm Leibniz 7 vols., ed por C. J. Gerhardt, Hildesheim/New York , Georg Olms, 1978.

O G. W. Leibniz. Escritos filosóficos, ed. por Ezequiel de Olaso, Buenos Aires, Charcas, 1982.

L G. W. Leibniz. Philosophical Papers and Letters, 2ª ed., ed. por Leroy E. Loemker, Dordrecht/Boston, Reidel, 1969.

PA G. W. Leibniz Sämtliche Schriften und Briefe, ed. por la Academia de Ciencias de Prusia, después de 1945 Academia Alemana de Ciencias, Darmstadt y Leipzig: 1923 –(inconclusa).

HERRERA Ibáñez, Alejandro, "Leibniz y el concepto de materia" en Mauricio Beuchot, Laura Benítez *et al.*, *El concepto de materia*, México, Colofón, 1992.

MATES, Benson, *The Philosophy of Leibniz*, New York, 1986.

C Opuscules et fragments inedits de Leibniz, ed. por L. Couturat, Hildesheim, Georg Olms, 1961.
SCHRÖDINGER, Erwin, *Ciencia y humanismo*, trad. de Francisco Martín, Barcelona, Tusquets, 1985. Orig. *Science and Humanism*, 1951.

EL PROBLEMA MENTE-CUERPO EN DESCARTES Y HUME

Margarita Costa*

Para Hume, Descartes ejemplifica en forma paradigmática el escepticismo que llama *antecedente*, previo a toda reflexión filosófica, como un medio para evitar el error y alcanzar un principio absolutamente evidente. Pero, concebida de modo tan radical, esta forma de escepticismo es para Hume absurda e imposible de alcanzar y, en caso de que lo fuera, sería incurable y no nos permitiría avanzar un paso en el camino del conocimiento[1] ni nos proporcionaría esa tranquilidad de espíritu que era la meta de los antiguos escépticos y, en general, de todo hombre prudente que quisiera ocuparse de los asuntos de su vida y lograr el mayor bienestar posible. Suspender la creencia en la existencia del mundo que nos rodea y en la de los demás hombres es una ilusión que no puede mantenerse sino por un lapso relativamente breve. El mundo del puro pensamiento es como una atmósfera enrarecida en la que la mente no puede respirar por mucho tiempo. El propio Descartes describe su experiencia como la de quien está sumergido en aguas profundas sin poder tocar fondo ni mantenerse en la superficie. La razón por la cual Descartes duda de todo es su intento de alcanzar un principio original, que no pueda en modo alguno ser erróneo o engañoso. Para Hume no existe tal principio "ni, de existir, po-

* Universidad de Buenos Aires.

1 D. Hume, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, Reimpresión de la edición de 1777 con introducción e índice analítico de L.A. Selby-Bigge, 3ª Ed. con anotaciones de P.H. Nidditch, Oxford, Clarendon Press, 1988, p. 116 (en adelante *Enquiries*).

dríamos avanzar un paso más allá de él –según se ha señalado más arriba–, excepto usando esas mismas facultades en las que ya supuestamente no confiamos".[2]

Sin embargo, pueden hacerse algunas observaciones a la crítica de Hume. En primer lugar, no es rigurosamente verdadero que dicho escepticismo sea "*antecedente* a todo estudio y filosofía",[3] puesto que Descartes, al elegir la duda, está filosofando. De hecho, ningún no-filósofo emprendería ese camino. Por otra parte Descartes declara haber estudiado, no sólo en las aulas de las escuelas sino en el gran libro del mundo y es *la falta de certeza* de las opiniones recibidas, no la falta de estudios –de los que nos confiesa que su mente estaba como abarrotada– lo que lo conduce por el camino de la duda. Descartes ha dudado y vacilado antes de empezar a filosofar por cuenta propia y esa duda reflexiva y consciente que ahora emprende es el punto de partida de su filosofía. Se trataría entonces de lo que Hume admite como una forma de escepticismo moderado, que nos permite "una adecuada imparcialidad en nuestros juicios y libera nuestra mente de todos aquellos prejuicios que podemos haber adquirido por la educación o la opinión precipitada."[4] Si bien en este caso no hace referencia a Descartes, podemos afirmar que la modalidad que éste adopta al comienzo de su filosofía respondería, más bien, a la razonabilidad que Hume concede a una forma moderada de escepticismo antecedente. Es más adelante, al pautar los pasos de la duda, que Descartes responde a la versión humeana de un escepticismo antecedente radical.

Nuestra segunda observación se refiere a la primera certeza a la que llega Descartes y a su incorregibilidad. Se trata de la experiencia del *cogito*, en principio un puro acto de pensamiento.

2 *Ibidem, loc. cit.* (las versiones castellanas de los textos en inglés son mías).
3 *Loc. cit.*
4 *Loc. cit.*

Agrega la certeza de la *existencia* del *cogito*, sin que esto agregue una nueva idea al *cogito* mismo. Podemos señalar una primera coincidencia con Hume, quien afirma que: "La idea de existencia,..., se identifica totalmente con la idea de lo que concebimos como existente".[5] Por el momento, pues, la primera certeza no significa nada más que efectivamente "se piensa", "alguien piensa", "yo pienso". El orden de estas expresiones obedece a mi deseo de subrayar otra semejanza con Hume. Éste comienza hablando de percepciones –impresiones e ideas– que también, pese a no emplear formas verbales, pueden ser interpretadas como actos. Interpretarlas como cualidades implicaría el reconocimiento de algo en lo que las cualidades inhieren y en esto Hume se muestra ambiguo: sabemos que niega la certeza de un sujeto de inherencia y en esto es claro, ya que se aparta radicalmente de la metafísica de la sustancia; pero habla de las "percepciones de la mente humana", de "su aparición en la mente". De modo que entre Descartes y Hume habría hasta aquí sólo la diferencia –fundamental, por supuesto, pero de consecuencias análogas en cuanto al punto de partida de un filosofar auténtico, original y desprejuiciado– en el carácter de esa intuición original: pensamiento en Descartes, sentimiento o impresión en Hume. El pensar (las ideas) lo agrega Hume a continuación del sentir, mientras que para Descartes el pensar es lo original y las impresiones humeanas son como modos de ese pensamiento. Precedencia del sentir sobre el pensar en Hume; cuasi-identificación en Descartes, con predominio de lo intelectual. El pensamiento de Descartes se "despliega", por así decir, en múltiples modos: conceptos, juicios, sensaciones, emociones, imágenes, etcétera.

5 D. Hume, *Tratado de la naturaleza humana. Acerca del entendimiento*, trad., pról. y notas de Margarita Costa, Buenos Aires, Paidós, 1974, p. 123 (en adelante *Tratado*)

Pero la segunda semejanza entre Hume y Descartes que hemos querido subrayar es la *incorregibilidad* de la evidencia de que cada uno parte. Así, dice Hume: "Puesto que todas las acciones y sensaciones del espíritu nos son conocidas por la conciencia, deberán aparecer tal como son en todos sus detalles y ser tal como aparecen".[6] Daniel Shaw considera que en Hume encontramos "una versión plausible de la tesis de la incorregibilidad: la versión que afirma que si realizamos una inspección cuidadosa de nuestros deseos conscientes en cualquier momento, no podemos equivocarnos respecto de ellos."[7] Lo afirmado puede extenderse a todo tipo de experiencia consciente. Passmore agrega que "pensar de otro modo sería suponer que cuando estamos íntimamente conscientes podríamos estar equivocados".[8]

No obstante, no podemos dejar de subrayar una diferencia fundamental entre ambos filósofos. En Hume hay una acentuación de la nota "impersonal": impresiones, ideas –en la mente, mi mente–, pero tratadas en cuanto fenómenos o actos sin que esa mente desempeñe, a lo largo del Libro I del *Tratado*, ningún papel protagónico. En Descartes, en cambio, el pensamiento en sus distintos modos aparece en primera persona: yo pienso, yo siento, yo imagino. No dice "existen mis impresiones", "existen mis ideas", aunque esto también sea verdadero, ya que los "modos" participan de la misma evidencia del *cogito*. Sólo un rasgo común al latín y al español, que no se da en otros idiomas, permite la elisión del sujeto en la oración, en este caso el pronombre "yo", lo que no significa sino que en las primeras enunciaciones cartesianas ese yo está sobreentendido. Aquí aparece,

6 *Tratado*, p. 290.

7 D. Shaw, "Hume's Theory of Motivation", en *Hume Studies*, vol. XV, núm. 1, abril de 1989, p. 180.

8 J. Passmore, *Hume's Intentions*, Londres, Gerald Duckworth & Co. Ltd., 1968, p. 117.

pues, en Descartes, una primera concepción del yo como *mente*, aunque no se afirme nada aún acerca de su naturaleza. Pero el yo es sin duda la mente (*mens sive anima sive intellectum*), aun cuando se trate de lo que Husserl interpretará como un yo trascendental. Pero la pregunta que se plantea aquí es de otra índole: ¿es esta mente un hombre, un hombre completo? Según la interpretación de un filósofo humeano contemporáneo, no puede afirmarse esto en absoluto: "Si tomamos en serio la duda cartesiana, no podemos por cierto probar la existencia de personas, porque las personas son objetos físicos y no podemos probar la existencia de objetos físicos si la ayuda de Dios nos es inasequible",[9] dicho esto último en tono irónico. Si tenemos en cuenta esta tesis materialista o al menos fenomenista, cabe todavía preguntarse si la mente no es algo que integra la persona de modo de poder distinguirla de otros aspectos o partes de ésta. Recordemos que para Hume todo lo distinguible es separable. Si la mente se distinguiera de otros aspectos o partes de la persona, de lo que pudiéramos tener conocimiento, sería, según esta tesis de Hume compartida por Descartes, separable de ellos –y así nos encontramos nuevamente con un punto de coincidencia entre ambos filósofos. La mente no sería para Descartes un hombre completo, pero sí algo de lo que no sólo podemos hablar independientemente, sino que además ella *debe poder existir separadamente*. No obstante, contentémonos por ahora con poder hablar de ella como de algo de cuya existencia no dudamos, sin arriesgar aún su separabilidad de todavía no sabemos qué.

El *cogito* es, pues, una mente, algo que ni siquiera Hume negaría en este primer sentido, aunque no admitiría que nuestra primera experiencia sea pensar y no sentir. Pero insisto en que

9 P. Árdal, *Passion and Value in Hume's Treatise*, Edimburg University Press, 2ª ed., 1989, p. XXII.

la incorregibilidad de esa evidencia, sea racional o empírica, es atribuible a ambas formas de lo mental. Por su parte, si el hombre es su mente, lo que estaría en cuestión aquí son las distintas concepciones de la mente en cuanto a su "naturaleza original" y el acceso a su conocimiento, no su existencia.

De una primera concepción del yo como mente en Descartes, puede sacarse otra conclusión. Lo que está en cuestión aquí no es si el *yo* es una mente, no si el hombre en su integridad lo es. Veamos qué sucede a medida que avanzamos en nuestro análisis.

En la *Segunda meditación* Descartes se ocupa de la naturaleza del yo. Hemos visto que aquello cuya existencia pone (*positum*) es el *cogito*, no un *cogitare*. Y el *cogito* implica un sujeto de predicación, en este caso un *ego*, cualquiera sea la naturaleza de éste. La *mente* de Hume, o aun el *haz de percepciones* como lo llamará más adelante, no es un yo identificable desde un comienzo, es un puro fluir de impresiones e ideas sin un punto de referencia *del cual emanen*. El haz o colección de percepciones se constituye por las leyes de asociación, como toda idea compleja, pero no hay una apreciación, un polo que determine *a priori* ese fluir de actos mentales. En Descartes sucede lo contrario: todos los actos de pensamiento por los cuales sé que existo, aun cuando se trate de dubitaciones, refieren a mi yo. Todos los verbos empleados en este contexto están en primera persona. ¿Qué soy –se pregunta por último– ahora que sé que soy?

La respuesta a la pregunta de Descartes planteada al final del párrafo anterior es tan terminante como su primera afirmación del *cogito*, aunque no presente el mismo grado, ni siquiera la misma forma de evidencia e incorregibilidad. Descartes afirma sin vacilar: "yo no soy, pues, hablando con precisión, más que una cosa que piensa, es decir, un espíritu, un entendimiento o una razón, que son términos cuyo significado antes me era desconocido. Así, yo soy una cosa verdadera y

verdaderamente existente; pero, ¿qué cosa? Ya lo he dicho: una cosa que piensa".[10]

Podríamos a partir de este punto comenzar a dudar nuevamente, ya que esta segunda afirmación, que pretende reforzar la primera y brindar al filósofo la tranquilidad que buscaba en el mar de la duda es, sin embargo, una que no todos pueden suscribir con la misma certeza y que ha sido blanco de numerosas críticas. Se nos ha pedido no precipitarnos y abandonar nuestros prejuicios, pero ninguna de esas exigencias del método cartesiano se cumple en este segundo paso hacia la constitución de un sistema. De pronto, como un supuesto que la duda no ha logrado desarraigar, resurge la sustancia, la cosa, aquello que verdaderamente es, según se enseñaba en las escuelas. La realidad es un predicado de la esencia y la esencia es aquello que la sustancia es, su naturaleza. Las sustancias, pues, son lo que es; las esencias son el qué de las sustancias. La sustancia que ocupa a Descartes es la *cogitante*, la pensante. El yo ya no es un mero punto de referencia gnoseológico o lingüístico, ha cobrado consistencia ontológica. Tenemos, entonces, una mente que piensa, aunque esto pueda considerarse redundante, ya que el pensamiento no le es accidental a la mente, sino esencial, constituye su naturaleza. Pero, a su vez, la mente como sustancia no le es accidental al pensamiento. Por eso se ha podido cuestionar la diferencia entre la sustancia y su atributo esencial. "La sustancia –dice Descartes– se reconoce, ciertamente, por cualquier atributo, pero, sin embargo, una propiedad principal de cada sustancia constituye su naturaleza y esencia, y a ella se refieren todas las demás".[11] Tenemos, pues, el conocimiento más claro y distinto de nuestra mente al que según Descartes podemos aspirar.

10 R. Descartes, *Meditaciones metafísicas*, en *Obras escogidas*, Buenos Aires, Ed. Charcas, 1980 (trad. de Ezequiel de Olaso y Tomás Zwanck), p. 226.
11 R. Descartes, *Los principios de la filosofía*, LIII, ed. cit., p. 334.

La concepción cartesiana de la mente o alma difiere de la escolástica. Para esta última, en efecto, el alma es la forma de una sustancia, no una sustancia completa en sí misma como para Descartes. Sin embargo, éste no introduce nada nuevo en el sentido estrictamente metafísico. Encuentra no sólo el término acuñado para ese algo que ha intuido al cabo de la duda, sino también un concepto correspondiente que no cuestiona ni analiza. No tenemos más remedio que reconocer el traspié dogmático de Descartes. En ningún momento se desdice de él o intenta una justificación. En rigor, por esta vía Descartes inaugura una nueva escolástica que se prolongará hasta el siglo XVIII y tendrá su mayor desarrollo en Alemania hasta Kant.

Existo, es decir, existe mi yo, es decir, existe mi mente y esa mente es algo, una sustancia, algo que no necesita de otra cosa para existir, excepto "del concurso de Dios", según aclarará Descartes a continuación, dado que mi mente es una sustancia finita. La necesidad de Dios está ya implícita en la definición de la mente humana, antes de que Descartes la emplee para demostrar la existencia de un mundo exterior a ella, en el cual deberá encontrarse, entre otras cosas, mi propio cuerpo.

También encontramos en Descartes referencias al cuerpo antes de que haya expuesto sus pruebas de la existencia de Dios, que son las que a su vez le permitirán, como acabamos de afirmar, demostrar la existencia del mundo exterior. En la *Segunda meditación* admite una idea clara y distinta del cuerpo –no de *mi* cuerpo, por cierto, aunque éste pudo haber sido elegido como ejemplo si su transformación molecular fuese tan rápida y fácil de observar como la de la cera sometida al calor del fuego. Esta idea, de algo extenso, flexible y mudable, nos da a conocer algo de una naturaleza totalmente distinta de la del alma, no por los sentidos ni por la imaginación, que por estar ligada a los sentidos comparte su carácter engañoso, sino por el entendimiento o razón. Pero dado que Descartes sospecha ahora que ésta –la razón-

podría engañarnos, pese a la claridad y distinción con que nos presenta sus objetos, a diferencia de la oscuridad y confusión con que se me dan a los sentidos, no bastará con poseer la idea para poder *demostrar* con certeza la existencia del *ideatum.* La existencia del cuerpo no tiene la evidencia de la del alma, pese a que su esencia sí goza de ella. El criterio de verdad proporcionado por el *cogito* comienza a funcionar, pero necesita una *garantía*, que es requerida por la existencia de una sustancia distinta del yo, no por su esencia, cuya idea es tan clara y distinta en el caso del cuerpo como en el de la mente.

Que el alma sea más fácil de conocer que el cuerpo significa, ante todo, que se me da *primero* y, en segundo lugar, que al dárseme el *cogito*, lleva implícito en él su inherencia en un sujeto o yo. Pero bien podría afirmar aquí Descartes, sin incurrir aparentemente en contradicción, que conocemos el cuerpo como conocemos la mente, por el solo hecho de captar su esencia, en este caso la extensión en lugar del pensamiento. En efecto, lo que nos queda del conocimiento de la cera, una vez dejados de lado los datos sensibles, es "algo extenso, flexible y mudable". Sin embargo, señala Descartes, podría suceder que ese *algo* no existiese si un genio nos engañara.[12] Esto requiere una aclaración. Descartes ha afirmado que no puedo dudar de que pienso, pero al parecer tampoco de que ese algo que piensa sea una sustancia, lo que hemos considerado como una trasgresión al concepto de evidencia. ¿Cuál es, pues, la diferencia respecto de la evidencia del cuerpo? Creo que la hallaremos si dejamos de lado el discutible concepto de sustancia.

En el caso de los cuerpos, hemos visto que puedo dudar de que haya en el mundo algo extenso, pese a que la extensión se me da como un dato evidente a la razón. Pero puedo dudar de la extensión misma como esencia –que nada sería sin su inhe-

12 *Ibidem*, p. 333.

rencia a una sustancia– ya que un genio podría engañarme y el hecho de *dudar* de la extensión no la haría indubitable, como hace indubitable el pensar el propio acto de dudar de que pienso. En ningún momento puede ponerse en duda la validez del recurso metodológico cartesiano respecto de la evidencia de nuestro puro acto de pensar.

Ahora bien, si ese Dios existe, según pretende demostrar Descartes, y ese Dios no me engaña, según implica su concepto –otra *idea* que encuentro en mí, como la del *cogito* y la de la extensión, que goza de la misma evidencia que todas las *cogitationes* claras y distintas en cuanto tales– será la garantía de que la extensión no es una mera idea perteneciente al mundo de los posibles, sino de que existe algo extenso, que por tener una esencia distinta del pensamiento, no pertenece a la mente sino que la trasciende. Nuevamente incurre Descartes en su aludida trasgresión dogmática: la afirmación de otra sustancia. No sólo encuentro en mi mente las ideas claras y distintas de extensión y movimiento, a partir de las cuales pueden construirse una matemática y una física *a priori*, sin recurrir a la experiencia sensible, sino que existen sustancias a las que esos conocimientos son aplicables: los cuerpos.

Hasta aquí hemos puesto de manifiesto algunas confusiones en el pensamiento cartesiano. La evidencia, según su propia concepción de ella, que la vincula con la verdad en la primera regla del método, debe aplicarse en rigor a las ideas de esencias. Es su consecuencia o *adaequatio* con las esencias, no con supuestas sustancias, lo que les da su carácter de verdaderas. Pero de aplicarse este único criterio, nunca saldríamos de un mundo ideal, matemático. La existencia del *cogito* es, en realidad, una excepción a lo que corrientemente se llama existencia: es la de un puro pensamiento; mientras que la afirmación de la existencia de la *sustancia pensante* es en rigor una violación a la primera regla del método cartesiano. Para admitir la existencia de la

mente, no es necesario hacer de ella una cosa; en cierto sentido, es más bien desvirtuarla. Por eso Descartes, sin recurrir a Dios, no podrá trascender nuevamente el mundo de los posibles o las esencias, sólo una de las cuales *es* en el sentido más estricto de la *actualitas* –Dios mismo– y otra en sentido analógico, en cuanto no existe por sí misma, o en virtud de sí misma, como Dios, pero *es conocida* como existente al conocerse su esencia: el *cogito.* Aquí pondremos fin a la interpretación idealista de Descartes que, de mantenerse, nunca nos permitiría alcanzar el cuerpo, con el que la mente (*mens sive anima*) debe tener alguna relación no mediada para que puedan cumplirse los actos de nuestra vida, en particular los vegetativos y sensitivos.

Antes de pasar a una interpretación más realista de Descartes, debemos volver a Hume en el Libro I del *Tratado*, para ver si encontramos alguna referencia a la mente que nos permita compararla con la sustancia pensante cartesiana. Pero de ahí en adelante, Hume se aparta totalmente de Descartes; podríamos decir más bien que éste es el blanco de sus más acérrimos ataques. La sustancia pensante es para Hume una de esas ficciones de las que debe liberarnos la filosofía, con el mismo vigor con que rechaza toda superstición. Es una falsa creencia en la que los hombres incurren, cuyo origen puede explicarse, como el de otras creencias, por los principios de la naturaleza humana establecidos por Hume, pero que debe desecharse en pro de la auténtica filosofía. Así afirma Hume que "lo que llamamos una *mente* no es sino un haz o una colección de diferentes percepciones unidas entre sí por ciertas relaciones, y al que se supone, aunque esto sea falso, dotado de una perfecta simplicidad e identidad".[13]

No se le escapa a Hume el fundamento religioso de la noción de mente o alma como algo capaz de existir por sí, ya que ello era necesario para poder sostener su inmortalidad, que Descartes

13 D. Hume, *Tratado*, p. 331.

suponía probada por el solo hecho de haber demostrado que teníamos de ella una idea distinta de la del cuerpo[14] y que, en consecuencia, podía existir separada de él. Por el contrario, apoyándose en la opinión de eminentes teólogos, Hume considera que "aunque el vulgo no posea principios formales de infidelidad, en realidad es infiel en su corazón y no posee nada semejante a lo que podemos llamar una creencia en la eterna duración de sus almas."[15] Nos resulta llamativa esta apelación a "eminentes teólogos", ya que la creencia en la inmortalidad del alma es uno de los fundamentales artículos de fe de la religión cristiana, y había sido avalada desde la Antigüedad, especialmente por Platón. Para Hume el problema reside simplemente en explicar por qué creemos que la mente es una sustancia y lo considera una cuestión epistemológica más que parece resolverse reduciendo la mente a pura actividad.

Hemos visto, pues, en el Libro I del *Tratado*, disolverse el yo o la mente en un simple haz de percepciones, unidas en virtud de las leyes de la asociación como cualquier otra de las llamadas "cosas" o "sustancias". No entraremos aquí en la cuestión de la identidad personal ni en los múltiples problemas que plantea en torno a la filosofía de Hume. Se han escrito muchos y muy valiosos trabajos sobre el tema, que si bien no logran tal vez dar una respuesta satisfactoria, permiten explicar nuestra creencia en un yo que es exclusivamente nuestro y al que atribuimos toda clase de cualidades: tanto virtudes como vicios, talentos como imperfecciones, que actúa también como observador y que nos permite elaborar una tesis epistemológica cuestionadora de todo dogmatismo, pero no escéptica en sus últimas consecuencias, al menos en lo que respecta a las enseñanzas del sentido común.

14 R. Descartes, *Second Réponses* aux *Méditations Métaphysiques*, París, Garnier-Flammarion, 1979, pp. 251-252.
15 D. Hume, *Tratado*, p. 188.

Pero Hume no ha podido prescindir del yo, ya no como observador, como "pasivo espejo de la naturaleza", sino como agente cuyas acciones afectan a los demás y modifican el entorno, tanto físico como histórico-cultural. Como señala Capaldi, "Hume niega que podamos conocer un *yo* como una sustancia mental permanente aunque, por otra parte, podemos *pensar* ese yo, es decir, podemos tener una *idea* de él y de hecho necesitamos y a la vez adquirimos esta idea de nuestras posiciones como *agentes*. La idea del *yo* surge entre las pasiones y, cuando Hume explica los hábitos de la mente, presupone un yo como entidad que adquiere y ejercita tales hábitos."[16]

En una palabra, el yo de Hume es un yo activo, no sólo el yo de las pasiones y de la acción moral y política, sino también el de las rápidas inferencias que responden a un instinto casi infalible. Nuestras pretensiones ontológicas o epistemológicas pueden quedar frustradas, pero seguiremos considerando nuestro yo como un sujeto o agente de nuestra vida, "aquél que siempre va conmigo", según diría Machado. Consideramos que todas nuestras creencias, incluso aquéllas que como filósofos admitimos que producen los mundos de semi-fantasía en que habitamos, se fundan en ese sistema que es mi yo, que con todas sus variantes nos acompaña a lo largo de nuestra existencia, ya que su conciencia, según señalan los psicólogos, se adquiere tempranamente. Solamente podríamos ejemplificar un sucederse de percepciones impersonales que no nos atribuimos como agentes, en el caso del niño pequeño que habla de sí mismo en tercera persona. Pero esto sería más bien objeto de una epistemología genética. En cuanto a la pérdida de la propia identidad, que al parecer una teoría como la humeana podría explicar, sería considerada por

16 N. Capaldi, "The Problem of Hume and Hume's Problem", en *McGill Hume Studies*, compilado por D. Norton, N. Capaldi y W. Robinson, San Diego, Austin Hill Press, 1979, p. 14.

todos como un fenómeno patológico contrario a las leyes de la naturaleza. El problema crucial, según señala Passmore, residiría en que "Hume nunca realmente elaboró la relación entre su epistemología y su teoría de las pasiones".[17]

La actividad mental ocupa, sin lugar a dudas, un lugar preponderante en la filosofía de Hume. En el *Tratado* comienza hablando de nuestras percepciones como si éstas no necesitaran siquiera un "escenario" en el cual desarrollar su acción. Todo lo que de alguna manera tiende a sustancializar la mente es considerado como mera metáfora. Y de todas las metáforas empleadas por Hume, quizás la que nos da una idea más aproximada de su concepción de la mente es su comparación con una república. Si se nos pregunta qué es una república, no bastará definirla como una forma de gobierno, sino que para hacer suficientemente clara nuestra idea a quien no sepa qué es, tendremos que referirnos a sus miembros y leyes y a los cambios que pueden operarse en unos y otros. Así, dirá Hume: "En este sentido no se me ocurre nada más apropiado que comparar el alma con una república o estado cuyos diversos miembros están unidos por los vínculos recíprocos del gobierno y la subordinación, y dan origen a otras personas que, a su vez, expanden esa misma república por el cambio constante de sus partes. Y así como la misma república individual puede no sólo cambiar sus miembros sino también sus leyes y su constitución, la misma persona puede modificar su carácter y su disposición, al igual que sus impresiones e ideas, sin perder su identidad."[18] Las relaciones, en este caso, no son las que constituyen la república, sino las leyes básicas de la naturaleza humana: las leyes de asociación entre las que, naturalmente, desempeña un papel primordial la de causalidad.

17 J. Passmore, *op. cit.*, p. 126.
18 D. Hume, *Tratado*, p. 376.

Pero al mismo tiempo que dicha ley o principio es fundamental para la constitución de la mente, Stroud ha señalado que la *mente* es, por su parte, esencial para la segunda definición de 'causa' que da Hume: "...una cosa puede cumplir las condiciones de la segunda 'definición' sólo porque hay mentes; y las cosas cumplen las condiciones de la segunda 'definición' siempre que observamos que cumplen las condiciones de la primera sólo porque esas mentes son como son".[19] La aparente circularidad de este argumento, si comparamos las afirmaciones de Stroud con la cita del *Tratado*, se despeja fácilmente si tenemos en cuenta que llegamos a la idea de *mente* aplicando la primera definición de causalidad –al menos en aquello que concierne a la descripción 'externa' del proceso como sucesión de hechos semejantes por otros hechos semejantes– y luego esta mente, así constituida en una primera aproximación, al incluir sentimientos y pasiones (así como propensiones, que Hume no trata muy explícitamente) nos hace tomar conciencia del elemento esencial –'subjetivo'– del proceso causal, que a la vez nos confirma con evidencia la acción de dicha mente, no meramente como observadora de la relación, sino como agente de su constitución.

En cuanto a nuestro cuerpo, encontramos en los textos de Hume numerosas referencias a él, aunque no trate específicamente el tema en ninguna parte. Por un lado, podríamos pensar que la creencia en la existencia de nuestro propio cuerpo se explica como nuestra creencia en el mundo exterior y que, por tanto, no tiene sentido preguntarnos si nuestro cuerpo existe, sino sólo por qué *creemos* que existe, y que los argumentos de la constancia y la coherencia de algunas de nuestras percepciones serían la respuesta adecuada para dar cuenta de esa creencia. Sin embargo, en su teoría acerca del mundo exterior y las razones de nuestra creencia en él, Hume se ocupa de objetos *externos* a

19 Stroud, *Hume*, trad. Antonio Zirión, México, UNAM, 1986, p. 135.

nuestro propio cuerpo, como mesas, leños, escaleras y cuerpos de otras personas, como el del conserje que sube a traernos una carta, y el problema del propio cuerpo no queda englobado en el tratamiento escéptico de la cuestión, sino más bien deliberadamente excluido. El estudio de nuestro cuerpo parece formar parte de su ciencia, la *ciencia del hombre*, no de la que se ocupa de la naturaleza exterior. Sin embargo, la primera referencia al cuerpo que encontramos en Hume aparece en la Sección II del Libro I del *Tratado*, donde señala que: "El análisis de nuestras sensaciones atañe más bien a los anatomistas y filósofos de la naturaleza que a los filósofos morales y, por tanto, no será efectuado aquí".[20] Es decir que Hume se dispone, al parecer, a estudiar la menta humana y a dejar de lado, al menos por el momento, lo relativo al cuerpo. Al respecto señala P. Ardal: "Hume cree que hay una parte interna y otra externa en esos objetos físicos que llamamos personas... Pero... no puede probarse que las creencias naturales sean verdaderas".[21] J. Passmore es más asertivo al respecto: "Esto no significa que Hume dudara jamás de que las impresiones eran producidas en la mente por alguna clase de mecanismo fisiológico".[22] También Kemp Smith sostiene que Hume "considera que la 'naturaleza humana' incluye el organismo físico entre sus elementos constitutivos",[23] aunque este conocimiento sería meramente conjetural. Su escepticismo, al parecer, no se extiende explícitamente a la creencia en el propio cuerpo.

Por otra parte, dado que su teoría de la mente es tributaria de su teoría de la necesidad física, no debe extrañarnos que Hume

20 D. Hume, *Tratado*, p. 40.
21 P. Árdal, *op. cit.*, p. XXII
22 J. Passmore, *op. cit.*, p. 93.
23 N. Kemp Smith, *The Philosophy of David Hume*, Londres, Macmillan, 1966, p. 240.

intente explicar los fenómenos que atañen a nuestro propio cuerpo, así como aquéllos que denotan interacción con la mente, aplicando su teoría general de la naturaleza humana.

Por su parte, Stroud observa que: "Hume trata de esbozar algunos principios generales. Sus afirmaciones se apoyan primordialmente en fenómenos psicológicos, entendidos conforme a la teoría de las ideas, pero no encontramos en su programa general nada que excluya a la fisiología de la teoría total del hombre".[24] Emplea, curiosamente, la palabra 'máquina' para referirse al cuerpo, lo que nuevamente lo acercaría a Descartes... Considera que "un cuerpo humano es una máquina enormemente complicada: que muchos poderes secretos se ocultan en ella, que están completamente más allá de nuestra comprensión: que a menudo debe aparecérsenos como muy incierta en sus operaciones."[25]

Otra alusión al cuerpo del hombre, que nos muestra que es tenido en cuenta en las más diversas circunstancias, aparece en la Parte II del Libro III del *Tratado*, la que se refiere al origen de la justicia y la propiedad. Refiriéndose a las circunstancias externas que coadyuvan en el establecimiento de la justicia como virtud artificial, Hume menciona tres clases de bienes, entre los que se encuentran "las ventajas externas de nuestro cuerpo", que según él pueden sernos arrebatadas pero sin beneficio alguno para quien nos produce el daño. La existencia del cuerpo es aceptada, pues, de manera casi dogmática, o al menos incluida entre aquellas creencias que no pueden ponerse seriamente en duda, pese a que, como señala Passmore: "Las esencias reales de los cuerpos nos son tan desconocidas como las esencias reales de las mentes."[26]

24 B. Stroud, *Hume*, *op. cit.*, p. 227.
25 D. Hume, *Enquiries*, p. 67.
26 J. Passmore, *op. cit.*, p. 11.

La cuestión de la interacción entre el alma y el cuerpo es un problema crucial que Descartes no puede eludir al llegar a la *Sexta meditación*. Hasta ahora nos ha hablado de ideas "claras y distintas" de la razón, de demostraciones a partir de ellas y de las nociones comunes, que le permiten explicar el mundo, tanto mental como físico, por medio de argumentos que incluyen la intuición racional, una deducción inmediata bi-direccional de tipo matemático y la intervención de su Ser Supremo, sin el cual carecerían de auténtico valor cognoscitivo nuestra "claridad y distinción" mentales. Pero no ha podido negar la existencia del cuerpo (quizás habría preferido prescindir de él o como le recrimina Pascal *"pouvoir se passer de Dieu"*,[27] pero al parecer ambos le son imprescindibles para construir su sistema). La existencia de Dios parece demostrada y, lo que es casi más importante para Descartes, se ha probado que Dios *no quiere* engañarnos.

Por eso en la *Sexta meditación* debe volver a ocuparse del "engaño de los sentidos". Todo lo que nos pertenece, lo que pertenece a nuestra naturaleza, nos ha sido dado por Dios y por tanto debe tener un fin bueno. No somos espíritus puros: nos alimentamos de *nourritures terrestres*, nos desplazamos, percibimos aquello que nos rodea y esa percepción nos permite elegir lo bueno y evitar lo malo. Lo bueno y lo malo no sólo para el cuerpo, como si en su accionar automático sólo estuviera previsto librarnos del mal físico y proporcionarnos el bien físico. Esto, evidentemente, no es así. Descartes ha introducido en el cuerpo del hombre lo que podríamos considerar como un "alma platónica" o como un espíritu o fantasma, como dirá Gilbert Ryle, que no permanece indiferente al funcionamiento de la máquina. Dice Descartes que la naturaleza que Dios nos ha dado, no puede engañarnos respecto de lo bueno y lo nocivo para el "compuesto alma-cuerpo". Esta admisión nos lleva muy lejos de lo que nos

27 B. Pascal, *Pensées*, París, Garnier s.f., p. 94.

enseñaban las ideas claras y distintas acerca de dos sustancias esencialmente heterogéneas. Descartes puede continuar admitiendo su distinción, pero tendrá que dar cuenta del fenómeno de la interacción entre ambas.

Para ello Descartes se vale de dos subterfugios: el primero consiste en modificar el sentido de la palabra "naturaleza" tal como la empleara en las meditaciones anteriores. Ésta es ahora entendida como "Dios mismo" y las disposiciones que Dios nos ha dado, avaladas por su veracidad, no como la engañosa tendencia a creer que nuestras percepciones "proceden de fuera" simplemente por una fuerte inclinación. ¿Por qué Dios permitía que nos engañáramos antes y no lo permitirá ahora? Simplemente, porque la cuestión no es ahora el conocimiento claro y distinto de la realidad que Dios con su infinita bondad nos ha otorgado por medio de la razón, sino nuestra supervivencia, a la que Dios también ha debido proveer. Ya no importa que las cosas no sean tal como nuestros sentidos nos las representan sino sólo como la razón nos lo muestra y demuestra, ya que la función de los sentidos no es cognitiva, al menos no en el sentido en que Descartes define el conocimiento.

En la *Sexta meditación*, en la que Descartes trata explícitamente el problema de la relación entre la mente y el cuerpo y sus consecuencias, comienza afirmando algo que ya sabíamos a partir de la *Segunda meditación*. A saber, que la existencia del cuerpo es menos posible, ya que su idea no encierra contradicción y Dios, que encarna la racionalidad suprema, es capaz de producir todo lo que yo concibo con claridad y distinción. Recordemos que la extensión es percibida con evidencia, pero las demás cualidades de las cosas extensas, las que nos permitirían distinguir, por ejemplo, un trozo de cera de un trozo de mármol, hemos debido dejarlas de lado como engañosas. De lo posible deberemos pasar a lo realmente existente, pero antes Descartes recurre a otra facultad, la imaginación, como una suerte de intermediario

entre el intelecto y los sentidos. Aquello que concibo como posible puedo también imaginarlo; pero para ello es necesaria "cierta contención de mi espíritu", debo poner límites a lo universal de mi concepción para dar de ella una representación que debe ser individual, sensible, aun con conciencia de la ausencia del objeto. Para ello la imaginación debe "volverse hacia el cuerpo", mostrarnos algo corporal correspondiente a la concepción del espíritu. Naturalmente, el ejemplo que da Descartes es de índole matemática, ya que sólo en términos matemáticos puede concebir claramente algo corporal y es de esa idea clara que la imaginación toma la configuración o estructura que traslada al cuerpo para presentárnosla de manera más accesible a nuestra mente limitada. Para ello no basta que los cuerpos sean posibles; hemos dado un paso más: la imaginación nos muestra la "probabilidad" de la existencia de los cuerpos. ¿Cómo nos sería posible representarnos algo corporal sin que se diese al menos la probabilidad de su existencia? Si bien se trata de un juicio conjetural, está mucho más próximo a la afirmación de una existencia real que la mera concepción de su "posibilidad". Podríamos comparar con la división que establece Hume entre el "conocimiento" propiamente dicho, que es una relación entre puras ideas y no nos informa acerca de los hechos, y la "probabilidad", fundada en nuestras inferencias causales acerca de los hechos, que nunca alcanza la certeza absoluta. Por otra parte, el propio Hume emplea en este contexto la palabra "imaginación", a menudo como sinónimo de entendimiento, y reserva el término "razón" para las demostraciones matemáticas.

De la imaginación de objetos no matemáticos, Descartes pasa a considerar otras cosas de naturaleza corporal, como los colores, sabores, placeres, etc., y de allí a afirmar que "percibo mucho mejor esas cosas por los sentidos, por medio de los cuales y de la memoria parecen haber llegado hasta mi imaginación".[28] Pero

28 R. Descartes, *Meditaciones metafísicas*, ed. cit., p. 272.

si bien reconoce que "me había servido de los sentidos antes que de la razón",[29] como la duda metódica había anulado su confianza en ellos, debe tratar ahora de demostrar: 1) que no pueden ser un simple y gratuito mecanismo de engaño; 2) que se sentía estrechamente unido a su propio cuerpo más que a ningún otro y que "sentía en él y por él todos mis apetitos y todas mis afecciones".[30] Esto lo vuelve más cauto: ya no emprenderá para el conocimiento la vía de los sentidos y aceptará las enseñanzas de su razón como las más valederas, pero al mismo tiempo deberá quizás salvar de la duda alguna de las cosas en que creía, pues su vida se tornaría absurda sin ellas. Al llegar a este punto, Descartes podría haber suscrito la afirmación de Hume: "Podemos muy bien preguntar *qué causas nos inducen a creer en la existencia de los cuerpos*, pero es en vano interrogar *si existen o no*. Éste es un punto que debemos dar por sentado en todos nuestros razonamientos."[31]

Dios debe ser absuelto nuevamente de la posibilidad de que nos engañe, ya que si bien nos engañamos por cuenta propia, es decir, por nuestra misma naturaleza, esa naturaleza y todas nuestras facultades han sido puestas en nosotros por Dios y su bondad no permitiría que nos condujeran *indefectiblemente* a error, sin habernos dado los medios de comprender cómo deben entenderse cada una de esas facultades y el alcance de cada una de ellas. Dado que a menudo las percepciones de los sentidos son oscuras y confusas, no deberemos guiarnos por ellas para conocer la realidad tal cual es, pero dado que creemos firmemente que proceden de los cuerpos, "es necesario concluir que existen las cosas corporales".[32] Es decir, mientras que la posibilidad de su

29 *Ibidem*, p. 274.
30 *Loc. cit.*
31 D. Hume, *Tratado*, p. 287.
32 R. Descartes, *op, cit.*, ed. cit., p. 278.

existencia depende de la razón y su probabilidad de la imaginación, la *certeza* de su existencia depende de los sentidos. Esto ha hecho señalar a Hume: "Recurrir a la veracidad del Ser Supremo, a fin de probar la veracidad de nuestros sentidos, es por cierto hacer un rodeo muy inesperado. Si su veracidad tuviera algo que ver con esta cuestión, nuestros sentidos serían enteramente infalibles, porque no es posible que él pueda jamás engañar."[33] Pero Descartes va aún más lejos: los sentidos no sólo nos enseñan que los cuerpos existen, sino que cada mente está muy estrechamente unida a *su* cuerpo. Ya no es un alma platónica en la que el cuerpo sólo puede provocar la reminiscencia de las únicas realidades, que son las Ideas, sino que hay en ellas sentimientos, si bien confusos, de hambre, sed, dolor, etc., –sentidos por el alma, no por el cuerpo– que atestiguan la realidad de mi cuerpo y de su extraña unión con mi alma a lo largo de mi vida.

Como señala Desmond Clarke: "Para explicar nuestra conciencia de las experiencias sensoriales, Descartes introduce una teoría del cuerpo humano, como un mecanismo para trasmitir información al cerebro, y el concepto de interacción mente-cuerpo para explicar la advertencia de tales experiencias perceptivas".[34] Lo que resulta sin duda extraño es que un proceso puramente mecánico, como el contacto de los cuerpos extraños con nuestros órganos de los sentidos y su transmisión al cerebro a través de los nervios por medio de los espíritus animales (lo que no resulta incoherente, dados los rudimentarios conocimientos fisiológicos de la época), dé como resultado un fenómeno o fenómenos mentales no mecánicos, como la imagen perceptiva –que no reproduce fielmente el objeto que originara el proceso–, así como la concepción de la razón y el juicio, que nos permite

33 D. Hume, *Enquiries*, p. 120.

34 D. Clarke, "The Concept of Experience in Descartes's Theory of Knowlesedge", en *Studia Leibnitiana*, Band VIII, Heft 1, 1976.

establecer, si la voluntad no toma precipitadamente la delantera, qué es lo verdaderamente real acerca de esas cosas externas y, al mismo tiempo, aprovechar las confusas imágenes perceptivas para guiarnos en esta vida vegetativa y sensitiva en la que nuestro cuerpo desempeña un papel tan importante. El "realismo" cartesiano ha dado un gran paso, pero no ha logrado unir, no ya la mente y el cuerpo, sino las dos teorías acerca de su nítida separación y de su confusa unión e interacción, que se superponen en su obra.

En mi opinión, lo que permite a Hume una explicación más coherente de la interacción entre la mente y el cuerpo, sin verse obligado a forzar los principios de su filosofía como lo hace Descartes, es el hecho de haber desechado desde un primer momento la noción de sustancia. Sin embargo, también encontramos dificultades y contradicciones en los textos humeanos. Así, leemos en la Parte II del Libro I del *Tratado* una explicación de cómo la mente puede evocar aquellas ideas que necesita en un momento determinado:

> Hubiese sido fácil efectuar una disección imaginaria del cerebro y mostrar por qué, al concebir una idea, los espíritus animales se transmiten a todas las huellas contiguas y despiertan todas las otras ideas relacionadas con la primera. Pero aunque he desdeñado cualquier ventaja que pudiera haber obtenido de este asunto, al explicar las relaciones entre las ideas, me temo que aquí deba recurrir a él a fin de dar cuenta de los errores que surgen de esas relaciones. Por tanto, he de observar que, como la mente está dotada del poder de suscitar cualquier idea que le plazca, siempre que dirija los espíritus animales hacia la región del cerebro en que aquélla está ubicada, dichos espíritus despiertan infaliblemente una idea cuando aciertan con precisión la huella correspondiente e indagan la célula que corresponde a esa idea. Pero como su movimiento rara vez es directo y se desvía naturalmente a un lado o a otro, el impulso da con las

huellas contiguas y presenta otras ideas relacionadas en lugar de la que la mente deseaba primitivamente examinar.[35]

Según este texto, las ideas dejan "huellas en el cerebro", es decir, parecen ocupar un lugar en él. Pero podemos suponer que se trata de una simple correspondencia que no afecta el carácter *mental* de la idea.

Kemp Smith atribuye este tipo de afirmaciones de Hume al "punto de vista ingenuamente realista que caracteriza las primeras partes del Libro I del *Tratado*." Puede decirse que paulatinamente va modificando su posición, a fin de distinguir con mayor precisión lo mental de lo material. Como ejemplo de esa modificación, Kemp Smith señala los cambios que Hume introduce respecto de la relación de contigüidad: "Primero sustituye 'precedente y contiguo', 'relación de prioridad y contigüidad' por 'contigüidad en el tiempo o en el espacio'. La contigüidad, en adelante, se limita a la contigüidad en el espacio. Aquí se reconoce que es más adecuado describir la contigüidad en el tiempo simplemente como 'secuencia', en el sentido, por supuesto, de 'secuencia inmediata'".[36] Observa finalmente que en la *Investigación* "la contigüidad objetiva en el espacio es totalmente eliminada en su consideración de la causalidad".[37] La contigüidad entre las ideas y sus huellas en el cerebro parece una hipótesis desechable, sobre todo teniendo en cuenta que más adelante Hume rechaza la "conjunción local" de la mente y el cuerpo, fundándose en una máxima que formula al respecto: "Después de esto, no causará sorpresa que enuncie una máxima condenada por diversos metafísicos y considerada contraria a los principios más ciertos de la razón humana. Esta máxima sostiene que un objeto *puede*

35 D. Hume, *Tratado*, pp. 114-115.
36 N. Kemp Smith, *op. cit.*, p. 36.
37 *Ibidem*, *loc. cit.*

existir y sin embargo no estar en ninguna parte, y yo afirmo que esto no sólo es posible sino que la mayor parte de los seres existen y deben existir en esta forma".[38]

Parecería que, si bien estas modificaciones en la teoría de Hume permiten dar cuenta con mayor facilidad de todo tipo de inferencias causales, sean éstas de naturaleza moral o física, o aun de aquellos casos en que unas y otras "forman una sola cadena de argumentos",[39] no logran sin embargo dar cuenta satisfactoriamente de la *interacción mente-cuerpo*. Considerando que Hume no admite una acción real de las causas, sean naturales o morales, que pueda ser percibida por el observador, esta consecuencia parece inevitable.

El problema de la interacción mente-cuerpo es tratado por Hume casi exclusivamente en relación con la voluntad, particularmente la influencia que ésta parece ejercer sobre los órganos del cuerpo. Ella "sólo puede ser conocida por la experiencia, como todo otro hecho natural, y nunca puede ser prevista a partir de una aparente energía o poder en la causa, que la conecta con su efecto, y hace que una sea consecuencia infalible de la otra. El movimiento de nuestro cuerpo sigue a la orden de la voluntad. De esto somos conscientes en todo momento. Pero los medios por los cuales esto se efectúa, la energía por la cual la voluntad realiza una operación tan extraordinaria, de esto estamos tan lejos de ser inmediatamente conscientes, que escapará siempre a nuestra investigación más cuidadosa."[40] A continuación Hume se apoya en distintos argumentos, como el de que no podemos mover a voluntad todo sino sólo algunos órganos de nuestro cuerpo y que la anatomía nos enseña que para lograr el movimiento deseado deben producirse otros intermedios, de nervios,

38 D. Hume, *Treatise*, p. 345.
39 D. Hume, *Enquiries*, p. 70.
40 D. Hume, *Enquiries*, p. 52.

músculos y quizás de partículas tan pequeñas que nos son absolutamente desconocidas.

Quiero terminar con dos citas de Nicholas Capaldi, en una de las cuales señala en Hume la misma dificultad que en la mayor parte de los filósofos modernos, mientras que en la segunda nos muestra por qué Hume, como hemos tratado de demostrar en este trabajo, al limitarse a la explicación de nuestras inferencias y de nuestras creencias comunes, nos ofrece lo más próximo a una filosofía sólo moderadamente escéptica respecto de los problemas tratados:

1. ...podemos decir que Hume comparte con la mayoría de los filósofos modernos, de Descartes a Kant, el punto de vista de que hay un mundo físico externo, que hay un mundo interno "mental", que hay una interacción entre ambos y –el gran misterio no resuelto de la ciencia y la filosofía– que nadie ha explicado aún satisfactoriamente la interacción.[41]
2 Ahora bien, no incumbe a Hume resolver el problema físico de la interacción mente-cuerpo (y reiteradamente admite que no puede hacerlo) a fin de creer en un mundo físico externo que influye causalmente sobre su mente y las mentes de los demás. Hume nunca pretendió refutar la creencia en la existencia de ese mundo; más bien, rebatió las interpretaciones que otros filósofos daban acerca de esa creencia. Hume, de hecho, dio su propia explicación. Puesto que cree en ese mundo, y puesto que tiene una interpretación instrumental de dicha creencia, es libre de especular sobre cómo opera ese mundo y cómo puede haber a la vez explicaciones mecánico-fisiológicas de la actividad del *cerebro* y explicaciones del tipo selección natural del uso que hace nuestra *mente* de ciertas reglas de razonamiento. Hume puede hablar legítimamen-

41 N. Capaldi, *op. cit.*, p. 8.

te de causas secretas, tanto en el sentido de aquéllas que no han sido descubiertas aún como en el sentido de mecanismos tan pequeños que puede que nunca logremos descubrir.[42]

42 *Ibidem*, p .9.

EL NATURALISMO HUMEANO Y EL PROBLEMA MENTE-CUERPO

Carmen Silva*

Introducción

El objetivo de este trabajo, no es hacer un análisis pormenorizado de los argumentos humeanos sobre el tema mente-cuerpo, mi intención es más bien, sugerir algunos antecedentes de la propuesta humeana que nos ayuden a comprender porque este tema tan clásico y fundamental en los siglos XVII y XVIII tiene el tratamiento que tiene en Hume, el cual parece salirse totalmente de los canones según los cuales se abordaban estas cuestiones en su época. Para ello dividí el trabajo en dos partes. En la primera hablaré de cierta actitud filosófica propia del empirismo británico (que tiene su origen en el *Ensayo* de Locke), la cual considero un antecedente importante de las propuestas humeanas sobre la mente y el cuerpo. En la segunda expondré, sin problematizarlas, las tesis centrales humeanas sobre nuestra creencia en el mundo externo.[1] Por últi-

* Facultad de Filosofía y Letras, UNAM

1 Me centraré en nuestra creencia en el mundo externo, aunque el tema original sea mente-cuerpo; por dos razones: la primera, es porque a grandes rasgos el análisis que hace Hume en el *Tratado* tanto de la idea del yo como de nuestra creencia en el mundo externo son muy semejantes, sobre todo en un aspecto que quiero resaltar del empirismo británico, que es la tendencia que tiene; por un lado, de eliminar el elemento sustancial de sus explicaciones, y por el otro, el de reducir todo a una teoría de las ideas. La segunda, es una cuestión práctica, es decir, abreviar el trabajo.

Todas las citas y referencias del *Tratado* son de David Hume, *A Treatise of Human Nature* (ed. L.A. Selbby-Bigge. Texto revisado y notas de P.H. Nid-

mo, esta segunda parte tendrá como marco fundamental (de mi lectura) una interpretación naturalista de la filosofía de Hume.

I El empirismo británico y la metafísica (sustancialista)

El problema mente-cuerpo no parece un problema propio del empirismo británico

Considero que tanto en Locke como en Hume esté no es un tema fácil de abordar. El fundamento de mi intuición es que, en estos autores, el tema, tal y como se plantea, siguiendo la tradición cartesiana (racionalista), no satisfará nuestras espectativas. De ellos no se puede decir que elaboren una teoría de la mente y una del cuerpo y, menos aún, una sobre la relación entre ambas, pues me parece que inicialmente no elaboran una metafísica, en el caso de Locke es claro que la presupone como veremos a continuación. De esto último, es de lo que pretendo hablar en las páginas siguientes.

Locke, de hecho, no elaboró una teoría *propia* de la materia; lo que hizo[2] fue adoptar dos teorías de ámbitos diferentes y de autores diferentes; en el caso de la materia, por ejemplo, la metafísica cartesiana y la teoría corpuscular de Boyle propia, esta última, de la filosofía natural; ambas propuestas aparecen en diferentes lugares en el *Ensayo*. La mayoría de los estudiosos lockeanos está de acuerdo en sostener que Locke es un dualista

---

ditch), Oxford University Press, Oxford, 1978. Todas las citas y referencias del *Ensayo* son de John Locke, *An Essay Concerning Human Understanding* (ed. P.H. Nidditch), Oxford University Press, Oxford, 1972.

2 Como sugerí en mi artículo "El concepto de materia en Locke" en *El concepto de materia* Colofón, México, 1992. Locke no elaboró una teoría *propia* de la materia.

y un representacionista. Sin embargo, Locke jamás fundamenta este dualismo; es uno de sus supuestos pues, su preocupación central es sobre el origen de las ideas. Por otra parte, la mayoría de los estudiosos lockeanos está de acuerdo en que Locke es un dualista, y que tiene una teoría propia de la identidad personal. Si bien se puede hablar de una teoría de la identidad personal en Locke, al igual que en Hume, la peculiaridad de ambas frente a la cartesiana racionalista es que, en aquellas, la base de la identidad no se funda en un sustrato o en una sustancia, como en la cartesiana (racionalista). En el caso de la de Locke, la identidad se da por la memoria y en el de Hume en otra *facultad*[3] del entendimiento, a saber, la imaginación.

En resumen, no perdamos de vista que si bien racionalistas y empiristas británicos pertenecen al mismo periodo y el problema de la mente y el cuerpo es uno de los centrales y característicos de la época, los empiristas, no hablaron sobre él en la misma forma en que lo hicieron los cartesianos racionalistas. Por momentos me atrevería a afirmar que aquellos no pretenden tanto dar una solución diferente al problema del dualismo (como es el caso del monismo de Spinoza, o el paralelismo o armonía prestablecida de Leibniz) sino que más bien, pretenden no intentar resolver los problemas por la vía de la metafísica (sustancialista).[4]

3 El énfasis que quiero poner aquí, es el de que para Locke y Hume, el criterio de la identidad personal es una facultad del entendimiento, en vez de uno metafísico. Con la intención de trazar la diferencia entre estas teorías de la identidad y la de Descartes. Ahora bien, alguien podría objetar esto diciendo que para Descartes el *pensar*, que también es una facultad del entendimiento (la cual abarcaría a toda actividad del mismo, como desear, creer, recordar, imaginar, querer, etc.), es fundamental para la identidad, pues Descartes, cuando formula la pregunta qué soy, responde "soy una cosa que piensa". Mi respuesta a esto es que aun así, permanecen las diferencias, pues para Descartes el *pensar* es un atributo esencial de la sustancia mental.

4 Alguien podría preguntar: ¿y el espiritualismo de Berkeley? ¿No es un monismo, no es una sustancia espiritual o mental de la que habla el filósofo irlan-

En particular, sostendré que Locke y Hume tendrán como objetivo central el análisis del entendimiento humano[5] y, como el objeto primordial de éste son las ideas, el desarrollo de su filosofía se centrará de manera preferente en la teoría de las ideas. A partir de un planteamiento semejante de éste surgen los psicologismos, funcionalismos, naturalismos, en vez de los dualismos, monismos, realismos, etc.; conceptos, estos últimos, que se aplican a diferentes teorías propias de la reflexión metafísica y no así los primeros.

Teoría de las ideas *vs.* explicaciones metafísicas

En esta segunda sección de la primera parte del trabajo, desarrollaré mis intuiciones en relación con la tendencia de los empiristas británicos de eliminar el elemento sustantivo o sustancial de sus explicaciones, pues me resulta por demás interesante, y servirá, supongo, para aclarar las propuestas de Hume sobre el tema.

---

dés? Sí, pero la diferencia con los autores arriba mencionados es que en la filosofía berkeleyana, el problema central no es de carácter metafísico, sino epistemológico, es decir, al igual que en Locke, el problema a explicar es el origen de las ideas. No olvidemos que Berkeley, al eliminar a la sustancia material de las explicaciones, lo que trata de evitar es el escepticismo, la contradicción, el ateísmo y la ininteligibilidad que genera su postulación.

Aun cuando en Locke y en Berkeley encontramos dos posiciones metafísicas, dualista la del primero y monista la del segundo, supuestas o argumentadas, mi intención en este trabajo es la de ofrecer algunos elementos de interpretación que permitan percibir lo que llamo *actitud* propia del empirismo británico, la cual es una tendencia que tienen estos pensadores a darle un mayor peso a las cuestiones epistemológicas o de la teoría de las ideas, además de ir eliminando de las mismas el elemento sustancial, en otras palabras, la metafísica.

5 Éste es el objeto de estudio del Libro I del *Tratado*, donde Hume intenta explicar el origen de nuestra creencia en el mundo externo.

Considero que sólo con esta introducción general es posible entender por qué estos autores dan el tipo de respuestas que dan, en el caso que nos ocupa. La crítica por parte de Berkeley y Hume al sustancialismo la origina Locke [6] con su crítica al concepto de sustrato del que Locke afirma que es un "no sé qué". Aquí quiero sugerir que esta crítica al sustancialismo es central y definitoria del empirismo británico.

Pienso que Locke no intentó enfrentarse conscientemente al dualismo cartesiano, pero su crítica al sustrato fue el punto de partida de las tesis posteriores del empirismo británico.

Si la tendencia del empirismo británico, a partir del "no sé qué" de Locke, es la de eliminar el sustancialismo cartesiano, cuando estos autores hablen de la mente y el cuerpo, por lo menos, lo harán de una manera muy diferente a la de Descartes que tiene como fundamento de esta reflexión la metafísica sustancialista. Si bien los británicos pretenden reducir la influencia cartesiana al intentar eliminar la sustancia de sus tesis, siguen siendo cartesianos, al mantener las ideas dentro de aquellas. Llegan a tal extremo en esto, que su intención parece ser la de eliminar la metafísica como forma de explicación y abogar por la teoría de las ideas. Debido a ello veremos cómo, al menos Locke y Hume, no nos hablarán de mente y cuerpo al estilo cartesiano (racionalista), i.e. metafísico, sino desde la perspectiva de las ideas. De esto, Hume es un magnífico ejemplo, pues no hablará de cuerpo, mente, mundo externo, sino de las ideas del yo, del mundo externo, etc. Locke, que parece

6 Coincido con José A. Robles en que el antecedente inmediato de las propuestas humeanas en muchas cuestiones es Berkeley. Robles afirma que por lo menos hay tres elementos en los cuales se ve clara la relación entre ambos, los elementos son los siguientes: 1) la teoría del haz de percepciones (mejor conocida como la *bundle theory*) 2) la propuesta de que no existe la relación causal.3) Los dos autores coinciden en la propuesta que considero a su vez propia del empirismo británico: eliminar el elemento sustantivo o sustancial de sus explicaciones.

estar más lejos del ideal que él mismo originó, también cumple con esta característica al hablar, por ejemplo, (de manera preferente) de idea de sustancia y no de la sustancia. En resumen, el movimiento empirista es tal que tiene como intención ir eliminando la sustancia de sus explicaciones en favor de las ideas; en otras palabras, ir eliminando la metafísica como forma de explicación, por la teoría de las ideas y por la filosofía natural (en el caso de Hume), como veremos posteriormente.

Ejemplos de esto son los siguientes: Locke, en la *Introducción* al *Ensayo*, afirma explícitamente que su objetivo principal es analizar qué sea el entendimiento humano desde la perspectiva del conocimiento; por ello indaga cuál es su alcance y los límites y cuál es el origen de los contenidos mentales o ideas. No se pregunta qué sea la mente y qué el mundo externo, ni elabora una teoría metafísica, como sí lo hizo Descartes, con la finalidad de responder las preguntas arriba mencionadas. Con esto, Locke genera una ruptura con la tradición cartesiana (racionalista), al plantear una teoría del conocimiento en la cual el problema central es el de la investigación del origen de los contenidos mentales. En otras palabras, todo se centrará en un análisis de la naturaleza del entendimiento, siendo las ideas su objeto primordial, el estudio del entendimiento nos lleva directamente al de las ideas. Esta teoría de las ideas, tiene como supuesto la existencia tanto de la mente como la del cuerpo o mundo externo además de la relación entre ambos. Para apoyar todo esto cito el siguiente pasaje de la Introducción del *Ensayo*:

> Puesto que es mi intención *investigar* los origenes, alcance y certidumbre del *entendimiento* humano, junto con los fundamentos y grados de las creencias, opiniones y sentimientos, *no entraré aquí en consideraciones físicas de la mente, ni me ocuparé de examinar en qué pueda consistir su esencia*, o por qué alteraciones de nuestros espíritus o de nuestros cuerpos llegamos a tener sensaciones en

nuestros órganos, o ideas en nuestros entendimientos, ni tampoco si en su formación esas ideas dependen, o no, algunas o todas, de la materia. (Introducción, *Ensayo* el subrayado es mío)

Locke en este pasaje del *Ensayo* ofrece dos cuestiones importantes en relación al tema: 1) No hablará de la mente ni a nivel físico ni a nivel metafísico. 2) No investigará la relación entre la mente y el mundo externo. La existencia de la mente y su naturaleza además de la relación con la materia son dos de los supuestos fundamentales de su teoría de las ideas como dije anteriormente.

Dado que las ideas es de lo único que podemos hablar, pues son lo único que realmente conocemos, los empiristas británicos sólo hablarán de ellas; Locke, por ejemplo, así lo hace; Berkeley hablará de éstas y de los espíritus (sustancias). En el caso de Hume, claramente se ve que de lo que habla, por lo menos en todo el Libro I del *Tratado*, es de ideas, no de cosas ni de mentes, sino de ideas, de diversos tipos de ellas, de memoria, de sensación, de la idea del yo, la de causalidad, de la imaginación. A tal grado quiere Hume resolver todo vía las ideas, que las identifica con las facultades mismas que las explican. Así encontramos que en el caso de este autor, en vez de elaborar una metafísica para hablar del mundo externo, elaborará una teoría de la creencia que, a su vez, forma parte de la de las ideas para explicar por qué creemos en un mundo externo; ésta es su preocupación fundamental, *la creencia y no la existencia.*

Otra cuestión importante, que también ejemplifica de una manera clara este ideísmo británico, es la caracterización del objeto de conocimiento que Locke propone en el libro IV del *Ensayo*,[7] que dice:

7 Berkeley, según Robles, parte de este "no se qué" lockeano, para quitarle sustancialidad a la materia. Su crítica va claramente dirigida contra el repre-

Puesto que la mente, en todos sus pensamientos y razonamientos, no tiene otro objeto inmediato que no sean sus propias ideas, las cuales sólo ella contempla o puede contemplar, es evidente que nuestro conocimiento se ocupa únicamente de esas ideas. (*Ensayo* IV,I,1)

Y la siguiente definición de conocimiento:

Me parece, pues, que el conocimiento no es sino *la percepción de la conexión y acuerdo, o del desacuerdo y rechazo entre cualesquiera de nuestras ideas.* (*Ensayo* IV,I,2).

Para cerrar esta sección, quiero agregar que el propósito fundamental de la filosofía de Hume, por lo menos del *Tratado,* es elaborar una ciencia de la naturaleza humana, debido a lo cual veremos que prefiere optar por el camino de la ciencia o filosofía natural, que por el de la metafísica. Es éste un elemento más que muestra que el análisis del problema mente-cuerpo no se dará desde una perspectiva sustancialista, sino de una muy diferente. La respuesta humeana será una que llamaremos naturalista, en la cual el concepto clave explicativo no serán las sustancias, sino las ideas y las facultades.[8] Como se mencionó líneas atrás.

---

sentacionismo, de Locke y esto significa, a su vez, un avance frente a la filosofía lockeana, en el sentido de reducción de entidades. En realidad, la propuesta berkeleyana consiste en afirmar (1) todo lo que percibo son ideas y (2) no existe algo que se oculte a mi entendimiento. Con la teoría berkeleyana, al eliminarse el dualismo, se elimina el representacionismo, y con ello el velo de la percepción que lo acompaña. Es tal la identificación de Berkeley entre las ideas y lo que existe, que llega a identificarlas con las cosas. De la propuesta berkeleyana de eliminar "el no sé qué" lockeano, surge la de Hume, en el sentido de que él, a diferencia de Locke, ya no se preguntará a qué se refieren las ideas, o si existe algo más que ellas. En resumen, Hume ya no pretende salir de su propia mente, no pretende pasar de las ideas a las cosas, sabe que esto no tiene sentido.

8 En Hume parecería haber una especie de sustancialismo implícito, al estilo berkeleyano, en el sentido de que el punto de partida es el sujeto que tiene una

II Hume y el problema mente-cuerpo

El naturalismo humeano y el problema mente-cuerpo

No entraré aquí en la discusión sobre las diferentes interpretaciones de la filosofía humeana, sino que partiré de la concepción naturalista para leer las tesis humcanas sobre la mente y el cuerpo. Esta concepción considera que el proyecto fundamental de la filosofía de Hume, por lo menos del *Tratado*, es elaborar una ciencia de la naturaleza humana; no defenderé aquí esta interpretación, pues ya lo hice en otros trabajos.[9]

---

diversidad de percepciones en el entendimiento. Pero esto no es así.Cuando Hume analiza la idea del yo y propone la metáfora de que el yo es como un teatro "La mente es una suerte de teatro, en el que sucesivamente hacen su aparición diversas percepciones..." lineas abajo afirma "La comparación con el teatro no debe engañarnos. Son únicamente las percepciones sucesivas las que constituyen la mente; no tenemos tampoco la más remota noción del sitio donde se representan estas escenas ni de los materiales de que esta compuesto." (Los dos pasajes son de *Tratado* Libro I,Parte IV, sección vi)

9 Para quiénes no conozcan los trabajos anteriores, resumiré las ideas centrales de esta propuesta interpretativa.

El naturalismo humeano, tiene como meta elaborar una ciencia de la naturaleza humana. Ahora bien, ¿qué es tal ciencia? 1) El objetivo central es el hombre, qué sea y cuál su naturaleza: "...no somos tan sólo seres que razonamos, sino también uno de los objetos sobre los que razonamos."(*Tratado* Introducción, p.xix). Es curioso ver cómo Hume sigue muy de cerca a Locke, pero con un matiz muy importante: en el caso de Locke vemos que el objetivo del *Ensayo* es estudiar, analizar qué es el entendimiento humano, pues éste, según Locke nos distingue de los animales o del resto de las criaturas. Para Locke, este elemento es el que nos definiría como seres humanos, cito: "Puesto que el *Entendimiento* es lo que sitúa al hombre por encima del resto de los seres sensibles...es ciertamente, un asunto, ..., que amerita ser investigado." (*Ensayo* I,I,1). En el, caso de Hume, el objeto de estudio no es sólo el entendimiento porque éste nos distingue del resto de los seres sino que es la naturaleza humana en su totalidad, lo que sea el ser hombre.

La concepción naturalista de la mente y el cuerpo

Esta concepción se separa totalmente de la sustancialista, pues su punto de partida es el siguiente:

> Podemos muy bien preguntarnos *qué causas nos inducen a creer en la existencia de los cuerpos*, pero es inútil que nos preguntemos *si hay o no hay cuerpos*. Esto es algo que debemos dar por supuesto en todos nuestros razonamientos. (*Tratado* I,IV,ii)

Con este pasaje inicial del capítulo correspondiente a su análisis de la idea de mundo externo, Hume establece de manera muy clara cuál es su objetivo, a saber, no hará metafísica, sino que intentará averiguar cuales son las causas de nuestra creencia en la existencia de los cuerpos.

---

2) Hume la llama ciencia, pues pretende abordar las cuestiones humanas desde la perspectiva científica, elaborando una ciencia de lo humano, esto es, de las cuestiones humanas (o morales como las llamaban en su época) en vez de elaborar un sistema filosófico, que era y siguió siendo la forma más generalizada como los filósofos resolvían la problemática humana. Un rasgo fundamental de esta propuesta humeana, y por la cual la llama ciencia, es porque pretende aplicar a la investigación de la naturaleza humana el método experimental de razonamiento que Newton aplicó con tan buenos resultados a la naturaleza física.

Siguiendo a Stroud, podemos afirmar que "Su tema es todo lo humano" (Barry Shoud, *Hume*, UNAM, México, 1986, p.13). Ahora bien, otro elemento interesante de Stroud (*cf. op. cit* p.13) es el que señala que la preocupación por lo humano es tan antigua como la naturaleza humana misma; lo interesante de Hume es que propone una forma de analizarla que en sus días era totalmente novedosa, por ejemplo pretende hacer una ciencia de lo humano. Que significa lo siguiente: 1) No aceptará en su investigación ningún elemento *a priori* 2) Esto, porque aplicará el método experimental de razonamiento. 3) Pretende buscar los principios que rigen esa naturaleza humana. 4) Intenta dar una explicación general de la misma. "...Hume quiere una teoría completamente general de la naturaleza humana, que explique porqué los seres humanos actúan, piensan, perciben y sienten como lo hacen." (Barry Shoud, *op.cit*, p.14-15).

Esta propuesta humeana, a mi modo de ver, representa una doble alternativa a la filosofía tradicional: por un lado, es un intento por separarse de la metafísica sustancialista, en el sentido de que aquí, al ofrecer la explicación de nuestra creencia tanto en el yo, como en el mundo externo, Hume no se apoyará en el concepto de sustancia.

Por otro lado, es una alternativa a la concepción epistemológica tradicional en dos niveles: a) no intentará fundamentar a la teoría del conocimiento en la metafísica y b) no se preguntará por el fundamento (la verdad) de las creencias, sino sólo de su origen. Sobre esto último, B. Stroud dice:

> Los filósofos se han interesado especialmente por la acreditación epistemológica de lo que llaman nuestra creencia en el "mundo externo", pero *Hume no se preocupa en absoluto por la verdad o la racionalidad de esta creencia.* No comienza preguntándose si hay cuerpos o no, o si sabemos o creemos razonablemente que los hay. Como científico del hombre, se pregunta por qué tenemos la creencia o cómo llegamos a tenerla. "¿Qué causas nos inducen a creer en la existencia de los cuerpos?"(*T.* p.187). El hombre constituye su interés primordial, no los cuerpos.[10]

La siguiente cuestión que quiero comentar aquí sobre el análisis humeano tanto de la idea del mundo externo como la del yo, es que, el análisis y la conclusión de ambas es muy semejante al que se da en el caso de la idea de causalidad. Puede haber varias razones para la coincidencia entre el análisis y las conclusiones relativas a estas tres ideas en el *Tratado.* Debido a que Hume pretende elaborar una ciencia de la naturaleza humana y, parte de lo que entiende por *ciencia*, es encontrar los principios

10 *op. cit.*, 1986, p. 144 (el subrayado es mío).

que la rigen para así poder dar una explicación general, el tema que nos ocupa el de la mente y el cuerpo o el del mundo externo y el yo, se explicará, según la filosofía humeana con base a los principios de asociación, lo cual da una uniformidad a estas creencias. En otras palabras, la explicación que ofrece Hume tanto en un caso como en otro será similar, pues cualquiera de los tres es sólo una instancia de un único fenómeno humano *la creencia*.[11]

En el análisis de Hume de la creencia, sea ésta la que sea, (es decir, la de causalidad, la del yo o la del mundo externo) aparecen cosas sorprendentes como, por ejemplo, su concepción de la creencia como algo inevitable (*cf. Tratado* I,IV,i),y sobre ello tiene un pasaje muy claro en relación a nuestra creencia en el mundo externo; que es el siguiente:

> La naturaleza no le ha dejado a este respecto opinión alguna [habla del escéptico] pensando sin duda que se trataba de un asunto demasiado importante para confiarlo a nuestros inseguros razonamientos y especulaciones. (*Tratado* I,IV,ii)

En este trabajo, no se estudiará en detalle ninguna de las propuestas humeanas sobre las ideas arriba mencionadas.[12] Otro de los rasgos de la explicación humeana de estas ideas es lo que llamo una especie de ideísmo, que consiste en llevar a su extremo la teoría cartesiana de las ideas, pero en el caso de nuestro autor, sin el representacionismo propio del dualismo cartesiano; por ejemplo, al analizar el origen, esto es, la causa de nuestra idea de mundo externo, Hume afirma lo siguiente:

11 Hume las llama creencias, a diferencia de conocimiento porque, en todas ellas, el fundamento es la experiencia, lo cual nos lleva al terreno de lo probable y no de lo necesario.

12 Hume y sus intérpretes las llaman ideas, pero el problema es que también son creencias, inferencias, suposiciones, etcétera. Con esto lo que quiero señalar es que son algo más que *ideas*, por lo menos en los propios términos humeanos.

...nada hay realmente presente a la mente sino sus propias percepciones, sean impresiones o ideas, ... (*Tratado* I,II,vi)[13]

Es tan sorprendente lo consecuente que es Hume con la Teoría de las ideas, en el sentido de que éstas son los únicos elementos que conocemos y por ello de los que podemos hablar que, cuando intenta explicar el origen de nuestra creencia en el mundo externo, les atribuye a las propias impresiones los rasgos que antes eran propios de la sustancia. Los elementos a los que estoy haciendo mención son: la constancia y la coherencia, que en el caso de la explicación humeana son rasgos que encontramos en las impresiones mismas, rasgos que posteriormente tomará la imaginación y los convertirá en una creencia o suposición.(Acerca de esto *cf.Tratado* I,IV,ii)

En el caso de la idea de mundo externo, Hume inicia su investigación con la misma pregunta que en el caso de la de causalidad, es decir, ¿cuál es el origen de esta idea, la sensación, la razón o la imaginación? La respuesta será: la imaginación. (Respuesta que será la misma (explicación) para las tres ideas, sea ésta sobre la casualidad, el yo o el mundo externo). Como vemos, esta facultad juega un papel fundamental en las explicaciones humeanas; esto se debe, en gran medida, a que esta facultad, y sólo ésta, es la que se rige por los principios de asociación, que son en realidad la verdadera explicación de las creencias. Nunca

13 Véanse, además,los siguientes pasajes sobre el yo: "...cuando más íntimamente penetro en lo que llamo *yo mismo*, tropiezo siempre con una u otra percepción particular, de calor o frío, luz o sombra, amor u odio, dolor o placer. Nunca puedo aprehenderme *yo mismo* sin una percepción, y nunca puedo observar más que la percepción."(*Tratado* I,IV,vi) O bien,cuando nos preguntamos qué es el yo, lo único que encontramos debido a lo que se describe en el pasaje anterior es: "...un haz o colección de percepciones diferentes, que se suceden unas a otras con una rapidez inconcebible, y que están en flujo y movimiento perpetuos"(*Ibidem*)

se encontrará la explicación en la sensación o en la experiencia, pues si bien ésta es necesaria no es suficiente. Pues, de ser la causa la sensación deberíamos de tener una impresión correspondiente y eso es lo que nunca se encuentra; por ejemplo, nadie tiene una impresión de un yo, y luego o junto, todas las demás impresiones que la acompañan; tampoco tenemos una de causalidad o de conexión necesaria.

En realidad, sólo tenemos el conjunto de impresiones que junto con el hábito y los principios de asociación, explican la creencia. En todos estos casos se requiere de algo más que explique ese elemento diferente de las percepciones que es la creencia y que Hume a veces llama ficciones de la imaginación o suposición. La razón en esta filosofía humeana, no juega ninguna papel relevante, al contrario. Tradicionalmente se pensaba que el fundamento de este tipo de creencias estaba en la razón, pero uno de los rasgos de la filosofía humeana, además de los que ya he mencionado como el de ser una alternativa al sustancialismo y a la teoría del conocimiento tradicional, también lo es su concepción de la naturaleza del hombre pues, para Hume, la esencia del hombre no está en su razón. Por ello, no será esta facultad la que funcione como fundamento o explicación de las cuestiones humanas.

Así pues, el problema mente-cuerpo desde la perspectiva emprirista, muestra una serie de características diferentes a las del resto de los filósofos que conforman el periodo de la filosofía moderna. La característica que considero central es:

1 La tendencia de convertir problemas metafísicos tradicionales en problemas de teoría del conocimiento o de teoría de las ideas.
2 En el caso particular de Hume, vemos lo siguiente:

a) En su tratamiento del problema mente- cuerpo Hume se separa de la metafísica sustancialista, al no preguntarse por la existencia sino por la creencia.

b) Su objetivo es, dar una explicación del origen de nuestras creencias, por ello, no elabora un a metafísica sino una teoría de las ideas.
c) El análisis que hace de las ideas es estrictamente genético, por ello no se preocupa por la acreditación epistemológica de las mismas.
d) Por último, un elemento novedoso y peculiar humeano es la crítica de la razón humana como la facultad primordial de conocimiento.[14]

14 Agradezco a los miembros del Seminario sus comentarios a este trabajo, en especial a Laura Benítez y a José Antonio Robles.

LA UNIDAD TRASCENDENTAL DE LA APERCEPCIÓN

Q. Cassam*

- I -

Wittgenstein sostenía que el uso de 'yo' en reportes de experiencia inmediata es superfluo. Puesto que la experiencia de sentir dolor no es la experiencia de qué una persona "yo" tiene algo, el uso de 'yo' para representar la experiencia no es "esencial a los hechos".[1] En contraste, el punto de vista de Kant era que para que uno estuviera en absoluto consciente de "representaciones", debe ser posible que uno se percate de ellas como perteneciéndoles a uno mismo. Así, lejos de ser superflua, la representación "yo", que Kant llama 'apercepción pura', es "el correlato de todas nuestras representaciones en la medida en que haya de ser posible en absoluto que estemos conscientes de ellas".[2]

El trasfondo de la teoría kantiana de la auto-conciencia o "apercepción" es su aceptación del punto de vista de Hume de que un yo perdurable no puede encontrarse como objeto del sentido interno.[3] Sin embargo, Kant no concluye que no hay tal cosa como la auto-conciencia aparte de la conciencia de un flujo constantemente cambiante de apariencias internas. ¿Cuál era entonces la explicación que Kant daba de la auto-conciencia? Esta

* Wadham College. University of Oxford. [Trad. de A. Tomasini y N. Lara]

1 L. Wittgenstein. *Philosophical Remarks*. Basil Blackwell. Oxford (1975), p. 88.

2 I. Kant. *Critique of Pure Reason*. MacMillan. London (1929), A123.

3 Véase D. Hume, *A Treatise of Human Nature*, Libro I, Parte IV, Sec. VI, y la discusión de Kant en A107.

es la cuestión que nos ocupará aquí. El reto que tiene que enfrentar todo lector de Kant es el de dotar de sentido a lo que él dice de manera que ello sea compatible no sólo con su aceptación del resultado de Hume, sino también con su crítica de la doctrina racional del alma. Yo argumentaré que el intento de Kant por desarrollar una explicación que no fuera ni humeana ni cartesiana en última instancia fracasa, pero que cualquier teoría aceptable de la auto-conciencia debe acomodar algunos aspectos de su punto de vista.

Para entender debidamente la posición de Kant es esencial reconocer y comprender su compromiso con las siguientes tesis:

1) Debe distinguirse nítidamente entre la auto-conciencia y el auto-conocimiento.
2) Uno está consciente de sí mismo como del sujeto de sus propias representaciones *porque* uno puede adscribírselas a uno mismo.
3) Ser auto-consciente es ser consciente de sus representaciones como representaciones de "objetos".

La tesis (1) es una consecuencia directa del idealismo trascendental de Kant. Puesto que el idealismo trascendental no puede permitir el conocimiento del yo como es en sí mismo, el precio de hacer a la auto-conciencia dependiente del auto-conocimiento en este sentido sería hacer problemática a la auto-conciencia misma. La nítida solución de Kant consiste en insistir en que "la conciencia del yo está ...muy lejos de ser un conocimiento del yo".[4] Es importante ver que no es ésta una tesis que deba toda su plausibilidad al idealismo trascendental. Neo-kantianos recientes han tendido a subestimar la fuerza de la tesis de Kant. Según P.F.Strawson, por ejemplo, la auto-conciencia involucra el ser capaz de auto-adscribirse experiencias y ello a su vez requiere que uno tenga

4 I. Kant. *Ibid*, B158.

una aprehensión de los criterios empíricos de singularidad e identidad para los sujetos de experiencia. Aunque dichos criterios de hecho no sean empleados en la auto-adscripción de experiencias, los vínculos entre ellos y las auto-adscripciones que uno hace "en la práctica no se cancelan".[5] De la misma manera, Gareth Evans sostuvo que la auto-conciencia debe "ocurrir en el contexto de ciertas clases de conocimiento y de comprensión",[6] incluyendo, según parece, el conocimiento de lo que uno es. La propia posición de Kant sobre estas cuestiones es preferible, al grado de que reconoce la implausibilidad de negarle auto-conciencia al sujeto sólo sobre la base de que tiene creencias falsas acerca de su propia naturaleza. Esto sigue siendo implausible inclusive si la naturaleza del yo no es incognoscible.

La significación de la tesis (2) radica, en parte, en el modo en que invierte lo que podría parecer el orden natural de la explicación. El punto de vista del sentido común sería que es porque estoy consciente de mí mismo como del sujeto de representaciones dadas que estoy autorizado a llamarlas 'mis representaciones'. Si mis auto-adscripciones de representaciones no están basadas en una conciencia independiente de mí mismo como su sujeto, entonces, como observa Wittgenstein, el uso de 'yo' en reportes de experiencia parece estar injustificado. Kant pone de cabeza este argumento cuando escribe: "Estoy consciente del yo como idéntico respecto a la multiplicidad de las representaciones que me son dadas en una intuición, porque las llamo a todas y a cada una de ellas mis representaciones".[7] Si soy auto-consciente en virtud de que llamo mías a ciertas representaciones, entonces no se plantea que el uso de 'yo' vaya más allá de lo que es obtenible en la conciencia.

5 P.F. Strawson. *The Bounds of Sense*. Methuen. London (1966), p. 165.
6 G. Evans. *The Varieties of Reference*. Oxford University Press (1982), p. 232.
7 I. Kant. *Op. cit.*, B135.

La tesis (3) contiene la clave para la "deducción trascendental del las categorías". La tesis es que el que las representaciones caigan bajo conceptos de objetos (las categorías) es no sólo necesario sino también suficiente para la auto-conciencia. En palabras de Kant: "La unidad trascendental de apercepción es esa unidad a través de la cual toda la multiplicidad dada en una intuición queda unida en el concepto del objeto".[8] La tesis de la suficiencia es casi con certeza indefendible, pero se ha argumentado que algo se puede hacer con la tesis de la necesidad. La sugerencia de Strawson, por ejemplo, es que no podemos "dotar de sentido a la noción de una conciencia única a la que se supone que pertenecen 'experiencias' sucesivas",[9] a menos de que las experiencias mismas construyan "un cuadro de un mundo objetivo unificado a través de las cuales las experiencias mismas constituyan colectivamente una ruta de experiencia única y subjetiva".[10] El que las experiencias construyan un cuadro así equivale a estar "unidas en el concepto del objeto". De ahí que esto último sea, como Kant sostenía, una condición necesaria para la unidad de apercepción.

Esta reconstrucción del argumento de Kant falla en dos puntos. Asume erróneamente que Kant estaba tratando de explicar la propiedad de las experiencias antes que la conciencia de ellas como las propias de uno. Strawson presupone esto porque él considera que para que una experiencia sea mía yo debo ser capaz de adscribírmela. En cambio, Kant esta dispuesto a conceder que representaciones de las que no estoy en absoluto consciente, y que por lo tanto no puedo áutoadscribirme, de todos modos podrían ser mías. En segundo lugar, puesto que el punto de vista de Strawson es que las experiencias en general no pueden

8 I. Kant. *Ibid*, B139.
9 P.F. Strawson. *Op.cit.*, p. 100.
10 P.F. Strawson. *Ibid*, p. 104.

ser "des-poseídas", se sigue que inclusive las impresiones desconectadas de una "experiencia pura de *sense-datum*" tiene que "pertenecer" a un sujeto.[11] Así, no está claro cómo el hecho de que tales impresiones no representen objetos podría cuestionar la noción de que hay una "conciencia idéntica" a la que pertenecen.[12]

En la siguiente sección, expondré el argumento del propio Kant en favor de la tesis (3). Dicho argumento no tiene éxito, porque está apoyado –por lo menos en parte– en la tesis (2). Se argumentará más abajo que la tesis (2) es inaceptable y que ello es fatal no sólo para la tesis (3), sino también para el todo de la teoría kantiana de la auto-conciencia. Hacia el final de la discusión quedará en pie sólo la tesis (1).

- *II* -

Una distinción explícita entre percepción y apercepción fue trazada por vez primera por Leibniz, quien escribió: "Es tan correcto distinguir entre percepción, que es un estado interno de la mónada y que representa a las cosas externas, y apercepción, *que es la conciencia* o el conocimiento reflexivo de este estado interno y que no está dado a todas las almas ni en todos los tiempos a la misma alma".[13] Kant recoge la terminología leibniziana, pero traza la distinción de un modo ligeramente diferente. El concede que hay *representaciones* inconscientes, pero

11 Strawson critica en el capítulo III de *Individuals* (Methuen, London, 1959) la idea que las experiencias pueden existir sin ser poseídas.

12 Para una más amplia discusión del argumento de Strawson, véase mi "La Auto-conciencia Trascendental", por aparecer en *The Philosophy of P.F. Strawson*, P.K. Sen y R.R. Verma eds.

13 G. Leibniz. "Principles of Nature and of Grace, Founded on Reason" en *Philosophical Writings*, editado por G.Parkinson. Dent, London (1973), p. 197.

no percepciones inconscientes, puesto que él define una *percepción* como una "repesentación con conciencia".[14] En cuanto a la apercepción, Kant traza una distinción entre apercepción empírica y apercepción trascendental; la "apercepción" en sí misma, sin la distinción "trascendental-empírica", es definida como "conciencia del yo".[15]

En algunas ocasiones Kant identifica la apercepción empírica con el sentido interno (A 107), pero no puede considerarse que esto represente su punto de vista, porque también insiste en que sentido y apercepción (junto con imaginación) son facultades distintas del alma (A 115). La posición de Kant parece ser esta: estar consciente en absoluto de una serie de representaciones es (por lo menos potencialmente) estar consciente de ellas *como* "perteneciendo" a uno mismo. Supóngase entonces que R1 y R2 son representaciones sucesivas con conciencia y que estar consciente de una representación es estar consciente de ella como perteneciéndole a uno mismo. Esto todavía parece conceder, tanto para cada una de las R1 como para cada una de las R2, que hay conciencia de que pertenecen a *un* yo, sin conceder que hay conciencia de ellas como pertenecientes al mismo yo. La conciencia de las representaciones como pertenecientes a un yo pero no a uno y al mismo yo es lo que Kant piensa que es la apercepción empírica: "la conciencia empírica que acompaña a diferentes representaciones es en sí misma diversa y no mantiene relación con la identidad del sujeto".[16] La apercepción empírica debe ser una forma de *auto*-conciencia si ha de pasar en absoluto como una forma de *a*percepción, pero da lugar a "un yo tan colorido y diverso como representaciones de las cuales soy consciente tengo".[17]

14 I. Kant. *Op. cit.*, A320/B376.
15 I. Kant. *Ibid*, B68.
16 I. Kant. *Ibid*, B133.
17 I. Kant. *Ibid*, B134.

Podría objetarse que no merece pensarse en absoluto en la apercepción empírica, como se le acaba de describir, como en una clase de *auto*-conciencia y que sería mejor pensar en ella simplemente como "conciencia". Porque el punto de partida de Kant es la tesis, absolutamente fundamental, de que nuestras representaciones están sometidas al tiempo, la condición formal del sentido interno, y parece una condición mínima para cualquier cosa que merezca ser considerada como auto-conciencia que incluya el que uno se percate de las representaciones que ocurren en tiempos diferentes como pertenecientes a uno y el mismo sujeto. De ahí que para ser auto-consciente, así va el argumento, uno debe estar consciente de uno mismo como *extendido temporalmente*; la conciencia que no mantiene relación con la identidad del sujeto a través del tiempo no es en absoluto auto-conciencia.

Es esta una línea de argumentación que Kant habría visto con gran simpatía, si bien él no estaba dispuesto a abandonar por completo su noción de apercepción empírica. Con lo que él está de acuerdo es en que una auto-conciencia vigorosa comporta la idea de representaciones diversas, que ocurren en tiempos diferentes, como pertenecientes a uno y el mismo sujeto. La apercepción trascendental es entonces esta conciencia de la identidad numérica del sujeto de diversas representaciones. Es apercepción *trascendental* porque la conciencia de la auto-identidad es, sostiene Kant, una fuente posible del conocimiento *a priori*.

¿A qué equivale la conciencia de la auto-identidad, dado que ningún yo perdurable puede presentarse a sí mismo en el flujo de las apariencias internas? Según la tesis (2), uno está consciente de sí mismo como sujeto de una representación dada en virtud de la habilidad que uno tiene para llamarla su propia representación. Llamar a una representación mi representación es, de acuerdo con Kant, adjuntarle un "yo pienso". La apercepción

trascendental es aquella auto-conciencia que genera el "yo pienso", el cual es descrito tanto como la "representación universal de la auto-conciencia"[18] como representando (pero sin "distinguir un objeto particular") al "sujeto trascendental de los pensamientos = X".[19] Puesto que la conciencia de uno mismo como sujeto de una representación dada es un asunto de ser capaz de adjuntarle a dicha representación un "yo pienso", Kant concluye que ser consciente de que uno es el sujeto idéntico de una multiplicidad de representaciones es ser capaz de adjuntarle un "yo pienso" a cada una de ellas, consciente de que es uno y el mismo "yo pienso" que es adjuntado a cada representación.[20] En otras palabras, puesto que Kant niega que un yo perdurable sea algo que puede simplemente darse en el sentido interno, necesita postular algún otro modo de conciencia de la auto-identidad y su propuesta es que la conciencia de la identidad de la *representación* de la auto-conciencia y el sujeto trascendental actúa como un reemplazante de la conciencia como sustancia perdurable.

La clave para la explicación que Kant da de la identidad de la apercepción o del "yo pienso" es la noción de síntesis, el "acto de juntar diferentes representaciones y de captar lo que es múltiple en ellas en un [acto de] conocimiento".[21] Kant escribe que la relación de representaciones diferentes con la identidad del sujeto proviene "no simplemente por mi acompañar con conciencia a cada representación, sino únicamente en la medida en que yo conjugo una representación con otra y que estoy consciente de su síntesis. Sólo en la medida, por lo tanto, en que

18 I. Kant. *Ibid*, A343/B401.
19 I. Kant. *Ibid*, A346/B404.
20 I. Kant. *Ibid*, B132-5.
21 I. Kant. *Ibid*, A77/B103.

puedo unir una multiplicidad de representaciones dadas en una conciencia es posible que yo me represente a mí mismo la *identidad de la conciencia en [i.e., a lo largo de] esas representaciones*".[22] El argumento es, pues, este: ser capaz de representarse a uno mismo como el sujeto perdurable e idéntico de diversas representaciones es ser capaz de representarse a uno mismo la identidad de la *apercepción* o de la *conciencia* de las representaciones como siendo de uno –en otras palabras, la identidad del "yo pienso" acompañante– y esto a su vez requiere la conciencia de la *síntesis* de las representaciones.

El segundo paso de este argumento es ambiguo de un modo que muchos comentadores han señalado. Es posible pensar en la síntesis como en una *actividad* (de combinar representaciones), de manera que la conciencia de la identidad es conciencia de una actividad. Alternativamente, es posible pensar en la síntesis de las representaciones como en el *producto* de un acto de combinación, como lo que Kant llama una 'unidad sintética'. En esta interpretación, lo que requiere la identidad de apercepción es conciencia de una multiplicidad de representaciones *como* formando una unidad sintética. Una manera de resolver esta ambigüedad es examinando más detenidamente lo que Kant quiere decir mediante 'combinación'. Es tentador pensar en la combinación como un "acto" de juntar diferentes representaciones, pero de hecho Kant lo define como la "representación de la unidad sintética de la multiplicidad".[23] Así, cuando Kant escribe que la combinación de una multiplicidad no puede nunca llegar a nosotros a través de los sentidos y es con exclusividad un asunto del entendimiento,[24] el punto que establece no es que el entendimiento literalmente junta diferentes representaciones a

22 I. Kant. *Ibid*, B133.
23 I. Kant. *Ibid*, B130-1.
24 I. Kant. *Ibid*, B129.

la manera de un rompecabezas, sino más bien la tesis menos extraña de que los sentidos no pueden fundar la *representación* de una unidad como una unidad. Puesto que la síntesis está definida en terminos de combinación, ello sugiere que la mejor interpretación de la posición de Kant es verla como la tesis de que la identidad de apercepción descansa no en la conciencia de una actividad, sino en la representación de la unidad de la multiplicidad como una unidad.[25]

De hecho, el asunto es ligeramente más complicado de lo que esto sugiere, pues Kant también insiste en que "no podemos representarnos a nosotros mismos nada que esté combinado en el objeto que no hayamos nosotros combinado previamente".[26] Así, pues, la representación de la unidad sintética de la multiplicidad acarrea consigo el pensamiento de que soy capaz de unir representaciones, un pensamiento del cual Kant dice que presupone la posibilidad de la síntesis, pero que no es en sí mismo la conciencia de la síntesis de representaciones.[27] Sin embargo, lo que realmente está haciendo el trabajo en la explicación de Kant es la conexión entre la conciencia de la identidad de la apercepción y la representación de la multiplicidad de intuiciones que constituyen una unidad sintética. En palabras de Kant: "La unidad sintética de la multiplicidad de intuiciones, en tanto que generada *a priori*, es pues el fundamento de la identidad de apercepción misma.... En verdad, el principio de unidad necesaria de apercepción es en sí mismo una proposición idéntica y, por lo tanto, analítica; no obstante, revela la necesidad de una síntesis de la multiplicidad dada en la intuición y sin la cual la identidad completa de la auto-conciencia no puede pensarse".[28]

25 Para una más amplia discusión de la teoría kantiana de la síntesis, véase mi "Kantand Reductionisms", en *Review of Metaphysics*, 1989.

26 I. Kant. *Op.cit.*, B130.

27 I. Kant. *Ibid*, B134.

28 I. Kant. *Ibid*, B134-5.

Así, la posición es esta: la apercepción en general es conciencia del yo. La apercepción trascendental es conciencia de la identidad del yo a través del tiempo. La conciencia de la auto-identidad es el pensamiento o la conciencia de la identidad de la representación del yo y de la auto-conciencia (el "yo pienso"). Por último, el pensamiento de la identidad del "yo pienso" descansa en la representación o pensamiento de la unidad sintética de la multiplicidad. Es porque la apercepción trascendental requiere conciencia de una unidad que Kant habla de la *unidad* de apercepción, así como de la *identidad* de la apercepción. El punto importante es que la identidad de la auto-conciencia es posible únicamente a través de la conciencia de una unidad.

¿Cuál es el carácter de la unidad de la multiplicidad de la cual uno debe estar consciente si ha de pensar en sí mismo como extendido en el tiempo? Para Kant, síntesis es conceptualización; representar la multiplicidad como una unidad sintética es reunir la multiplicidad bajo conceptos. Los conceptos son reglas para las intuiciones en la medida en que representan, en cualquier apariencia dada, la "reproducción necesaria de su multiplicidad".[29] De ahí que la conciencia de la auto-identidad sea la conciencia de la "unidad necesaria de la síntesis de toda apariencia según conceptos, es decir, de acuerdo con reglas que no sólo las hacen necesariamente reproducibles, sino que también al hacerlo determinan un objeto como su intuición, es decir, el concepto de algo dentro del cual están necesariamente interconectadas".[30] Kant está autorizado a hablar de conceptos que determinan un objeto como su intuición debido a su explicación de lo que es el que las representaciones tengan una carga objetiva. Puesto que un "objeto" es "aquello en el concepto de lo cual la

29 I. Kant. *Ibid*, A106.
30 I. Kant. *Ibid*, A108.

multiplicidad de una intuición dada está *unido*",[31] es necesario y suficiente para que las representaciones tengan una carga objetiva el que se les haga caer bajo conceptos. Las representaciones que son unidas bajo conceptos son representaciones que forman una unidad y formar una unidad de esta clase es precisamente, insiste Kant, tener una carga objetiva.

Podría ser útil a estas alturas resumir como un todo el argumento de Kant:

a) Un sujeto auto-consciente debe percatarse, o potencialmente percatarse, de diferentes representaciones como pertenecientes a él mismo, es decir, a uno y al mismo sujeto.
b) De ahí que un sujeto auto-consciente deba estar consciente de sí mismo como del idéntico sujeto de diversas representaciones. Esto es "apercepción trascendental".
c) Ser consciente de uno mismo como del sujeto de diversas representaciones es ser consciente de, o ser capaz de, "pensar" la identidad del "yo pienso" o, para decirlo de otro modo, la identidad de la apercepción de la auto-conciencia. La conciencia de la auto-identidad es la conciencia de la identidad de la auto-conciencia.
d) Una condición necesaria de la conciencia de la identidad de apercepción es la síntesis o combinación de la multiplicidad, es decir, la representación de la multiplicidad como formando una unidad sintética.
e) Es necesario para la representación de la unidad de la multiplicidad que la multiplicidad quede reunida bajo conceptos (las categorías).
f) Ser consciente de la multiplicidad como de una unidad gobernada por conceptos –o reglas– es ser consciente de ella como teniendo carga objetiva.

31 I. Kant. *Ibid*, B137.

El propio modo de Kant de enunciar la conclusión de su argumento consiste en sostener que la unidad trascendental de apercepción es aquella unidad a través de la cual la multiplicidad de la intuición queda unida en un concepto del objeto (B 139). De hecho, el argumento de seis pasos no garantiza la sugerencia de que es suficiente para la conciencia de uno mismo como de un sujeto de representaciones extendido en el tiempo el que las representaciones tengan una carga objetiva, pero Kant tiene un argumento en favor de la tesis de que es una condición necesaria.

¿Es exitosa la defensa de Kant de la tesis de la necesidad? Lo que debe concederse es que la plena auto-conciencia debe, como él insiste, involucrar la conciencia de uno mismo como extendido en el tiempo. Así, pues, deberían aceptarse los dos primeros pasos del argumento de Kant. También debería dejarse pasar los dos últimos pasos del argumento, puesto que en alguna medida simplemente reflejan las propias definiciones que Kant da de 'objeto', 'concepto' y 'unidad'. Son los pasos (c) y (d) los que parecen más vulnerables. ¿Es correcto igualar la conciencia de la auto-identidad con la conciencia de la identidad de la auto-conciencia? Si no lo es, entonces (c) debe ser rechazado. ¿Es la representación de la unidad sintética de la multiplicidad el fundamento de la conciencia de la identidad de la auto-conciencia? Si no lo es, entonces (d) debe ser rechazado.

Se hizo más arriba un esfuerzo por explicar (c) con referencia a la idea de que lo que Kant está buscando es un *sustituto* para la conciencia de la identidad. La dificultad que Kant enfrenta es la siguiente: si uno se limita a sí mismo a la perspectiva del sentido interno, de todo de lo que uno está consciente es del flujo eternamente mutante de percepciones. Como observa Kant, "Lo que tiene *necesariamente* que ser representado como numéricamente idéntico [el yo] no puede ser pensado como tal

a través de datos empíricos como esos".[32] No hay ningún objeto de cuya identidad uno pueda decir que está directamente consciente y no hay nada en la conciencia a partir de lo cual uno pueda *inferir* la identidad del yo como es en sí mismo. En efecto, puesto que el yo nouménico es atemporal, el concepto de identidad a través del tiempo no se le aplica. Puesto que no hay ningún objeto cuya identidad sea dada en el flujo de las apariencias internas, podría parecer que Kant no tiene otra alternativa que la de abrazar la conclusión de Hume de que la identidad adscrita a la mente del hombre es ficticia.

Esto habría sido para Kant una conclusión intolerable y él la evita argumentando que puesto que la conciencia de la autoidentidad no es conciencia de la identidad del sujeto como objeto del sentido interno, ella es conciencia de la identidad de la *representación* de un sujeto, el "yo pienso". La conciencia de la identidad del "yo pienso" *suple* a la conciencia del sujeto como objeto. Puesto que "yo" suple a la conciencia del sujeto trascendental, es particularmente tentador considerar a rasgos de esta representación como rasgos del yo tal como es en sí mismo. Este es precisamente el error del cual se acusa al psicólogo racional. Puesto que la identidad del "yo" es constitutiva de lo que Kant llama la "unidad de la conciencia" y puesto que la conciencia de la identidad del "yo" acta como un suplente de la conciencia del sujeto como un objeto, se vuelve ahora inteligible el que el psicólogo racional se incline a confundir la unidad de la conciencia "por una intuición del sujeto como objeto".[33]

Aunque es posible ver qué subyace al compromiso de Kant con (c), es más difícil creer en la corrección de su posición. La dificultad central se deriva de la parte que desempeña en el argumento la tesis (2). Esta tesis es cuestionable desde diferentes

32 I. Kant. *Ibid*, A107.
33 I. Kant. *Ibid*, B421.

puntos de vista. El primero es este: la tesis (2) fue introducida más arriba en el contexto de la afirmación de Wittgenstein de que nada en el contenido de la experiencia o de la conciencia misma justifica el uso de 'yo' en reportes de experiencia inmediata. Se observó que Kant elude esta dificultad al sostener que uno está consciente de sí mismo porque uno se auto-adscribe representaciones, más que a la inversa. Esto es ingenioso pero contra-intuitivo. Si no pudiera decirse de la conciencia que produce auto-conciencia independientemente de la auto-adscripción, parece fraudulento decir que porque yo me auto-adscribo representaciones yo estoy después de todo consciente de un yo. La tesis (2) parece ser un cínico esfuerzo de obtener algo por nada. Quienes niegan que la auto-conciencia es posible sin auto-conocimiento podrían tratar de defender a Kant de esta acusación diciendo que no se trata simplemente de representaciones auto-adscriptivas, sino de auto-adscribirse representaciones en el contexto de un auto-conocimiento apropiado. Desafortunadamente, esta defensa es totalmente inútil tanto por el compromiso de Kant con la tesis (1) como porque es totalmente oscuro cómo puede suponerse que el hecho de que uno tenga auto-conocimiento ayude a bloquear la acusación de que al representar el contenido de la conciencia en términos de primera persona uno va más allá de lo que literalmente es accesible en la conciencia. Los recursos suplementarios accesibles a un opositor de la tesis (1) resultan ser igualmente ajenos al contenido de la conciencia misma que el uso de 'yo' sin auto-conocimiento.

Podría suceder que lo que subyace a la tesis (2) es el escepticismo acerca de la idea misma de que lo que "literalmente" es accesible en la conciencia es, inclusive en principio, independiente de la forma en la que se le reporta. Si no se les puede separar, entonces no hay lugar para la objeción de que al llamar a una representación 'mía', voy "más allá" de algo "dado" supuestamente impersonal. La sugerencia sería que la estructura

de lo que es accesible en la conciencia está determinada parcialmente por la forma de los reportes acerca de ello. Pero inclusive si se aceptara esta controvertible y difícil sugerencia, hay otro problema con la tesis de que es la posibilidad de vincular el "yo" de la apercepción con las representaciones lo que es que uno esté consciente de un yo. Esta tesis es tan sólo plausible si está claro por qué el "yo" cuenta como una representación de un yo. Si la conciencia de la auto-identidad es conciencia de la identidad del "yo", Kant debe explicar qué hace a este último una representación de sí mismo (A 355). Pero él de hecho describe el "yo" como "completamente vacío" (A 346/B 404), el concepto de un "mero algo" (A 355). ¿Cómo puede el concepto de un mero algo representar un sujeto? ¿Cómo puede la posibilidad de que se vincule una representación completamente vacía a otras representaciones justificar el que se hable de auto-conciencia?

Cuando Kant describe el "yo" de la apercepción como vacío, él no pretende negarle toda significancia representacional. El punto que él establece es simplemente el ya familiar de que un concepto sin una intuición correspondiente tiene únicamente un significado "puramente lógico" (A 147/B 186). No hay, pues, ningún conflicto entre decir que el "yo" es vacío y que representa a un sujeto. ¿Qué le da al "yo" del "yo pienso" su significancia representacional? Lo que Kant parece pensar es esto: en la concepción cartesiana tradicional del sujeto, el yo es una sustancia inmaterial simple que es numéricamente idéntica a sí misma a lo largo del tiempo. El "yo" de la apercepción *arremeda* estos rasgos y es una representación del yo en virtud de ese hecho.

La idea de que el "yo" arremeda los rasgos tradicionalmente adscritos al yo de la psicología racional parece subyacer al argumento de los Paralogismos. La sustancialidad del yo brota en el primer Paralogismo como la tesis de que "en todos nuestros pensamientos el 'yo' es el sujeto en el cual inhieren pensamientos

sólo como determinaciones".[34] En el segundo Paralogismo, en lugar de la simplicidad del sujeto mismo, tenemos la "simplicidad de la representación de un sujeto",[35] es decir, del "yo" que "no contiene en sí mismo la menor multiplicidad".[36] En el tercer Paralogismo, el "yo" arremeda la identidad del yo en virtud del hecho de que "acompaña en *mi* conciencia, y en verdad con identidad completa, todas las representaciones en todos los tiempos".[37] Es precisamente porque el "yo" tiene estos rasgos que, según Kant, el psicólogo racional comete el error de considerarlos como rasgos del yo tal como éste es en sí mismo. Como Kant lo expone en el segundo Paralogismo, "la simplicidad de la representación de un sujeto no es *eo ipso* conocimiento de la simplicidad del sujeto mismo".[38]

Una objeción a la explicación que Kant da de la significancia representacional del "yo" corre como sigue: puesto que, según Kant, no puede saberse si el sujeto trascendental de hecho tiene las características adscritas a él por la psicología racional ¿cómo puede el hecho de que él arremede estas características justificar la aseveración de que es una representación de un sujeto? Por ejemplo, el hecho de que el "yo" sea simple sólo sería relevante para la cuestión de si "designa" al sujeto trascendental si el sujeto es simple, pero eso es precisamente lo que no puede saberse. Tampoco funcionará el decir que basta que el "yo" arremede las características que se piensa que son características del yo, porque sólo se piensa que éstas son características del yo debido a una inferencia falaz a partir de la naturaleza del "yo". Kant tiene que renunciar a algo.

34 I. Kant. *Ibid*, A349.
35 I. Kant. *Ibid*, A355.
36 I. Kant. *Ibid*.
37 I. Kant. *Ibid*, A362-3.
38 I. Kant. *Ibid*, A355.

Un rasgo objetable suplementario de esta explicación de la significancia representacional del "yo" es que hablar acerca de la sustancialidad, simplicidad e identidad del "yo" es reificarlo y esto hace a la explicación vulnerable a una versión de la idea de Hume. En el tercer Paralogismo, por ejemplo, Kant inclusive describe el "yo" como "la única apariencia permanente que encontramos en el alma".[39] Esta es la base de la tesis de que el "yo" arremeda la identidad del yo, pero un humeano podría ahora objetar que no es más plausible decir que ninguna apariencia permanente se presenta a sí misma en el flujo de la apariencia interna que decir que lo hace un alma cartesiana.

Podría parecer que esta objeción descansa en una incomprensión de la posición de Kant. Aunque algunas de sus formulaciones efectivamente tienden a reificar el "yo pienso", él hace una clara advertencia en el primer Paralogismo de que no deben tomarse literalmente tales formulaciones: "En verdad, el 'yo' está en todos los pensamientos, pero no hay en esta representación la menor traza de intuición tal que distinga al 'yo' de otros objetos de intuición. Así, podemos efectivamente percibir que esta representación está invariablemente presente en todos los pensamientos, pero no que es una intuición duradera y continua...".[40] Este es un pasaje importante, pero suscita una cuestión concerniente a la transición de (b) a (c) en el argumento de seis pasos. Dado que el punto de vista oficial de Kant era que no es el caso que uno tenga una *intuición* de la simplicidad y de la identidad de una representación concreta que está siempre presente en la conciencia, ¿a qué equivale la *conciencia* de la simplicidad y la identidad del "yo"? Una postura sería que cuando él se refiere a la conciencia de la simplicidad de la representación de un sujeto, nosotros deberíamos entenderlo como

39 I. Kant. *Ibid*, A364.
40 I. Kant. *Ibid*, A350.

refiriéndose a la representación del sujeto como simple. Cuando él escribe del pensamiento o de la conciencia de la identidad de la auto-conciencia y del "yo" que acompaña "con identidad completa" todas mis representaciones, ello es simplemente un modo vívido de establecer el punto de que debo ser potencialmente consciente de diversas representaciones como mías, como pertenecientes a uno y el mismo yo. Si ello es así, sin embargo, entonces (c) se vuelve simplemente una variante notacional de (b). A menos de que hablar de ser capaz de "pensar" la identidad del "yo" no sea simplemente un modo de parafrasear lo que se dice al hablar acerca de la conciencia de la auto-identidad, la transición de (b) a (c) no representa ningún progreso explicativo genuino. Si hay aquí progreso explicativo y si el precio de tal progreso es la reificación del "yo", ello parece ser un precio demasiado elevado.

Esta objeción invita a la respuesta de que no es el caso ni que el argumento de Kant no efectúa ningún progreso explicativo genuino ni que depende de la reificación del "yo pienso". La propuesta de Kant es que pensar la identidad de la auto-conciencia es unir las representaciones que uno tiene bajo conceptos de objetos. Así, decir que la conciencia de la identidad es la identidad de la auto-conciencia es decir que la conciencia de la identidad del sujeto de diversas representaciones es la conciencia de las representaciones que uno tiene como representaciones de objetos independientes. Lo menos que se puede decir es que esta no es una tesis trivial.

Esta no sería una dirección muy prometedora por la cual encausar el argumento de Kant. La conexión entre la conciencia de la auto-identidad y la carga objetiva de las representaciones es precisamente lo que se suponía que Kant estaba tratando de establecer. Él apenas argumenta en favor de la tesis, en (d), de que la identidad de la auto-conciencia requiere la unidad de la multiplicidad en el concepto del objeto, y argumenta menos aún

que pueda identificarse con esta unidad.[41] Él ciertamente asevera que la unidad de la multiplicidad es el fundamento de la identidad de la apercepción, pero es difícil encontrar un argumento para esta aseveración. Sin un argumento así es difícil evitar la sospecha de que (c) es más o menos superfluo. Con lo que Kant se queda es con la aseveración hueca de una conexión entre la auto-conciencia y la objetividad, apoyada únicamente por una serie de maniobras definicionales nada convincentes.

III

Lo que Kant estaba buscando era una explicación de la autoconciencia que respetara la tesis de Hume acerca del carácter "inencontrable" del yo. Su idea era que la identidad de la apercepción, y por lo tanto la conciencia de la carga objetiva de las representaciones, actúa como un sustituto para la clase de autoconciencia eliminada por la tesis de Hume. Puesto que la alternativa de Kant es imposible de elaborar, podría valer la pena preguntar si la tesis de Hume es tan plausible como lo han supuesto tantos escritores. El blanco inmediato de las observaciones de Hume al principio de su capítulo sobre la identidad personal era Locke, quien sostenía que "aquella conciencia que es inseparable del pensar" es la fuente de la idea que uno tiene de sí mismo como "la misma cosa pensante, en diferentes tiempos y lugares".[42] Es a través de la conciencia misma que uno está consciente de sí mismo como extendido en el tiempo y ubicado

41 La tesis de que Kant realmente no tiene un argumento en favor de B139 está defendida en mi "Transcendental Arguments, Transcendental Synthesis and Transcendental Idealism", en *Philosophical Quarterly*, 1987.

42 D. Hume. *An Essay Concerning Human Understanding*, Libro II, Cap. 27, Sec. 9.

espacialmente. No se necesita buscar un sustituto para la conciencia de la auto-identidad, porque la identidad de uno mismo está dada en la conciencia. Desde este punto de vista no se plantea la cuestión de aceptar la sugerencia lichtenbergiana de que la conciencia es conciencia de que el pensamiento fluye, no de un "yo" que piensa.

¿Cuál es la naturaleza de este "yo" que piensa? No se necesita aceptar el supuesto de que uno está consciente de un yo al modo como lo sugiere Locke, el cual tendría que ser un alma-sustancia cartesiana. Porque el modo en que de hecho tenemos experiencias de nosotros mismos es como cosas que persiten físicamente como seres humanos, y es en términos de esta conciencia de nosotros mismos como seres materiales que debe comprenderse la auto-conciencia.[43] Esta forma de auto-conciencia no descansa en la posesión que uno tenga de las creencias apropiadas acerca de nuestra propia naturaleza –por lo que no hay conflicto con la tesis ()1 y sin embargo es genuinamente una forma de auto-conciencia. Una vez que nos sacudimos la influencia de Hume y Lichtenberg y que estamos preparados para asumir más o menos literalmente el discurso acerca de la identidad del yo –comprendido como una cosa material– la búsqueda kantiana por un substituto parece totalmente desencaminada.

Es claro que se necesita decir mucho más en defensa de la explicación de la auto-conciencia que se acaba de proponer. Para nuestros propósitos, sin embargo, la siguiente idea bastará: la teoría "materialista" de la auto-conciencia no era para Kant una opción, porque su idealismo trascendental le impedía aceptar que el yo tal como es en sí mismo pudiera ser material. El resultado es, como observa Michael Ayers, que "Kant no parece

43 Para una defensa de esta concepción de la auto-conciencia, véase el libro de Michael Ayers, *Locke*. Routledge and Kegan Paul, London (1991), Vol. 2, pp.285-288.

haber sido menos proclive a subestimar la tesis de que el yo corpóreo o empírico es generalmente, y se presenta como, el sujeto de la experiencia y de la conciencia, que el dualista dogmático".[44] Por otra parte, no se trata ciertamente de una constricción impuesta a una teoría aceptable de la auto-conciencia que debiera ser compatible con el idealismo trascendental.[45]

44 M. Ayers. *Op.cit.*, p. 287.

45 Agradezco a los participantes del Simposio "Mente-Cuerpo", organizado gracias al proyecto IN-600789-UNAM y el cual tuvo lugar en el Instituto de Investigaciones Filosóficas, México D.F., por sus comentarios a una versión anterior de este artículo.

MATERIALISMO, INTERACCIONISMO Y ANÁLISIS GRAMATICAL

Alejandro Tomasini Bassols*

El objetivo del presente trabajo es hacer ver cómo y por qué, esto es, debido a qué clase de confusiones, el famoso problema *mente-cuerpo* es, más que otra cosa, un enredo que puede ser visto como un auténtico nudo gordiano. La técnica filosófica del último Wittgenstein, sin embargo, me parece ser el Alejandro Magno que permitirá su disolución definitiva. En lo que sigue, aunque no intentaré siquiera practicar esa técnica, sí trataré de recoger algunos de los liberadores resultados a los que, creo, conduce. Para ello, empero, será necesario empezar por reconstruir algunas teorías filosóficas de renombre. Una vez hecho esto, las criticaré. En la tercera parte, intentaré articular una concepción de la "problemática" que sea tal que permita entender por qué la gran mayoría de los esfuerzos han sido mal encaminados.

- *I* -

Quisiera dedicar esta primera sección a la presentación, lo más concisa posible, de lo que se conoce como 'teoría de la identidad'. Quiero dejar aclarado que no es mi propósito examinar uno por uno a todos (o a los más conocidos) adherentes de la escuela, sino que más bien intentaré hacer una presentación global, recuperando diversas contribuciones y tratando de mostrar que efectivamente, a pesar de las variaciones que se pueden encontrar en los

* Instituto de Investigaciones Filosóficas, UNAM.

diversos escritos, hay un núcleo común de pensamiento y que puede haber una versión unificada o integrada de él.

Aunque el *status* de la tesis de la identidad no es obvio, su contenido (a primera vista, por lo menos) sí es claro. Lo que se sostiene es simplemente que los estados y procesos así llamados 'mentales' **son** estados cerebrales. (Obsérvese que los defensores de esta tesis no se comprometen explícitamente con un punto de vista acerca de un tema que no pueden eludir, *viz.*, el de la naturaleza de los **objetos** de la mente y del cerebro. Sobre este asunto, empero, regresaré posteriormente). La tesis es importante y merece que se le ejemplifique. Está implicado por ella que si alguien tiene, digamos, una creencia, eso que nosotros identificamos mediante un término mentalístico es en realidad un evento ubicado en el cerebro. Lo mismo acontece, *mutatis mutandis*, con los deseos, intenciones, pensamientos, percepciones, sensaciones, esperanzas, etc. En otras palabras, nuestros enunciados acerca de estados o procesos mentales son en el fondo enunciados acerca de procesos y estados que tienen lugar en el cerebro.

Es evidente que el esclarecimiento de la tesis requiere que se aclare el *status* del 'es' involucrado, o sea, del 'es' de la afirmación de que un estado (proceso) mental **es** un estado (proceso) cerebral. En relación con este asunto se han ofrecido por lo menos dos interpretaciones, en función de las cuales variará la concepción misma de la tesis, y tal vez no esté de más señalar que no es fácil determinar, ni *a priori* ni *a posteriori*, cuál de las dos interpretaciones es la más representativa de la teoría. U.T.Place, por ejemplo, sostiene que el 'es' en cuestión es el 'es' de lo que él denomina la 'composición'. Su idea es la siguiente: hay un uso de la partícula 'es' que permite que se afirme, cuando se dice algo de algo, que eso es todo lo que está en juego, que ya no hay nada más que decir. Por ejemplo, puede decirse que saber sumar es saber dar la respuesta correcta siempre y **nada más** o que un cuadrado es una figura geométrica de cuatro

lados y cuatro ángulos rectos y **nada más**, etc. Pero también podemos querer caracterizar un objeto mediante una descripción individualizadora de modo que podamos decir del objeto así descrito que es eso y **nada más**. Por ejemplo, podemos decir de una computadora que es una mezcla de *hardware* y de *software* y **nada más** o del automóvil de mi vecino que es un montón de fierros y cables y **nada más**. Este es el uso de 'es' de lo que Place llama la 'composición' y es precisamente el que se aplica al problema mente-cuerpo: los estados de conciencia, mentales, etc., son estados cerebrales y **nada más**. En esto coinciden el 'es' de la identidad con el 'es' de la composición y en esto se diferencian de otro uso de la cópula, a saber, el de la predicación. En este último caso, no podemos añadir '*y* **nada más**': siempre se pueden predicar nuevos atributos del objeto de que se hable. Por otra parte, el 'es' de la identidad se diferencia del de la composición por el hecho de que el primero da lugar a un juicio meramente analítico, como cuando decimos que 'Todo soltero es un no-casado'. Dicho juicio sería, en caso de ser verdadero, necesariamente verdadero. Ahora bien, difícilmente podría rechazarse que juicios como 'Todo estado mental es un proceso cerebral' tienen más bien la apariencia de enunciados sintéticos. Por lo tanto, un juicio así **podría** en principio ser verdadero y **podría** ser falso. Place sostiene, claro está, que es verdadero, pero lo que es interesante es percatarse de la clase de enunciados a la que él lo hace pertenecer: dado que el 'es' en cuestión es el de la composición y dado que el enunciado es contingente, la tesis de la identidad no puede ser, desde su perspectiva, otra cosa que una **hipótesis científica** y su verificación es, si es obtenible, empírica. De este modo, Place cree refutar a quienes pretenden eliminar con argumentos puramente *a priori* la tesis de la identidad. "Yo sospecho que aquellos que sostienen que el enunciado 'La conciencia es un proceso cerebral' es lógicamente insostenible basan su afirmación en la suposición equi-

vocada de que si los significados de los dos enunciados o expresiones están por completo desconectados, entonces no pueden proporcionar una caracterización adecuada del mismo objeto o estado de cosas: si algo es un estado de conciencia, no puede ser un proceso cerebral, puesto que no hay nada auto-contradictorio en la suposición de que alguien siente un dolor cuando nada está sucediendo dentro de su cráneo." [1] Si Place tiene razón, este argumento es irremisiblemente falaz.

Una concepción diferente de la tesis es la que encontramos en J.J.C.Smart. Para éste, el 'es' de la tesis es precisamente el 'es' de la identidad. Aquí se vislumbra una diferencia importante. Para Place, la tesis es, como se dijo, una hipótesis científica. Smart objeta a esto que si se tratara de una hipótesis científica, entonces debería resolverse en principio mediante experimentos. Es claro, sin embargo, que el problema no es un asunto de laboratorio. "Porque no hay ningún experimento concebible que pudiera decidir entre el materialismo y el epifenomenalismo".[2] Luego la tesis de la identidad debe estar expresando una verdad de otra clase, a saber, una verdad conceptual. Por consiguiente, el 'es' es el de la identidad. La identidad en cuestión, como veremos más abajo, es ontológica, no semántica o lingüística.

El otro gran abogado de la teoría de la identidad que deseo considerar aquí, *i.e.*, D.M.Armstrong, opta por otra versión del asunto. Su posición es que la tesis de la identidad no expresa una verdad lógicamente necesaria, sino contingente, lo cual lo acerca a Place. "Si algo es cierto en filosofía, ello es que el enunciado 'La mente es el cerebro' no es una verdad lógicamente necesaria".[3]

1 U.T.Place. "Is Consciousness a Brain Process?" en V.C.Chappell (comp.) *The Philosophy of Mind*. Dover Publications, Inc. New York (1981), p.103.
2 J.J.C.Smart. "Sensations and Brain Processes" en V.C.Chapell, p.172.
3 D. M. Armstrong. *A Materialist Theory of Mind*. Routledge and Kegan Paul; Londres, 1968, pp.76-77.

No obstante, lo que se afirma expresamente es una identidad. Con este reconcimiento, Armstrong se acerca a Smart. Luego su posición es que la tesis afirma o enuncia una identidad contingente. Esto, según Armstrong, no tiene nada de extraño: la ciencia está plagada de ellas. La prueba es que enunciados como 'el calor es la aceleración de las partículas', 'el oro es el metal de peso atómico 79', etc., que son afirmaciones típicamente científicas, son afirmaciones contingentes de identidad. Parecería, pues, que la posición común podría quedar formulada como sigue: hay enunciados contingentes de identidad, esto es, enunciados que si bien aluden a una y la misma cosa fue necesaria una investigación empírica para poder establecerla.

Ahora bien, independientemente de estas diferencias respecto al modo como conciben la tesis que defienden (diferencias que al parecer pueden ser subsanadas), es claro que todos los defensores de la tesis de la identidad coinciden en lo que es el punto crucial, a saber, la convicción de que no hay dos clases distintas de fenómenos, unos físicos y otros mentales (por ejemplo, una creencia por una parte y un estado del cerebro por la otra). Cada creencia, por ejemplo, **es** un estado cerebral determinado. No se trata de dos estados o procesos numéricamente diferentes conectados de algún misterioso modo. Aquí, sin embargo, se plantea otro problema, que es imposible pasar por alto.

Una pregunta que sin duda alguna se tiene derecho a plantear es la siguiente: ¿enuncia la tesis la identidad entre los estados y procesos mentales con clases específicas de estados cerebrales **en general** (es decir, creencias con cierta clase de estados cerebrales, deseos con otra, pensamientos con otra y así sucesivamente) o más bien es **cada** estado mental individual el que queda asociada con un estado cerebral particular (**esta** creencia con **este** estado cerebral, **este** deseo con *este* otro estado cerebral, etc.)? Armstrong aborda directamente el tema y ofrece razones aceptables en favor de la segunda vía. El punto importante es

que el defensor de la teoría de la identidad es un materialista, pero un materialista no puede cancelar *a priori* el que por medio de otros elementos que los puramente neurofisiológicos se puedan producir pensamientos, creencias, dolores, etc. Si lo que ellos dicen es cierto, entonces ciertamente algunos procesos que tienen lugar en el cerebro son simultáneamente eso que llamamos 'estados mentales', pero ello no implica que sea **únicamente** en los casos en los que están involucrados elementos orgánicos y procesos biológicos a los que dan lugar que podríamos hablar de estados y procesos mentales. "Podría resultar que nada que no sean objetos fisiológicos son capaces de jugar los papeles causales indefinidamente complejos que deben jugar los fenómenos mentales para que sean mentales. Pero deberíamos igualmente evitar el dogmatismo de asumir de antemano que ello es definitivamente así. Lo mental puede estar o no ligado a lo fisiológico." [4] El caso de las computadoras es pertinente: podría suceder que llegáramos a la conclusión de que también ellas tienen, por ejemplo, pensamientos, realizan cálculos, aprenden, etc., sólo que su *hard ware* sería completamente distinto (chips y circuitos en lugar de sistemas neuronales). Ahora bien, si se concede que en principio es legítimo hablar de, digamos, el dolor de una máquina, entonces la única relación que se puede intentar establecer es entre **este** estado físico, sea el que sea y esté conformado como esté conformado, con **este** estado mental (dolor, pensamiento, etc.). Una versión *prima facie* más viable de la teoría de la identidad, por consiguiente, parece tender a comprometerse más bien con una relación *token-token* antes que una *type-type*.

4 D.M. Armstrong. "Consciousness and Causality" en *Consciousness and Causality. A Debate on the Nature of Mind.* Basil Blackwell. Oxford, 1984; p.161.

La teoría de la identidad está obviamente ligada a muchas otras, pertenecientes a diversas áreas de la filosofía. Aquí quiero rápidamente mencionar dos particularmente importantes: una teoría causal y una teoría semántica. Los teóricos de la identidad coinciden en general en que la identidad de lo mental con lo físico no entraña identidad de significado. Esto es claro: el que yo diga que algo me duele **no significa** lo mismo que la afirmación de que estoy en tal o cual estado cerebral. La razón es en parte epistémica: aunque se trata de un mismo fenómeno, no lo conozco simultáneamente bajo sus dos modos de ser. Yo conozco mi dolor introspectivamente, esto es, como un fenómeno mental, pero no físicamente, es decir, no me examino el cerebro para determinar si tengo o no un dolor. U.T. Place llega tan lejos que sostiene que ello es **lógicamente** imposible. "Las operaciones requeridas para verificar enunciados acerca de la conciencia y enunciados acerca del los procesos cerebrales son fundamentalmente diferentes".[5] Si ello es así o no es algo que examinaremos más abajo pero, independientemente de lo que se decida a este respecto, lo que sí parece subyacer a la posición expuesta es la semántica fregeana: estamos en presencia de un caso típico de diferencia de sentido e identidad de referencia. 'Dolor' y 'estimulación de la fibra C', por ejemplo, nos remiten a uno y el mismo estado o proceso, conocido y verificado de modo diverso, sólo que el significado de las expresiones es distinto y no hay reducción, traducción, eliminación, etc., posibles de uno de ellos en favor del otro. Se trata de dos modos de presentación de uno y el mismo fenómeno.

La otra tesis relacionada con la tesis de la identidad es la teoría funcionalista de los procesos y estados mentales. Quien mejor ha desarrollado esta conexión entre causalidad e identidad psico-física es, creo, Armstrong. La idea es la siguiente: cuando nosotros hablamos de cosas como dolores o creencias estamos

5 U.T.Place. *Op.cit.*, p.105.

hablando, sin saberlo, de estados y procesos que tienen lugar dentro de los sujetos y, más específicamente, en el cerebro. Ahora bien, el evento físico (en nuestro caso, neurofisiológico) en cuestión es detectado por nosotros los hablantes no por sus características físicas, sino por sus características funcionales. Lo que esto quiere decir es que ciertos eventos físicos son identificados a través de sus efectos, es decir, del papel que desempeñan en la conducta general del sujeto. Esta función causal desempeñada por el proceso o estado "mental" es descrita por Armstrong (y hay razones para ello, como veremos más abajo) como una 'descripción formal'. La descripción formal en cuestión **apunta** a una propiedad física, cuyas características nosotros no conocemos, pero que la ciencia sí estudia y explica. De ahí que Armstrong sostenga que el lenguaje de lo mental es una especie de lenguaje proto-científico: hablar de creencias, deseos, intenciones, etc., es postular entidades teóricas de un modo un tanto primitivo. Place sostiene algo muy parecido. De acuerdo con él, observamos "por dentro" y "por fuera" el dolor, pero ¿cómo sabemos que las dos clases de observaciones son de los mismos procesos? Eso es algo que, por medio de correlaciones sistemáticas, la ciencia natural establece: "tratamos a los dos conjuntos de observaciones como observaciones del mismo evento en aquellos casos en los que las observaciones técnicas, científicas, ubicadas en el contexto del cuerpo apropiado de teoría científica proporcionan una explicación inmediata de las observaciones hechas por el hombre de la calle".[6]

Armstrong hace un serio esfuerzo por mostrar que ni el solipsismo ni la cuestión de las otras mentes representan un problema real para la teoría de la identidad. Parte de su solución consiste en que su teoría parece bloquear la distinción entre primera y tercera personas. El primer paso en su argumentación

6 *Ibid.*, p.106.

consiste en la eliminación del sujeto de las experiencias, pensamientos, etc., en nuestros propios casos. Siguiendo a Hume, Armstrong acepta la idea de que la introspección no revela nada especial en este sentido. La mente es una construcción teórica del lenguaje y el pensamiento comunes. Pero, por otra parte, es claro que si no podemos detectar nuestro propia mente menos aún podremos hacerlo en el caso de otras personas: "se nos tiene que enseñar que tenemos una mente. Al atribuirles mentes a otros estamos siendo doblemente teóricos. Les estamos atribuyendo estados mentales, internos, etc., que no observamos directamente. Pero estamos también asumiendo que esos estados están organizados en una mente".[7] La experiencia muestra que no hay necesidad de semejantes postulaciones. Armstrong llega tan lejos que afirma que hay una cierta uniformidad de los estados mentales en el sentido de que ni siquiera en nuestros propios casos son percibidos **como** mentales, sino simplemente como estados que desempeñan un cierto papel causal. Ahora bien, ese papel causal puede ser detectado de igual modo en el caso de otras personas. De este modo uniformizamos a lo mental y nos libramos de las asimetrías provocadas por las primera y tercera personas.

Antes de que por nuestra parte examinemos críticamente esta doctrina, creo que sería útil recoger algunas de las objeciones a las que sus defensores ya se enfrentaron y para su satisfacción respondieron. No consideraré sino algunas que me parecen ilustrativas. Para esto creo que lo mejor será cederle la palabra a Smart.

La primera objeción que Smart intenta refutar es la siguiente: hablar de estados mentales no es lo mismo que hablar de cerebro y procesos neuronales, puesto que un iletrado puede hablar de los primeros sin saber absolutamente nada acerca de los segundos. *Ergo* no son lo mismo. La respuesta es evidente: hay enun-

7 D.M. Armstrong. *Op. cit.*, p.145.

ciados contingentes de identidad y nadie tiene por que conocer de igual modo todos los modos posibles de presentación de un objeto. No hay ninguna contradicción en que alguien conozca A, pero no sepa al mismo tiempo que A = B.

Una segunda objeción consiste en decir que, como los teóricos mismos de la identidad aseveran, la relación entre el estado mental y el cerebral es contingente. Luego hablar del primero no puede ser lo mismo que hablar del segundo. La respuesta es, a estas alturas, adivinable: lo que tenemos son dos sentidos distintos, pero una misma denotación. Lo primero no obsta para lo segundo. "Si el significado de una expresión fuera lo que la expresión nombra, entonces desde luego que se seguiría del hecho de que 'sensación' y 'proceso cerebral' tienen significados diferentes que no pueden nombrar una y la misma cosa".[8] La teoría de Frege, empero, asumida tácita o explícitamente por los teóricos de la identidad, anula esta objeción.

Una tercera objeción es que los estados mentales no son espaciales, en tanto que los cerebrales sí lo son. Luego no pueden ser idénticos. La respuesta de Smart es simplemente que la objeción descansa en una confusión: no son las creencias, las imágenes, las esperanzas, etc., lo que son procesos o estados cerebrales, sino las experiencias de tenerlos. Es "la *experiencia* que se reporta en el reporte introspectivo"[9] lo que es el estado físico.

Una última objeción digna de ser mencionada apunta al carácter privado de las "experiencias mentales" y lo contrasta con el carácter público de los eventos físicos. Smart responde diciendo que lo único que esto muestra es que la lógica del lenguaje de los reportes introspectivos es diferente de la del lenguaje de los objetos físicos. Pero esto, obviamente, no anula la identidad deseada.

8 J.J.C.Smart. *Op. cit.*, p.165.
9 *Ibid.*, p.168.

Tal vez podamos ahora resumir nuestra exposición así: la tesis de la identidad psico-física es una tesis materialista y reduccionista. Dejando de lado la intuición fundamental (tal vez mejor expresada por Smart de este modo: "Que todo sea explicable en términos de física (...) salvo la ocurrencia de sensaciones me parece francamente increíble"),[10] la tesis se desarrolla en toda una doctrina por sus implicaciones, ramificaciones y conexiones con puntos de vista en otras áreas de la filosofía. De particular importancia son la semántica filosófica, la teoría del conocimiento y la filosofía de la ciencia. En principio, se trata de una teoría coherente y es obviamente una teoría filosófica sumamente ambiciosa. Quizá debamos ahora pasar a examinarla críticamente.

- II -

Antes de que presente yo lineas de pensamiento inspiradas por los teóricos de la identidad y de criticar algunos de sus argumentos, quisiera recoger algunas objeciones, a mi modo de ver decisivas, que han sido ya ofrecidas por algunos filósofos insignes. En particular, deseo reconstruir críticas a la teoría de la identidad elaboradas por N.Malcolm y por S.Kripke. Examinaré algunas de las cosas que dicen en ese orden.

Aunque no tiene mayor importancia, tal vez no esté de más empezar por señalar que la expresión misma 'tesis de la identidad' fue acuñada por N.Malcolm. Esto es, por lo menos, lo que se puede inferir de la respuesta de Malcolm al artículo de Smart sobre el materialismo.[11] Independientemente de ello, lo cierto

10 *Ibid.*, p.161.

11 N.Malcolm. "Scientific Materialism and the Identity Theory" en C.V.Borst (comp.): *The Mind-Brain Identity Theory*. Macmillan, St.Martin's Press; Nueva York, 1977; p.172.

es que en unos cuantos trabajos Malcolm presenta argumentos poderosos en contra de la "tesis" de la identidad. Adelantándose a Kripke y por medio de otro aparato conceptual y de otra técnica filosófica, Malcolm centra su ataque en dos puntos particularmente sensibles en el planteamiento del teórico de la identidad. Éstos son:

1) el carácter contingente de la identidad en cuestión, y
2) ciertas características del pensamiento.

Hay varios problemas para los teóricos de la identidad. Uno de ellos es que la noción misma de identidad es problemática en grado sumo y otro es que, inclusive si la aceptamos como una noción viable, lo que ellos parecen querer hacer es conciliar dos cosas que son incompatibles, a saber, la identidad y la contingencia. Sin pretender adentrarnos mayormente en el tema, parecería que podríamos emplear la noción de identidad en dos clases de casos: el de la identidad de una cosa consigo misma y el de la identidad de dos cosas distintas. Pero estos usos de la identidad ya fueron severamente criticados en el *Tractatus*: "Dicho sea de paso: decir de dos cosas que son idénticas es un sinsentido y decir de *una* cosa que es idéntica a sí misma no es decir absolutamente nada".[12] Wittgenstein sostiene que la identidad no es ni una propiedad ni una relación entre objetos. Lo único que podemos hacer es emplear los mismos o diferentes signos (con sus respectivos significados, convenciones de aplicación, etc.), según sea el caso, es decir, según queramos hablar de uno y el mismo objeto o de otro. Por eso Wittgenstein dice que "La identidad de un objeto la expreso mediante la identidad del signo y no usando un signo para la identidad. La diferencia

12 L. Wittgenstein: *Tractatus Logico-Philosophicus*. Routledge and Kegan Paul; Londres, 1978; 5.5303.

de objetos la expreso mediante diferencia de signos".[13] Así, pues, hay *prima facie* una seria dificultad con el uso filosófico de la noción de identidad (digamos, el asociado con el principio de identidad de los indiscernibles) y es claro que dicho principio es relevante para la tesis de los materialistas. No obstante, no quisiera basar mi argumentación sobre esta idea de Wittgenstein por la sencilla razón de que entonces bloqueamos la problemática misma. Concedamos, pues, que no hay una dificultad de raíz y concentrémonos en el otro aspecto de la tesis, *viz.*, la contingencia de la conexión (la identidad) que se establece.

La objeción de Malcolm es clara: dejando de lado las propiedades modales e intencionales de los objetos (*i.e.*, las "propiedades" atribuidas a los objetos mediante descripciones no identificadoras, como cuando digo que un objeto tiene la propiedad de ser conocido, descrito, explicado, pensado, imaginado, etc. por mí de tal o cual modo), las genuinas propiedades del objeto deben aparecer independientemente de cómo se le describa. Supongamos que Juan es el ciudadano más connotado y que Juan es el hijo de mi vecina. Se sigue que el hijo de mi vecina tiene la propiedad de ser el ciudadano más connotado. O sea, los objetos **Juan** y **el hijo de mi vecina**, dado que son uno y el mismo, tienen la propiedad de ser el ciudadano más connotado. En cambio, el que Juan sea conocido por mí o no es una propiedad intencional de Juan y, por lo tanto, no es una genuina propiedad. Lo que Malcolm objeta es que lo mismo debería suceder con cualquier estado "mental" que fuera idéntico a algún estado cerebral y el problema es que ello no sucede. Los "objetos" en cuestión tienen propiedades diferentes. Por ejemplo, los primeros no ocupan espacio, en tanto que los segundos sí. Ya vimos que Smart cuestiona esto. Es tiempo, pues, de determinar si su respuesta (bastante tortuosa, dicho sea de paso) es aceptable o no.

13 *Ibid.*, 5.53.

Smart afirma que la objeción surge de una equivocación: lo que es físico es la experiencia, no el contenido de la experiencia (digamos la experiencia del pensar, el acto o estado mismo de pensar, ubicable temporalmente, no acerca de lo que es). La respuesta es insatisfactoria por lo menos por dos razones:

a) la admisión de que hay imágenes mentales, proposiciones, etc., es incompatible con el materialismo que Smart defiende. Más importante, empero, para nuestros propósitos es la otra razón.
b) la objeción de Malcolm sigue sin ser respondida. Para salir del problema, Smart tendría que hacer ver que un mismo objeto, identificado de dos modos distintos (en contextos extensionales), tiene propiedades distintas **por el mero hecho de ser descrito o pensado de modo diferente**. Pero eso es algo que él dista mucho de haber mostrado. Si efectivamente un estado mental y uno cerebral son una y la misma cosa, por más que difieran en cuanto a la modalidad de acceso a ellos (propiedades intencionales), lo que valga para uno debería valer para el otro. Sin embargo, hablar de la ubicación espacial de un pensamiento, una imagen, etc., es emitir un sinsentido.

Infiero que está aquí esbozada una linea de crítica en contra de la teoría de la identidad que es potente y que los materialistas no han neutralizado. En particular, sirve para refutar también la cuarta objeción discutida por Smart.

La segunda crítica de Malcolm requeriría mucho espacio para ser presentada debidamente. De modo inequívoco, está inspirada (por no decir extraída) de los escritos de Wittgenstein (en especial de *Zettel*). El núcleo de la objeción es el siguiente: un pensamiento es algo que requiere de un contexto para que pueda gestarse. Un pensamiento no es un objeto lingüístico, mental o lógico. La pregunta orientadora acerca del pensamiento no es '¿qué es un pensamiento?', sino '¿cuándo decimos de alguien que emi-

tió un pensamiento?' Parte de la respuesta es, evidentemente, que se requiere de un contexto. La diferencia entre 'Hola que tal' dicho por una grabadora y dicho por una persona es que en el primer caso sólo por cortesía, similitud, etc., podemos hablar de un pensamiento. Se le aplica la noción, si se le aplica, de forma derivada. La persona no repite la frase a tontas y a locas, como lo hace la grabadora cada vez que se enciende, sino en las circunstancias apropiadas (hablo desde luego de casos normales). Sin esas circunstancias, el signo complejo (el nombre complejo, como diría Frege) 'Hola que tal' no dice nada, no expresa un pensamiento.

Ahora bien, si el contexto es esencial para la gestación de los pensamientos, entonces se vuelve a aplicar la linea de argumentación anterior: hay propiedades o características que tienen los "procesos de pensamiento" (válgaseme la expresión) que no tienen los procesos cerebrales. En particular, estos últimos no requieren de ningún contexto para que se produzcan. Dado que el argumento vale para toda clase de estado, proceso, evento, etc. mental, podemos deducir que procesos cerebrales y procesos mentales no son lo mismo (*i.e.*, no son, como quieren ciertos filósofos, idénticos).

Otro argumento relacionado con esta linea de crítica es el siguiente: si realmente un estado mental fuera un estado cerebral, entonces sería lógicamente posible que, excitando del modo adecuado las neuronas de, por ejemplo, el hombre de Cro-Magnon, éste pudiera pensar (estar en el estado mental de pensar) en, *e.g.*, la utilidad de las computadoras. Pero esta suposición, validada por la tesis de la identidad, es claramente absurda. La razón es la siguiente: el pensamiento es algo cuya identidad depende de algo externo al sujeto cognoscente. Por lo tanto, nada de lo que pueda pasarle a este último puede en principio generar un pensamiento si no tuvo la(s) experiencia(s) relevante(s) para ello. Salvo en el caso de las sensaciones, la tesis de que todos los

estados de conciencia, actos de pensamiento, etc., pueden ser identificados con procesos o estados mentales conduce a absurdos como el mencionado más arriba. La tesis, por consiguiente, no se sostiene.

Consideremos rápidamente ahora la igualmente devastadora crítica desarrollada por Kripke en la última de sus conferencias del texto conocido como 'El Nombrar y la Necesidad'. Yo creo que su crítica tiene dos partes, complementarias pero lógicamente independientes e igualmente demoledoras. La primera, que podemos presentar de modo sucinto, recoge una inquietud que ya había sido expresada por Malcolm en relación con el carácter supuestamente contingente de la identidad que se afirma, sólo que ahora es elaborada de un modo más formal (Dicho sea de paso, se trata de una objeción inspirada en Russell, puesto que parece ser una generalización en términos de mundos posibles de un argumento russelliano aplicado a los nombres propios en sentido lógico). Sean, pues, M y C dos designadores rígidos, el primero de los cuales es un nombre de un estado mental en tanto que el segundo lo es de un estado cerebral y sea la afirmación 'C = M'. Por definición, ambos designadores nombran o denotan al mismo objeto en todos los mundos posibles. Ahora bien, los enunciados de identidad, como muchos otros, se dividen en contingentes y necesarios. ¿Cuándo nos las habemos con un enunciado contingente de identidad? Sólo cuando por lo menos uno de los designadores no es rígido (*i.e.*, cuando fijamos la referencia de un nombre por medio de una descripción, puesto que nunca será auto-contradictorio que la descripción no se aplica al objeto nombrado). Pero cuando los dos designadores son rígidos, se puede mostrar que, puesto que no se está describiendo objetos, los enunciados de identidad sólo pueden ser necesarios. Así que, contrariamente a lo que sostienen los materialistas reduccionistas, la identidad entre estados cerebrales y mentales no puede ser contingente, sino que tiene que ser necesaria. Afirmar, em-

pero, que necesariamente un estado mental es un estado cerebral es palpablemente falso. Si este argumento es válido, la tesis de la identidad queda refutada.

Consideremos ahora el segundo argumento o la segunda parte de la argumentación de Kripke. El problema es el siguiente: Kripke reconoce que hay enunciados de identidad en los que aparecen descripciones o designadores no rígidos pero que, a pesar de ello, dan lugar a enunciados necesarios. Son enunciados necesarios *a posteriori.* Son enunciados así muchos enunciados científicos, como por ejemplo la aseveración de que el calor **es** la aceleración de las moléculas, el agua **es** H_2O, el lobo **es** el animal con tales o cuales cromosomas, etc. (Kripke, recordemos, intenta mostrar que los objetos tienen propiedades esenciales). Introduzcamos ahora la noción de **situación de evidencia idéntica o diferente**. Así, es perfectamente posible que alguien que está observando a lo que él llama un 'tigre' afirme con verdad que ese animal tiene tales y cuales cromosomas, pero también podría suceder que otro observador en una situación diferente pero cualitativamente semejante se equivocara porque afirmara del animal que contempla que es un tigre puesto que tiene, junto con todas sus propiedades fenomenales usualmente asociadas con los tigres, cromosomas diferentes. En la medida en que las características fenomenales pueden ser reemplazadas por descripciones científicas, esta situación de error en la identificación de objetos está abierta.

Llegamos ahora al núcleo del argumento: así como es lógicamente posible tener un acceso cognitivo contingente a una propiedad esencial de un objeto vía sus propiedades fenomenales, es imposible que se produzca una situación semejante en el caso de los estados mentales. En este segundo caso, no hay la posibilidad de un doble acceso al estado mental en cuestión. Supongamos que alguien afirma que el dolor (D) no es sino la estimulación de las fibras C(M), es decir, que D = M. Dado el argumento

anterior, no podría ser el caso de que hubiera dolor sin que fuera al mismo tiempo estimulación de fibras C. Así, pues, para que la identidad establecida por el teórico de la identidad pudiera ser contingente, tendría que existir la posibilidad de que el dolor, que quedó identificado contingentemente con M, pudiera ser identificado **de otro modo** que como dolor. Sólo en un caso así, la identidad podría ser contingente. Pero ¿cómo podría el dolor ser identificado de otro modo que como dolor, esto es, con las características fenomenológicas que de hecho tiene? No hay una situación epistémica tal que permita percibir un dolor de otro modo que como dolor. La posibilidad que existe en el caso de las verdades de la ciencia no existe en el caso de los estados mentales. Por lo tanto, o bien la identidad en cuestión no puede ser contingente o bien no hay tal identidad.

Resumiendo: no es posible que la identidad de un estado mental con uno cerebral, expresada (como debe ser) por medio de dos designadores rígidos, pueda ser contingente. Contrariamente a lo que **tienen** que sostener los teóricos de la identidad, dicha identidad sólo puede ser necesaria. Pero ¿cómo pueden dos objetos, que son el mismo, ser idénticos a sí mismos de modo puramente contingente? Si Kripke tiene razón (y por mi parte simpatizo con lo que dice), la tesis de la identidad es incoherente.

Creo que podemos sostener que hay elementos suficientes que nos permiten afirmar con relativa confianza que la tesis de la identidad es imposible de defender. Debemos ahora pasar a considerar una posición alternativa. Tengo en mente el interaccionismo, cuyo principal representante es probablemente Karl Popper.

- III -

Dos son los textos que tomaré en cuenta para reconstruir y criticar las tesis de Popper:

a) el artículo 'Language and the Body-Mind Problem' y
b) las secciones relevantes del libro *The Self and its Brain,* del cual es co-autor.

Son en verdad muchas las objeciones que, si no me equivoco, se podrían elevar en contra de lo que Popper afirma. A diferencia de lo que sucede con la gran mayoría de los filósofos a quienes critica, él es, por ejemplo, sumamente descuidado al hablar: intercambia términos que no son equivalentes ('ego' y 'mente'), se contradice (por introspección sabemos que somos **uno** y al mismo tiempo el yo no es un ego), concluye con demasiada precipitación, es vago (de **algún modo** la conciencia y el cerebro están conectados, etc.). Empero, no es mi objetivo ni mi interés analizar críticamente a Popper. A lo único que aspiro es a presentar y discutir la posición que él supuestamente representa en relación con el problema que aquí nos incumbe.

Tal vez lo más conveniente sea exponer a la luz los fundamentos de su teoría. Lo que está a la base de esta última es una cierta metafísica, una cierta concepción de la realidad, *viz.*, la idea de que la realidad se divide en tres grandes reinos del ser: el Mundo 1, de los objetos y procesos físicos, el Mundo 2, de los objetos y actividades mentales, y el Mundo 3, de los objetos lógicos, significados, opiniones, teorías, argumentos, verdades, etc. Aunque son diferentes (no todos son abstractos, por ejemplo), estos tres mundos son igualmente objetivos, irreducibles y autónomos. Y la idea de Popper (que da la impresión de ser un argumento sacado directamente de la obra de Platón) parece ser que si no postulamos dichos mundos, entonces sencillamente no podremos explicar satisfactoriamente ni la acción humana ni el cambio en general. El mundo físico y el cambio dentro de él se explican sin tener que apelar a otra dimensión de ser, pero mucho del cambio que tiene lugar se realiza por la intervención o mediación del hombre. Esto es así porque el hombre piensa, apre-

hende verdades, descubre cosas (entidades abstractas) que luego se aplican al mundo físico. Claramente, por lo tanto, se produce un proceso de interacción entre el Mundo 3 y el Mundo 1, y, de igual manera, entre el Mundo 2 y el Mundo 3. Lo que Popper desea sostener es que también hay interacción entre los Mundos 1 y 2. Ahora bien ¿cómo se realiza eso?

Lo primero que hay que decir es que Popper se declara enfáticamente en contra de toda idea de sustancia, no sólo física, sino también mental. "...no me parece que haya razón alguna para creer en un alma inmortal o en una sustancia psíquica que puede existir independientemente del cuerpo".[14] De un modo más claro, en otra parte afirma lo siguiente: "Lo importante de todo esto es que no estoy de acuerdo con la teoría del 'yo puro'. El término filosófico 'puro' se le debe a Kant y sugiere algo como 'previo a la experiencia' o 'exento de experiencia'; y de este modo el término 'yo puro' sugiere una teoría que creo que es equivocada: que el ego estaba allí previamente a la experiencia, de modo que toda experiencia está, desde el principio, acompañada por el 'Yo pienso' cartesiano y kantiano. (...). En contra de esto, yo sugiero que ser un yo es en parte el resultado de disposiciones innatas y en parte el resultado de la experiencia, especialmente de la experiencia social." [15]

Nada de esto, evidentemente, cancela la realidad de los actos y las facultades mentales. Lo que de todo esto podría inferirse es que la interacción, se de como se de, no se dará entre sustancias, sino estados mentales y estados físicos. Sin embargo, ello no es tan simple, por la sencilla razon de aunque no sigue a Eccles en todo, Popper de todos modos parece aceptar que hay algo que surge de la compleja organizacion de la materia y sus

14 K.R.Popper y J.C.Eccles: *The Self and its Brain*. Springer International. Springer-Verlag, Berlin, 1977; p.146.

15 K.R.Popper *Ibid.*, p.111.

actividades y que ordena e integra las experiencias de una persona en un todo unificado. Respecto al punto crucial de cómo se de dicha relación Popper, como veremos, tiene mucho cuidado en no decir absolutamente nada. (Dicho sea de paso, me resultó imposible comprender cómo y por qué Popper admite la existencia de estados mentales. Lo que él dice es únicamente que él "conjetura"[16] que los hay. Posteriormente, empero, simplemente los da por supuestos). Lo que por lo pronto emerge de sus escritos como importante a este respecto es:

a) aprehendemos verdades, significados, etc.
b) no hay un órgano intelectual, sino una facultad, que se desarrolla en el tiempo
c) la conciencia está íntimamente ligada al lenguaje
d) somos yoes y nos experimentamos como tales, sin por ello ser sustancias.

Creo que estos son los elementos centrales de su doctrina. Veamos ahora cómo se desarrolla ésta.

No me parece en lo más mínimo descabellada la sugerencia de que hay ciertos puntos en relación con los cuales Popper debería haber dicho **algo**. Por ejemplo, cabe preguntar: ¿cómo se efectúa este proceso de "aprehensión de verdades" y cómo sabemos que, siendo esta facultad falible, no nos equivocamos permanentemente y confundimos falsedades con verdades? En algún lugar él dice: "sugiero que es más fácil comprender como hacemos objetos del Mundo 3 que comprender cómo los comprendemos, los aprehendemos o los 'vemos'".[17] Su respuesta es entonces, dicho de modo brutal, un simple 'No sé'. Lo único que se puede hacer es "conjeturar" (noción apropiada para reem-

16 *Ibid.*, p.36.
17 *Ibid.*, p.44.

plazar a la de postular, que puede suscitar animadversión) que hay "intuiciones intelectuales". Me parece que ni el más fanático de los popperianos podría negar que todo esto es altamente insatisfactorio. El problema es, como iremos viendo, que este esquema de razonamiento se reproduce de modo alarmante en sus escritos y, en particular, en puntos cruciales.

Quizá debamos preguntar cómo caracteriza Popper al "yo". Recuérdese que, generosamente, él nos hace saber que "desea enunciar claramente y sin ambigüedades" que él "está convencido de que los yoes existen".[18] Este aspecto de su trabajo es, probablemente, el más original de su labor en esta área de la filosofía. Según él, de lo que se parte es de un organismo, esto es, una cosa física, objeto de estudio de la biología, ciencia que para ciertos efectos tiene respecto a él la última palabra. Este organismo, empero, está dotado de ciertas facultades que permiten el surgimiento de actividades cognitivas, siendo tal vez la más importante de todas la facultad de hablar. Siguiendo en esto a Chomksy (aunque sin citarlo), Popper distingue entre capacidad y aprendizaje (lo que Chomsky más bien distinguiría como 'competencia' y '*performance*'). La capacidad está programada genéticamente, pero el aprendizaje es un asunto empírico. "Así, el aprendizaje del lenguaje es un proceso en el que disposiciones basadas genéticamente, que han evolucionado mediante selección natural, de algún modo se traslapan e interactúan con un proceso consciente de exploración y aprendizaje, basado en la evolución cultural. Esto apoya la idea de una interacción entre los Mundos 3 y 1".[19] El organismo lingüístico en cuestión, en tanto que cuerpo, interactúa con otros y, en tanto que agente con facultades mentales y cognitivas, con objetos del Mundo 3. Es de este estar anclado en los tres mundos que surge el yo. "Me

18 *Ibid.*, p.101.
19 *Ibid.*, p.48.

parece de considerable importancia el que no nazcamos como yoes, sino que tengamos que aprender que somos yoes; de hecho tenemos que aprender a ser yoes. Este proceso de aprendizaje es un proceso en el que aprendemos acerca del Mundo 1, del Mundo 2 y, en especial, del Mundo 3".[20] El problema ahora es el de determinar cómo se identifican o se individualizan los yoes. Esto involucra el problema de la naturaleza de la conciencia y de la relación con sus bases físicas o, más precisamente, neurofisiológicas.

Popper asume que el yo tiene un "fundamento" físico. "Este", nos dice, "parece centrarse en nuestro cerebro".[21] Hay una relación **fuerte** entre el yo y el cerebro, pero la naturaleza de la relación no es clara, puesto que puede haber mutilaciones cerebrales que lo afectan y mutilaciones que lo dejan intacto. Y aquí es donde Popper ofrece lo que, hasta donde yo logré ver, es su único argumento en favor del interaccionismo. Veamos cómo queda articulado.

La idea es la siguiente: podemos imaginar un trasplante cerebral total. Un trasplante así equivaldría a un trasplante de la mente o del yo. El cuerpo, por lo tanto, no es identificable con la mente. En todo caso, lo sería el cerebro. Pero este tampoco es idéntico con una mente. Para la identidad de la persona, cuerpo y mente son lógicamente independientes y la mente, más no el cuerpo, es crucial o decisivo para su identidad. El cuerpo funciona sólo mientras se asume que el cerebro sigue siendo el mismo. "Así, en condiciones *ordinarias* podemos considerar la identidad del cuerpo como un criterio de la identidad de la persona, y del yo. Pero nuestro experimento de pensamiento, el trasplante (que espero que nunca sea efectuado sobre un ser humano), muestra que la identidad del cuerpo es un criterio sólo

20 *Ibid.*, p.109.
21 *Ibid.*, p.115.

en la medida en que implica la identidad del cerebro; y el cerebro a su vez juega este papel solo porque conjeturamos su *'liaison'* con la mente, porque conjeturamos que debido a esta *liaison* el cerebro es el portador de la auto-identidad de la persona".[22] Así, pues, en última instancia toda la grandilocuente labor de Popper se reduce a una inmensa petición de principio.

Hay un punto muy importante que queda por considerar, *viz.*, el de la naturaleza de la conciencia. Desafortunadamente, lo que Popper tiene que decir en relación con el tema es excesivamente oscuro. Su opinión es que la conciencia está ligada, primero, al cerebro y, segundo, al lenguaje **en sus funciones descriptivas**. Por otra parte, la conciencia no es identificable ni reducible ni explicable en términos de sensibilidad y actividad intelectual, como ya vimos. La conciencia, por lo tanto, aparece como algo intermedio y unificador de funciones biológicas y actividades cognitivas, fundadas en el lenguaje. No podrá sorprendernos, por consiguiente, que Popper en repetidas ocasiones hable de la conciencia como de "el más grande de los milagros".[23] Ahora bien, este "milagro" no tiene nada de misterioso. Para dilucidarlo, lo que hay que hacer es conjugar una perspectiva evolucionista con el conocimiento de la formidable actividad cerebral. Y entonces, según Popper, se vuelve "comprensible".

Por otra parte, sería un error identificar a la conciencia con el yo. El yo no tiene interrupciones o cortes, la conciencia sí. Esta última es más bien la vía de tránsito entre lo mental y lo físico y entre lo mental y los objetos del Mundo 3, en tanto que el yo es sencillamente el organismo que se incrusta en los tres mundos simultáneamente. Esto, empero, no pasa de ser una mera conjetura mía, puesto que en ocasiones Popper se permite intercambiar las expresiones 'mente' y 'yo', creando con

22 *Ibid.*, p.118.
23 K.R.Popper. Véanse, por ejemplo, pp.129 y 144.

ello un auténtico caos y la imposibilidad de comprender su pensamiento.

Intentemos ahora, en unas cuantas palabras, atrapar el contenido del mensaje popperiano. Por vías más bien dudosas, Popper rechaza tanto el conductismo como el materialismo. El lenguaje, por lo tanto, es decisivo en su concepción. Desafortunadamente, tiene poco que decirnos acerca de él. No está en lo más mínimo claro si, de acuerdo con él, las operaciones de la mente están estructuradas por el lenguaje o no, pero él más bien sugiere que hay actitudes críticas, de teorización, comprensión, etc., independientes del uso del lenguaje (el yo puede pensar musicalmente, figurativamente, etc.). Si esto es así, podemos entonces preguntar:

1) ¿cómo se relacionan el pensamiento y el lenguaje y cómo puede la mente aprehender verdades que nos estén vestidas, por así decirlo, en palabras? Dicho de otro modo: ¿cómo se relacionan los Mundos 2 y 3?
2) Asumamos que, aunque compleja, hay una relación entre las actividades y funciones cognitivas, es decir de la mente o el yo (en una de las acepciones de la palabra), por una parte, y las actividades cerebrales, por la otra. ¿Cómo se relacionan? ¿Qué relación hay entre, *e.g.*, un deseo y los estados o funciones del cerebro? Sobre esto, como sobre la pregunta anterior, Popper no dice absolutamente nada.

Lo primero que hay que decir es, pienso, que ni el interaccionismo ni ninguna otra corriente de pensamiento puede quedar establecido sobre la base de un rechazo de otras posiciones. Esta, sin embargo, parece ser la estrategia de Popper. Por ejemplo, él sostiene cosas como la siguiente: "El viejo argumento de que cosas tan diferentes no pueden interactuar estaba basada en una teoría de la causación que ya fue desde hace

mucho superada".[24] Esta clase de argumentos por eliminación son en general sospechosos, sobre todo cuando, como en el caso en cuestión, no está para nada claro qué sostenía la teoría supuestamente desbancada, con qué se le superó y si la nueva propuesta es realmente tan efectiva como lo asegura su abogado. En segundo lugar, y más importante, es preciso señalar que el fundamento de las "conjeturas" de Popper, a saber, su concepción de la realidad y su visión del lenguaje, es no sólo endeble, sino abiertamente falso o, si se prefiere, filosóficamente inútil. Las charlas entre él y Eccles, recogidas en el libro común, dan en ocasiones la impresion de desembocar en auténticos cuentos de hadas (especialmente cuando Eccles habla). La concepción popperiana del lenguaje es realmente insensata e inútil. Es en parte esto lo que le permite a Popper mantener puntos de vista casi infantiles acerca del pensamiento, como algo lógicamente independiente de la manipulación coordinada de signos. Hay fuertes elementos de un misticismo trasnochado en mucho de lo que Popper insinúa (cuando habla, por ejemplo, acerca de la muerte, la belleza, etc.). Y, lo peor de todo quizá, cuando se plantea la cuestión de desarrollar la tesis interaccionista, Popper guarda un reservado silencio. De ahí que el único comentario que se me ocurre sea que si el interaccionismo tiene en Popper su mejor o más elocuente representante, entonces probablemente sea la menos defendible de todas las escuelas o corrientes en filosofía de la mente.

- *IV* -

Llego ahora a lo que en principio deberá ser la parte constructiva y aportativa del trabajo, esto es, la enunciación de mis propios pun-

24 K.R.Popper. "Language and the Body-Mind Problem" en *Conjectures and Refutations*. Routledge and Kegan Paul; Londres, 1963; p.298.

tos de vista, inspirados naturalmente no sólo en las discusiones aquí recogidas, sino pretendidamente basados en la labor filosófica de pensadores como L. Wittgenstein y N. Malcolm. No deseo descartar de entrada la idea de que puede haber problemas genuinos en relación con el par *<mente-cuerpo>* . Lo que sostengo es que si son genuinos, entonces son de carácter empírico. Lo que sí quiero es negar que hay problemas auténticos de carácter filosófico, además de ciertas elusivas (y, por otra parte, apasionantes) confusiones conceptuales.

Me parece que lo más conveniente será enumerar las "tesis" que deseo defender. En realidad, como podrá apreciarse con mayor facilidad posteriormente, no se trata de tesis en sentido estricto, sino de meras constataciones que, en el fondo, equivalen casi a trivialidades. La utilidad de un recordatorio de verdades que todos conocen consiste en que el tenerlas presente debería en principio evitarnos caer en el pantanoso terreno de la teorización filosófica. Las principales "tesis" son, pues, las siguientes:

1) Hay dos planos que es de primera importancia distinguir: el de la formación de conceptos y el de la explicación causal
2) Las explicaciones causales comportan o acarrean modificaciones semánticas
3) Hay una variedad de nociones de mente (así como no hay una única noción de materia) y, por ende, relaciones de diversa índole entre lo referido por predicados "mentales" y lo referido por predicados "materiales". Estas nociones, empero, son nociones técnicas, derivadas de la noción madre u original.

Abogaré en favor de estas tesis en el orden en que fueron enunciadas.

A) Conceptos mentales

Por razones que he dado en otra parte y que no repetiré, partiré de la idea de que 'mente' no es un nombre propio, es decir, su función en el lenguaje no es la de designar un objeto, a la manera como, *e.g.*, 'Juan' denota a Juan. Mediante 'mente' englobamos toda una variedad de acciones, lingüísticas y extra-lingüísticas, que asociamos de modo unificado con un cuerpo determinado. Huelga decir que no estoy sosteniendo que 'mente' **significa** ese conjunto de acciones, actividades, etc. Hablar de mente o de posesión de mente es hablar de un sinnúmero de prácticas y lineas de acción, sutilmente diferenciadas.

Hay un sentido originario o principal de 'mente' o, dicho de otro modo, hay un concepto *Mente* que es básico: el que los hablantes normales empleamos. Sinceramente, no creo que nadie quiera sostener que disponemos de una noción de mente que sea lógica y epistemológicamente previa a la incorporada en el uso del lenguaje natural. De ahí que resulte difícil de negar que el problema "mente-cuerpo", consista en lo que consista, no podrá resolverse de modo satisfactorio si no se tiene una aprehensión perspicua del concepto Mente (asumo que no hay mayores problemas con el concepto *cuerpo*). Por consiguiente, la pregunta con la que debemos iniciar nuestra labor debería ser: ¿para qué sirve dicho concepto, es decir, qué nos permite hacer? En terminología wittgensteiniana: ¿qué utilidad real tiene en nuestras vidas el juego de lenguaje de lo mental?

A mí me parece que la función central de *Mente* y sus derivados es la de recoger lo que no es físico de una persona pero que, no obstante, es de algún modo objeto de experiencia por parte de otros. Preguntémonos ¿qué es lo no físico de una persona y que, no obstante, es experimentable por parte de otros? La única respuesta que a mí se me ocurre es: sus acciones y reacciones **significativas**, es decir, los movimientos que, por estar insertos

en contextos determinados, están cargados de intenciones, significado, etc. Es el todo constituído por contextos y acciones que de algún modo aprendemos a diferenciar y clasificar y son ellos lo que nos lleva a hablar de sentimientos, emociones, pensamientos, creencias, deseos y demás. Dado que 'mente' es propiedad pública y no es el nombre de un **algo**, lo experimentable aunque no físico no puede tratarse de algo interno, oculto a los demás. No queda, pues, más que una opción: la de ver en lo mental algo que es esencialmente de carácter social. Esto me parece comprensible: los usuarios del lenguaje no se ven a sí mismos como meros cuerpos, sino como **personas**, como **seres humanos**. La razón de este modo de vernos no estriba en que estemos dotados de modo innato de un magnánimo sentimiento de solidaridad hacia el resto de nuestros congéneres, sino simplemente en que cada uno de nosotros, en tanto que usuarios del lenguaje, **tenemos** que considerar en algún sentido a los otros, esto es, a los co-hablantes, de modo semejante a nosotros, por más que no podamos concebirlos de exactamente el mismo modo como nos concebimos a nosotros mismos.

Esto me obliga a hacer otro recordatorio elemental: un rasgo inevitable del lenguaje natural es el peculiar desdoblamiento de nuestros conceptos de experiencia. Hay "yo" y hay "el otro", que puede ser "tú"o "él" ("ella"). Yo siento (de eso no tengo dudas), pero como tu eres tan hablante como yo, entonces no me queda más remedio que considerar que tu eres (en algún sentido fundamental) como yo y que, por lo tanto, tú también sientes. Ahora bien, a pesar de esta igualdad de planos, la semejanza impuesta por el carácter colectivo del lenguaje no conduce a uniformidades en cuanto a los usos, sino que tiene sus matices y límites y los marca de modo ineludible. El lenguaje natural está estructurado de tal modo que me impide decir 'yo sientes' para decir que yo siento lo mismo que tu, ni tiene sentido decir 'tu siento', para indicarle a alguien que tu sientes lo mismo

que yo. El hecho es que todos sentimos lo mismo cuando hablamos de lo mismo. Esto en un sentido es trivial, pero en otro sentido puede ser aclaratorio. Lo que estoy diciendo es que la forzosa asimetría lingüística recién mencionada no acarrea cambios de significado. Así, cuando yo hablo de mi dolor **no** hablo de algo distinto que cuando hablo de tu dolor o cuando otro habla de su dolor. 'Dolor' tiene siempre el mismo sentido, sólo que se predica de modo distinto en función de las respectivas posiciones ocupadas por los hablantes o usuarios de uno y el mismo sistema compartido de comunicación. Y lo que vale para *dolor* vale para el todo del lenguaje de "lo mental". Todos comprendemos lo que es la adscripción y la auto-adscripción de estados mentales y sólo los filósofos ven en ello un misterio.

Recapitulando: lo que la asimetría señalada si implica es una diferencia en cuanto al modo de predicar estados mentales, en función de si se trata de la primera persona o no. En el caso de las segunda o tercera personas recurrimos a **criterios** para la adscripción de estados mentales; en nuestro propio caso la auto-adscripción es más bien el resultado de un proceso de reemplazos de reacciones naturales o espontáneas por su correspondiente formulación lingüística.

En el caso del concepto que podríamos llamar *Mente* 1, es decir, el concepto normal de mente, la aplicación a otros es claramente de orden conductista, en el sentido más amplio posible del término, esto es, de modo que incluya no sólo acciones y reacciones, sino también trasfondos, contextos, costumbres, etc. Lo que está asociado con la mente, en este primer sentido, es el cuerpo, pero no el cuerpo en tanto que posible objeto de estudio de la física, sino el cuerpo del hablante, es decir, del ser social con el que se interactúa. Así, decir que alguien cree algo es como abrir una caja de switches, cada uno apuntando a una acción posible. Por ejemplo, si digo

que A cree que México firmará un Tratado de Libre Comercio, puedo querer decir cosas tan diversas como:

1) que si le preguntan a A si cree que *p*, A asentirá
2) que A actúa desde ahora en concordancia con p, es decir, toma medidas de orden financiero, comercial, etc.
3) que en las pláticas A afirma constantemente que *p* acaecerá
4) que si alguien quiere apostar con él, A estaría dispuesto a hacerlo en favor de que *p*, etc.

Es importante comprender que el conjunto de lineas de acción posibles, lingüísticas o extra-lingüísticas, es un conjunto **difuso**, quizá infinito. Este asunto es indecidible *a priori*. En todo caso, es muy importante darse cuenta de que el lenguaje de lo mental sirve aquí para hacer **descripciones** (atribución y auto-adscripción de estados mentales), las cuales se efectúan gracias al ofrecimiento de **razones**. Aquí es importante comprender que el lenguaje no determina nada de modo nítido.

Mientras nuestra vida se desarrolle dentro del marco generado por el lenguaje natural, pienso que lo mental queda acotado como dije. Parte del problema surge cuando, paulatina e imperceptiblemente, introducimos en nuestro lenguaje términos teóricos. Voy a considerar rapidamente el caso de la noción neurofisiológica de mente.[25]

25 Por otra parte, puede uno confundirse y pensar que el 'porque' de las adscripciones normales de estados psicológicos es el 'porque' de las explicaciones científicas. Alguien, en efecto, puede querer afirmar que cuando decimos 'A hizo x *porque* creía que *p*' estamos dando la **causa** de su acción. Pero esto es un error y ello puede mostrarse. Obsérvese que las causas parecen excluirse de un modo como las razones no parecen hacerlo. La prueba de ello es que el que alguien **demuestre** que A hizo *x* porque **creía** que X (es decir, ofrece una creencia como razón de su acción) no impide en ningún sentido que alguien afirme significativamente que A hizo x porque se **imaginó** que Y. En cambio,

B) Causas y significados

No sorprenderé a nadie, supongo, si afirmo que el correlato evidente de la mente para los neurofisiólogos es no el cuerpo, sino el cerebro. Lo que es importante detectar, empero, es que los científicos, que son ante todo forjadores y manipuladores de conceptos, no pueden evitar hacer en relación con el cerebro lo que en el lenguaje natural se hizo en relación con el cuerpo. En el fondo, a lo que ellos aspiran es a construir un nuevo concepto de mente, *Mente* 2, idéntico o casi al concepto originario, esto es, *Mente* 1, pero asociado ahora con el cerebro y sus funciones, en lugar del cuerpo y sus movimientos. Independientemente de que ello se logre o no, el intento representa un cambio conceptual drástico, pero también un salto semántico en el vacío, entre otras razones porque en este nuevo nivel semántico no se ve que haya nada peculiar en la adscripción o auto-adscripción de estados mentales que lo distinga del nivel anterior: si el neurofisiólogo se auto-atribuye un deseo, inclusive si se observa durante una operación de su cerebro, la auto-adscripción se logra apelando al concepto normal de deseo. No hay otro. Por el lado de "lo mental" no hay, pues, cambio. El problema es que sí lo hay por el lado de los presupuestos de aplicación de los conceptos mentales, puesto que el neurofisiólogo pretende estar estableciendo una nueva correlación entre su antiguo concepto de deseo y "lo otro", que en este caso sería el cerebro o las

---

si alguien demuestra que la **causa** de que A haya hecho *x* es X, en la medida en que encuadra la acción dentro de un contexto de leyes o reglas generales, automáticamente anula el que otra persona pueda apuntar otra causa como explicación de la acción de A. Es, pues, claro que, en lugar de causas, lo que en general damos son las **razones** de las acciones. Como asunto de uso normal del lenguaje, me parece más apropiado hablar de "causas de acción" cuando aludimos a estados físicos que afectan al sujeto (*e.g.*, lo hizo porque lo atropellaron, porque le cayo un rayo, etc). En lo que sigue, me ajustaré a esta "convención".

neuronas o las funciones de sistema nervioso central, etc. Se trata, pues, de intentar una aplicación **nueva** de un concepto viejo por medio de una aparentemente inocua sustitución de contexto. Yo estoy convencido de que eso es imposible. La sustitución en cuestión automáticamente acarrea consigo una transmutación del concepto original. Cabría suponer que, dado que el nuevo correlato de *Mente* 1 cambió, lo que de hecho se nos está dando es un nuevo concepto de mente, digamos *Mente* 2, semejante quizá pero lo suficientemente diferenciado del original de modo que no entrañara confusiones conceptuales. El problema es que no se nos está dando nada nuevo, puesto que este nuevo concepto no ha sido construído todavía y por otra parte, como ya dije, el neurofisiólogo pretende seguir hablando de sus creencias, esperanzas y demás **como antes**. Pero hay diferencias decisivas respecto a la situación inicial. En primer lugar, se ha conformado un nuevo contexto para la atribución de estados mentales en un terreno impropio para ello, *viz.*, el de la neurofisiología, constituido por las neuronas y sus "actividades" (y no, como antes, la conducta contextualizada del individuo) y, en segundo lugar, lo que con esta nueva correlación se pretende es proporcionar no ya razones, sino **causas** y no son ya meras descripciones lo que estaría en juego, sino **explicaciones**. Pero el cambio contextual aunado al cambio de operación lingüística que se lleva al cabo implica inmensos cambios de significación tan grandes que de hecho equivalen a un cambio de temática, el cual queda encubierto tan sólo por el perseverante uso de las mismas palabras. La interpretación filosófica de la labor del neurofisiólogo, por ello, no pasa de ser un espejismo semántico, puesto que no está en lo más mínimo claro qué relación hay o vale entre descripciones de la forma 'A cree que *p* porque lo leyó' y explicaciones de la forma 'A cree que *p* porque esta en el estado cerebral C'. De lo que sí estamos seguros es de que no dicen lo mismo. Intentemos justificar esto último.

A mí me parece claro que entre la descripción de la creencia y la explicación causal de la creencia no puede haber ninguna relación de reducción semántica, de implicación, de presuposición, etc. Es posible que exista alguna relación entre ellas, pero aparte de ser contingente es muy poco lo que podemos decir. Ahora bien, este abismo semántico algo debe estar indicando. A mi juicio, lo que indica es, simplemente, que nos las habemos, por una parte, con una noción ya formada y, por la otra, con algo que parece ser una noción en vías de formación. Por consiguiente, si lo que se pretende es poner en relación la *Mente* 1 con el cerebro y sus funciones, a lo que directamente se va es al fracaso, puesto que se está buscando una conexión que no sólo no existe, sino que no puede existir, a menos de que se nos de *Mente* 2, lo cual no ha sucedido todavía. La justificación de esta predicción es, me parece, obvia: *Mente* 1 **ya está conectado con algo, a saber, el cuerpo**. Pretender ahora, de modo abrupto, conectar ese concepto con algo distinto es transformarlo, transmutarlo. No es que sea difícil establecer dicha conexión: es simplemente que no tiene el menor sentido intentar hacerlo.

Por lo que he estado diciendo, creo (aunque no estoy seguro de ello) que entiendo lo que los filósofos a la Smart pretenden. En mi opinión, inquietudes como las de él se resolverían en principio del siguiente modo: se trataría de extraer, partiendo de la descripción del cerebro y sus funciones, es decir, partiendo de la teoría del cerebro, **un nuevo concepto de mente**, esto es, lo que yo llamé *Mente* 2. Por el momento, si no me equivoco, se carece de dicho concepto. Lo que es claro es que éste, independientemente de qué forma finalmente revista, habrá quedado conformado **del mismo modo como quedó conformado el concepto originario**: habrá un uso en primera persona y uno en tercera, desde luego con contextos diferentes, expectativas distintas, etc. Si ya se hubiera construído el concepto teórico *Mente* 2, entonces sí se podría pasar a intentar establecer alguna clase

de relación entre él y *Mente* 1. Sólo entonces podría intentarse establecer alguna conexión sistemática entre la *Mente* 1 y el Cuerpo 2, esto es, el cerebro. Esto, empero, me parece practicamente un sueño.

Desde mi punto de vista es absolutamente crucial, para quedarnos tranquilos frente a este gran pseudo-problema de la filosofía, no olvidarse de varios resultados wittgensteinianos. Dos, por lo menos, son particularmente pertinentes para nosotros:

1) lo "mental", entendido como sea, surge sólo cuando hay manipulación de un simbolismo, esto es, de un lenguaje, y
2) el lenguaje es esencialmente de carácter social.

Apliquemos rápidamente esto, en conjunción con lo que hasta aquí hemos dicho, a la ciencia-ficción de las máquinas pensantes.

Si en efecto los resultados de Wittgenstein son correctos, entonces se vuelve evidente que es un error garrafal buscar el pensamiento, creencias, deseos y demás en **una** entidad. Por sofisticada que sea la computadora, el mero análisis de ella no puede darnos lo que es el pensamiento o de cómo se relaciona con el *hard ware*. Imaginemos mejor esta otra situación, similar pero diferente. Disponemos de un **sistema** de máquinas, tales que una vez prendida una, ésta prende a las demás. Las máquinas están programadas de modo tal que se comunican, se dan instrucciones, responden, se rebelan, se equivocan, se corrigen, etc. Lo quiero defender es la idea de que es en esa interacción maquinal que surge el pensamiento o, mejor dicho, que aplicamos el concepto de pensamiento. El pensamiento no es, desde luego, un proceso, y menos aún algo concentrado en una de ellas. Por más que se analicen chips o se abran cráneos, no se encontrará allí "pensamiento", como no se encontrará significado en un signo, aunque se le examine con lupa. Es un error categorial, como bien lo dice Malcolm, pretender buscar el pensamiento en alguna parte.

Esta sugerencia pone de relieve con claridad, creo, la insuficiencia del materialista: dado su planteamiento y su enfoque, exactamente el mismo problema que él trataba de resolver en el caso de los seres humanos se le vuelve a plantear a las máquinas de su imaginada solución, inclusive concibiéndola del mejor modo posible. En efecto, siempre será posible preguntar ¿en dónde está el pensamiento de la máquina pensante: en el chip, en el circuito, etc., o en algo que "interactúa" con el *hard ware*? Es claro que no hemos avanzado ni un milímetro respecto a lo que era nuestro problema anterior. Lo único que se logró fue reemplazar una formulación de un problema por otra. Esta falla no implica, desde luego, ninguna ventaja para el mentalista internalista. La inutilidad de la propuesta de este último es tan obvia (en psicología, por ejemplo) y ha sido tan rebatida que no creo que valga la pena adentrarse en ella. A manera de diagnóstico general, creo que puede ofrecerse el siguiente: el materialista comete el error de darle prioridad a 'tú' o 'él' frente a 'yo', en tanto que el mentalista comete el error contrario: se olvida de que hay un uso legítimo de 'tú' o 'él' y trata de reducir todo al uso de la primera persona. En la medida en que ambos atentan contra el lenguaje natural, ambos están equivocados.

Quisiera recalcar que lo que he dicho no afecta o influye en demérito de la labor científica. Es un hecho que los científicos no están jugando o perdiendo el tiempo y la prueba de ello es que pueden manipular, vía alteraciones en las funciones del cerebro, en las neuronas, etc., al individuo. ¿Vamos a inferir que la neurofisiología es susceptible de manejar o manipular la vida **mental** de los individuos? La respuesta dependerá en gran medida de si dotamos a 'vida mental' de un sentido genuino o si más bien le adjudicamos un pseudo-sentido. Si se insiste en seguir viendo a lo mental como algo interno, privado, oculto, etc., entonces la pregunta misma ni se entiende y, por ello mismo (en caso de aceptarla como genuina), no nos libraremos de toda

clase de especulaciones, sabrosas pero estériles. Si, por el contrario, entendemos lo mental a la manera que lo he sugerido en este trabajo, como algo que, por así decirlo, tiene necesariamente una doble faceta o personalidad (*yo* y *tú* o *él*) y un carácter comunitario, entonces se nos aclara un poco la clase de cambio, factual y conceptual, que se podría estar gestando. Veamos esto con un poco más de detalle.

En primer lugar, hay que observar que la relación mente-cerebro parece ser, en ambos sentidos, una relacion "uno-muchos". Con un cerebro, en efecto, puede estar asociada conducta de lo más variado (*i.e.*, muchas lineas de conducta son compatibles con una y la misma conformacion neuronal) y, a su vez, con un estado mental dado puede estar asociada una multitud de estados cerebrales. Hasta ahora por lo menos, según sé, nadie ha demostrado lo contrario. La situación es, pues, poco alentadora. Por otra parte, empero, es claro que el neurofisiólogo puede, mutilando un cerebro, alterar drásticamente las reacciones, los estados mentales (excitación, irritación, intensidad de las sensaciones, etc.) de un sujeto. Esta facultad de manipular al individuo a través de alteraciones de su sistema nervioso sí puede producir cambios en nuestro orden conceptual: si la conducta normal del sujeto se modifica, entonces sencillamente ya no podremos aplicar de modo normal, *e.g.*, el concepto de creencia. Si alguien no se conduce como una persona que cree algo, es decir, no tiene las manifestaciones propias de quienes tienen creencias, entonces no podremos decir de él, porque no tendrá sentido hacerlo, ni que cree ni que no cree. Aquí el problema consiste en que no es posible determinar **cuántas** son las manifestaciones de las creencias, puesto que, por lo que dije anteriormente, el conjunto de criterios es difuso. Pero hay otro punto importante, a saber, que a lo más que puede aspirar el neurofisiólogo es a relacionar de modo sistemático estados o funciones neuronales con manifestaciones estrictamente físicas, esto es, no con conducta en el

sentido humano del término, sino con movimientos físicos o mecánicos. El neurofisiólogo no puede **alterar conducta**, en el sentido pleno de la expresión, por la sencilla razón de que la conducta es algo conformado en parte por elementos externos al sujeto de la conducta y por fuerte o eficaz que sea la neurofisiología de seguro que nunca podrá serlo al grado de englobar lo externo al sujeto. Lo único que podemos hacer es poner en relación sistemática funciones cerebrales, estados neuronales y demás con movimientos corpóreos generalmente asociados con la conducta correspondiente y gracias a la cual le atribuímos a alguien estados mentales. Cualquier otra pretensión me parece una inmensa (y cómica, si el asunto no fuera tan serio) ilusión.

Pero, de todos modos, preguntémenos: ¿qué se ganaría con esa nueva correlación sistemática entre, digamos, teoría de las neuronas y descripciones de movimientos físicos, posteriormente asociados con los factores que nos permiten hablar de estados o procesos así llamados 'mentales'? La respuesta es: mucho o nada, dependiendo de lo que se pretenda hacer con esas correlaciones. No se gana nada respecto a la **comprensión** de la **naturaleza** de lo **mental**. Sí se gana, en cambio, en cuanto a poder de manipulación de las condiciones o factores que permiten la atribución de estados psicológicos, sólo que estos son físicos. Pero si lo que se pretende es reemplazar el antiguo concepto por uno nuevo, más técnico, científico, etc., lo único que se estará haciendo será mutilar el concepto mentalista anterior y crear un auténtico enredo conceptual. En este caso, el supuesto nuevo concepto mental, es decir, lo que sería el auténtico concepto mental *M 2*, sólo surgiría si nos las hubiéramos con una multiplicidad de funciones de, *e.g.*, las neuronas, consideradas en, por ejemplo, redes y de modo tal que esas actividades no pudieran ser descritas por medio de la terminología físico-química. Quizá esté equivocado, pero puede sostenerse que el funcionamiento de las neuronas no pasa de ser, en última instancia, consumo de

glucosa y descargas eléctricas. Ello puede ser sumamente interesante, pero parece obvio que por minuciosa que sea la descripción de esas actividades, de ella no podrá brotar ningún concepto mental. *Mente* 2 es un fantasma semántico y *Mente* 1 no puede quedar directamente vinculado a ninguna teoría neuronal.

C) ¿Variedades de mente?

Quisiera terminar con unas cuantas consideraciones acerca de la noción de mente. Creo que podemos legítimamente hablar de una noción básica, a saber, la que todos los usuarios emplean. Esta nos permite distinguir entre movimientos mecánicos y movimientos "intencionales" y separar, así, a los seres hablantes o humanos del resto. Esto **no** quiere decir que el concepto de mente no pueda evolucionar. Lo que sucede es que para que evolucione, como sucedió con, por ejemplo, el conductismo, se tiene que elegir ciertos rasgos de las personas, rasgos que quedan perfectamente especificados en función de objetivos previamente determinados (*e.g.*, el estudio de la personalidad, de las reacciones, del ego, etc.). Una vez determinados los rasgos que nos interesan de las personas, se articula un nuevo aparato conceptual y se inicia el proceso de produccion de teorías, predicciones, etc. Ya no será el todo de la persona sino simplemente, *e.g.*, su conducta, sus actitudes, su coherencia, etc., lo que se tomará en cuenta. Así surgen y se desarrollan las diversas escuelas psicológicas más o menos bien establecidas. Aparentemente, con la neurofisiología sucede algo semejante: se elige un rasgo de la persona (su vida cerebral) en función de la cual se pretende estudiar su vida "mental". Pero eso es una confusión total, por la sencilla razón de que no hay ningún vínculo directo u obvio entre conceptos de experiencia y conceptos teóricos. No hay ni puede haber un nuevo concepto, un concepto neurofisiológico de mente, lo que llamé aquí *Mente* 2. En el fondo, el esfuerzo por

construirlo constituye una intromisión inaceptable de una ciencia en un ámbito que no le corresponde. Dicho del modo más breve posible: es tan posible un concepto neurofisiológico de mente como un concepto psicoanalítico de materia.

D) Conclusiones

En este trabajo intenté hacer una presentación sumaria, pero fidedigna, de dos influyentes posiciones con respecto al así llamado 'problema mente-cuerpo', *viz.*, la de los teóricos de la identidad y la de interaccionistas como K.Popper. Vimos que, aunque la discusión puede extenderse *ad infinitum*, los argumentos centrales de ambos bandos resultan ser increíblemente frágiles y comportan deficiencias de diversa índole (no siendo la menos importante de ellas la coherencia interna y la inteligibilidad misma de las tesis defendidas). Por otra parte, traté de ofrecer algo más que mera crítica y avancé algunas razones, que están todavía en estado embrionario y desde luego lejos de ser exhaustivas, en favor de la idea de que es sólo gracias a descripciones detalladas de cómo funcionan las diversas ramas de nuestro lenguaje (esto es, de cómo se aplican en ellas las palabras) que podemos librarnos de lo que no es sino un auténtico rompecabezas. Las descripciones que seamos capaces de proporcionar deberían en principio hacernos conscientes de lo que tiene sentido preguntar, de lo que se puede y no se puede afirmar, etc. y, de este modo, deberían evitarnos plantear preguntas como '¿Piensa una membrana?', ¿perciben las neuronas?', '¿Puede soñar una máquina?', etc. Lo que este enfoque revela es que, así como está, el "problema" mente-cuerpo es efectivamente imposible de resolver, pero la razón es que no se trata de un problema real, sino de una inmensa confusión conceptual, el resultado de una profunda incomprensión de la gramática de ciertos conceptos clave. Es relativamente claro que el problema no puede resolverse

por medios puramente *a priori* y tampoco puede recibir una solución "empírica". Más bien, lo que hay que hacer es disolverlo y ello se logra mediante la consideracion paciente y cuidadosa de los modos como empleamos nuestros signos, tanto en el lenguaje natural como en los lenguajes de las ciencias. Este, creo, era el enfoque favorecido por L.Wittgenstein. Como todos sabemos, en más de una ocasión Wittgenstein nos puso en alerta en contra de tendencias como las de la generalización precipitada, la consideración de una única y más bien escuálida clase de ejemplos y la asimilación de unos casos en otros. Parte de su labor, como él mismo decía, era enseñarnos a ver o detectar diferencias, más que semejanzas. En relación con el carácter de lo mental y de su relación con lo físico, estos llamados de atención son particularmente pertinentes. La distinción misma insinúa ya que lo que no cae dentro de un lado cae en el otro y esto hace pensar que los elementos de cada lado son objetos de una clase completamente diferente. En *Zettel* y en otros escritos de filosofía de la psicología, Wittgenstein mostró que esto es un grave error y que el lenguaje de lo psicológico es mucho más variado de lo que los filósofos quieren que sea. El hizo ver, por ejemplo, cómo la gramática de las palabras para emociones no es idéntica a la de sensaciones ni ésta a la de nociones como *tener intenciones, creer, tener esperanzas, pensar, comprender*, etc. Yo creo que él tenía razón, a pesar de que en la literatura filosófica estándar todas las nociones psicológicas, por haber quedado etiquetadas como "mentales", sigan recibiendo exactamente el mismo tratamiento. En los escritos de Wittgenstein está más que delineado todo un programa de una nueva filosofía de la psicología, esto es, una filosofía de la psicología radicalmente diferente de la prevalece en nuestros días. No es éste el lugar, empero, para iniciar una discusión en torno a ese liberador modo de pensar, el cual parece ser completamente ajeno al que prevalece en esta "edad oscura".

PROTOTIPOS Y ANÁLISIS CONCEPTUAL*

William Ramsey**

Sinopsis: En este escrito exploro las implicaciones de la investigación empírica reciente sobre el concepto de representación para la empresa filosófica del análisis conceptual. Arguyo que el análisis conceptual, tal como se practica usualmente, está comprometido con ciertos supuestos acerca de la naturaleza de nuestra categorización intuitiva de los juicios. Luego intento mostrar cómo es que estos supuestos chocan con versiones contemporáneas de la representación conceptual en psicología cognitiva. Luego de examinar una objeción a mi argumento, concluyo considerando maneras en las que podría alterarse el análisis conceptual para concordar mejor con el trabajo empírico.

I. Prototipos y análisis conceptual. Cuando los filósofos escriben sobre la importancia filosófica de las ciencias cognitiva y del cerebro, típicamente presentan la ingerencia que esta investigación tiene en *cuestiones* y *temas* en una u otra área de la filosofía. Así pues, los filósofos que trabajan en ciencia cognitiva generalmente se preocupan de la pertinencia de la investigación cognitiva en cuestiones filosóficas tales como, por ejemplo, si la percepción tiene o no carga teórica o si son racionales las

* Reconocimientos: Debo agradecer a John Bickle, Miriam David, Terence Hogan, Stephen Stich, John Tienson, Paul Weithman y a un árbitro anónimo varios comentarios y sugerencias útiles. Versiones anteriores de este escrito se leyeron en Central Michigan University, The University of Memphis, The University of Mississippi y en la Segunda conferencia internacional de ciencia cognitiva en San Sebastián, España. En estos diversos foros obtuve útiles sugerencias.

** University of Notre Dame.

estrategias de formación de creencias. En este ensayo pretendo hacer algo un poco diferente. Más que analizar la importancia de la investigación cognitiva con respecto a algún *tema* filosófico tradicional, quiero concentrar mi atención en las implicaciones de esta investigación con respecto a la *manera* como, con frecuencia, se hace filosofía. De manera más específica, quiero lanzar una mirada a lo que muchos psicólogos están diciendo ahora acerca de cómo nos representamos los conceptos y estudiar algunas consecuencias de sus tesis con respecto a la añeja tarea filosófica del análisis conceptual –esto es, la búsqueda de definiciones precisas que especifiquen condiciones necesarias y suficientes para nociones abstractas. Un psicólogo ha concluido recientemente que, "sobre este problema, hay un hueco entre el trabajo psicológico y el filosófico (Smith, 1990; p. 51). Aquí, mis esfuerzos los voy a dedicar a achicar ese hueco. Mi meta será la de convencerlos de que si es verdad mucho de lo que dicen estos psicólogos, entonces los filósofos analíticos deben de repensar seriamente este *modus operandi.*

Para hacer esto, procederé de la siguiente manera. En la siguiente sección daré una breve caracterización del tipo de empresa filosófica que estoy denominando análisis conceptual e intentaré hacer explícitos algunos de los supuestos en los que se basa. En la sección III, consideraré algunos de los recientes hallazgos empíricos en psicología y proporcionaré un esbozo de las versiones más populares de representación de conceptos que han propiciado estos resultados. En la sección IV, examinaré las implicaciones de las teorías psicológicas tratadas en la sección III con respecto a la empresa filosófica tratada en la sección II. Aquí desarrollo y defiendo mi aseveración de que los supuestos que impulsan el análisis conceptual no están fundados empíricamente y chocan con las teorías psicológicas prevalecientes de representación de conceptos. Si es verdad lo que están diciendo los psicólogos, entonces es una empresa seriamente desencami-

nada la búsqueda, mediante un examen de nuestras intuiciones, de definiciones nítidas. En la sección V, 1'11 considero una objeción en contra de mi argumento y muestro por qué ésta no funciona. Finalmente, en la sección VI, 1'11, ofrezco una breve conclusión y estudio formas posibles como la empresa del análisis conceptual podría alinearse con la investigación psicológica.

Empero, antes de iniciar, son pertinentes un par de comentarios preliminares. El primero es que la tesis que defenderé está condicionada con respecto a lo que se sigue *si* son correctas ciertas versiones de la representación de conceptos. Puesto que mi aseveración es acerca de la *relación* entre una rama de la investigación empírica y el análisis filosófico, no haré ningún real esfuerzo por defender las versiones empíricas mismas. En segundo lugar, aun cuando, hasta donde sé, nadie ha intentado formular el tipo de argumento detallado que voy a presentar, vale la pena hacer notar que muchos de los temas que aquí estaré estudiando, han estado flotando por algún tiempo en torno a los círculos filosóficos. Quizás la expresión mejor conocida de estos sentimientos es el examen que hace Wittgenstein de los conceptos de parecido familiar en las *Investigaciones*, aun cuando pueden encontrarse ideas similares en los escritos de otros filósofos, incluyendo a Hilary Putnam (1962), Peter Atchinstein (1968), Harold Brown (1988), Terence Horgan (1990) y, en particular, Stephen Stich (1990, por aparecer), cuya obra reciente inspiró este escrito. Lo siguiente puede leerse como un intento por fundar estas ideas en una base empírica y darles un argumento firme.

II. El análisis conceptual y sus supuestos subyacentes. Sería un tanto eufemista sostener que el análisis conceptual ha sido un aspecto importante de la filosofía occidental. Desde los escritos de Platón, en los que Sócrates y su cohorte intentan repetidamente discernir la esencia verdadera de materias tales como la piedad y la justicia, los filósofos han estado en la tarea de proponer y (de manera más típica) atacar definiciones para una

gran cantidad de nociones abstractas. Éstas incluyen conceptos tales como conocimiento, causación, racionalidad, acción, creencia, persona, justificación y moralidad (para nombrar sólo unas cuantas). Como lo señala un autor, "lo que han buscado los filósofos a lo largo de su historia son esas características de lo que examinan, sea esto *conocimiento, verdad, necesidad, mente, imprudencia, valor* o *tiempo*, en virtud de las cuales eso es lo que es; esas características que le son necesarias y le dan su esencia" (White, 1975; p. 103). Sin embargo, incluso los filósofos que no consideran como su meta *principal* el análisis de algunos conceptos, emprenden, con frecuencia, explícita o implícitamente, una forma de análisis conceptual como un paso para lograr algún otro fin. Por tanto, poco se puede dudar acerca de la importancia de esta empresa para la filosofía analítica. Pero, ¿cómo es que se lleva a cabo esta empresa y, quizás más importante, cuáles son sus supuestos subyacentes acerca de la manera como nos representamos los conceptos?

II. A. Dos criterios para las definiciones. Responder la primera pregunta –esto es, ¿cómo es que se lleva a cabo el análisis conceptual?– es, a primera vista, relativamente fácil, a saber, los filósofos proponen y rechazan definiciones para cierto concepto abstracto, pensando en serio acerca de casos intuitivos del concepto y tratando de determinar cuáles podrían ser sus propiedades esenciales. Sin embargo, esta caracterización es muy vaga para decirnos algo útil. Quizás, una mejor manera de obtener una comprensión del análisis conceptual es tener en cuenta lo que normalmente se espera de las definiciones que se presentan. Considerando los criterios que usan los filósofos para las definiciones, podemos captar, de manera más firme, lo que pretenden y quizás descubrir algunos de los supuestos que se ocultan tras esta empresa.

Naturalmente, hay muchos criterios diferentes que usualmente evocan los filósofos en busca de definiciones. Aquí me concen-

traré sólo en dos que, aun cuando poco se mencionan, han tenido, sin embargo, una influencia grande en la manera como se hace con frecuencia el análisis conceptual. El primero de estos requisitos es que las definiciones sean relativamente directas y simples. En efecto, una forma sintáctica asumida para las definiciones es la de un pequeño conjunto de propiedades consideradas como individualmente necesarias y conjuntamente suficientes para el concepto en cuestión. Por tanto, con mayor frecuencia que lo contrario, las definiciones filosóficas toman una forma sintáctica en la que al distintivo (al menos entre los correctores) "ssi" lo sigue una breve *conjunción* de propiedades. Así, X es conocimiento si y sólo si X es una creencia verdadera justificada o X está actuando libremente si y sólo si X está haciendo lo que él o ella quiere. Como con las teorías explicativas en la ciencia, un popular supuesto subyacente es que son defectuosas o *ad hoc* las definiciones excesivamente complejas y difíciles de manejar, aun cuando no esté inmediatamente disponible ninguna definición mejor. Si un análisis proporciona una definición que es muy disyuntiva, está fuertemente calificada o comprende varias condiciones, un sentir común es que el filósofo aún no ha dado con la correcta. Conforme a esto, diferentes análisis se consideran típicamente como *competidores* y, en su mayor parte, poca gente toma en serio la idea de que el análisis correcto podría ser uno que comprendiese una combinación disyuntiva de estas definiciones alternas. Para tomar prestada una frase de Jerry Fodor, los análisis de este tipo complejo comúnmente se consideran como "yucky". Para muchos filósofos, una definición propuesta debe ser corta y simple.

Un segundo criterio que generalmente se espera que cumplan las definiciones es una preocupación, no acerca de su forma, sino de su grado de robustez. Si una definición ha de considerarse como una definición *real* entonces generalmente se asume que no puede admitir ningún contraejemplo intuitivo. Por tanto, como

todos lo aprendimos en la filosofía introductoria, la forma normal de eliminar un análisis propuesto es encontrar o bien un no ejemplo del concepto que posea las propiedades definitorias en cuestión –mostrando, con ello, que las propiedades definitorias son insuficientes para capturar el concepto– o un caso del concepto que no posea las propiedades definitorias –mostrando, con ello, que tales propiedades no son necesarias. Si pueden encontrarse contraejemplos de este tipo, entonces, típicamente, la definición propuesta se considera como inadecuada. Este sentir bien lo expresa el siguiente pasaje de un texto popular sobre método filosófico:

> [T]entativamente, nosotros consideraremos satisfactoria una definición si, luego de una cuidadosa reflexión, no podemos pensar en ningún posible ejemplo en el que o bien la palabra definida se aplique verdaderamente a algo, pero no las palabras definientes o las palabras definientes se apliquen verdaderamente pero no la palabra definida. Cuando podamos pensar en un ejemplo así, entonces habremos encontrado un contraejemplo para la supuesta definición, que muestra que no tenemos una definición informativa precisa. Si no podemos encontrar ningún contraejemplo para una definición, entonces podemos considerarla como inocente, hasta que se encuentre un contraejemplo que pruebe lo contrario. (Cornman *et al.*, 1982; p. 18)

Por tanto, las definiciones buscadas por los filósofos comprometidos con el análisis conceptual, deben típicamente pasar dos pruebas: deben ser relativamente simples –generalmente una conjunción de propiedades individualmente necesarias y conjuntamente suficientes y no deben admitir ningún contraejemplo intuitivo. Con esto en mente, podemos ahora pasar a la cuestión de las presuposiciones psicológicas.

II.B. Presuposiciones psicológicas del análisis conceptual. A primera vista, podría parecer un poco extraño suponer que el

análisis conceptual comprende *cualquier* presuposición acerca de la manera como trabaja nuestra mente. Después de todo, si la gente está interesada en definir nociones como justicia o causación, entonces es la justicia o la causación la que le preocupa –no la psicología humana. Empero, cuando miramos con más atención los criterios para las definiciones que acabo de esbozar, ciertamente podemos encontrar, asomándose en el trasfondo, algunos supuestos acerca de la cognición humana. Quizás la forma más fácil de ver esto es considerando el importante papel que desempeñan los juicios de categorización intuitiva en este tipo de filosofía. Nótese, por ejemplo, que para que realmente cuente cualquier tipo de contraejemplo como tal, tiene que haber intuiciones bastante fuertes y ampliamente compartidas de que alguna cosa o situación particular o bien es o no es un caso del concepto en cuestión. En otras palabras, el proceso de evaluar las definiciones, requiere comparar y contrastar el conjunto definitorio de propiedades con casos y con no casos, juzgados *intuitivamente*, del concepto al que se apunta. Sin estos juicios de categorización intuitiva, nunca podría despegar el análisis conceptual, como una práctica.

Debido a este papel importante de los juicios intuitivos, el análisis conceptual no puede dejar de estar comprometido con ciertos supuestos acerca de la naturaleza de nuestro sistema cognitivo. Uno de esos supuestos es que hay un traslape considerable en los tipos de juicios de categorización intuitiva que hacen diferentes personas. Sin este consenso, un contraejemplo intuitivo para un individuo no lo sería para otro y no se podría estar de acuerdo acerca de ninguna definición única. Además, dado que se espera que las definiciones expresen conjunciones simples de propiedades esenciales y que no acepten ningún contraejemplo intuitivo, también parece que hay un supuesto bastante fuerte de que nuestros juicios de categorización intuitiva coincidirán perfectamente con la presencia o la ausencia de un conjunto

pequeño pero específico de propiedades. En otras palabras, asomándose en el trasfondo de esta empresa, está el supuesto de que nuestras intuiciones convergirán a un conjunto cuyos miembros son todas y solamente las cosas que poseen alguna colección particular de rasgos. Dado que los filósofos esperan encontrar nítidas definiciones conjuntivas y dado que emplean *intuiciones* como sus guías en esta búsqueda, el supuesto parece ser que nuestros juicios de categorización intuitiva corresponderán precisamente con cúmulos simples de propiedades.

En esta etapa, parece razonable formular la siguiente pregunta, ¿de donde vienen estos supuestos? Después de todo, *prima facie*, uno supondría que es una pregunta abierta si nuestros juicios intuitivos pueden guiarnos o no a un conjunto relativamente simple de característica definitorias de conceptos abstractos. ¿Qué es lo que exactamente garantizan estos presupuestos?

Una posible respuesta a esta pregunta se basa en la idea de que los filósofos están empleando una teoría popular de la mente subyacente o quizás, de manera más específica, una teoría popular acerca de la manera como representamos conceptos. Desde Platón, un tema recurrente en filosofía ha sido que poseemos conocimiento "tácito" de algún dominio que puede no ser directamente accesible a nuestra conciencia pero que, sin embargo, se manifiesta mediante juicios intuitivos y puede alcanzarse analizando estas intuiciones. De manera similar, fuera de la filosofía, Chomsky (1965, 1972) y otros lingüistas, han propuesto que los hablantes poseen conocimiento tácito de su lenguaje nativo que toma la forma de un gramática, "psicológicamente real", representada internamente. Tanto Platón como Chomsky reconocen que este conocimiento tácito no es inmediatamente accesible a la conciencia. Pero, puesto que se piensa que esta competencia subyacente mueve nuestros juicios intuitivos, un supuesto común es que puede "descubrirse" o hacerse explícita estudiando y explotando estas intuiciones. Por ejemplo, los chomskianos supo-

nen que podemos determinar un conjunto de reglas sintácticas para el inglés considerando con detenimiento los juicios lingüísticos intuitivos de hablantes competentes del inglés porque, de acuerdo a los primeros, estos juicios los genera una real gramática del inglés, cognitivamente representada. Conforme a esta propuesta, los juicios intuitivos sirven como datos, con respecto a los cuales podemos contrastar hipótesis acerca de la naturaleza de las estructuras subyacentes que los producen.

Siguiendo líneas similares, parece que la manera más fácil de dar cuenta de los supuestos que subyacen al análisis conceptual es suponer que acechando en el trasfondo hay una teoría similar –aun cuando, quizás, tácita ella misma– acerca de la manera como nos representamos conceptos. En otras palabras, se supone que las intuiciones de categorización nos conducen a conjuntos nítidos de propiedades necesarias y suficientes porque, se supone además, a estas intuiciones las generan representaciones subyacentes *de* propiedades necesarias y suficientes. Conforme a esta propuesta, se supone que tenemos conocimiento tácito de la "esencia" de conceptos abstractos, que la esencia es un conjunto pequeño de condiciones necesarias y suficientes y que podemos descubrir este conocimiento apelando a nuestros juicios de categorización intuitiva. Llamaré a esta teoría de la representación de conceptos la tesis 'clásica'. Aun cuando relativamente pocos filósofos dedicados al análisis conceptual apoyan explícitamente la tesis clásica, ésta me impresiona como la manera más plausible y caritativa de darle sentido a esta empresa. Apelar a nuestros juicios intuitivos de categorización para apoyar o minar un análisis dado de un concepto es casi lo mismo que apelar a nuestros juicios gramaticales intuitivos para apoyar o minar una versión dada de la gramática de un lenguaje. Si los lingüistas pueden hacerlo, ¿por qué no los filósofos?

Empero, aun cuando pueda ser verdad que esta tesis de la representación de conceptos sea la manera más natural de ex-

plicar los supuestos del análisis conceptual, es importante tener en mente que el análisis conceptual no *requiere* de esta tesis de la representación de conceptos. Sólo requiere del supuesto más débil de que el esquema de representación para conceptos –sea cual sea la forma que pueda tomar– produce juicios intuitivos que corresponden con conjuntos pequeños de condiciones necesarias y suficientes. Mientras que la tesis clásica de representación de conceptos puede ser la forma más fácil de apoyar esta hipótesis más débil, se podría tratar de defenderla por otros medios. En efecto, yo supongo que se podría evitar defenderla del todo y simplemente tomar como un artículo de fe que nuestros juicios intuitivos se comportarán de esta manera. Puesto que toda empresa filosófica debe comenzar con algunos supuestos, ésta no sería una estrategia falta por completo de razón. No sería falta de razón, esto es, siempre que no hubiese alguna motivación imperiosa para dudar de estos supuestos.

La aseveración central de este escrito es que, de hecho, *hay* una motivación para dudar tanto de la teoría clásica de la representación como de los supuestos de que nuestros juicios intuitivos proporcionarán definiciones simples y plenamente intuitivas de conceptos abstractos. Si resultan ser correctas las clases de historias que ahora nos cuentan los psicólogos acerca de representaciones de conceptos y de juicios de categorización, entonces no pueden satisfacerse conjuntamente los dos criterios de los filósofos para las definiciones –que son directamente simples y que no admiten contraejemplos. Para tener una sensación mejor de dónde yacen las dificultades, debemos considerar ahora esas explicaciones psicológicas.

III. Explicaciones empíricas de la naturaleza de los conceptos. La tesis clásica de representación de los conceptos que hemos estudiado en la sección anterior ha sido popular no sólo en filosofía. Hasta hace muy poco, también fue la tesis aceptada de nuestro sistema conceptual en psicología cognitiva. Sin embargo,

comenzando con la obra revolucionaria de Eleanor Rosch (1973, 1975, 1978; Rosch y Mervis, 1975), ha habido un número creciente de psicólogos que creen que la tesis clásica da un cuadro impreciso de la manera como organizamos información conceptual. Como lo señala una fuente, "los psicólogos están actualmente abandonando por docenas la tesis clásica" (Smith y Medin, 1981, p. 33). Quizás la manera más fácil de explicar cómo a la versión clásica le llegaron tiempos difíciles en la psicología cognitiva, es ver primero diversos hallazgos con respecto a la naturaleza de los juicios de categorización intuitiva. Después de esto, estaremos en una posición mejor para examinar aspectos importantes de las teorías más populares de representación de conceptos que han suplantado la tesis clásica.

III.A. Hallazgos empíricos acerca de los juicios de categorización. En un número de estudios diferentes, los psicólogos han demostrado el que ahora se considera como un fenómeno cognitivo muy sólido, a saber, que en una amplia gama de juicios de categorización –incluyendo los de conceptos abstractos (Hampton, 1981)–, la membresía de clase no parece ser una cuestión directa de todo o nada. En lugar de esto, parece que, para la mayoría de los conceptos, algunos casos se considera que son ejemplos mucho mejores que otros del concepto al que apuntamos. En otras palabras, nuestras intuiciones de categorización parece que reflejan un sistema taxonómico en el que la mayoría de las categorías tienen una membresía graduada.

Este aspecto de nuestros juicios de categorización, al que se conoce comúnmente como "tipicalismo" o "efectos prototipo", se ha demostrado en diversos diseños experimentales. Cuando a los sujetos se les ha preguntado qué tan buen ejemplo de un concepto dado es un ítem particular, ellos tienden a asignar valuaciones diferentes a ítems diferentes de una manera bastante consistente (Rips *et al.*, 1973; Rosch, 1973). Así, para un concepto dado (PAJARO, p.ej.), algunos casos se evaluarán como

mejores ejemplos (petirrojo, p.ej.) que otros casos (lechuza o avestruz, p.ej.). Además, esta gradación de casos de los típicos a los atípicos, se refleja en otros estudios diversos. Por ejemplo, cuando simplemente se pide una lista de casos de alguna categoría, los sujetos enlistan regularmente, primero, los casos típicos y los atípicos al final (Rosch, 1975; Rosch *et al.*, 1976). De manera similar, los sujetos son mucho más rápidos para juzgar los casos típicos, que los atípicos, como miembros de la categoría a la que apuntamos (McCloskey y Glucksberg, 1979; Rips *et al.*, 1973; Rosch, 1973). Otros estudios revelan que los casos típicos de un concepto se aprenden con mayor facilidad (Rosch, 1973; Rosch *et al.*, 1976), tienen una influencia mayor en la formulación de decisiones (Cherniak, 1984; Rips, 1975), se citan con mayor frecuencia que otros casos (Barsolou, 1983, 1985) y se considera que son *menos* similares a los miembros atípicos que éstos lo son con aquéllos (Tverski y Gati, 1978).[1] Cuando se les pide a los sujetos que enlisten atributos de diferentes casos de un concepto dado, se ha encontrado que las listas de ejemplos con gradaciones de alta tipicalidad tienen significativamente más atributos, que son comunes a otros miembros del concepto que los casos con gradaciones bajas de tipicalidad. En cambio, casos muy típicos comparten menos atributos con miembros de otros conceptos que lo que sucede con casos menos típicos. Para algunos conceptos, ningún atributo se aplicó a todos los casos considerados (Rosch y Mervis, 1975). También, hay una evidencia considerable de que diversos factores diferentes –tales como la frecuencia de ejemplificación, la similitud a algún ideal, una familiaridad completa y el contexto de presentación– desempeñan todos un papel que influye en nuestros juicios de

1 Aun cuando usualmente pensamos en la similitud como en una relación simétrica, en diversos episodios cognitivos no exhibe este carácter. Para un examen mayor, véase Tversky y Gati (1978).

categorización y que el rejuego de estos factores genera mucha flexibilidad e indeterminación en nuestras intuiciones. Por ejemplo, varios investigadores han proporcionado evidencia que sugiere que la importancia de cierto rasgo para los juicios de categorización puede variar en diferentes contextos, proporcionando, así, diferentes intuiciones acerca del mismo tipo de caso en diferentes contextos (Barsalou, 1985; Rosch, 1978; Roth y Shoben, 1983).[2]

¿Qué es lo que hallazgos como éstos nos dicen acerca de la naturaleza de nuestra estructura conceptual? Algunos psicólogos, como Rosch, han dudado mucho en apoyar cualquier modelo particular de representación de conceptos sobre la base de estos hallazgos, sugiriendo, en su lugar, que estos resultados limitan, pero no determinan modelos de representación y de procesamiento. Sin embargo, incluso el cauto Rosch, concede que, "por sí sola, [una] representación de categorías, en términos de atributos necesarios y suficientes, probablemente sería incapaz de manejar todos los hechos hasta ahora conocidos ..." (1978, pp. 40-41). No es difícil de encontrar la fuente de estos y otros sentimientos anticlásicos, ahora muy difundidos en la psicología cognitiva. Asumiendo que los juicios de categorización son un producto de nuestro esquema conceptual, uno esperaría hallazgos muy diferentes si ese esquema conceptual representase categorías a la manera sugerida por el marco clásico. Por ejemplo, uno esperaría encontrar que todos los casos de un concepto se juzgasen intuitivamente como ejemplos igualmente buenos y no encontrar ninguna disparidad en los tiempos de reacción de los juicios de categorización, puesto que los criterios para la membresía en una categoría se aplicarían igualmente bien a todos y

2 Para una útil visión general acerca de estos hallazgos, véanse Smith y Medin (1981), capítulo 30, Medin y Smith (1984) y Lakoff (1987, capítulo 2).

sólo a aquellos casos.[3] Los estudios de categorización han mostrado que esto no es así. De manera similar, uno esperaría encontrar que, para cualquier concepto dado, los sujetos tenderían a enlistar un conjunto común de rasgos definitorios para todos los miembros de la categoría. Los estudios de asignación de atributos minan también esta espectativa.

Debido a estas y otras consideraciones similares, muchos psicólogos se han hecho cada vez más escépticos acerca de la tesis clásica. Claro está que, como en cualquier dominio empírico, refutar una teoría de representación de conceptos, de una vez y para siempre, no es una tarea fácil y hay diversas cosas que pueden decirse, y se han dicho, a favor de la tesis clásica como respuesta a esta contraevidencia *prima facie*.[4] Sin embargo, como con frecuencia sucede en ciencia, aquí la situación exige apelar a la *mejor* explicación y los esfuerzos por salvar la teoría clásica los ha reducido en gran medida el éxito que los psicólogos cognitivos han tenido para construir explicaciones alternativas que concuerdan mucho mejor con los datos. Es seguro decir que, con unas cuantas excepciones, los psicólogos, hoy, creen que el cuadro clásico de la representación de conceptos no funciona más.

¿Qué aspecto tienen estas versiones alternativas de la representación de conceptos? Puesto que el espacio no permite un examen exhaustivo de los muchos modelos no clásicos que gozan ahora de popularidad, delinearé, en lugar de eso, algunos de los temas estándar y principios comunes a estos modelos, en un esfuerzo por dar un breve esbozo de su forma general.

3 Claro está que esto supone que no hay problema para discernir si el caso en cuestión tiene o no alguna propiedad definitoria.

4 Para un examen interesante de algunas de las formas como se podría intentar rescatar la tesis clásica y de las dificultades con estas estrategias, véase Smith y Medin (1981, capítulo 3).

III.B. Teorías contemporáneas de categorización y de estructura de conceptos. Para dar cuenta de los efectos de tipicalidad y de otros hallazgos mencionados en la sección III.A., muchos psicólogos han adoptado la tesis de que los conceptos son representados por *prototipos* y que los juicios de categorización surgen por la comparación de casos individuales con esta representación de prototipos (o, quizás, de manera más precisa, comparando una representación del caso con la representación del prototipo). Mientras que hay varios modelos diferentes que emplean este supuesto básico, podemos entender la idea general concentrándonos en un tipo estándar de teoría, llamado comúnmente la versión "probabilística" o "fundada en rasgos" (Smith y Medin, 1981). La aseveración central de esta tesis es que el prototipo es un conjunto abstracto de rasgos y los juicios de categorización son una función del número e importancia de los rasgos compartidos entre el prototipo y el caso en cuestión. Si alguna cosa o un suceso y la representación prototipo comparten un número suficiente de rasgos sobresalientes, entonces esto da origen al juicio intuitivo de que la cosa o el suceso es un caso del concepto al que se apunta. Si los dos no comparten un número suficiente de rasgos, entonces el caso no se juzga así. Otras variaciones sobre el tema incluyen la propuesta de que a los rasgos prototípicos se les asignan diferentes pesos o grados de distinción. Así, por ejemplo, un modelo podría asignar el peso de 30 para "emplumado", 25 para "alado", 15 para "vuela" y así sucesivamente y luego asumir que cualquier cosa que posea bastantes rasgos pertinentes para aproximarse o superar una suma de, digamos, 50, se consideraría como un caso del concepto PAJARO. Conforme a esta tesis, un juicio positivo podría generarlo cualquiera de las diferentes combinaciones de propiedades que se aproxime o exceda algo como un umbral. De esta manera, el enfoque probabilístico apoya una tesis altamente disyuntiva de la ejemplificación de conceptos, aun cuando el grado

de disyuntividad pueda variar entre conceptos. La tesis la resumen adecuadamente Medin y Smith de la siguiente manera:

> La tesis probabilística asume que los conceptos son abstracciones o representaciones sumarias, pero arguye que para que una propiedad se incluya en el resumen, ésta sólo necesita tener una probabilidad sustancial de figurar en casos del concepto, esto es, sólo necesita ser característica del concepto, no definitoria. Entonces, un objeto se categorizará como un caso de algún concepto. Así, por ejemplo, si posee algún número señalado de propiedades o suma de propiedades pesadas, incluidas en la representación sumaria de A. La categorización es, así, un asunto de evaluar la similitud más bien que de aplicar una definición (Medin y Smith, 1985, p. 117).

Naturalmente, esta caracterización simple hace surgir, de inmediato, diversas cuestiones, como ¿qué es lo que determina qué propiedades forman el prototipo sumario? ¿Cuántas propiedades deben compartir el prototipo y el caso para un juicio de ejemplificación? Y, ¿cómo han de asignarse los pesos a los rasgos incluidos en la representación prototipo? Afortunadamente, no deben aquí preocuparnos demasiado estas preguntas; baste decir que los intentos por enfrentarse a estas y otras preocupaciones similares han producido múltiples florecientes programas de investigación en ciencia cognitiva. Por ejemplo, en una versión de esta historia, los rasgos que se incluyen en la representación sumaria son los más destacados que tienen una probabilidad sustancial de figurar en conjunción con otras propiedades. Conforme a esta tesis, los rasgos que conforman el prototipo sumario –y que, por tanto, conforman la representación del concepto–, son los que, a la vez, son comunes y únicos de un cúmulo dado de entidades (Rosch, 1978; Rosch y Mervis, 1975; Smith y Medin, 1981). Otras versiones sostienen que la pertinencia de una propiedad dada para los juicios de categorización se debe a di-

versos factores, algunos de los cuales son sensibles a las variaciones de contexto. Conforme a esta variación, la posesión de ciertas propiedades nos inclinaría hacia un juicio de categorización positivo en algunos contextos pero no en otros (Barsalou, 1985, 1987; Lakoff, 1987; Malt y Smith, 1982; Smith y Medin, 1981).

Estas y otras numerosas variantes sobre el tema del prototipo están ahora muy extendidos en la psicología cognitiva. Si, una vez más, consideramos los hallazgos estudiados en la Sección III.A., podemos ver por qué se les considera tan comúnmente superiores a las versiones clásicas. Virtualmente, todos los diversos efectos de tipicalidad los acomoda fácilmente la versión del prototipo. La tesis del prototipo da cuenta del hallazgo de que algunos casos de un concepto se consideren mejores ejemplos que otros, puesto que esto es exactamente lo que uno podría esperar si la categorización está determinada por el compartir rasgos entre un prototipo y un caso. Naturalmente, algunos ejemplos, y no otros, van a tener más rasgos (o rasgos a los que se les da más peso) en común con el prototipo y esto dará origen a la gradación en los juicios de membresía en el concepto. De manera similar, estos modelos pueden dar cuenta de las diferencias en tiempos de reacción. Si los juicios de categorización surgen del empate de los rasgos de un caso con los de una representación de prototipo, entonces habría que esperar obtener un número suficiente de empates para un juicio positivo mucho más rápidamente cuando el caso y el prototipo tienen un amplio traslape. Los modelos probabilísticos también pueden explicar por qué hay casos de un concepto no claros intuitivamente. Esto sucede o bien porque el valor sumado de rasgos del caso se acerca mucho a lo que podría considerarse como un umbral para un juicio positivo o son casos en los que el conjunto de rasgos excede los necesarios para más de un concepto (Smith y Medin, 1981, p. 69). Como McCloskey y Glucksberg lo señalan,

> Para una categoría así considerada, el grado de membresía de un ejemplar potencial,[5] lo determina la proporción de las propiedades de la categoría que comparte. Los ítems que comparten una amplia proporción de las propiedades de la categoría, tienen un alto grado de membresía en la categoría, mientras que ítems que comparten pocas de estas propiedades, tienen un bajo grado de membresía. Así, el grado de membresía varía continuamente y ninguna frontera precisa separa los miembros de los no miembros de una categoría (McCloskey y Glucksberg, 1978, p. 467).

Puesto que mi argumento es condicional, les dejaré una defensa mayor de los modelos prototipo a los psicólogos cognitivos. Para nuestros propósitos, la cosa importante que hay que señalar es que todos estos modelos comparten la idea wittgensteiniana central de que no hay ninguna propiedad o conjunto de propiedades que un ítem debe poseer para que se le considere como un caso de un concepto al que apuntamos. En cambio, la posesión de un número importante de rasgos es suficiente para generar intuiciones de que es un caso de una categoría dada. Además, la pertinencia o distinción de estas propiedades no es preciso que sea una constante fija y puede variar según los contextos. Mientras que este esbozo apenas si rasca la superficie de todo lo que puede decirse acerca de esta familia de teorías,[6] el mismo debe bastar para dar su sabor general.[7]

5 Aquí el término 'ejemplar' se refiere a un caso de un concepto al que apuntamos.

6 Hay que notar que los recientes modelos conexionistas de la memoria y la cognición caen también dentro del paradigma de prototipo. Véanse, por ejemplo, Rumelhart *et al.* (1986) y Smolensky (1988).

7 Aun cuando algo anticuado, Smith y Medin (1981) es todavía el tratamiento clásico de las teorías de prototipo.

IV: Análisis conceptual y teorías de prototipo. Luego de haber reunido, en las últimas dos secciones, todos los elementos clave para mi caso en contra del análisis conceptual, estamos ahora en una posición de ver precisamente cómo marcha el argumento. En la Sección II sugerí que una motivación subyacente posible para el análisis conceptual es la tesis clásica de la representación de conceptos –que nosotros representamos conceptos abstractos al representar un pequeño conjunto de propiedades individualmente necesarias y conjuntamente suficiente. Puesto que las teorías de prototipo de la representación de conceptos, del tipo que acabamos de esbozar, contradicen directamente las versiones clásicas, podría pensarse que esto solo basta para mostrar la incompatibilidad de las teorías de prototipo y el análisis conceptual. Pero, como lo señalé en la Sección II, el análisis conceptual no *requiere* realmente que sea correcta la teoría clásica. De aquí que, si hemos de desarrollar un caso en contra del análisis conceptual sobre la base de teorías de prototipo, entonces se necesita un poco más de trabajo. Pero no mucho más.

Como lo vimos en la Sección II, lo que la práctica del análisis conceptual *sí* necesita es el supuesto de que nuestros juicios intuitivos acerca de categorización y de membresía en una clase pueden dar una definición de conceptos abstractos que 1) tome la forma de un pequeño conjunto conjuntivo de propiedades esenciales y 2) sea coherente con todas las intuiciones pertinentes y no permita contraejemplos. Pero, si son correctas las teorías de la representación de conceptos, del tipo de las que estudiamos en la última sección, entonces es muy dudoso que los dos criterios para las definiciones puedan jamás satisfacerse conjuntamente. La razón es que, *ex hypothesi*, nuestros juicios de categorización están favorecidos por un esquema taxonómico que genera intuiciones de categorización que son demasiado variadas y diversas como para que las capturen definiciones simples y no disyuntivas. Si ser un caso intuitivo de X es tan sólo cuestión de tener

un cúmulo de propiedades que sea suficientemente similar a alguna representación prototipo y si hay un número de formas diferentes como esto puede hacerse, algunas de las cuales pueden variar según los contextos, entonces nunca pasará la prueba de la intuición ninguna definición precisa que comprenda algún subconjunto de estas propiedades y que las trate como necesarias y suficientes. *Siempre* admitirá contraejemplos intuitivos, porque será mucho más amplio el rango de diversidad sancionado por nuestro esquema de representación conceptual, que el que permita cualquier definición precisa y directa. Por tanto, la búsqueda de una definición simple, no disyuntiva, de un concepto filosófico dado que concuerde con todas nuestras intuiciones y no admita contraejemplos, es una empresa inútil –simplemente no hay nada así.

Otra forma de ver este asunto es considerar los apoyos psicológicos de la filosofía del contraejemplo, a la luz de la teoría prototipo de los conceptos. Como antes lo señalamos, los filósofos apelan a dos tipos de contraejemplos para demoler un análisis propuesto, a saber, uno que ilustra que las propiedades en el *analysans* no son necesarias, el otro que ilustra que tales propiedades no son suficientes. Consideremos el primer tipo de contraejemplo. Puesto que, conforme a la versión del prototipo, uno cualquiera de entre diversos conjuntos de propiedades bastará para generar un juicio de ejemplificación, siempre debemos ser capaces de descubrir casos intuitivos del concepto que no posean todas las supuestas propiedades necesarias de una simple definición conjuntiva. En efecto, puesto que los filósofos están dispuestos, en general, a considerar casos imaginarios como contraejemplos (sin importar cuán grotescos), un filósofo sagaz puede meramente *construir* un contraejemplo inventando un caso con un conjunto de rasgos que excluya una o más de las propuestas propiedades definitorias pero cuyos pesos sumados bastan, sin embargo, para producir un juicio de ejemplificación.

Conforme a la tesis del prototipo, siempre serán posibles contraejemplos intuitivos para cualquier definición que especifique rasgos necesarios.

Pasando ahora al otro tipo de contraejemplo –del tipo que ilustra que la definición no es suficiente– las cosas son un poco más complicadas. *Prima facie*, se podría pensar que si fuese correcta la historia del prototipo, sería muy difícil presentar un contraejemplo de este tipo. La razón es que en tanto que la versión del prototipo puede no comprender ninguna propiedad *necesaria*, ciertamente parece que acepta que siempre podría resultar que varios cúmulos se mostrase que son *suficientes* para juicios de ejemplificación del concepto. Sin embargo, hay aspectos de la versión del prototipo que sugieren que también contraejemplos de este tipo son relativamente baratos. Un aspecto así es la dependencia contextual de la importancia de un rasgo dado para los juicios de categorización. Si la distinción de rasgos para los juicios de categorización es sensible al contexto, entonces debe de ser posible hallar o construir ejemplos encajados en ciertas circunstancias en las que estén presentes todos los supuestos rasgos definitorios pero que, debido al contexto, tengan un peso sumario que sea demasiado bajo para un juicio de ejemplificación. Otra posibilidad es que, para algunos conceptos, la distinción de algunos rasgos en un caso podría cancelarse por la posesión de otros rasgos más cercanos a un concepto diferente. De aquí que, aun cuando un caso dado podría contener un conjunto de rasgos que normalmente lo llevarían a uno a juzgarlo como miembro de una categoría particular, podría acabar siendo juzgado como un caso de una categoría diferente, porque resulta que posee propiedades que son mucho más importantes para la última (Smith y Medin, 1981).

Es importante ver que la cuestión aquí no es que estos contraejemplos intuitivos no sean contraejemplos reales de las definiciones propuestas. Más bien, de lo que se trata es que desde la pers-

pectiva de tesis contemporáneas de representación de conceptos, los contraejemplos a las definiciones clásicas son baratos y abundantes. Mientras que los contracasos ciertamente reflejan lo inadecuado de la definición propuesta, esto no es por las razones tradicionalmente supuestas por los practicantes del análisis conceptual. La definición es defectuosa, no porque haya algún otro conjunto simple de condiciones necesarias y suficientes que puede hacer una tarea mucho mejor para definir el concepto en cuestión. Más bien, la definición falla porque, conforme a las versiones de prototipo de los conceptos, *no hay ninguna colección precisa y simple de propiedades que posean todos y sólo los casos intuitivos del concepto; cualquier definición que se exprese de esta manera está condenada desde el principio.* Conforme a esta tesis, simplemente no es posible dar con una definición que sea, a la vez, una simple conjunción de propiedades 'esenciales' y que, al mismo tiempo, capture todas nuestras intuiciones.

V. Una objeción potencial y su respuesta. Antes de concluir me gustaría considerar una objeción, en contra de este tipo de argumento, que he oído expresarse en diversas ocasiones. Aun cuando no pienso que la objeción funcione, es útil porque ayuda, sin embargo, a precisar más lo que está en disputa en este asunto.

La objeción es que yo soy culpable de una forma de psicologismo ilícito –que he confundido cuestiones acerca de *qué* es realmente algo con cuestiones acerca de nuestros *pensamientos* ordinarios *acerca* de lo que es. Los filósofos a los que les preocupa el análisis conceptual, realmente les interesa descubrir las esencias verdaderas de cosas tales como el conocimiento y la causación, no la manera como sucede que las representamos. De aquí que las teorías psicológicas acerca de nuestro esquema conceptual en gran medida no son pertinentes para esta empresa *metafísica.*

En efecto, una versión de esta queja se ha presentado, no sólo en contra de mi argumento, sino en contra de teorías del prototipo, muy anticlásicas, en las que se funda mi argumento. En una

reciente comunicación, Georges Rey (1983, 1985), arguyó que los psicólogos que apoyan la tesis prototipo, han "confundido, inútilmente, asuntos metafísicos de *identidad* conceptual, con (en términos generales) asuntos epistemológicos de acceso conceptual" (1983, p. 238). Como yo la entiendo, la objeción de Rey se desarrolla de la siguiente manera. Los teóricos del prototipo que atacan la tesis clásica, no logran apreciar una importante distinción entre, por una parte, rasgos que ordinariamente cumplen una función pragmática de facilitar los juicios de categorización cotidianos y, por otra parte, rasgos definitorios más 'esenciales' que determinan lo que la cosa es *realmente*. Además, sabemos por las obras de Kripke y de Putnam que no es preciso que los rasgos definitorios de un concepto estén realmente representados en la mente\cerebro excepto, quizás, en la de unos cuantos de los expertos importantes. Sin embargo, a pesar de su ignorancia de los rasgos definitorios, la gente ordinaria asume con frecuencia, empero, que tales rasgos existen y que podrían añadirse a su representación tras una investigación mayor. Aun cuando Rey no arguye que esta propuesta rescata plenamente la tesis clásica, uno podría, concebiblemente, salvar la tesis clásica arguyendo que los efectos de tipicalidad y de inestabilidad surgen porque los sujetos emplean las propiedades identificatorias no esenciales cuando hacen juicios intuitivos de categorización. Esto deja abierta la posibilidad de que haya un conjunto más definicional de rasgos que podría ser (o ha sido) descubierto por los expertos adecuados y que *realmente* podría decirnos si algo era o no un caso de algún concepto. Ciertamente, versiones recientes de la narración del prototipo admiten tal posibilidad, donde se sugiere que nuestra representación prototipo inicial se aumenta con ciertas propiedades "núcleo" adquiridas de los expertos pertinentes y que toman un *status* definitorio.

Pero, si todo esto es correcto, entonces ¿no he exagerado el caso en contra de los filósofos analíticos preocupados con el

análisis conceptual? Pues si la tesis Kripke-Putnam es correcta y los expertos pueden descubrir rasgos verdaderamente esenciales para las categorías de clases naturales, entonces, ¿por qué no suponer que ellas pueden también descubrirse para conceptos filosóficos? Si los científicos pueden determinar un conjunto de propiedades necesarias y suficientes para los conceptos de clases naturales, a pesar de que tengamos una representación basada en prototipos, entonces, ¿por qué no suponer que los filósofos pueden hacer lo mismo para las nociones abstractas de la filosofía? Parece que yo he confundido lo metafísico con lo epistemológico, exactamente del modo como nos lo previene Rey.

El problema con esta objeción, tal como yo la veo, es que culpa al sujeto equivocado. Pues si hay cualquier conexión ilegítima entre asuntos metafísicos y asuntos epistemológicos –esto es, entre lo que una cosa es "realmente" y la manera como la representamos–, es el practicante del análisis conceptual y no yo, quien forja este eslabonamiento. Esto puede verse con mayor claridad si consideramos las maneras diferentes en las que proceden el científico y el filósofo en la tarea de determinar rasgos definitorios.

Conforme a Kripke (1972) y Putnam (1975), los investigadores en las ciencias naturales intentan descubrir propiedades definitorias "ocultas" de las clases naturales –tales como cierta estructura física, química o genética– mediante varios tipos de investigaciones empíricas. Exactamente por qué es que estos tipos de rasgos nos impresionan como definitorios de los conceptos de las clases naturales es una cuestión interesante de suyo que aquí no consideraré.[8] Lo que quiero señalar es que mientras

8 Una respuesta podría ser que a estos tipos de propiedades se les concede un peso mayor, en nuestras representaciones de prototipo de conceptos de clases naturales, debido a su pertinencia con respecto a poderes causales. Para un examen interesante de la relación entre taxonomía, estructura física y poderes causales, véase Fodor (1987, capítulo 2).

que los *tipos* de propiedades definitorias que busca la ciencia podrían tener un fundamento intuitivo, las mismas propiedades definitorias –las propiedades específicas atómicas, químicas o genéticas– no se basan en la intuición. La aseveración de que el agua es H_2O ciertamente no tiene aura intuitiva ninguna. Como Rey mismo tiene cuidado en señalar, estos tipos de propiedades esenciales llegan a representarse sólo como una *consecuencia* de la investigación científica y, para la mayoría de la gente, no tienen representación alguna. Pero esto significa que las definiciones de la ciencia no dependen completamente de nuestros juicios de categorización intuitiva y pueden trascender, de diversas maneras, nuestras representaciones de prototipo para las clases naturales. En efecto, no es nada extraño que el análisis científico proporcione clasificaciones que son muy contraintuitivas. Aun cuando en alguna ocasión clasificamos las ballenas y los delfines como peces, los esquemas contemporáneos de clasificación nos dicen que esto no es así. Pero, generalmente, nosotros no consideramos estos casos contraintuitivos como *contraejemplos* de estas taxonomías científicas. En lugar de esto, adoptamos la posición de que la ciencia nos ha revelado que estaban equivocados nuestros juicios de categorización ordinarios. De aquí que, en su búsqueda de definiciones, se le permita al científico salirse de los límites que proporcionan nuestros juicios intuitivos de categorización.

Pero ahora contrástese este modo de investigación con el del análisis conceptual en filosofía. Para los principiantes, a diferencia de la situación con las clases naturales, lejos está de ser claro qué *tipos* de propiedades podrían considerare como más o menos "definitorias" de los diversos conceptos abstractos en filosofía. No es como si nociones tales como causación y conocimiento tuviesen una microestructura o una conformación genética. Más aún, el filósofo dedicado al análisis conceptual, está operando bajo un conjunto muy diferente de supuestos que el

científico. Como lo vimos en la Sección II, el filósofo asume con frecuencia que las propiedades definitorias pueden descubrirse analizando nuestras intuiciones existentes, preteóricas (las que, con frecuencia, se denominan "incorruptas"). De aquí que cualesquiera propiedades que gobiernen nuestros juicios prístinos, intuitivos de categorización, van a tener un peso mayor en las evaluaciones de análisis en competencia. Si una definición dada no concuerda con nuestros juicios intuitivos (esto es, admite contraejemplos intuitivos), entonces, usualmente es la definición y no las intuiciones la que se considera que falla. Debido a que el análisis conceptual toma los juicios intuitivos de categorización como sus datos centrales, el mismo está limitado por nuestra representación ordinaria de conceptos, de una manera en que no lo está la investigación científica. Así que es el análisis conceptual, y no mi argumento, el que nos obliga a basar *qué* es lo que es una cosa en la manera como la representamos. Y puesto que, *ex hypothesi*, es una representación prototipo la que gobierna los juicios de categorización cotidianos, entonces será tal representación la que guíe y limite cualquier búsqueda de definiciones realizada a través del análisis conceptual.

En consecuencia, cualquier distinción entre una representación prototipo que facilite los juicios ordinarios de categorización y una representación más definitoria compuesta de rasgos descubiertos y divulgada por los investigadores empíricos –a pesar de cualquier legitimidad que pudiera tener para los conceptos de clases naturales–, no va a acarrear mucho peso cuando consideremos conceptos definidos mediante el análisis conceptual tradicional. Para estos últimos conceptos los *filósofos* son los expertos pertinentes. Y si los filósofos buscan propiedades definitorias considerando los juicios de categorización intuitivos, entonces sus definiciones nunca rebasarán las propiedades que comprende la representación prototipo ordinaria que gobierna estos juicios. Pero, si esto es así, como lo he argüido

en la sección IV, será imposible descubrir una definición que sea, a la vez, simple y que no admita ningún contraejemplo.[9]

VI. Observaciones finales. Antes de concluir, vale la pena reflexionar por un momento cuán poco sorprendente debe ser esto para los filósofos. El fracaso de la filosofía analítica para producir un análisis no controversial, completamente satisfactorio, de la gran mayoría de los conceptos abstractos, debe sugerir, por sí mismo, que algo falta. Además, incluso dentro de la literatura filosófica, ha habido casos que sugieren dificultades con intuiciones del tipo predicho por la tesis del prototipo. Considérese el bien conocido estudio de Bernard Williams (1970) de la identidad personal. Aquí, un conjunto propuesto de condiciones para la identidad personal proporciona juicios intuitivos positivos cuando se describe de una manera pero, como Williams lo señala,

9 Aun cuando el definir clases científicas no está sujeto a todas nuestras intuiciones normales de categorización, la siguiente definición podrá eventualmente llegar a la cultura popular y así influir de diversas maneras en nuestros juicios intuitivos, junto con la vieja representación de prototipos (y ésta dará origen a efectos típicos; Rey, 1983; Armstrong *et al.*, 1983). Nuestra comprensión contemporánea de sustancias simples tales como el agua o los diversos conceptos de la matemática, puede funcionar de esta manera. Empero, la cuestión crucial que, en estos casos, hay que tener en mente, es que la definición se *aprende* de otros que o bien la han descubierto por investigación empírica o simplemente la han estipulado de entrada. Aun cuando, ciertamente, podemos representar definiciones directas y simples de muy diversos conceptos, esto solo no le ayuda al abogado del análisis conceptual si lo hacemos simplemente porque alquien más ha descubierto o decretado empíricamente las propiedades definitorias en cuestión. Puesto que cualquier técnica da entrada a contraejemplos intuitivos para las definiciones, ninguna respeta las reglas en las que insiste el análisis conceptual. Claro está que si en todos nosotros ya existen definiciones nucleares (quizás junto con una representación de prototipo usada con propósitos de identificación) que pudiesen descubrirse reflexionando en nuestras intuiciones de categorización, entonces, después de todo, sería correcta algo así como la tesis clásica y sería falso el antecedente de mi argumento condicional (esto es, que son correctas las teorías de prototipo).

básicamente el mismo conjunto de condiciones tiende a producir juicios contrarios cuando se describe de manera algo diferente. Esto, claro está, es justamente lo que se habría de esperar si los juicios de categorización son sensibles al contexto de la manera que lo sugiere la teoría del prototipo. Como se señaló antes, lejos está de ser claro qué es lo que garantizaba el supuesto de que nuestros juicios intuitivos de categorización habrían de converger de una manera que produjese un nítido conjunto de condiciones necesarias y suficientes. Puesto que ésta es una cuestión empírica acerca de asuntos de cognición y de representación conceptual, ciertamente no es algo que se preste a sí mismo a una prueba *a priori.*

Pero ahora, ¿qué hemos de hacer del análisis conceptual a la luz de estas consideraciones? Como veo las cosas, los filósofos tienen dos opciones principales. La primera es tratar de minar las versiones prototipo de la representación de conceptos que, según he sugerido, son un desafío a las presuposiciones del análisis conceptual. Puesto que, en gran medida, es empírica la cuestión acerca de cómo representamos conceptos y categorías, es mejor que los filósofos que adopten esta estrategia estén preparados a llevar a cabo una investigación empírica seria y cuidadosa, comparable a la que motivó a los psicólogos a abandonar la tesis clásica, en primer lugar. Creo que una opción mucho más razonable para que la sigan los filósofos es que revisen sus supuestos y expectativas básicos acerca de las definiciones de conceptos abstractos. Como lo he intentado mostrar, si son correctas las historias de prototipos, entonces las dos expectativas primarias que típicamente sitúan los filósofos en los análisis de conceptos –simplicidad y evitar todos los contraejemplos– no pueden satisfacerse a la vez. De aquí que, mínimamente, parece que habría que abandonar uno de estos criterios o, al menos, relajarlo de manera significativa. Por una parte, los filósofos que exigen análisis en la forma de propiedades esenciales 'conjuga-

das', necesitarán abandonar el supuesto de que tal análisis dará cabida a *todas* nuestras intuiciones y estar dispuestos a aceptar la existencia de contraejemplos intuitivos. Por otra parte, si los filósofos van a insistir en tener análisis herméticos que no admitan contraejemplos intuitivos, entonces deben de abandonar la esperanza de tener definiciones nítidas y desarrollar una tolerancia para definiciones altamente disyuntivas y muy matizadas. Naturalmente, cualquiera que sea la estrategia que adoptemos ésta dependerá, en gran medida, del tipo de trabajo que deseamos que realicen para nosotros nuestros análisis. Empero, incluso aquí, al decidir cuál de los dos criterios debemos relajar, el trabajo empírico reciente tiene un papel que desempeñar. Por ejemplo, aun cuando parece que hay un traslape considerable entre los juicios intuitivos de diferentes sujetos con respecto a casos muy típicos y muy atípicos, recientemente se ha demostrado (Barsolou, 1987) que, entre diversos individuos, hay mucha diferencia en la gradación en casos moderadamente típicos. En efecto, para individuos singulares, las gradaciones de tipicalidad para un concepto dado se ha mostrado que cambian progresivamente durante un periodo extenso de tiempo (con poco cambio en el periodo de un día y un cambio importante en el periodo de un mes). Algunos investigadores, tales como Barsolou, creen que esto refleja considerable inestabilidad en nuestro esquema de representación conceptual. Si esto es correcto, entonces cualquier definición fija –incluso una que sea altamente disyuntiva– puede no llegar a capturar todas las intuiciones todo el tiempo, pues estas intuiciones no serían universales o estáticas. En consecuencia, abandonar la simplicidad a fin de acomodar todas las intuiciones puede ser una estrategia equivocada y un enfoque más prometedor podría ser el de preservar la simplicidad y reconocer que no podemos capturar todos los juicios intuitivos. Vale la pena señalar que esta última táctica no es totalmente ajena a la filosofía. Por ejemplo, en ciertos aspectos es muy

similar al elocuente estudio de Carnap (1950) del concepto ***explicación***, donde se apoya explícitamente la ***revisión*** de nuestros juicios intuitivos de categorización, por razones de simplicidad. De manera similar, en respuesta a los diversos contraejemplos intuitivos a los utilitaristas, Smart (1965, 1990) sugiere que han de cuestionarse o ignorarse nuestras intuiciones y no la versión utilitaria de la moralidad. Si es correcta la historia prototipo de la representación de conceptos, entonces las estrategias sugeridas por estos filósofos –en particular, reducir la importancia de los contraejemplos intuitivos– merece una consideración y una aplicación mayores que las que ha recibido.

Trad. de José A. Robles

Bibliografía

Armstrong, S.L., Gleitman, L.R. y Gleitman, H. (1983): 'What some concepts might not be', *Cognition* 13, 263-308

Atchinstein, P. (1968): *The Concepts of Science*, Johns Hopkins University Press, Baltimore, Maryland.

Barsalou, L. (1983): 'Ad-hoc categories', *Memory and Cognition* 10, 82-93.

_____ (1985): 'Ideals, central tendency, and frequency of instantiation', *Journal of Experimental Psychlogy: Learning, Memory, and Cognition* 11, 629-54.

_____ (1987): 'The instability of graded structure: implications for the nature of concepts', en *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, U. Neisser (comp.), Cambridge University Press, Cambridge.

Brown, H. (1988): *Rationality*, Routledge, Londres.

Carnap, R. (1950): *Logical Foundations of Probability*, University of Chicago Press, Chicago.

Cherniak, C. (1984): 'Prototypicality and deductive reasoning', *Journal of Verbal Learning and Verbal Behavior* 23, 625-42.

Chomsky, N. (1965): *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.

_____ (1972): *Language and Mind*, Harcourt, Brace and Jovanovich, Nueva York.

Cornman, J.W., Lehrer, K. y Pappas, G.S. (1982): *Philosophical Problems and Arguments: An Introduction* (3a. ed.), MacMillan, Nueva York.

Fodor, J. (1987): *Psychosemantics*, Bradford/MIT Press, Cambridge, MA.

Hampton, M.A. (1981): 'An Investigation of the Nature of Abstract Concepts', *Memory and Cognition*: 9(2), 149-56.

Horgan, T. (190): 'Psychologistic semantics, robust vagueness, and the Philosophy of language', en *Meanings and Prototypes: Studies in Linguistic Categorization*, Routledge, Londres, 535-57.

Kripke, S. (1972): *Naming and Necessity*, Harvard University Press, Cambridge, MA.

Lakoff, G. (1987): *Women, Fire and Dangerous Things*, University of Chicago Press, Chicago, Ill.

Malt, B.C. y Smith, E.E. (1982): 'The role of familiarity in determining typicality', *Memory and Cognition*, vol. 10 (1), 69-75.

McCloskey, M. y Glcuksberg, S. (1978): 'Natural categories: Well defined or fuzzy sets?', *Memory and Cogniton* 6, 462-72.

Medin, D.L. y Smith, E.E. (1984): 'Concepts and concept formation', *Annual Review of Psychology* 35, 113-38.

Putnam, H. (1962): 'The Analytic and the Synthetic', en *Minnesota Studies in the Philosophy of Science, vol. III*, H. Feigl y G. Maxwell (comps.), University of Minnesota press, Minneapolis, MN, 350-97.

_____ (1975): 'The meaning of "meaning"', en *Mind, Language and Reality: Philosophical Papers, vol. II*, Cambridge University press, Cambridge, 215 -71.

Rey, G. (1983): 'Concepts and stereotypes', *Cognition* 15, 237-62.

_____ (1985): 'Concepts and conceptions: A reply to Smith, Medin and Rips', *Cognition* 19, 297-303.

Rips, L.J., Shoben, E.J. y Smith, E.E. (1973): 'Semantic distance and the verification of semantic relations', *Journal of Verbal Learning and verbal Behavior* 12, 1-20.

Rips, L.J. (1975): 'Inductive judgements about natural categories', en *Cognitive Development and the Acquisition of Language*, T.E. Moore (comp.), Academic press, Nueva York.

Rosch, E. (1973): 'Cognitive representation of semantic categories', *Journal of Experimental Psychology: General* 104, 192-233.

_____ (1975): 'Principles of categorization', en *Cognition and Categorization*, E. Rosch y B. Lloyd (comps.), Lawrence Erlbaum, Hillside, Nueva Jersey, 27-48.

Rosch, E. y Mervis, C. (1975): 'Family resemblances: studies in the internal structures of categories', *Cognitive psychology* 8, 382-439.

Rosch, E., Simpson, C. y Miller, R.S. (1976): 'Structural bases of typicality effects', *Journal of Experimental Psychology: Human Perception and Performance* 2, 491-502.

Roth, E.M., Shoben, E.J. (1983): 'The effect of context on the structure of categories', *Cognitive psychology* 15, 346-378.

Rumelhart, D., Smolensky, P. McClelland, J. y Hinton, G. (1986): 'Schemata and sequential thought processes in PDP models', en *Parallel Ditributed Processsing*, vol. II, MIT/Bradford, Cambridge, MA, pp. 7-57.

Smart, J.J.C. (1965): 'The Methods of ethics and the methods of science', *Journal of Philosophy* 62, pp. 344-349.

_____ (1990): 'Integrity and squeamishness', en *Utilitarianism and its Critics*, J. Grover (comp.), MacMillan, Nueva York, pp. 170-174.

Smith, E.A. (1990): 'Categorization', en *Thinking: An Invitation to Cognitive Science*, vol. III, D. Osherson y E. Smith (comps.), MIT, Cambridge, MA, pp. 33-53.

Smith E.A. y Medin, D.L. (1981): *Categories and concepts*, Harvard University Press, Cambridge, MA.

Smolensky, P. (1988): 'On the proper treatment of connectionism', *The Behavioral and Brain Sciences* 11, pp. 1-74

Stich, S.P. (1990): *The Fragmentation of Reason*, Bradford/MIT Press, Cambridge, MA.

_____ (por aparecer): 'What is a Theory of Mental Representation?'.

Tversky, A. y Gati, I. (1978): 'Studies of Similarity', en *Cognition and Categorization*, E. Rosch y B. Lloyd (comps.), Lawrence Erlbaum Associates, Hillsdale, N.J.

White, A. R. (1975): 'Conceptual analysis', en *The Owl of Minerva*, C.J. Bontempo y S.J. Odell (comps.), McGraw Hill, Nueva York.

Williams, B. (1970): 'The self and the future', *Philosophical Review*, vol. LXXIX (2), 161-180.

Wittgenstein, L. (1953): *Philosophical Investigations*, Mac-Millan, Nueva York.

FILOSOFÍA Y CIENCIAS DE LA CONDUCTA

Nydia Lara Zavala*

Francisco Cervantes-Pérez*

Es factible sostener que la filosofía puede jugar tres papeles distintos en relación con la gestación y el desarrollo de una ciencia:

a) Puede enfrentarse a un problema filosófico tradicional para deslindar terrenos y permitir así una formulación clara de la clase de preguntas que pueden ser abordadas por la ciencia,
b) puede contribuir a la gestación de nuevas herramientas metodológicas para abordar el estudio de una nueva cuestión, o
c) puede utilizar los recientes avances que se han generado en alguna o algunas áreas de la ciencia para explorar si la nueva evidencia puede ofrecer información que le sea útil para resolver un problema filosófico.

No es exagerado afirmar que durante la gestación y el desarrollo de las ciencias que se ocupan de la conducta (natural y artificial) bajo el paradigma de procesamiento de información, actualmente denominadas 'ciencias cognitivas', la filosofía de hecho ha jugado estos tres papeles en distintos momentos de la historia de dichas ciencias. En este trabajo vamos a tratar de mostrar cómo en efecto la filosofía ha desempeñado estos tres papeles y lo que se puede lograr gracias a ella.

La filosofía ha jugado un papel importante en la conformación de las ciencias cognitivas, en particular gracias al desarrollo y

* Laboratorio de Neurocomputación, Centro de Instrumentos, UNAM.

la aplicación de dos de sus ramas: la lógica y la filosofía de la mente. Desde ambas perspectivas se abordó, sobre todo durante el siglo XIX y, podríamos decir, hasta Frege, lo que parecía ser un mismo problema, a saber, la organización y el contenido del lenguaje, puesto que las así llamadas 'leyes del pensamiento', de las cuales se ocupaban los psicólogos, no eran otra cosa que las leyes lógicas que rigen al lenguaje. El estudio de la lógica y el estudio de la mente perecían, pues, ser dos caras de una misma moneda.

El estudio del lenguaje absorbió la atención no sólo de muchos psicólogos, sino también la de no pocos matemáticos y de algunos lógicos, por la sencilla razón de que, como se sugirió más arriba, el lenguaje era visto como el instrumento que permitía procesar y transmitir un pensamiento objetivo, el cual podía quedar identificado como el resultado de alguna clase de actividad interna al sujeto (mental o física). Sin embargo, el estudio del lenguaje en tanto que expresión de la actividad interna de pensar muy pronto dio lugar al surgimiento de dos nociones antagónicas del pensamiento: la psicologista y la anti-psicologista . En la postura psicologista (iniciada por gente como Fries y Beneke) se concebía al pensamiento como esencialmente una actividad de la mente. La postura anti-psicologista, en cambio, concebía al pensamiento como una especie de cálculo, es decir, como una técnica o procedimiento formal, más que como un producto de la mente. Ahora bien, contrariamente a lo que muchos podrían pensar, este antagonismo no ha perdido vigencia, puesto que la misma tensión entre estas dos nociones opuestas de pensamiento se vuelve a manifestar en la actualidad justamente en las ciencias cognitivas. Esto es lo que ahora pasaremos a hacer ver.

La tradición cartesiana legó a la posteridad la idea de que la esencia de la mente es el pensamiento, por lo que una teoría del pensamiento tenía que estar inevitablemente ligada a una teoría

de la mente. Para los empiristas ingleses, sin embargo, el pensamiento era más bien un poder de la mente, no su esencia, y resultaba de una organización particular de ideas que surgían o provenían de la experiencia, recurriendo para ello a leyes de asociación. De ese modo, para los empiristas una teoría del pensamiento no podía ser equivalente a una teoría de la mente, sino más bien a una teoría de las operaciones "mentales". Por su parte, Kant propuso la existencia de un esquema conceptual *a priori*, gracias al cual se lograba imponer un orden en los datos de la experiencia. De otro modo, lo que normalmente llamamos 'nuestra experiencia' no sería inteligible. En este sentido, para Kant la mente tenía que ser considerada como una estructura regulada por las formas puras de la intuición sensible, las categorías y los *data* generados por su propia actividad. De ahí que en su caso el conocimiento de las operaciones de la mente sí desembocaba en una teoría de la mente.

Durante el siglo XIX, la psicología era un estudio en el que se mezclaban o amalgamaban diversas nociones de mente. La mayoría de quienes se ocupaban de estos temas parecía aceptar que era el alma lo que dotaba de unidad a los procesos mentales, que eran de lo más variado. No obstante, el material empírico de la psicología tenía inevitablemente que ver con la manera como se asociaban las sensaciones y las ideas en concordancia con ciertas leyes generales comunes a todos los seres pensantes, a todas las mentes. El pensamiento se entendía, por consiguiente, como el resultado de una actividad de la mente lógica y si bien lo que el pensamiento conectaba eran ideas, lo hacía, una vez más, en concordancia con ciertos principios *a priori*. Dichos principios no sólo no eran un producto de la experiencia, sino que eran el presupuesto de la experiencia. La lógica, que también se suponía que se ocupaba del pensamiento, no era considerada como una disciplina completamente autónoma en relación con la psicología. Se creía que el conocimiento de las leyes del pensamiento

era relevante para la justificación de las combinaciones de ideas. Las leyes abstractas del pensamiento eran las mismas para todos, pero las correlaciones de ideas efectuadas con base en dichas leyes variaban de mente en mente. De esa manera, aunque las leyes formales del pensar daban los modos de combinación de ideas, de todos modos no podían decir nada sobre el contenido de pensamientos particulares. Desde esta perspectiva, a la psicología le correspondía dar cuenta no sólo de aquello que se obtenía como medio mental al combinarse las ideas, sino también de la manera como la mente imponía sus leyes de pensamiento para la combinación de las ideas.

Por otra parte, era normal que la forma de abordar el estudio de la organización y los contenidos de la mente se planteara también en términos de introspección. Los piscologistas entrenaban a sus sujetos de estudio a "mirar dentro de ellos" y a reportar verbalmente lo que les pasaba cuando se les pedía que llevaran a cabo actividades de diversa índole. Los reportes verbales se interpretaban como expresiones lingüísticas de los procesos mentales, por lo que el estudio de la mente empezó a concebirse de manera natural en términos de la estructura y las propiedades semánticas del lenguaje. Esto tenía que ser así, puesto que el lenguaje parecía ser como una especie de puente capaz de permitir extraer las leyes abstractas del pensar (dado que el contenido de las actividades mentales a través de símbolos y combinaciones de símbolos representaban ideas) y en parte porque era gracias a los símbolos lingüísticos que los hablantes podían tanto referirse a como describir la conducta y las acciones de las personas. Gracias a la introspección, el sujeto podía dar cuenta no sólo de cómo se realizaban sus actividades mentales, sino también de cómo realizaba o planeaba la ejecución de ciertas actividades conductuales, recurriendo para ello a la generación de modelos internos del mundo externo. La percepción sensible, la fantasía o la imaginación, unidas a los deseos, las creencias,

las intenciones y demás actitudes proposicionales, podían ser analizadas y comprendidas no sólo por medio de las conductas que lleva a cabo un individuo, sino también por medio de la verbalización introspectiva y en la que intervenían todas aquellas variantes que el individuo podía pensar o imaginar sin que necesariamente se actualizaran.

Por la manera como habían sido concebidas las relaciones entre, por una parte, el lenguaje y, por otra, las actividades mentales y conductuales, la mayoría de las investigaciones se abocaron a investigar simultáneamente el funcionamiento y las interacciones de tres clases distintas de procesos:

a) los mentales, en principio sólo accesible al sujeto cognoscente,
b) los conductuales, abiertos al conocimiento público, y
c) los lingüísticos, considerados como el canal a través del cual era posible comunicar y conocer los procesos mentales y vincular así lo público con lo privado.

En este sentido, el lenguaje era considerado como el instrumento público de expresión de las actividades mentales, en tanto que las actividades mentales se concebían como la causa del desempeño tanto de las actividades lingüísticas como de las conductuales.

La convicción de que las actividades mentales no sólo eran la causa sino que eran lo que dirigía la actualización de las otras dos clases de actividades obligó a tratar de dar cuenta, primero, del contenido y la organización de la mente para, sobre esa base, tratar de explicar la conducta y el lenguaje. En todo caso, lo importante es notar que se tendía a pensar que tanto el estudio del lenguaje como el de la conducta estaban inevitablemente subordinados al de la mente. El problema de este enfoque es que el conocimiento del contenido y de la organización de la mente sólo podía ser abordado o bien a través de la introspección, lo

cual no era un método fiable, puesto que se trataba de un "método" por completo privado, o bien a través del lenguaje. Además, parte de la dificultad era que no se podía determinar con certeza hasta qué punto efectivamente se podía comunicar sin distorsionar, a través del empleo de signos, lo que acontecía en la mente del sujeto. De alguna manera, se pensaba, averiguar esto implicaba poder contrastar las representaciones de una mente con otra para poder determinar si había o no alguna clase de correspondencia entre signos e ideas, lo cual, como ya se dijo, sólo podía hacerse por la vía de la introspección, confinada sólo al conocimiento en primera persona. El acceso a lo mental por la vía del lenguaje ya era en sí mismo un rompecabezas filosófico, pero vale la pena notar que, por la manera como estaban concebidas las conexiones entre los tres tipos de procesos, aun cuando se pudiera dar por sentado que las palabras correspondían de manera sistemática a ideas y que la organización de una oración reflejaba una determinada organización mental, de todos modos quedaba pendiente el otro problema, de explicar cómo se conectaba la actividad mental con la corpórea.

El lenguaje como la expresión de las actividades mentales no sólo exigía la explicación concomitante de tres diferentes clases de procesos distintos (*i.e.*, mentales, lingüísticos y conductuales), sino que también exigía la explicación de sus conexiones. El lenguaje podía ser comprendido como el puente que permitía vincular a las actividades mentales con las conductuales, pero todo parecía indicar que mientras se concibiera el lenguaje como un mero instrumento para la expresión de actividades mentales, el estudio del lenguaje por sí solo no podía proporcionar ninguna clase de conocimiento de lo que acontecía en la mente, además de que tampoco permitía ofrecer una explicación clara de cómo los porcesos mentales podían ser traducidos primero en palabras y oraciones y, posteriormente, retraducidos otra vez a ideas o procesos mentales, los cuales (se suponía) son lo que daba origen

a conductas concretas de individuos concretos. Así, este complejo entramado de relaciones e interconexiones entre actividades mentales, lingüísticas y conductuales hace ver que, desde esta perspectiva, se estaba tratando de abordar cuestiones diferentes como si se tratara de una y la misma. En efecto, un asunto era el estudio del lenguaje como una vía para comunicar dos diferentes clases de "actividades", las mentales y las conductuales, y otro era el estudio de las relaciones entre diferentes clases de sustancias, esto es, la mente y el cuerpo. Esta segunda clase de estudio empezaba a revelarse como completamente redundante.

El problema mente-cuerpo es un enigma filosófico tradicional para el cual, como es de esperarse, no hay una solución universal satisfactoria. Hay que notar, por otra parte, que el estudio del lenguaje como estudio del medio idóneo para la expresión de la conducta no equivale al estudio de aquello que la causa. Sea como fuere, el lenguaje aun visto como expresión del pensamiento en el sentido psicologista del término, no puede ser visto como la causa directa de ninguna actividad, sino sólo como la expresión de una línea de conducta posible. Si se dice, por ejemplo, "Trae el lápiz que está sobre la mesa" o "Suma dos y dos", aun cuando es fácil darse cuenta de que la primera oración expresa un movimiento corpóreo, esto es, una actividad física, y la segunda expresa una "actividad" que puede ser catalogada como mental, las dos oraciones aluden a clases de conducta que pueden realizarse sólo a través del empleo de palabras. Ninguna de las dos oraciones apunta a algo más que a una conducta posible, por lo que cualquiera de estas dos oraciones puede ser comprendida simplemente como la expresión de actividades de clases diferentes y que se pueden o no llevar a cabo. En otras palabras, la significatividad del lenguaje se deriva de su poder para expresar diferentes clases de actividades, esto es, conducta específica. Empero, ni lo que expresa el lenguaje ni la clase de

conducta que está indicada en la oración refieren a aquello que los genera. Esto es, cuestiones tales como qué o quién profiere la oración, qué entidad "se esconde" detrás del cuerpo, etc., son cuestiones que dejan de tener relevancia en este estudio, ya que de lo que se trata es simplemente de comprender cómo puede el lenguaje funcionar y dar lugar a actividades conductuales específicas. Desde esta perspectiva, el estudio de la clase de conducta que puede ser expresada por un trozo de lenguaje o el estudio de la manera como una conducta puede desencadenarse a partir de lo que indica una oración, se vuelven objetos de estudio perfectamente objetivos y no involucran ninguna otra clase de consideraciones. Esto es así porque esta clase de estudio sólo versa sobre los efectos que puede generar el empleo de una oración y de las actividades a las que puede dar lugar. Una vez que se aclara que es posible estudiar el lenguaje como si se tratara de una herramienta que contiene mecanismos para expresar alguna clase de actividad o de conducta, el siguiente paso es trazar las líneas de investigación que pueden derivarse de un estudio así. Normalmente, este es el punto donde la ciencia toma la estafeta e inicia su labor por su cuenta.

No estará de más recordar que originalmente se habían planteado simultáneamente alrededor del estudio del lenguaje dos preguntas:

a) la primera, como ya se dijo, concernía a las leyes del pensamiento, entendidas como leyes de la lógica, y
b) la segunda se refería a las leyes del pensamiento, entendidas como leyes de la mente.

Ambas clases de investigación convergían en el estudio del lenguaje y, por ende, en el de las relaciones entre el lenguaje y la mente, por una parte, y el lenguaje y la conducta, por la otra. Es claro que difícilmente podía la segunda cuestión separarse

de la primera. Sin embargo, pasaron muchos años antes de que se comprendiera que la psicología era, más que el estudio de la mente, el estudio de la conducta vía el estudio del lenguaje y sus descripciones, usos, etc., y que el estudio de las leyes formales del pensamiento desligadas de los simbolismos concretos y sus aplicaciones no pasa de ser un mero juego. El estudio del lenguaje era cada vez más vital y era el canal para la conexión entre el pensamiento, entendido como algo visible o palpable, y la conducta, humana o no.

Sin duda, en los lógicos matemáticos del siglo XIX poco a poco cristalizó la idea de que las leyes del pensamiento, en el sentido anti-psicologista del término, podían concebirse como una serie de combinaciones simbólicas susceptibles de ser conocidas con total independencia de los supuestos contenidos del pensamiento a los que los signos aluden. Sin embargo, los psicologistas tendían a ver en la lógica una disciplina subordinada a la psicología, ya que se pensaba que los signos que se usaban eran forzosamente nombres de ideas y porque se creía que divorciar los signos de sus significados mentales era cancelar el conocimiento de la mente. El problema para ellos era que el recurso a la introspección no podía ser un buen método científico para acceder a nada (los contenidos de la mente o lo que fuere). Pero, además, la única forma de acceso a los contenidos de otras mentes eran las manifestaciones públicas del sujeto y un análisis más minucioso de los reportes verbales que trataban de dar cuenta de las cuasas o motivaciones de sus conductas permitía poner en duda la necesidad de postular que los signos lingüísticos tenían que estar conectados con ideas para ser significativos. En última instancia, cuando se le pedía a un sujeto traer, por ejemplo, un lápiz rojo, lo que menos importaba era si su representación mental de lápiz o de rojo coincidía o no con las representaciones mentales de lápices rojos de otras personas. De hecho, ni siquiera importaba si el sujeto en cuestión efectivamente requería de una

representación mental para llevar a cabo correctamente la tarea que se le indicaba. Lo único que importaba era que el sujeto fuera capaz de procesar las información contenida en la oración para ejecutar correctamente la tarea y realizarla.

Si lo que se ha dicho es acertado, entonces puede apreciarse que el avance consistió en darse cuenta de que la propiedad de las expresiones lingüísticas de expresar una actividad o de desencadenarla poco o nada tiene que ver con lo que acontece en la mente del sujeto, que es independiente de ella. La función del sujeto se puede limitar a simplemente seguir paso a paso y de manera mecánica la regla incorporada en la oración. El contenido de la oración no refiere a sus supuestos significados en la mente del sujeto o al modo como operan las entidades psíquicas, sino sólo a la manera como los signos están conectados unos con otros para expresar o generar una conducta concreta. Esto posibilita dos cosas:

a) desprenderse de las especulaciones metafísicas, tan preciadas por los psicologistas, en torno a los contenidos mentales, y
b) generar una teoría del lenguaje no meramente formal, sino también de sus modos de significación.

La distinción entre contenido mental y contenido lingüístico es una dicotomía que permite orientar la búsqueda de los principios que nos van a permitir eliminar los problemas de las interpretaciones psicologistas y construir las bases metodológicas de toda una serie de trabajos que culmina en la teoría de lo computable y que es, de hecho, la herramienta metodológica que hasta la fecha utilizan las ciencias cognitivas. Esto nos lleva a otro aspecto del tema general abordado en este ensayo.

Durante siglos los filósofos se han enfrentado al problema mente-cuerpo y han tratado de explorar la relación que vale entre la materia y la mente desde prácticamente todas las perspectivas,

sin que haya sido posible obtener una forma de acoplarlas de manera enteramente satisfactoria. De este estudio han brotado las más diversas teorías metafísicas en torno a las clases de entidades involucradas. Por ejemplo, se ha pensado en la mente y en el cuerpo como dos sustancias distintas que pueden tener o no una relación directa, se ha planteado la posibilidad de reducir una a la otra, de convertir a una en causa de la otra, etc. Lo cierto es que hasta la fecha y a pesar de todos los esfuerzos que el desarrollo de cada una de estas posibles combinaciones acarrea, puede afirmarse que la relación mente-cuerpo sigue siendo un problema insoluble. Es, pues, sumamente extraño que el desarrollo de la tecnología computacional y los avances en el área de las neurociencias, actualmente exploradas con entusiasmo por una multitud de filósofos, incite a pensar que las nuevas evidencias obtenidas en estas nuevas líneas de investigación den la pauta para la solución definitiva del problema mente-cuerpo. Nosotros pensamos que esto es un error y que se está sosteniendo en una nueva terminología una tesis que ya había sido superada, *viz.*, la tesis de que el lenguaje no es más que el instrumento para la expresión de algo diferente a él, esto es, algo "mental". Esto, obviamente, tiene que ver con la posición de las ciencias cognitivas en relación con el problema mente-cuerpo. Como se sabe, el nombre 'ciencias cognitivas' denota al grupo de disciplinas que estudian al hombre, a los animales, los cerebros, las máquinas, el lenguaje, etc., tomando como guía el paradigma de procesamiento de información, el cual se deriva directamente de las ciencias de la computación. Las ciencias de la computación surgen del estudio de la relación lenguaje-conducta, pero aunque se reconoce que el lenguaje es capaz de expresar tanto actividades mentales como físicas, no se incluye en esta propuesta ni una solución ni una reformulación del problema mente-cuerpo. Más bien, se promueve una forma de evadirlo. En este caso, lo que se acepta es que el lenguaje puede expresar, a través de palabras

y oraciones, actividades que pueden ser ejecutadas conductualmente, pero la cuestión de qué o quién las ejecute es lo único que no importa o que es irrelevante en este contexto.

El estudio del lenguaje como ligado a la conducta tiene la ventaja de que permite generar modelos de procesamiento de información que pueden ser aplicados, *inter alia*, a humanos, animales, máquinas, cerebros, códigos genéticos, etc. Se puede, por ejemplo, construir un algoritmo y diseñar una máquina que lo ejecute, pero también se puede estudiar la actividad de un órgano u organismo para tratar de analizar su "conducta" en términos del mismo paradigma, es decir, el de procesamiento de información. La inteligencia artificial y la robótica están haciendo lo primero; las neurociencias, las ramas derivadas de la fisiología de lo computable, la neuroetología, la psicología e incluso la lingüística están haciendo lo segundo. Pero lo importante es notar que no hay la menor razón para aceptar que cualquiera de ellas estudia o se ocupa de la mente o de algo "mental": lo que estudian es sencillamente la relación entre la conducta y la clase de procesamiento de información que suponemos que se requiere para ejecutarla.

Si todas las ramas de las ciencias cognitivas pueden aglutinarse alrededor del paradigma de procesamiento de información es porque dicho paradigma, al no tener ninguna clase de compromiso con la clase de entidad que procesa la información, abre la posibilidad de atrapar el análisis de un procesamiento de información tanto en abstracto como ligado a una entidad concreta, puesto que es obvio que la conducta asociada con un programa puede involucrar o no una actividad física. La expresión "Suma dos y dos" y la expresión "Trae el lápiz que está sobre la mesa" pueden ser identificadas simplemente como algoritmos que pueden ser ejecutados por cualquier tipo de entidad que esté capacitada para procesarlo o ejecutarlo conductualmente. Lo único que se asume es que el algoritmo (es decir, el lenguaje) expresa

actividades posibles tales que, si se ha construido la máquina adecuada para ello, a su vez puede desencadenar conductas específicas. Algunas de estas conductas exigen un movimiento corpóreo, otras no. Si a un tipo de conducta se le llama 'física' y a la otra 'mental', ello se debe a que no contamos con otro término para referirnos a la clase de conducta que no necesariamente implica un movimiento corpóreo. Pero ningún misterio está aquí implicado.

Es legítimo interpretar la conducta que llamamos 'mental' como algo que emerge de un algoritmo, siempre y cuando se comprenda que lo que emerge del algoritmo no es una entidad ni una actividad etérea, sino simplemente conducta. De hecho, el algoritmo se concibe como aquello que efectivamente puede generar una conducta, por lo que el mismo concepto de algoritmo acarrea consigo la idea de conducta, está lógicamente vinculado a ella. Por eso, un algoritmo puede ser objeto de experimentación: podemos probar lo que expresa un algoritmo por el efecto que produce como conducta. Pero suponer, como lo hacen muchos, que el algoritmo es la expresión de la actividad de un "algo", una mente, o que del algoritmo surge la entidad denominada 'mente', es tan sólo el resultado de una grave confusión. Ninguna mente está aquí involucrada: lo único que hay son signos, reglas, conducta. Ahora bien, parecería que no es así como los cognitivistas están tratando de reenfocar el problema mente-cuerpo. Empero, si lo que hemos dicho es correcto, entonces puede verse que los nuevos reduccionismos y emergentismos que se encuentran diseminados en todas las áreas de las ciencias cognitivas no son otra cosa que el resultado de una mala interpretación de sus propios conceptos y de los alcances de la metodología que utilizan.

Si vemos las cosas desde una perspectiva histórica podremos percatarnos de que el problema mente-cuerpo, en relación con el lenguaje y la conducta que puede ser expresada a través de o gracias a él, en el fondo ni se ha tocado. Lo que por nuestra

parte hemos tratado de hacer ver es simplemente que hay estrategias filosóficas que permiten abordar el estudio del pensamiento y de la conducta vía el estudio del lenguaje sin para ello vernos hundidos en el famoso problema mente-cuerpo. Este problema sigue pendiente y, por lo mismo, sigue siendo parte del quehacer filosófico. Nadie afirma que sea un error filosófico tratar de utilizar los nuevos resultados o avances de las ciencias para tratar de esclarecer un viejo problema filosófico. Lo que afirmamos es que la ciencia, al menos como la conocemos hasta ahora y por las razones que se han expuesto a lo largo de este trabajo, no tiene por qué abordar, para progresar, problemas que la misma filosofía aún considera como rompecabezas, esto es, como un conjunto de piezas sueltas que aún no sabemos siquiera cómo acomodar.

El problema de la relación mente cuerpo se terminó
de imprimir en agosto de 1993, en Grupo Editorial
Interlínea, S. A. de C. V., Irlanda 121 Bis,
col. Parque San Andrés, Coyoacán, tel. 689 4829.
Se imprimieron 1 000 ejemplares sobre papel cultural
de 36 gr; forros en cartulina couché de 210 gr; la
edición estuvo al cuidado de José Antonio Robles