

Revista RLF latinoamericana de filosofía

J. Comesaña, En defensa de un externalismo epistémico

E. Cresto, Creer, inferir y aceptar: una defensa de la inferencia a la mejor explicación apta para incrédulos

C. Sartorio, El comportamiento colectivo y la relación entre responsabilidad y causación

L. A. Rossi, El tedio y el terror: la crisis de 1929 en el curso de Heidegger
Los conceptos fundamentales de la metafísica

S. Pinto, Dos argumentos en contra del lenguaje de pensamiento

A. Cassini, Bunge, la lógica inductiva y el grado de confirmación de las teorías

D. Brauer, La "memoria productiva" - Acerca de la concepción de Hegel del recuerdo y la memoria

DISCUSIÓN - COMENTARIOS BIBLIOGRÁFICOS

COMITE DE REDACCION

María Julia Bertomeu
Leiser Madanes

Alejandro Cassini
(†) Ezequiel de Olaso
Secretaria: Cecilia Lastra

Osvaldo Guariglia
Mario A. Presas

CONSULTORES

Rafael Braun (Argentina)	Mario Bunge (Argentina-Canadá)
Danilo Cruz Vélez (Colombia)	Carla Cordua (Chile)
Jorge E. Gracia (EE.UU.)	Marcelo Dascal (Brasil)
Joao Paulo Monteiro (Brasil)	Alfonso Gómez-Lobo (EE.UU.)
O. Porchat Pereira (Brasil)	Francisco Miró Quesada (Perú)
David Sobrevilla (Perú)	C. Ulises Moulines (R. Fed. Alemana)
Roberto Torretti (Chile)	Ernesto Sosa (EE.UU.)
Luis Villoro (México)	Margarita Valdés (México)
Ernesto Garzón Valdés (Argentina-Alemania)	

Para suscripciones, pedidos, correspondencia o informaciones dirigirse a:

REVISTA LATINOAMERICANA DE FILOSOFIA

Casilla de Correo 5379

Correo Central

C1000WCB Buenos Aires - Argentina

Fax: (54-11) 4787-0533 E-mail: cifrlf@mail.retina.ar

Para suscripciones y pedidos en Estados Unidos y Canadá dirigirse a:

REVISTA LATINOAMERICANA DE FILOSOFIA

(dos direcciones)

BOX 1493
Birmingham, Alabama 35201-1943 - U.S.A.

15 Southwest Park
Westwood, Mass 02090 - U.S.A.

Argentina

Individuos	\$22
Instituciones	\$44

Exterior

Individuos	U\$S 25
Instituciones	U\$S 35

(más gastos de envío)

La Revista Latinoamericana de Filosofía se publica los meses de mayo y noviembre de cada año. Director Responsable: Ana María Mustapic, Presidente del Centro de Investigaciones Filosóficas, CIF, propietario de la publicación. Domicilio Legal: Miñones 2073, 1428 Buenos Aires. Copyright. Centro de Investigaciones Filosóficas, CIF. Queda hecho el depósito que marca la Ley N° 11723.

ISSN 0325-0725

Diseño de tapa: Buyi Presas

Noviembre 2002

REVISTA LATINOAMERICANA DE FILOSOFIA

PRIMAVERA 2002

VOLUMEN XXVIII

N° 2

SUMARIO

ARTÍCULOS

- | | |
|--|-----|
| J. Comesaña, En defensa de un externalismo epistémico | 173 |
| Concurso Ezequiel de Olaso 2003 | 198 |
| E. Cresto, Creer, inferir y aceptar: una defensa de la inferencia a la mejor explicación apta para incrédulos | 201 |
| C. Sartorio, El comportamiento colectivo y la relación entre responsabilidad y causación | 231 |
| L. A. Rossi, El tedio y el terror: la crisis de 1929 en el curso de Heidegger
<i>Los conceptos fundamentales de la metafísica</i> | 247 |
| S. Pinto, Dos argumentos en contra del lenguaje de pensamiento..... | 275 |
| A. Cassini, Bunge, la lógica inductiva y el grado de confirmación de las teorías | 303 |
| D. Brauer, La "memoria productiva" - Acerca de la concepción de Hegel del recuerdo y la memoria..... | 319 |

DISCUSIÓN

- | | |
|--|-----|
| M. Bunge, La lógica inductiva es inútil..... | 339 |
|--|-----|

COMENTARIOS BIBLIOGRÁFICOS

V. Palacios, *Lenguaje y pacto en Thomas Hobbes* (A. Pac). Z. Loparic, *A semântica transcendental de Kant* (H. Pringe). N. L. Cordero (comp.), *Ontologie et Dialogue. Hommage à Pierre Aubenque* (O. Kulesz). A. Salles y M. J. Bertomeu (edit.), *Bioethics: Latin American Perspectives* (S. Woods). M. Baumann, *El mercado de la virtud: moral y responsabilidad social en la sociedad liberal* (J. Montero)..... 343

COLABORADORES

DANIEL BRAUER, hizo su doctorado en filosofía en la Universidad de Erlangen-Nürnberg (1980). Ha sido becario del Servicio Alemán de Intercambio Académico y de la Fundación Alexander von Humboldt. Actualmente es titular de las cátedras de *Filosofía de la Historia* y de *Metafísica* en la Universidad de Buenos Aires, e Investigador Independiente del Consejo Nacional de Investigaciones Científicas y Técnicas. Sólo mencionaremos dos de sus publicaciones que resultan significativas de su labor: (1) el libro: *Dialektik der Zeit. Untersuchungen zu Hegels Metaphysik der Weltgeschichte*, Stuttgart 1982 y (2) "Significado y verdad en la narración histórica. Una re-visión de la objetividad historiográfica", en: *Revista Latinoamericana de Filosofía*, vol. XXVI, N° 1 (Otoño 2000).

M. BUNGE. Ver RLF, Vol. XXV N° 1 (otoño 1999).

A. CASSINI. Ver RLF, Vol. XXV N° 1 (otoño 1999).

JUAN COMESAÑA, Licenciado en Filosofía (UBA), 1996. Actualmente es estudiante de doctorado en la Universidad de Brown. Ha publicado el libro *Lógica informal, falacias y argumentos filosóficos* (Eudeba, 1998; segunda edición 2001), así como artículos sobre filosofía de la ciencia, semántica, y teoría del conocimiento.

Dirección: Juan_Comesana@Brown.edu.

ELEONORA CRESTO, Licenciada en Filosofía en la Universidad de Buenos Aires (1995). Ha publicado varios artículos sobre temas de lógica y teoría del conocimiento. Actualmente es estudiante de doctorado en la Universidad de Columbia. Se ha especializado en problemas de revisión de creencias y teoría de la decisión cognitiva.

Dirección: ebc17@columbia.edu

SILVIO PINTO, profesor de filosofía del lenguaje y epistemología en la Universidad Autónoma del Estado de Morelos (México) desde el año 2000. Obtuvo una estancia pos-doctoral en el Instituto de Investigaciones Filosóficas (UNAM) entre 1998 y el 2000. Su título de doctor lo obtuvo en el King's College London (Universidad de Londres) en 1998 con una tesis sobre la filosofía de la matemática de Ludwig Wittgenstein. Su maestría la hizo en filosofía en la Universidade Estadual de Campinas (Brasil) en 1991, con una tesis sobre la segunda antinomia de la razón pura. Sus intereses se concentran en la filosofía del lenguaje, filosofía de la mente, epistemología y filosofía de las matemáticas.
Dirección: silviopint@yahoo.com

LUIS A. ROSSI (1964), es egresado de la carrera de Filosofía de la Universidad de Buenos Aires. Desempeña tareas docentes en esta universidad y en la Universidad Nacional de Quilmes. Ha participado de grupos de investigación en ambas universidades y realiza sus estudios de doctorado acerca del nacionalismo en las filosofías de M. Heidegger y C. Schmitt bajo la dirección del Prof. Jorge Dotti. Ha publicado artículos sobre la historia de las ideas en la Argentina y sobre filosofía política.
Dirección: lrossi@unq.edu.ar

CAROLINA SARTORIO, Licenciada en Filosofía (UBA), 1996. Actualmente es estudiante de doctorado en el *Massachusetts Institute of Technology* y se especializa en metafísica y ética. Ha publicado el libro *Conjuntos e infinitos* (Eudeba, 2000), así como artículos sobre filosofía de la ciencia y semántica.
Dirección: caro@mit.edu.

EN DEFENSA DE UN EXTERNALISMO EPISTÉMICO*

Juan Comesaña

1. Introducción

El debate entre concepciones internalistas y externalistas de la justificación epistémica es uno de los problemas fundamentales de la teoría del conocimiento contemporánea. Como con otros problemas filosóficos fundamentales, no siempre es fácil saber cuáles son, exactamente, las posiciones en conflicto. Esta cuestión conceptual ha sido aclarada en gran medida recientemente, pero la cuestión doctrinal (para seguir con la terminología quineana) todavía queda en pie. Sabemos en qué consistiría ser externalista o internalista (más bien: conocemos varias maneras de ser internalistas o externalistas), pero todavía no sabemos qué debemos ser.

En este artículo argumento que, por lo menos con respecto a una de las distinciones, debemos ser externalistas. La estrategia general será la de un argumento por eliminación: ninguna de las alternativas internalistas puede solucionar un grave problema, problema que sí es solucionado por un externalismo de corte confiabilista.

2. Preliminares: conocimiento, justificación epistémica y garantía

Una primera cuestión que debe clarificarse al hablar de internalismo y externalismo epistémicos¹ es a qué se aplica la distinción. Dije al co-

* Una versión de este artículo fue presentada en una charla en el Instituto de Filosofía de la Universidad de Buenos Aires organizada por Eduardo Barrio. Mi sincero agradecimiento a Eduardo, así como a los asistentes a la charla por sus preguntas y comentarios. Muchas gracias también a Alejandro Cassini, Manuel Comesaña, Dave Matheson, Michael Pace, Carolina Sartorio, Joseph Shieber y Ernesto Sosa. Este trabajo fue realizado gracias al apoyo de la Fundación Antorchas.

1. El calificativo "epistémico" es necesario porque hay distinciones en filo-

mienzo que el debate gira en torno a la idea de justificación epistémica: quiero ahora aclarar qué es la justificación epistémica, y por qué creo que es con respecto a ella que la discusión es interesante.

En la sección siguiente voy a caracterizar distintas versiones de la distinción entre internalismo y externalismo. Pero para abordar la pregunta por el *locus* de dicha distinción hay que tener siquiera una idea vaga de en qué consiste. En términos generales, entonces, ser internalista es creer que el estatus epistémico de una actitud doxástica por parte de un sujeto depende exclusivamente de lo que pasa "adentro" de dicho sujeto, mientras que un externalista cree que dicho estatus puede ser determinado en parte por lo que sucede "afuera" del sujeto.² (Gran parte de la tarea clarificadora consiste en expresar de manera más literal lo que estas metáforas espaciales indican.)

La noción epistémica que tradicionalmente ha sido considerada como central es la noción de conocimiento. Ahora bien, puesto que conocimiento implica verdad, poco sentido tendría ser internalista con respecto al conocimiento. Si Eduardo sabe que la nieve es blanca, entonces la nieve es blanca; y como que la nieve sea blanca no es algo que está determinado exclusivamente por lo que pasa adentro de Eduardo, qué es lo que sabe Eduardo no es algo que esté determinado exclusivamente por lo que pasa dentro de él.³

Otra noción epistémica importante en la discusión contemporánea

sofía de la mente y psicología moral que también se marcan con las palabras "internalismo" y "externalismo". En lo que sigue, me referiré siempre a las posiciones en teoría del conocimiento, dejando el calificativo implícito.

2. Hablo de actitudes doxásticas y no simplemente de creencias porque también la suspensión del juicio con respecto a una proposición puede estar justificada o injustificada.

3. Hay quienes sostienen, es cierto, que que la nieve sea blanca *sí* está determinado por lo que pasa adentro de Eduardo –o, (sólo) un poco más plausiblemente, por lo que pasa al interior de la comunidad a la que pertenece Eduardo–. Si uno sostiene algo así, entonces puede ser internalista (o socio-internalista) con respecto al conocimiento. Pero esa posición con respecto a la verdad es, por decirlo suavemente, una teoría en busca de un argumento –y no, por lo tanto, un punto de partida plausible en la discusión epistémica–. En esta trabajo presupongo una concepción no epistémica de la verdad. Para una sugerencia de combinar el confiabilismo en teoría del conocimiento con una teoría epistémica de la verdad, ver Eleonora Cresto (1997).

es la noción de garantía. Es útil presentar esa noción con una breve historia, levemente mitológica. Según el mito, Platón definió el conocimiento como creencia verdadera justificada y convenció al resto de los filósofos que se interesaron por el tema durante unos veinticuatro siglos, cuando Edmund Gettier publicó una nota de tres páginas mostrando que la definición tenía contraejemplos.⁴ Uno de los contraejemplos tipo Gettier es el siguiente.⁵ Un empleado en la oficina de S, el señor Nogot, le ha dado a S datos que justifican que S crea "El señor Nogot, que está en la oficina, es dueño de un Ford",⁶ de lo que S deduce *p*: "Alguien en la oficina tiene un Ford". Pero, aunque S no lo sabe, el señor Nogot miente y *p* es verdadero sólo porque otra persona en la oficina, el señor Havit, es dueño de un Ford. En este caso, S tiene una creencia verdadera y justificada en *p*, pero sin embargo no es cierto que S sabe que *p*.

Tener una creencia justificada, entonces, no es suficiente para tener conocimiento, ni siquiera cuando la creencia en cuestión es verdadera. Pero, por supuesto, hay muchos casos en los que un sujeto sabe que una proposición es verdadera. Por lo tanto, tiene que haber *algo* (aunque no la justificación) que, sumado a una creencia verdadera, dé como resultado conocimiento. Llamemos a ese algo "garantía".⁷ "Garantía", entonces, es *definida* como "aquello que, sumado a una creencia verdadera, resulta en conocimiento", con lo cual no es posible formular contraejemplos tipo Gettier para la noción de garantía.

Es una tesis ampliamente difundida en la gnoseología contemporánea que la noción de garantía contiene aspectos externalistas. Trenton Merricks, por ejemplo, ha argumentado que cualquier análisis correcto de esa noción debe implicar que si una creencia está garantizada enton-

4. Cf. Gettier (1963). En ese artículo, Gettier dice que Platón consideró esa definición en el *Teetetus*, y que parece aceptarla en el *Menón*. La cuestión, obviamente, es disputada. Los "contraejemplos tipo Gettier", como se conocen en la bibliografía sobre el tema, fueron anticipados (aunque no como tales) por Russell (1948) y, según Chisholm (1977), Meinong. Para un detallado examen del estado de la cuestión hasta comienzos de los '80, ver Shoppe (1983).

5. Tomado de Lehrer (1965).

6. Los datos incluyen un certificado que dice que Nogot es dueño de un Ford, el hecho de que S ha visto a Nogot manejar un Ford varias veces, etc.

7. La terminología deriva de Alving Plantinga (1993).

ces es verdadera⁸ —y si eso fuera cierto, entonces tampoco tendría sentido ser internalista con respecto a la garantía—. Y, aunque uno tenga dudas con respecto al argumento general de Merricks, muchos de los análisis que distintos filósofos han ofrecido de la noción de garantía cumplen la condición de que una creencia puede estar garantizada sólo si es verdadera,⁹ lo cual ofrece razones inductivas en favor de la tesis.

Ni con respecto a la noción de conocimiento ni (aunque de manera menos segura) con respecto a la noción de garantía tiene sentido, entonces, ser internalista. ¿De qué hablamos, pues, cuando hablamos del debate entre internalismo y externalismo? Hablamos de internalismo y externalismo con respecto a la *justificación*, aquello que Gettier mostró que no es suficiente, ni siquiera con creencia y verdad, para darnos conocimiento. Pues aun cuando la justificación no sea tan poderosa como se creía antes de Gettier, muchos siguen pensando que tiene que cumplir un papel fundamental en la definición de conocimiento y que, en cualquier caso, es una noción interesante por mérito propio.

Se habla de justificación *epistémica* para distinguir el tipo de justificación que antes de Gettier se creía necesaria y, junto con creencia verdadera, suficiente para el conocimiento —y que, luego de Gettier, se considera todavía como centralmente relacionada con el conocimiento, aunque no suficiente junto con creencia verdadera y, quizás, ni siquiera necesaria—; para distinguir ese tipo de justificación, digo, de la justificación a veces llamada “práctica”, o “prudencial”. Alguien gravemente enfermo, por ejemplo, puede estar justificado en este sentido práctico en creer que no va a morir, puesto que si cree eso la probabilidad de que sane se eleva, aun sin estar justificado epistémicamente en creerlo. *Grosso modo*, la justificación epistémica está estrechamente relacionada con la verdad de la proposición creída, mientras que las justificaciones de otro tipo no lo están (aunque es notoriamente difícil transformar esta caracterización vaga en algo más preciso sin que se vuelva inaceptable).¹⁰

Lo que se discute, entonces, cuando se discute si hay que ser internalista o externalista, es *si el hecho de que una creencia esté justificada para*

8. Ver Merricks (1995).

9. Ver los trabajos de John Pollock (1986), Alvin Goldman (1976), James Tomberlin (1980), y Ernesto Sosa (2000a) y (200b).

10. En adelante, cuando se hable de justificación a secas debe entenderse que hablo de justificación epistémica.

un sujeto depende exclusivamente de lo que pasa dentro de ese sujeto o si factores externos desempeñan también un papel. En esta sección explicaré por qué la distinción se aplica a la noción de justificación y no a otras nociones epistémicas fundamentales. En la sección que sigue intento caracterizar de manera más precisa la distinción.

3. Cuatro tipos de internalismo/externalismo

Resulta que no hay sólo *una* distinción entre internalismo y externalismo, sino varias. Hay dos caracterizaciones positivas de internalismo, que responden a motivaciones para ser internalista que son ligeramente diferentes entre sí, y dos sentidos en los cuales uno puede ser internalista (de cualquiera de esos dos tipos) con respecto a la justificación. Hay, así, cuatro sentidos en los cuales uno puede ser internalista y, correspondiendo a ellos, cuatro sentidos en los cuales uno puede ser externalista. Con respecto a cada uno de esos sentidos, definiremos al externalismo simplemente como la negación del internalismo; pero es consistente ser internalista en uno de los sentidos y externalista en otro (aunque la mera coherencia de una posición no es en sí misma una recomendación muy fuerte en favor de esa posición).

Siguiendo a Conee y Feldman, llamaremos a las dos caracterizaciones positivas del internalismo “mentalismo” y “accesibilismo”.¹¹ Ambas versiones pueden definirse en términos de tesis de superveniencia. La idea general, recordemos, es que el estatus epistémico de una actitud doxástica por parte de un sujeto está enteramente determinado por configuraciones internas a ese sujeto. El accesibilismo explica la metáfora en términos que son en última instancia epistémicos:

ACCESIBILISMO: El estatus epistémico de las actitudes doxásticas de *S* sobreviene en factores a los cuales *S* tiene un tipo especial de acceso.

Una consecuencia del accesibilismo así definido es que si dos sujetos son idénticos con respecto a los estados a los que tienen el tipo especial de acceso en cuestión, entonces también son idénticos con respecto al estatus epistémico de sus actitudes doxásticas (por ejemplo, están justificados en creer las mismas proposiciones).

11. Ver Conee y Feldman (2001).

Es obvio que el accesibilismo no será una tesis totalmente inteligible hasta que se nos haya dicho qué es ese tipo especial de acceso, y en qué consiste "tenerlo". Conviene caracterizar la idea general de ese acceso en términos negativos: hay algunos factores que son externos al sujeto, en un sentido de "externo" que heredamos de Descartes –un sentido en el cual, por ejemplo, el hecho de que yo tenga manos es algo externo a mí–. Llamemos al tipo de acceso que tenemos a esos factores externos "acceso ordinario". El accesibilismo, entonces, es la tesis de que el estatus epistémico de las actitudes doxásticas está determinado por factores a los cuales el sujeto en cuestión tiene un acceso *no* ordinario. Es cierto que no se nos ha dicho todavía, en términos positivos, en qué consiste ese acceso no ordinario. Pero una caracterización negativa es mejor que ninguna caracterización, y dejar la definición de accesibilismo en ese nivel general nos permite abarcar un número más amplio de teorías que las que serían pertinentes si diéramos una definición positiva.

Vayamos ahora a nuestra segunda pregunta: ¿en qué consiste *tener* el tipo especial de acceso a un cierto factor? Aquí tenemos dos alternativas. El accesibilista puede sostener que tener acceso especial a un factor implica o bien *haber accedido* ese factor de esa manera, o bien sólo tener la capacidad de acceder a él de esa manera. Esta segunda alternativa es, al menos *prima facie*, la más plausible.

El mentalismo evade el desvío epistémico del accesibilismo y define al internalismo directamente en términos de qué tipos de factores determinan el estatus epistémico de las actitudes doxásticas:

MENTALISMO: El estatus epistémico de las actitudes doxásticas de *S* sobreviene en factores pertenecientes a la vida (puramente) mental de *S*.

El mentalismo implica que si dos sujetos son idénticos mentalmente, entonces también son idénticos con respecto al estatus epistémico de sus actitudes doxásticas (por ejemplo, están justificados en creer las mismas proposiciones). Queda por aclarar (cosa que haremos enseguida) por qué se dice que la base de superveniencia es la vida *puramente* mental del sujeto.

Uno puede ser internalista, entonces, en el sentido accesibilista o en el sentido mentalista –o, por supuesto, en los dos sentidos–.¹² Pero hay

12 Muchos serán internalistas en los dos sentidos por pensar que el menta-

todavía más opciones, pues uno puede ser internalista, en cualquiera de esos sentidos, *con respecto a* dos aspectos distintos de una teoría epistémica. Para presentar esos dos aspectos es conveniente suponer una concepción de lo que es una teoría epistémica que, aunque quizá polémica con respecto a otros temas, no prejuzga las cuestiones que nos preocupan aquí. La idea fundamental de esa concepción es que una actitud doxástica está justificada para un sujeto siempre y cuando el sujeto, al formar esa actitud, esté siguiendo las reglas epistémicas adecuadas. Hay mucho por decir con respecto a esa concepción de las teorías epistémicas, pero, puesto que no es nuestro tema principal, un simple ejemplo y una breve caracterización deberán bastar aquí. Supongamos que me parece visualmente que hay una mesa frente a mí y que, como consecuencia de ello, formo la creencia de que hay una mesa frente a mí. Puesto que es *porque* me parece visualmente que hay una mesa que formo la creencia de que hay una mesa, diremos que estoy siguiendo la siguiente regla:

PERCEPCIÓN: Si a uno le parece visualmente que *p*, entonces uno debe creer que *p*.

Mi creencia, entonces, estará justificada siempre y cuando PERCEPCIÓN sea una de las reglas epistémicas correctas. Lo mismo se aplica, *mutatis mutandis*, a todos los casos de formación de actitudes doxásticas.¹³ Toda formación de una actitud doxástica tendrá varias explicaciones, y si alguna de ellas se conforma a una de las reglas epistémicas correctas, entonces la actitud en cuestión estará justificada.¹⁴ Una teoría

lismo y el accesibilismo son (quizás necesariamente) coextensivos –aunque esa era, justamente, la cuestión que la definición del mentalismo de manera independiente del accesibilismo buscaba evitar–.

13 O a la retención, o cambio, de actitudes doxásticas. Aunque no me ocupo explícitamente de estos aspectos diacrónicos de la justificación, supongo que lo que digo aquí tiene una aplicación más o menos directa a estos casos.

14 Otro tema interesante, relacionado con el uso de reglas, que tampoco trataré aquí es la aplicación a la formación de actitudes doxásticas de reglas que el sujeto no está siguiendo, pero que *debería* seguir. Así, puede suceder que una cierta actitud doxástica esté justificada en el sentido estrecho definido en el texto, ya que se adecua a una de las reglas epistémicas correctas, pero que sin embargo no esté justificada en última instancia, ya que el sujeto debería

epistémica, entonces, consiste en la explicitación de las reglas epistémicas correctas.

Estamos ahora en condiciones de decir cuáles son los dos aspectos de una teoría epistémica con respecto a los cuales uno puede ser internalista (de cualquiera de las dos variantes antes definidas). Primero, uno puede ser lo que llamaré "internalista de antecedentes" si sostiene que los antecedentes de las reglas epistémicas correctas son internos –si uno es mentalista de antecedentes, sostendrá que los antecedentes son factores mentales, mientras que si uno es accesibilista sostendrá que los antecedentes son factores a los cuales el sujeto tiene un acceso especial–. Segundo, uno puede ser "internalista de corrección" si sostiene que lo que determina cuáles son las reglas epistémicas correctas son factores internos –factores mentales si uno es mentalista, y factores a los cuales el sujeto tiene un acceso especial si uno es accesibilista–. Suponiendo que las opciones son independientes entre sí, obtenemos dieciséis combinaciones de posiciones distintas con respecto al debate internalismo/externalismo –lo cual puede ayudar a explicar por qué muchas veces los participantes en los debates parecen estar hablando de distintas cosas (lo están)–. Las posibles posiciones pueden resumirse en la siguiente tabla (donde, por ejemplo, un "Sí" en la intersección de la columna "Antecedentes" con la fila "Mentalismo" significa que uno es mentalista de antecedentes):

	Antecedentes	Corrección
Mentalismo	Sí/No	Sí/No
Accesibilismo	Sí/No	Sí/No

Podemos ahora volver a la cuestión de por qué definimos al mentalismo en términos de la vida *puramente* mental del sujeto. Es posible sostener que PERCEPCIÓN no es una regla epistémica correcta y que,

formar una actitud doxástica distinta si tuviera en cuenta la situación en la que se encuentra y todas las reglas que se aplican a esa situación. Esta cuestión está estrechamente relacionada con el hecho de que las reglas epistémicas dan sólo razones *prima facie* para formar la actitud doxástica en cuestión.

además, no proporciona una descripción adecuada de qué es lo epistémicamente pertinente cuando creo que hay una mesa frente a mí. Es posible sostener que el estado epistémicamente pertinente es, no que me parece visualmente que hay una mesa frente a mí, sino que *veo* que hay una mesa frente a mí. Si uno cree eso, entonces creará también que la regla epistémica correcta es la siguiente:

PERCEPCIÓN*: Si uno ve que *p*, entonces uno debe creer que *p*.

Ahora bien, hay quienes quieren contar a estados como *ver que p* (y también *oír que p*, *sentir que p*, etc.) como estados mentales –estados mentales *factivos*, para distinguirlos de otras actitudes proposicionales que no implican la verdad de la proposición en cuestión–.¹⁵ Pero es justamente la factividad de esos estados lo que va en contra de contarlos como estados internos. Si definiéramos al mentalismo en términos de estados mentales (suponiendo que contamos a actitudes proposicionales factivas como estados mentales), entonces contaríamos a alguien que acepta PERCEPCIÓN* como a un internalista, y esa clasificación no responde a la idea preteórica de en qué consiste ser internalista. Cuando en la definición de mentalismo hablamos de la vida *puramente* mental del sujeto, entonces, buscamos excluir estados mentales factivos de la base de superveniencia.

En lo que resta de este trabajo dejaré a un lado el internalismo de antecedentes –no porque crea que no tiene interés sino más bien por todo lo contrario: merece un tratamiento por separado–.¹⁶ Me concentraré, entonces, en la cuestión de la corrección de las reglas epistémicas. En la próxima sección examino tres tipos distintos de internalismo de corrección y concluyo que todos ellos sufren un problema similar. Luego argumento que un externalismo de corrección, de tipo confiabilista, está especialmente bien equipado para solucionar ese problema.

15. Ver Timothy Williamson (2000).

16. Yo creo que el internalismo de antecedentes es una posición atractiva. Defenderlo implica, entre otras cosas, tener en cuenta la posición "disyuntivista" en filosofía de la mente –posición defendida, por ejemplo, por Mc Dowell (1982)–.

4. ¿Internalismo de corrección?

4.1. La racionalidad egocéntrica de Foley

La primera versión de internalismo de corrección que quiero considerar puede encontrarse en el trabajo de Richard Foley.¹⁷ La idea fundamental de Foley es que la justificación epistémica consiste en ser inmune a la autocrítica, en no ser un hipócrita epistémico. Más específicamente, Foley sostiene que una creencia por parte de un sujeto *S* está epistémicamente justificada si y sólo si esa creencia se adecua a principios epistémicos que el mismo *S* aceptaría como correctos si reflexionara de manera profunda.¹⁸ Si una creencia pasa esa prueba, Foley dirá que la creencia es “egocéntricamente racional”. Comentando los principios epistémicos de Chisholm, dice Foley:

Es en general racional para nosotros tener creencias que se conforman con los principios epistémicos de Chisholm porque en general los principios epistémicos de Chisholm, o al menos algo similar a ellos, reflejan nuestros propios estándares epistémicos profundos. Reflejan nuestra propia opinión profunda acerca de cómo alcanzar de la mejor manera nuestro objetivo intelectual primario –es decir, la verdad–.¹⁹

Dejando a un lado cuestiones de índole técnica, la propuesta de Foley, tiene dos problemas (o quizás un problema que surge de dos maneras distintas) que atacan la estructura general de la teoría y que son, en mi opinión, fatales.

Según Foley mismo, no hay *absolutamente ninguna* seguridad, por pequeña que sea, de que siguiendo nuestros propios principios epistémicos profundos llegaremos a la verdad. En una sección de uno de sus libros que lleva por título “La falta de garantías”, Foley dice que “no hay seguridad de que al ser egocéntricamente racionales evitaremos el

17. Ver Foley (1987) y (1993). Para ver como Foley mismo enmarca su teoría en una concepción basada en reglas epistémicas, ver Foley (1997).

18. La reflexión profunda con respecto a una proposición cualquiera *p* es, según Foley, una reflexión que da un veredicto con respecto a *p* tal que ninguna reflexión ulterior cambiaría ese veredicto.

19. Foley (1997), p. 263.

error. No hay seguridad de que evitaremos el error extremo. Ni siquiera es necesario que sea probable que evitemos el error extremo”.²⁰ Un poco más adelante, al hablar de la falta de garantías que no cometan petición de principios, Foley es incluso más explícito:

El problema surge sin que importe la fuerza de la supuesta seguridad, y surge en exactamente la misma forma en que lo hizo para Descartes. Surge si decimos que al ser racionales podemos asegurarnos de evitar el error generalizado. Surge si decimos de manera más cautelosa que al ser racionales podemos asegurarnos de evitar la probabilidad del error generalizado. Surge aun si decimos que al ser racionales podemos asegurarnos sólo de que la probabilidad de que evitemos el error es al menos tan grande como la que sería si no fuéramos racionales.²¹

Hay un problema grave en combinar, como lo hace Foley, la idea de que la racionalidad epistémica es egocéntrica con la idea de que el objetivo (la verdad) nos está radicalmente vedado –no en el sentido de que no podemos alcanzarlo, sino en el sentido de que nunca podremos, no ya saber, sino tampoco tener la más mínima indicación confiable de haberlo alcanzado–. Si alcanzamos la verdad, eso es, para Foley, un fenómeno que bien podríamos llamar “suerte epistémica”. Por supuesto, todos debemos aceptar que el conocimiento involucra suerte epistémica, en el sentido de que hay casos en donde la justificación no implica la verdad de la creencia en cuestión. En esos casos, si sabemos es porque tuvimos suerte epistémica. Pero el tipo de suerte epistémica involucrada en la propuesta de Foley va mucho más allá de este tipo benigno. Lo que Foley sostiene es que estar justificados en creer que *p* no sólo no implica la verdad de *p*, sino que no tiene *ninguna relación en absoluto* con la verdad de *p*.

Hasta aquí el problema no parece ser demasiado grave. Simplemente, la noción de racionalidad egocéntrica de Foley tiene una diferencia fundamental con la noción intuitiva de justificación (puesto que la primera está totalmente desconectada de la verdad, mientras que la segunda tiene una conexión que no por ser difícil de explicitar es menos real). Pero Foley puede convencernos o bien de que la noción intuitiva de jus-

20. Foley, (1993), p. 59.

21. Foley (1993), p. 61.

tificación está equivocada, o bien de que, dados nuestros objetivos teóricos, deberíamos reemplazar la noción intuitiva de justificación con la noción de racionalidad egocéntrica.

La gravedad del problema surge de manera especialmente clara cuando nos preguntamos si Foley mismo, o cualquier otro sujeto a quien Foley haya convencido de la verdad de la tesis de la falta de garantías, es racionalmente egocéntrico al tener las creencias que tiene. Recordemos que Foley cree que los principios epistémicos de Chisholm son básicamente correctos. Esos principios, y cualesquiera principios cuya estructura fundamental sea similar, implican que nuestras creencias derivadas de la percepción, la memoria y la introspección están justificadas. Supongamos ahora que Foley ve una mesa frente a él, y que la situación es perfectamente normal (por ejemplo, Foley no tomó hace poco ninguna droga que modifique su percepción, ni hay hologramas cerca, etc.). Puesto que Foley cree que la percepción es una manera adecuada de formar creencias, forma la creencia de que hay una mesa frente a sí. Ahora bien, Foley también cree que el hecho de que haya formado esa creencia en base a su percepción visual no ofrece la más mínima seguridad de que la creencia sea verdadera –ni siquiera puede Foley saber si formar la creencia de esa manera es al menos tan seguro como formarla como resultado de consultar los astros–. Pero si el objetivo intelectual que debemos tener en cuenta al formar creencias es alcanzar la verdad, y si, de acuerdo con Foley mismo, formar creencias basados en la percepción no nos acerca ni un ápice a ese objetivo, entonces, ¿por qué está Foley egocéntricamente justificado en creer que hay una mesa frente a él? ¿No debería, más bien, suspender el juicio con respecto a esa proposición? Esa parece ser la actitud doxástica recomendada por los principios epistémicos profundos de Foley, incluyendo el principio de la falta de garantías.²²

En resumen, el problema en combinar la racionalidad egocéntrica con el principio de la falta de garantías es que la única actitud doxásti-

22. Como me lo hizo notar Manuel Comesaña, la racionalidad egocéntrica de Foley es análoga a la corroboración popperiana, que evita la inducción al precio de que la evaluación de teorías no sea proyectable al futuro, con lo cual no tiene derecho a decir que la teoría mejor corroborada es la mejor teoría, y se desliza, así, del escepticismo con respecto a la inducción al escepticismo con respecto a la racionalidad de la ciencia.

ca racional que surge de esa combinación es una suspensión del juicio digna del mejor escéptico.

Pero si el problema surge de combinar la teoría de la racionalidad egocéntrica con la tesis de la falta de garantías, entonces es a lo sumo un problema para la coherencia interna de los compromisos de Foley. *Nosotros* podemos rechazar la tesis de la falta de garantías y seguir a Foley en lo que respecta a la pregunta por la corrección de los principios epistémicos. En otras palabras, podemos seguir a Foley al decir que los principios epistémicos correctos son aquellos que aceptaríamos luego de un proceso de reflexión profunda pero añadir que esos principios sí nos acercan a la verdad. Así evitamos el escepticismo acerca de la justificación al mismo tiempo que sostenemos una posición internalista con respecto a la corrección.

Es cierto que, como esta respuesta supone, la racionalidad egocéntrica es subjetiva en tanto y en cuanto qué actitudes doxásticas estén justificadas para un cierto sujeto depende de qué principios epistémicos profundos acepte ese sujeto. Pero es justamente esta característica de la teoría de Foley la que la hace vulnerable a un problema quizás más grave (aunque puede también verse como una versión del problema anterior). Este problema ha sido advertido por varios comentaristas,²³ y puede formularse de la siguiente manera: Aun cuando nosotros mismos podemos evadir el problema en el que se encuentra Foley al aceptar la tesis de la falta de garantías, la teoría de Foley sigue dando los resultados inadecuados *para Foley mismo*. Con respecto a Foley, su teoría implica que la única actitud doxástica justificada es la suspensión del juicio –y esto sigue siendo un resultado inaceptable por más que la teoría tenga resultados aceptables cuando la aplicamos a nosotros mismos–. O consideremos a alguien que sostiene principios epistémicos aun más claramente inadecuados, como alguien que crea que sólo las proposiciones necesariamente falsas están justificadas. Cuando este sujeto crea proposiciones necesariamente falsas, será egocéntricamente racional, pero entonces es claro que la racionalidad egocéntrica nada tiene que ver con la justificación epistémica. El problema con la teoría de Foley, puesto en términos generales, es que sólo da resultados correctos para aquellos sujetos que acepten las reglas epistémicas adecuadas –y

23. Ver, por ejemplo, James E. Taylor (1998), p. 272; Richard Feldman (1989), pp. 157-8; Marshall Swain (1989), pp. 163-5.

lo que queríamos de la teoría era, precisamente, que nos diga cuáles son las condiciones para que una regla epistémica sea adecuada—. La racionalidad egocéntrica es *en el mejor de los casos* extensionalmente equivalente a la justificación epistémica.

Foley ha respondido a este tipo de objeción.²⁴ En su respuesta Foley nos recuerda que su proyecto surge desde un punto de vista esencialmente cartesiano: el punto de vista de un sujeto que no tiene más que sus propios recursos cognitivos para interpretar el mundo. Si ésa es la situación, dice Foley, entonces

no ayuda que a uno le digan que crea algo si y sólo si es verdadero, y tampoco ayuda que a uno le digan que use sólo métodos confiables. Parte del problema es determinar, desde la propia perspectiva, qué métodos son confiables. De exactamente la misma manera, no ayuda que Alston, Feldman y Swain le sugieran a uno, como cada uno de ellos hace, que crea sólo aquello para lo cual uno tiene fundamentos objetivamente adecuados, fundamentos que de hecho hacen que lo que uno cree sea probable. Porque, nuevamente, parte del problema es determinar, desde la propia perspectiva, qué fundamentos son objetivamente adecuados. Y esto es algo que uno puede hacer más o menos racionalmente.²⁵

Y, podría haber agregado Foley, no ayuda que a uno le digan nada en absoluto, porque parte del problema es determinar, desde la perspectiva propia, si uno va a creer en lo que le dicen.²⁶

Quizás Foley tenga algo de razón. Si existe un problema tal como el que nos presenta en el punto de partida cartesiano, el problema de resolver qué creer contando sólo con los recursos cognitivos propios en ese instante, entonces quizás la mejor respuesta a ese problema sea que hay que creer una proposición si y sólo si al creerla uno sería egocéntricamente racional. Pero es dudoso que tal problema exista. Empezamos a creer sin intentar determinar qué debemos creer y así aprendemos no sólo hechos, sino también nuevos métodos de adquisición de información (como el método fundamental que consiste en preguntarle a alguien que sepa). Cuando llega el momento en que nos volvemos reflexivos y tratamos de determinar de manera explícita qué debemos creer

24. Ver Foley (1989).

25. Foley (1989), p. 171.

26. Pero entonces, ¿cómo puede ayudar lo que diga Foley mismo?

(si es que ese momento llega) estamos ya en *media res* con principios epistémicos que nos permiten usar no sólo anteojos, sino preguntarle a otros, apelar a la ciencia, etc. En ese momento, el "problema cartesiano" no parece ya tan urgente.

Pero quizá Foley presenta el punto de partida cartesiano no como una hipótesis acerca del proceso real de adquisición de información, sino como una "reconstrucción racional" de ese proceso. Sea ello como fuere, aun si le damos a Foley su punto de partida cartesiano, sigue siendo aplicable algo que es esencialmente la misma objeción. Quizás debamos admitir que la respuesta de Foley es la correcta si la pregunta es "¿Qué debo hacer si se me presenta el problema cartesiano?", pero es claro que esa pregunta no tiene ningún significado epistémico —porque la racionalidad egocéntrica, que es la respuesta a la pregunta cartesiana, no tiene nada que ver con la justificación epistémica—. Un sujeto enfrentado al problema cartesiano y armado con principios epistémicos ridículos claramente *no* estará justificado en creer las proposiciones que sean egocéntricamente racionales para él. Nuevamente, la racionalidad egocéntrica coincidirá con la justificación epistémica sólo en el caso de aquellos sujetos que cuenten con principios epistémicos correctos —pero entonces no puede ser la racionalidad egocéntrica lo que *explique* cuáles son los principios epistémicos correctos—.

4.2 El sentimiento de verdad de Pryor

En "The Skeptic and the Dogmatist", James Pryor argumenta que nuestras percepciones nos dan justificación *prima facie* para formar creencias de acuerdo con ellas. Así, Pryor cree que PERCEPCIÓN es una regla epistémica correcta. Pryor quiere permanecer oficialmente agnóstico acerca de *por qué* PERCEPCIÓN es correcta —acerca de la pregunta por el criterio de corrección—. En una nota al pie, de todas maneras, hace referencia a tres posibles posiciones, dos de las cuales rechaza y una de las cuales encuentra plausible. La posición favorecida por Pryor es que lo que explica por qué nuestras experiencias nos dan justificación *prima facie* es

la peculiar "fuerza fenoménica" o manera que nuestras experiencias tienen de presentarnos proposiciones. Nuestras experiencias representan proposiciones de manera tal que "se siente como si" pudiéramos ver que esas proposiciones son verdaderas —y que las estamos percibiendo como verdade-

ras— en virtud simplemente de que nos son representadas de esa manera. (Por supuesto, se necesita un alto grado de autoconciencia reflexiva para poder articular este “sentimiento”.) Yo creo que este “sentimiento” es parte de lo que distingue la actitud de experimentar que *p* de otras actitudes proposicionales, como la creencia y la imaginación visual. Las creencias y las imágenes visuales pueden venir a nosotros de manera irresistible, sin tener ese tipo de “fuerza fenoménica”. (...) Es difícil explicar en qué consiste esta “fuerza fenoménica”, pero creo que es una noción importante, y que necesita ser parte de la historia acerca de por qué nuestras experiencias nos dan la justificación que nos dan.²⁷

Confieso que el uso de las comillas por parte de Pryor en el pasaje citado me confunde un poco —no entiendo bien si él cree que la “fuerza fenoménica” es un “sentimiento” o no—. Pero supongamos que sí lo cree —llamémoslo el “sentimiento de verdad”—. En ese caso, es claro que Pryor está defendiendo una posición internalista con respecto a la corrección.

Debemos hacernos dos preguntas fundamentales con respecto al sentimiento de verdad de Pryor. Primero, ¿es cierto que sólo las experiencias nos presentan proposiciones con el sentimiento de verdad? Segundo, suponiendo que la respuesta a la primera pregunta es afirmativa, ¿explica por qué PERCEPCIÓN es correcta el hecho de que sólo las experiencias representen proposiciones con el sentimiento de verdad?

Para responder la primera pregunta tendríamos que saber un poco más acerca de en qué consiste el sentimiento de verdad. Por empezar, sea cual fuere el sentido en el cual mi experiencia actual representa la proposición de que hay un monitor de computadora frente a mí con el sentimiento de verdad, la “experiencia” que tengo cuando pienso que $7+5=12$ representa esa proposición con el mismo sentimiento. Pryor no tiene por qué negar esto: por el contrario, si su explicación va a tener algún grado de generalidad, *le conviene* que todas las fuentes de creencias que, preteóricamente, consideramos como dándonos justificación, representen proposiciones con el sentimiento de verdad. El problema surge cuando notamos que algunas fuentes de creencias que preteóricamente consideramos como *no* dándonos justificación también presentan proposiciones con el sentimiento de verdad. Consideremos un caso que Pryor mismo menciona, la imaginación visual. Cierro los ojos y me imagino que estoy en una playa desierta. Cierro los ojos nuevamente y me

27. Pryor (2000), p. 547, n. 37.

imagino que estoy cabalgando en Pegaso, el caballo alado. ¿Hay alguna diferencia significativa, con respecto al sentimiento de verdad, entre mi primera imaginación y una percepción normal? Si entiendo en qué consiste el sentimiento de verdad, tanto mi primera imaginación como una percepción normal lo comparten, mientras que mi segunda imaginación no lo tiene (admito que es posible que me haga falta ejercitar mi autoconciencia reflexiva).²⁸

Aun suponiendo que sólo las experiencias (o sólo aquellas fuentes de creencias que preteóricamente consideramos como dándonos justificación) presentan proposiciones con el sentimiento de verdad, nuestra segunda pregunta es incluso más seria: ¿por qué sería algo epistémicamente bueno para la percepción —o para las creencias formadas a partir de ella, o para la relación entre éstas— que presente proposiciones con el sentimiento de verdad? Supongamos que existe una comunidad de sujetos cuyo aparato perceptual funciona de la misma manera que el nuestro excepto por el hecho de que sus experiencias no presentan proposiciones con el sentimiento de verdad. ¿Serían las creencias que esos sujetos forman a partir de sus experiencias peores, epistémicamente hablando, que las nuestras?

Pryor podría objetar que el sentimiento de verdad es esencial a las experiencias, de modo que mis sujetos imaginarios no estarían percibiendo sino haciendo alguna otra cosa. Uno podría concederle a Pryor que el sentimiento de verdad es esencial a la percepción y aun así insistir en que no representa ninguna diferencia epistémica, pero ya no podría, al parecer, apelar al experimento mental para justificar esa posición. Aunque, ¿es tan claro que no podría? Supongamos que alguien sugiere que lo que explica por qué me gusta tanto el agua es el hecho de que contiene hidrógeno. Yo respondo que no puede ser eso, porque el agua me gustaría tanto como me gusta ahora aun si no contuviera hidrógeno. Ahora bien, puedo estar equivocado —podría ser cierto que es el hidrógeno lo que explica por qué me gusta tanto el agua, aunque yo no lo sepa—; pero, ¿es mi respuesta *ininteligible* simplemente por el hecho de que nada que no contenga hidrógeno sería agua? No lo parece, y no parece haber ninguna diferencia pertinente entre este caso y el caso del sentimiento de verdad.

28. Ello no significa, por supuesto, que yo crea que no hay *ninguna* diferencia entre mi primera imaginación y una percepción normal.

De todas maneras, no es necesario apelar a ese experimento mental para justificar la idea de que el sentimiento de verdad no es epistémicamente pertinente. Es claro que un sentimiento de otro tipo –por ejemplo, un sentimiento de *asco*– no sería epistémicamente pertinente, ya que el hecho de que una experiencia va acompañada de un sentimiento de asco no es explicación alguna acerca de por qué una creencia formada sobre la base de esa experiencia está justificada. Pero si el sentimiento de asco no es epistémicamente pertinente, ¿por qué lo sería el sentimiento de verdad? ¿Qué es lo que tiene el sentimiento de verdad, y no tiene el sentimiento de asco, que le permita explicar el poder justificatorio de las experiencias?

Pryor puede haber pensado que creer de acuerdo con experiencias que tienen el sentimiento de verdad es una manera de ser internamente *coherente*, y, puesto que la coherencia es una noción con obvia pertinencia epistémica, es así que el sentimiento de verdad puede desempeñar su papel explicativo.

Al evaluar esta sugerencia debemos recordar que el sentimiento de verdad no es una *creencia*. Pryor no está diciendo que tener una experiencia como si *p* es, entre otras cosas, creer que *p*. Si el sentimiento de verdad fuera una creencia, entonces decir que es incoherente no creer que *p* cuando una proposición nos es presentada con el sentimiento de verdad es quedarse corto: más que incoherente, es imposible.

Si recordamos que el sentimiento de verdad no es una creencia, entonces la apelación a la incoherencia no puede salvar al sentimiento de verdad de la objeción que estábamos presentando –la objeción de acuerdo con la cual que una proposición nos sea presentada con el sentimiento de verdad no es algo epistémicamente pertinente, no es algo que pueda explicar por qué las experiencias justifican creencias–. ¿Por qué sería incluso *prima facie* incoherente no creer una proposición que nos es presentada con el sentimiento de verdad mientras que no es incoherente no creer una proposición que nos es presentada con el sentimiento de asco?

Debemos notar que hay una respuesta a esta pregunta a la que ni Pryor ni ningún otro internalista puede apelar. Según esta respuesta, el sentimiento de verdad explica por qué las experiencias justifican creencias porque las proposiciones presentadas con el sentimiento de verdad *tienden a ser verdaderas*. Esto es, me parece, lo que explica el atractivo intuitivo que tiene la apelación al sentimiento de verdad: si fuera cierto que el sentimiento de verdad está positivamente correlacionado con la

verdad misma, entonces sería obvio que una apelación al sentimiento de verdad sí es epistémicamente pertinente. Pero ni Pryor ni ningún otro internalista puede aceptar esta respuesta porque lo que sería explicativo en ese caso sería, no el sentimiento de verdad mismo, sino su conexión confiable con la verdad.

4.3 El apriorismo de Wedgwood

Examinemos ahora el tercer (y, en lo que a este artículo concierne, último) intento de dar un criterio de corrección internalista para las reglas epistémicas. De acuerdo con Ralph Wedgwood,²⁹ lo que explica la corrección de las reglas epistémicas adecuadas es el (supuesto) hecho de que esas reglas vienen incluidas en la estructura de las capacidades cognitivas de los sujetos. Lo que Wedgwood quiere decir al hablar de reglas que vienen “incluidas en” capacidades cognitivas es, creo, que, por ejemplo, la capacidad de adquirir información perceptualmente *consiste en* (seguramente entre otras cosas) ser tal que si a uno le parece perceptualmente que *p*, entonces uno cree que *p*. Hay otras reglas que seguimos que no están incluidas en nuestras capacidades cognitivas –pero Wedgwood cree que seguimos estas reglas de manera derivada, al seguir reglas básicas que nos dicen que estas otras reglas son confiables–. De manera que la pregunta por el criterio de corrección para las reglas epistémicas se reduce, en el marco de Wedgwood, a la pregunta por el criterio de corrección de las reglas básicas –y es esta última pregunta la que Wedgwood responde apelando a la idea de que esas reglas vienen incluidas en nuestras capacidades cognitivas–.

La misma objeción que presenté en contra del sentimiento de verdad de Pryor se aplica al apriorismo de Wedgwood –y quizás de manera más clara aun en este último caso–. ¿Qué importa si hay algunas reglas que vienen incluidas en nuestras capacidades cognitivas? ¿Qué bien le haría eso a las reglas, o a las creencias formadas siguiendo esas reglas, o a la relación entre ambas? ¿Por qué sería mejor –mejor *epistémicamente*– formar creencias de acuerdo con reglas que vienen incluidas en nuestras capacidades cognitivas que de acuerdo con cualquier otra regla?

Ni Wedgwood ni ningún otro internalista puede apelar a lo que se-

29. Ver Wedgwood (1999) y el final de Wedgwood (a).

ría una excelente respuesta: debemos formar creencias siguiendo las reglas que vienen incluidas en nuestras capacidades cognitivas porque las creencias así formadas tienden a ser verdaderas. Ni Wedgwood ni ningún otro internalista puede apelar a esta respuesta porque localizarían el poder justificatorio de las reglas epistémicas en una propiedad paradigmáticamente externalista: la verdad.

Supongamos que le concedemos a Wedgwood que hay ciertas reglas que vienen incluidas en las capacidades cognitivas que tenemos. Hay algo acerca de dichas reglas que todos los que no somos escépticos debemos aceptar, a saber: que seguirlas resulta en una alta proporción de creencias verdaderas. Se sigue, entonces, que hay algo epistémicamente bueno acerca de esas reglas. Pero sería una clara falacia inferir de ello que lo que es bueno acerca de esas reglas es el hecho de que vienen incluidas en nuestras capacidades cognitivas.

4.4 El fracaso del internalismo de corrección

En esta sección hemos examinado tres intentos de dar un criterio de corrección internalista para las reglas epistémicas, y hemos concluido que todos ellos deben enfrentar serios problemas. Los problemas surgen, en todos los casos, de una misma fuente: los criterios internalistas propuestos no tienen la consecuencia de que satisfacer ese criterio es algo epistémicamente bueno –bueno para las reglas, o para las creencias formadas siguiendo esas reglas, o para la relación entre ambas–. Pero si el internalismo de corrección no funciona, ¿cuál es entonces la teoría adecuada sobre la corrección de las reglas epistémicas?

5. Confiabilismo de corrección

Propongo el siguiente como criterio de corrección de reglas epistémicas:

CONFIABILIDAD: Una regla epistémica es adecuada sólo si el antecedente de la regla está confiablemente relacionado con la verdad de la proposición mencionada en el consecuente de esa regla.

Para entender CONFIABILIDAD debemos precisar qué entendemos por “relación confiable”. En términos generales, un factor A está confiablemente relacionado con otro factor B si y sólo si hay una regularidad

entre A y B que fundamenta contrafácticos del tipo “Si A ocurriera, entonces B también ocurriría”.³⁰

Veamos cómo funciona CONFIABILIDAD aplicándolo a nuestro ejemplo de regla, PERCEPCIÓN. De acuerdo con el criterio, la regla será correcta sólo si hay una regularidad entre que le parezca a uno que *p* y la verdad de *p*. Puesto que esa regularidad se da, PERCEPCIÓN pasa la prueba de CONFIABILIDAD.

Nótese que presento a CONFIABILIDAD como una condición necesaria, pero no suficiente, para la corrección de reglas. Esto puede hacer surgir dos preguntas. Primero, ¿no estoy cargando los dados a favor del externalismo? Si las condiciones de corrección internalistas presentadas en la sección anterior se propusieran sólo como condiciones necesarias, ¿se seguirían aplicando las mismas objeciones que presenté en contra de ellas? Segundo, independientemente de las propuestas internalistas, ¿no es un criterio de corrección que sólo de condiciones necesarias radicalmente incompleto? En particular, queda en pie la pregunta, ¿qué más, además de satisfacer CONFIABILIDAD, hace falta para que una regla epistémica sea correcta?

Con respecto a la primera pregunta, la respuesta es que *no* estoy cargando los dados a favor del externalismo, ya que hay una asimetría entre internalismo y externalismo. Recordemos que caracterizamos al mentalismo y al accesibilismo como tesis de superveniencia de acuerdo con las cuales el estatus epistémico de las actitudes doxásticas sobreviene en estados internos –estados que son *todos ellos* internos–. Y recuerde-

30. En las semánticas de Lewis (1973) y Stalnaker (1968) para contrafácticos, la verdad de *p* y *q* garantiza la verdad del contrafáctico *Si p fuera verdadera, entonces q también lo sería*. Si adoptamos esta semántica, entonces la definición de *relación confiable* en términos de contrafácticos sería obviamente inadecuada, ya que resultaría que hay una relación confiable entre dos factores cualesquiera siempre y cuando ambos existan. Hay dos maneras de sortear esta dificultad. Primero, podemos apelar a una semántica no tradicional, como la de Nozick (1981), en donde un contrafáctico puede ser falso aun si tanto su antecedente como su consecuente son verdaderos. Segundo, podemos definir la noción de *relación confiable* directamente en términos de mundos posibles, sin tomar el desvío por los contrafácticos, de esta manera: hay una relación confiable entre dos factores A y B siempre y cuando en la mayoría de los mundos cercanos en donde A ocurre (incluyendo al mundo actual si es que A ocurre en el mundo actual, pero incluyendo también otros mundos), B también ocurre.

mos también que caracterizamos al externalismo simplemente como la negación del internalismo –es decir, como la tesis de que no es cierto que el estatus epistémico de las actitudes doxásticas sobreviene en factores que son *todos ellos* internos–. Para mostrar la falsedad del internalismo –y, por ende, la verdad del externalismo– alcanza con que *al menos uno* de los factores en los cuales sobreviene el estatus epistémico de las actitudes doxásticas no sea interno. Mis críticas a las propuestas internalistas en la sección anterior no se basaban en peculiaridades de esas propuestas que pudieran salvarse apelando a más condiciones internalistas, sino que apuntaban al hecho de que ninguna cantidad de condiciones internalistas va a darnos una respuesta satisfactoria a la pregunta por la pertinencia epistémica de factores internos. Si CONFIABILIDAD es un criterio adecuado –adecuado incluso sólo como condición necesaria–, entonces el internalismo es falso y el externalismo es verdadero.

Dicho sea de paso, el hecho de que CONFIABILIDAD no sea una condición suficiente de corrección nos permite responder una objeción común al confiabilismo. La objeción es que el confiabilismo no funciona para el caso de las verdades necesarias, ya que implica que estamos justificados en creer que $345 \times 27 = 9315$ sea cual fuere la razón por la que lo creemos (por ejemplo, aun si lo creemos porque lo soñamos). Esta objeción sería correcta si CONFIABILIDAD fuera una condición suficiente de corrección. En el caso de las proposiciones necesariamente verdaderas, existe una conexión confiable entre *absolutamente cualquier cosa* y el hecho de que esa proposición sea verdadera. En particular, el contrafáctico “Si soñara que $345 \times 27 = 9315$, entonces sería cierto que $345 \times 27 = 9315$ ” es verdadero –y, de manera más general, cualquier contrafáctico de la forma “Si soñara que p , entonces p es verdadera”, en donde p es una proposición necesariamente verdadera, será también verdadero–. Con lo cual, si CONFIABILIDAD fuera una condición suficiente de corrección, la siguiente sería una regla epistémica correcta (restringida a proposiciones necesariamente verdaderas):

SUEÑO: Si uno sueña que p , entonces uno debe creer que p .

Pero, puesto que CONFIABILIDAD es sólo una condición necesaria, la objeción no se aplica. El hecho de que SUEÑO satisfaga CONFIABILIDAD no implica que es una regla epistémica correcta.

El criterio CONFIABILIDAD tiene lo que hace falta para responder

a las objeciones en contra de los criterios internalistas presentadas en la sección anterior. Es obvio que si una regla epistémica satisface CONFIABILIDAD, entonces las creencias que formamos siguiendo esa regla tienen algo epistémicamente bueno en su favor. Tenemos que conceder, a la luz del ejemplo del sueño con verdades necesarias, que el valor epistémico que otorga satisfacer CONFIABILIDAD no es suficiente para que una regla epistémica sea correcta. Pero, al menos *prima facie*, ese valor epistémico sí es necesario para ser una regla epistémica correcta, lo cual es más de lo que se puede decir sobre las propuestas internalistas.³¹

6. Conclusión

En este artículo definí dos sentidos en los cuales uno puede ser internalista (y, correspondiendo a la negación de cada uno de ellos, dos sentidos en los cuales uno puede ser externalista) y dos aspectos de una teoría epistémica con respecto a los cuales uno puede ser internalista (o externalista) en cualquiera de esos dos sentidos. Consideré luego tres tipos de internalismo de corrección y argumenté que ninguno de ellos puede superar una dificultad fundamental, a saber, que no es para nada claro qué tendría de bueno satisfacer una condición internalista. Luego propuse un externalismo de corrección de tipo confiabilista y argumenté que en este caso sí es posible ver que tiene de bueno satisfacer la condición externalista. Mi conclusión es que debemos ser externalistas con respecto a la corrección.

Brown University

31. Hay varios problemas con el confiabilismo que habría que discutir para defenderlo de manera más cabal. Uno de los principales problemas se conoce con el nombre de “el nuevo problema del demonio maligno”. El problema es que una víctima del demonio maligno está tan justificado en creer lo que cree como nosotros lo estamos en creer lo que creemos, pero el confiabilismo diría que no lo está –ya que los procesos que usa para llegar a sus creencias no son confiables–. En Comesaña (en prensa) argumento que una versión del confiabilismo implica que los procesos usados por las víctimas de un demonio maligno sí son confiables en el sentido pertinente.

BIBLIOGRAFÍA

- CHISHOLM, Roderick (1977), *Theory of Knowledge*, 2da. ed. Englewood Cliffs, NJ: Prentice-Hall.
- COMESAÑA (en prensa), "The Diagonal and the Demon", *Philosophical Studies*.
- CONEE, Earl y Richard FELDMAN (2001), "Internalism Defended", *American Philosophical Quarterly*, 38, pp. 1-18.
- CRESTO, Eleonora (1997), "Escepticismo, verdad y confiabilidad", *Revista Latinoamericana de Filosofía* 23(1), pp. 93-125.
- FELDMAN, Richard (1989), "Foley's Subjective Foundationalism", *Philosophy and Phenomenological Research*, Vol. 50, No. 1, pp. 149-58.
- FOLEY, Richard (1987), *The Theory of Epistemic Rationality*, Cambridge, MA: Harvard University Press.
- FOLEY, Richard (1989), "Reply to Alston, Feldman and Swain", *Philosophy and Phenomenological Research*, Vol. 50, No. 1, pp. 169-188.
- FOLEY, Richard (1993), *Working Without a Net*, New York, Oxford University Press.
- FOLEY, Richard (1997), "Chisholm's Epistemic Principles", *The Philosophy of Roderick Chisholm*, Chicago, Ill.: Open Court, pp. 241-64.
- GETTIER, Edmund (1963), "Is Justified True Belief Knowledge", *Analysis* 23, pp. 121-3.
- GOLDMAN, Alvin (1976), "Discrimination and Perceptual Knowledge", *Journal of Philosophy* 73, pp. 771-91.
- LEHRER, Keith (1965), "Knowledge, Truth, and Evidence", *Analysis* 25, pp. 168-75.
- LEWIS, David (1973), *Counterfactuals*, Oxford: Blackwell.
- MCDOWELL, John (1982), "Criteria, Defeasibility and Knowledge", *Proceedings of the British Academy* 68.
- MERRICKS, Trenton (1995), "Warrant Entails Truth", *Philosophy and Phenomenological Research*, Vol. LV, No. 4, pp. 841-55.
- NOZICK, Robert (1981), *Philosophical Explanations*, Cambridge, MA: Harvard University Press.
- PLANTINGA, Alvin (1993), *Warrant: The Current Debate*, Nueva York, Oxford University Press.
- POLLOCK, John (1986), *Contemporary Theories of Knowledge*, Totowa, NJ, Rowman and Littlefield.
- PRYOR, James (2000), "The Skeptic and the Dogmatist", *Nous* 34, pp. 517-49.

- Russell, Bertrand (1948), *Human Knowledge: Its Scope and Limits*, Nueva York, Allen and Unwin.
- SHOPE, Robert (1983), *The Analysis of Knowing*, Princeton, Princeton University Press.
- SOSA, Ernest (2000a), "Skepticism and Contextualism", *Philosophical Issues* 10, pp. 1-18.
- SOSA, Ernest (2000b), "Replies", *Philosophical Issues* 10, pp. 38-42.
- STALNAKER, Robert (1968), "A Theory of Conditionals", en N. Rescher (ed.), *Studies in Logical Theory*, Oxford, Blackwell.
- SWAIN, Marshall (1989), "On Richard Foley's Theory of Epistemic Rationality", *Philosophy and Phenomenological Research*, Vol. 50, No. 1, pp. 159-68.
- TAYLOR, James E. (1998), Review of Richard Foley, *Working Without a Net*, *Nous* 32:2, pp. 265-75.
- TOMBERLIN, James (1980), "Critical Review of Carl Ginet's *Knowledge, Perception, and Memory*", *Nous* 14, pp. 239-50.
- WEDGWOOD, Ralph (1999), "The A Priori Rules of Rationality", *Philosophy and Phenomenological Research*, Vol. LIX, No. 1, pp. 113-31.
- WEDGWOOD, Ralph (a), "Internalism Explained", *Philosophy and Phenomenological Research*, en prensa.
- WILLIAMSON, Timothy (2000), *Knowledge and its Limits*, Nueva York, Oxford University Press.

ABSTRACT

In this paper I characterize two ways in which one can be an internalist, the mentalist and the accesibilist way, and two aspects of an epistemic theory with respect to which one can be an internalist, the antecedents of epistemic principles and the correctness conditions for epistemic principles. I then argue that we should be correctness reliabilists. My argument works by assessing several proposed internalist correctness conditions and noting that what is wrong with all of them suggests that reliabilism provides the right condition.



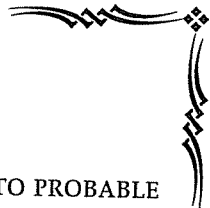
CONCURSO EZEQUIEL DE OLASO 2003

Para jóvenes filósofos de Iberoamérica que tengan menos de 40 años a la fecha de cierre del Concurso.

Los trabajos deben tener una extensión no mayor de 10.000 palabras y estar originalmente escritos en español o portugués. Se recibirán hasta el 31 de diciembre de 2003. Se tomará como válida la fecha del correo.

Deben estar redactados en forma de artículos, firmarse con seudónimo y en sobre aparte cerrado ha de figurar la identidad y el domicilio del concursante así como la constancia de que es menor de cuarenta años a la fecha de cierre del concurso, y que es nativo de Latinoamérica o de la Península Ibérica. Deberán enviarse dos copias en papel y un disquete a la RLF, Casilla de Correo 5379 W1000CB - Buenos Aires, Argentina.

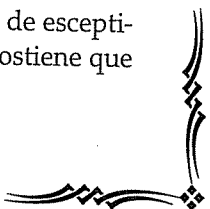
El jurado estará compuesto por Antoni Domèneche (España), Plinio Junqueira Smith (Brasil), C. Ulises Moulines (México-Alemania) y Ernesto Sosa (EEUU).



TEMA DEL CONCURSO: HISTORIA Y JUSTIFICACIÓN DEL CONOCIMIENTO PROBABLE

La concepción tradicional del conocimiento, heredada de Platón, no deja lugar para la noción de conocimiento probable. Todo conocimiento es conocimiento cierto, esto es, no puede ser falso. El probabilismo de los académicos se consideró siempre como una forma de escepticismo, precisamente porque implica que hay conocimiento incierto. La idea de que el conocimiento sea sólo probable y, consiguientemente, pueda ser falso, emerge muy lentamente durante la modernidad. Su desarrollo coincide, aproximadamente, con el de la teoría matemática de la probabilidad. La tesis de que el conocimiento es meramente probable se afirma cuando el uso de la teoría de la probabilidad se extiende a las ciencias naturales y sociales. No obstante, todavía en nuestros días muchos filósofos defienden la idea de que el conocimiento presupone certeza.

El tema del presente concurso es el análisis del concepto mismo de conocimiento probable y de sus implicaciones, tanto en sus aspectos históricos como sistemáticos. Esto incluye, entre otros, problemas tales como el significado de las probabilidades, el uso de teorías y modelos probabilísticos en las ciencias empíricas, y la construcción de teorías probabilistas de la confirmación de teorías científicas. También comprende el desarrollo de la noción de conocimiento probable y la justificación de esta manera de concebir el conocimiento frente a las diversas formas de escepticismo y frente al dogmatismo tradicional que sostiene que no hay conocimiento sin certeza.



**CREER, INFERIR Y ACEPTAR:
UNA DEFENSA DE LA INFERENCIA A LA MEJOR
EXPLICACIÓN APTA PARA INCRÉDULOS***

Eleonora Cresto

1. Introducción

Mi objetivo principal en este trabajo es ofrecer una elucidación de la idea de Inferencia a la Mejor Explicación (de ahora en más, IME). En la sección 2 propongo entender a una IME como un proceso complejo, que involucra el uso de herramientas de teoría de la decisión cognitiva, y que en general culmina con un cambio en el estado epistémico del agente, en un sentido que habrá que explicitar. Argumento además que la idea de IME debe distinguirse del concepto de abducción. En la sección 3 explico en qué consiste la teoría de la decisión cognitiva, y presento brevemente una de las versiones más conocidas de dicha teoría. En la sección 4 sugiero que toda investigación se lleva a cabo en un determinado contexto, y que, en los casos típicos, el objetivo de los agentes consiste en ampliar el conjunto de elementos aceptados para el contexto en cuestión. Esto, sin embargo, no necesariamente lleva a la adquisición de nuevas *creencias*. Con base en estas sugerencias, en la sección 5 esbozo una teoría formal de cambio epistémico basada precisamente en la distinción entre aceptar y creer. Argumento luego que la modificación epistémica que introduce una IME debe entenderse a la

* Una versión anterior de este trabajo fue presentada en las Conferencias de Invierno del Instituto de Filosofía de la Universidad de Buenos Aires, en julio de 2002. Agradezco a todos los asistentes, y especialmente a Eduardo Barrio por haberme invitado a participar en dichas conferencias. Deseo agradecer también a Alejandro Cassini, quien leyó cuidadosamente el primer manuscrito y me advirtió sobre diversas imprecisiones, así como a Gustavo Ortiz-Millan, por sus sugerencias sobre la naturaleza de las creencias. Durante el período de realización de este trabajo conté con un subsidio de la Fundación Antorchas.

luz de esta distinción: en los casos típicos, una IME lleva al investigador a abrazar una nueva aceptación contextual, pero no necesariamente una nueva creencia. Finalmente, en la sección 6 señalo algunas conclusiones interesantes que el presente análisis entraña para la perspectiva epistemológica conocida como "realismo científico".

2. La inferencia a la mejor explicación

¿Qué es una inferencia a la mejor explicación? Mucho se ha escrito desde que Gilbert Harman acuñó el término por primera vez en (1965).¹ Sin embargo, todavía no existe consenso acerca de cómo debe entenderse exactamente este concepto.

Si nos atenemos a un punto de vista filológico, una IME es, antes que nada, un tipo de inferencia. Pero aquí comienzan los problemas, porque la noción de inferencia es ambigua. Puede referir tanto a un *argumento*, como al *proceso* por el cual un agente pasa de creer o aceptar ciertas proposiciones, a creer o aceptar alguna otra. De manera correspondiente, una IME podría consistir o bien en un tipo especial de argumento, o bien en un proceso, tal vez complejo, por el cual un agente pasa de creer o aceptar que cierto fenómeno ha ocurrido, a creer o aceptar cierta explicación para dicho fenómeno. En el primer caso, se presupone que una IME puede ser capturada por una forma de razonamiento característica, o tal vez por un conjunto de tales formas. En el segundo caso, el énfasis está puesto en el conjunto de operaciones que el agente realiza para llegar a la mejor explicación. Dichas operaciones bien podrían recurrir a estructuras formales bien definidas, pero ello no es una característica esencial de la idea de IME como proceso.

Formalmente, podría pensarse que la manera más sencilla de capturar una IME sería a través de una falacia de afirmación del consecuente: si la hipótesis *H* fuera cierta, el fenómeno *F* se seguiría; *F* ocurre; por lo tanto, *H* es cierta. De acuerdo con otro análisis posible, una IME equivaldría a una inducción por eliminación, o tal vez a una falacia de afirmación del consecuente complementada con una inducción eliminativa.² En algunos autores encontramos también la sugerencia de que una

1. Véase también Harman (1968), para algunas aclaraciones importantes sobre su idea original.

2. Harman (1965), pp.88-89, afirma que una IME equivale aproximada-

IME encubre un silogismo disyuntivo con premisas inciertas, esto es, un razonamiento deductivo de la forma: $(H_1 \vee H_2 \dots \vee H_n) \& \neg H_1 \& \neg H_2 \dots \& \neg H_{n-1} \vdash H_n$, donde $H_1 \dots H_n$ son explicaciones posibles para cierto *explicandum*, pero tal que no tenemos total certeza de que la serie de disyunciones sea completa.³ Por otra parte, desde el punto de vista de una caracterización completamente informal, una IME simplemente describiría cómo, frente a un determinado fenómeno que resulta problemático por alguna razón, el ingenio del investigador combina de manera más o menos intuitiva los distintos aspectos que hacen a una buena explicación (por ejemplo, simplicidad, plausibilidad, etc.), hasta arribar a la mejor explicación de dicho fenómeno que es capaz de construir. Alternativamente, sería la operación según la cual el investigador, guiado otra vez de manera más o menos intuitiva por los aspectos mencionados, selecciona la explicación que cree mejor a partir de un grupo dado de explicaciones posibles.⁴

Tenemos pues al menos dos maneras de entender una IME: la idea de IME como forma (o formas) de razonamiento, y la idea de IME como proceso que puede involucrar tanto aspectos formales como informales. La identificación de IME con un conjunto de reglas formales es más común en el campo de la inteligencia artificial,⁵ mientras que las descripciones provenientes del terreno de la filosofía son más propensas a una caracterización en términos de procesos complejos.⁶ En este trabajo optaré por esta última perspectiva.

En segundo lugar, nuevamente desde un punto de vista filológico,

mente, entre otras cosas, a lo que otros autores han llamado "inducción eliminativa".

3. Véase Josephson y Josephson (1994), p.12.

4. Véase Thagard (1978) para la sugerencia de cómo combinar informalmente diferentes criterios de selección, a partir de algunas consideraciones pragmáticas. Lipton (1991) constituye un buen ejemplo de una perspectiva totalmente informal sobre las IME.

5. Tal identificación es harto común considerando el presupuesto adicional de que IME es un nombre alternativo para hablar de abducción. Véase más abajo, en esta misma sección, la discusión sobre los diversos significados de "abducción".

6. Para una pequeña muestra de cómo pueden combinarse factores formales e informales véase el artículo de Richard Fumerton, "Inference to the Best Explanation", en Dancy y Sosa (1993), pp. 207-209.

una IME debe proporcionar alguna manera de seleccionar la *mejor* explicación para cierto evento. (Obsérvese que esto descarta, entre otras cosas, la posibilidad de que una IME pueda ser capturada *con el solo recurso* de una falacia de afirmación del consecuente.) Claro está que nadie supone que pueda haber recetas para ofrecer la mejor explicación *simpliciter*; una "mejor explicación" forzosamente hace referencia a las posibilidades consideradas por cierto investigador en cierto momento. Por ello creo conveniente exigir que un análisis de la idea de IME sea capaz de ofrecer algún mecanismo explícito para elegir una hipótesis explicativa, o tal vez más de una, entre un conjunto pre-seleccionado de explicaciones potenciales. La presentación explícita del problema como un problema de selección de hipótesis equivale tan sólo a traer a la luz un elemento que está ya más o menos dado por supuesto en la mayoría de los estudios sobre IME. En este trabajo voy a proponer una elucidación de las IME que satisfaga este requisito básico.

Más precisamente, sugiero entender una IME como un proceso que consta de dos etapas bien definidas. En un primer momento el investigador propone un conjunto de hipótesis explicativas. Cada una de ellas brinda información específica que permite mejorar la comprensión que el agente tiene de cierto fenómeno; expresado de otra manera, cada una de ellas constituye una solución característica al problema (o perplejidad) que inició la investigación. En un segundo momento, el investigador selecciona una de tales hipótesis, o tal vez suspende el juicio entre dos o más de ellas. Al primer momento, siguiendo algunos escritos de Peirce,⁷ lo llamaré *abducción*; no me ocuparé demasiado de la etapa abductiva en este trabajo.⁸ A la segunda etapa la llamaré *etapa selectiva*, o etapa de selección de hipótesis, y haré una elucidación de la misma en términos de la teoría de la decisión cognitiva. Entiendo que dicha teoría proporciona la manera más eficiente de elegir elementos entre un conjunto predeterminado de hipótesis.

Hay otro ingrediente que es importante mencionar a la hora de ofrecer un análisis de la idea de IME. Desde Harman en adelante, todos los

7. Peirce (1908). Esto es, me refiero a la segunda época de Peirce; véase más abajo en esta misma sección.

8. Por ende, no me ocuparé de qué propiedades constituyen una *buena* explicación. Para un análisis de IME que se concentra precisamente en este problema, véase Lipton (1991).

autores que han escrito sobre este tipo de inferencia parecen estar de acuerdo en que el objetivo último de una IME es provocar algún tipo de cambio epistémico en el agente que la realiza. De allí el atractivo que esta idea despierta en los ámbitos de teoría del conocimiento o filosofía de las ciencias. Así pues, respetando el sentido habitual de IME, entenderé aquí que el resultado de la etapa de selección de hipótesis será, al menos en los casos típicos, la modificación del estado epistémico del agente, en un sentido que habrá que elucidar. En este punto voy a proponer un esquema de cambio epistémico, deudor de las llamadas "teorías de revisión de creencias",⁹ que hace uso de una distinción fundamental entre los conceptos de *aceptación* y *creencia*. La distinción entre creer y aceptar goza actualmente de cierta popularidad en diversos ámbitos de la filosofía,¹⁰ pero ninguna de las interpretaciones conocidas recoge las ideas que la presente propuesta intenta capturar. Retomaré este tema en las secciones 4 y 5.

Una pregunta posible a esta altura es por qué restringir el uso de "abducción" sólo a la primera etapa de una IME. En efecto, mi uso del término se separa de aquellos autores que tienden a *identificar* IME con abducción.¹¹ Creo que llamar abducción a todo el proceso inferencial es un error, o, para ser más exactos (puesto que cada uno es libre de usar las palabras como más le guste), es simplemente el fruto de una confusión histórica. Lamentablemente, es una confusión bastante extendida, de modo que tal vez sea apropiado dedicar algunos párrafos a este tema.

Virtualmente todos los autores que hacen uso del término abducción dicen rescatar el sentido usado por Peirce. Pero el asunto es complicado, porque Peirce dio por lo menos dos caracterizaciones diferentes, correspondientes a dos épocas distintas de su pensamiento. De acuerdo con el primer sentido, cronológicamente más antiguo, abducción o hipótesis es un tipo de razonamiento ampliativo que concluye con una

9. Para un panorama sistemático de tales teorías véase Gärdenfors (1988) o Hansson (1999).

10. Cf. por ejemplo Cohen (1992), van Fraassen (1989), o Maher (1993), entre muchos otros.

11. A título de ejemplo, véanse van Fraassen (1989), cap. 7; Josephson y Josephson (1994); o S. Psillos, "Abduction: between conceptual richness and computational complexity", en Flach y Kakas (2000), pp. 59-74.

explicación de algunos de los eventos descritos en las premisas.¹² Peirce entiende que es una manera posible de invertir el orden de premisas y conclusión en un razonamiento deductivo. Curiosamente, dentro de los deductivos también ubica lo que hoy llamaríamos silogismos estadísticos. Por ejemplo, estaríamos haciendo una abducción si tenemos premisas que dicen que “el 80% de los usuarios intensivos de computadoras presentan el síndrome del ojo seco, dolores metacarpianos, y contracturas cervicales”, y que “este individuo tiene los tres síntomas descritos”, y a partir de ellas concluimos que “este individuo es un usuario intensivo de computadora”. En otra terminología, una abducción en este sentido equivale o bien a una falacia de afirmación del consecuente, llamada también “*modus ponens inverso*” [*backward modus ponens*], o bien a la inversión de un silogismo estadístico que imite la forma en que dicha falacia invierte al *modus ponens*.

La idea de abducción como forma de razonamiento ampliativo sobrevive actualmente sobre todo en el campo de la inteligencia artificial, donde podemos encontrar numerosas ofertas de reglas abductivas que, al menos en principio, podrían ser implementadas por una máquina.¹³ En muchos de estos casos leemos además que lo que se ofrece es en verdad lo mismo que, en otros ámbitos, suele llamarse “IME”. Pero, según hemos visto, una IME es algo bastante más complejo que una mera falacia de afirmación del consecuente. Requiere que al investigador se le ocurran hipótesis explicativas apropiadas, y requiere también algún tipo de mecanismo de comparación entre las posibles hipótesis ya propuestas. Por supuesto, si por “reglas abductivas” entendemos un conjunto lo suficientemente variado y sofisticado de reglas, no habría en principio inconvenientes en llamar abducción a una IME; simplemente obtendremos una perspectiva puramente formal de IME al estilo de las que mencionamos al comienzo de esta sección. Pero es claro que esto ya no es “abducción en el (primer) sentido de Peirce”.

De todos modos, creo que la raíz del problema (esto es, el *origen* de la identificación entre abducción e IME) debe buscarse en el segundo

12. Peirce (1878).

13. Para un panorama general del tratamiento de la abducción en inteligencia artificial véase Flachs y Kakas (2000); en Pagnucco (1996) puede encontrarse un estudio detallado de la idea de abducción desde la perspectiva de las teorías de revisión de creencias.

sentido que da Peirce a esta palabra. En una segunda acepción, abducción es el proceso de formación de hipótesis a partir de una observación que requiere explicación. Pertenecer, por tanto, a lo que hoy llamaríamos “contexto de descubrimiento”. Es el sentido de abducción rescatao, entre otros, por Hanson (1958). Peirce afirma que dicho proceso, visto globalmente, responde a la siguiente forma lógica:

“El hecho sorprendente C es observado;
Pero si A fuera verdadero, C resultaría obvio,
Por ende, hay razones para sospechar que A es verdadero.”¹⁴

Es decir que nos topamos nuevamente con una falacia de afirmación del consecuente, esto es, con la forma que anteriormente había clasificado como abducción o hipótesis (ello explica, por supuesto, que utilice el mismo término¹⁵). Pero aquí hay una diferencia fundamental respecto de la primera época. Una abducción ahora *no es un procedimiento ampliativo*: las hipótesis resultantes de un proceso abductivo deben contrastarse primero, antes de estar en condiciones de incorporarse a un cuerpo de creencias. En esta segunda época Peirce entiende por *inducción*, justamente, el proceso de contrastación de hipótesis con la experiencia a través de predicciones;¹⁶ esta etapa, posterior a la abductiva, pertenece claramente a lo que hoy llamamos “contexto de justificación”.

Para Peirce, además, la formulación de hipótesis no es una actividad que esté sujeta a reglas, sino que depende de la intuición del investigador. Ahora bien, no cualquier hipótesis posible es admisible como resultado de una abducción: dicha hipótesis debe ser explicativa, económica y capaz de evaluación experimental. En verdad, ello significa que ya hay algún tipo de elección o descarte implícito entre las hipótesis lógicamente posibles que puede concebir un agente. Sin embargo, esta selección, como ya se ha dicho, no culmina con la aceptación definitiva de una hipótesis, sino con la propuesta de que la hipótesis en cuestión es

14. Peirce (1958), 5.189.

15. En sentido estricto, en sus escritos tempranos Peirce usa únicamente el término “hipótesis”. Sin embargo, cuando más tarde critique su clasificación temprana de razonamientos, se referirá a su anterior concepto simplemente como al de una idea confusa de abducción. Véase Peirce (1958), 2.102.

16. Véase Peirce (1908), pp. 367-373.

digna de ser puesta a prueba. Estas afirmaciones parecen haber confundido a muchos autores. En efecto, algunos análisis de abducción e IME parecen confundir los mecanismos por los cuales Peirce indica que una hipótesis puede ser considerada una buena *candidata* para explicar un fenómeno, con aquellos mecanismos que permiten decidir si cierta candidata es la mejor de un conjunto previamente constituido.¹⁷ Para evitar estos problemas terminológicos es que Laudan, por ejemplo, propone hablar del *contexto de prosecución* [*context of pursuit*], para referir al terreno intermedio entre descubrimiento y justificación.¹⁸

No está claro si la mencionada confusión puede atribuirse al propio Harman, aunque sí parece haber tocado a muchos de sus lectores. Harman (1965) expresa que "la IME corresponde *aproximadamente a lo que otros han llamado 'abducción', 'el método de hipótesis', 'inferencia hipotética', 'el método de eliminación', 'inducción eliminativa' e 'inferencia teórica'*" (pp. 88-89, énfasis mío). Pero, justamente porque todos estos términos están cargados de connotaciones indeseables es que, según nos dice, prefiere su propia terminología. De modo que no sería correcto culpar a Harman por haber identificado abducción con IME, porque, en sentido estricto, él no hace semejante cosa: la identificación, en todo caso, correría por cuenta de aquellos "otros" a quienes Harman refiere.

Independientemente de cuál haya sido la idea original de Harman, es claro que "IME" no puede ser un sinónimo de "abducción" en el segundo sentido de Peirce, a menos que se rechace la idea de IME como mecanismo de cambio epistémico, lo cual ningún autor parece estar dispuesto a conceder. Y, como ya hemos visto, tampoco puede equipararse a "abducción" en su primera acepción. Desde luego, cada uno es libre de estipular el uso que prefiera; mi intención en estos párrafos fue tan sólo la de observar que no sería correcto, sin embargo, atribuir dicho uso a Peirce.

Para recapitular, voy a entender a una IME en términos de un proceso que consta de dos etapas, y que, por lo general, culmina con la modificación del estado epistémico de un agente. Durante la primera etapa (fase abductiva) tiene lugar la generación de respuestas potenciales, y durante la segunda (etapa selectiva) tiene lugar la selección de hipótesis previamente propuestas en la primera etapa. En los casos típicos, el ele-

17. De manera paradigmática, véase Josephson y Josephson (1994), p. 5.

18. Laudan (1981).

mento seleccionado es *aceptado* por el agente para cierto contexto de investigación, lo cual modifica su estado epistémico en un sentido particular que habrá que clarificar. Para ello argumentaré que se hace necesario distinguir entre las nociones de aceptación y creencia.

Como ya he indicado, en este trabajo no me ocuparé ni del contexto de descubrimiento propiamente dicho, ni del de "prosecución" (en terminología de Laudan). Es decir que no me ocuparé de examinar qué hace que una hipótesis sea en verdad explicativa, ni tampoco de detallar qué otros rasgos, aparte de su capacidad explicativa, hacen que una hipótesis merezca ser considerada una posibilidad seria. Centraré mis esfuerzos, en cambio, en estudiar cómo procede la elección entre hipótesis rivales una vez que éstas han sido seriamente propuestas como resultado de una abducción. Más precisamente, analizaré el caso particular en que dichas hipótesis son todas empíricamente adecuadas, esto es, compatibles con la evidencia observacional disponible en un momento dado. ¿Cómo se realiza la selección en este caso? Entiendo que la teoría de la decisión cognitiva brinda una respuesta adecuada para esta pregunta. En la siguiente sección explicaré someramente en qué consiste dicha teoría, y presentaré una de las versiones más conocidas, debida a Isaac Levi.

3. La teoría de la decisión cognitiva

Supongamos que un agente considera un conjunto determinado de diferentes acciones posibles, y que debe decidir cuál de dichas acciones llevar a cabo. La teoría de la decisión indica cómo efectuar la elección de manera racional, dada una serie de supuestos básicos. Por ejemplo, se supone que el agente puede ordenar las consecuencias de sus acciones por grado de preferencia. Es decir que un agente preferirá que ocurran ciertas consecuencias más que otras. Esta escala de preferencias no es simplemente ordinal, sino que se supone, además, que el agente puede asignar utilidades específicas a cada consecuencia.¹⁹ También se presupone que cada agente es capaz de hacer asignaciones coherentes de

19. La teoría requiere que el agente asigne utilidades, debido a que ello permite medir la distancia relativa entre los elementos de la jerarquía. No se necesita ninguna otra información más allá del orden y la distancia relativa entre dichos elementos. Las escalas de utilidad son pues lo que se conoce como "escalas de intervalo": en sentido estricto, la escala de utilidades de un agente determi-

probabilidad subjetiva sobre los estados posibles del mundo, lo cual determina a su vez que cada una de las consecuencias posibles de una acción tendrá asociada cierta probabilidad. Con todos estos elementos, la utilidad esperada (UE) de cada acción considerada por el agente es la sumatoria de las utilidades asociadas con cada consecuencia, ponderadas por la probabilidad de cada consecuencia. Si c_i es la consecuencia de la acción A dado que ocurre el estado i , $u(c_i)$ la utilidad de c_i , y $P(c_i)$ la probabilidad de c_i , y además, según el agente, hay n estados posibles, tenemos que $UE(A) = \sum^n [u(c_i) P(c_i)]$. La teoría, en su uso normativo, indica que los agentes, para actuar racionalmente, deben actuar de modo tal que maximicen su utilidad esperada.²⁰

Pues bien, la teoría de la decisión cognitiva consiste en la aplicación de estos conceptos básicos al campo epistémico. Así, en lugar de calcular la utilidad esperada que resulta de llevar a cabo diversas acciones posibles, la teoría calcula la utilidad *epistémica* esperada de distintas hipótesis. La utilidad epistémica esperada de una hipótesis equivale en verdad a la utilidad esperada que tendría para el agente el acto de creer o aceptar la hipótesis en cuestión.

Posiblemente la versión más detallada de dicha teoría se deba a Isaac Levi.²¹ La versión de Levi será el punto de partida de mi análisis del contexto selectivo de una IME, de modo que la expondré con cierto detalle, aunque muchos aspectos deberán ser inevitablemente simplificados por razones de espacio.

Levi recurre a la idea de decisión cognitiva a la hora de caracterizar lo que él llama una *expansión inductiva*. En una expansión inductiva típica, el estado epistémico de un agente se ve ampliado por el agregado de un nuevo elemento H , tal que los elementos previamente presentes

na un conjunto infinito de funciones de utilidad, todas ellas equivalentes bajo transformación positiva lineal.

20. Es importante señalar que hay usos no normativos de la teoría de la decisión. Por ejemplo, los autores que se comprometen con alguna versión de la llamada "teoría de la preferencia revelada" proponen interpretar la acción de un individuo como aquella que maximiza su utilidad esperada, con lo cual las escalas de preferencia y probabilidad son reconstruidas *a posteriori*. De hecho una de las formulaciones estándar de la teoría, debida a Savage (1954), incurre en dicho compromiso.

21. Por ejemplo en Levi (1984), capítulos 5, 6 y 7, o Levi (2001).

hacen razonable (*i.e.*, apoyan inductivamente) la incorporación de H . H , sin embargo, no se *deduce* de los elementos previos, y por lo tanto, según la perspectiva del agente, aceptar H entraña cierto riesgo de error.

La oportunidad de realizar una expansión inductiva aparece cuando el agente desea obtener una respuesta para cierta pregunta. Sea K una representación del estado epistémico inicial del agente. La representación más usual de dicho estado epistémico consiste en un conjunto deductivamente cerrado de enunciados de un lenguaje formal debidamente regimentado. Es habitual referirse a K como el "conjunto de creencias" del agente. Con base en K , se define luego una *partición última* U_K , o conjunto de hipótesis $h_1...h_n$, tal que: (i) cada una de las hipótesis provee una respuesta a la pregunta que inicia la investigación; (ii) todas las hipótesis son excluyentes entre sí; (iii) todas son compatibles con K ; (iv) ninguna de ellas se deduce de K ; y (v) una y sólo una de ellas es verdadera dado K . Para la aplicación que nos interesa en este artículo, podemos interpretar a $h_1...h_n$ como explicaciones potenciales rivales, empíricamente adecuadas, y capaces de satisfacer cierta demanda de información por parte del agente.

El siguiente paso consiste en evaluar cuál es el elemento preferido de dicha partición última; en caso de duda, más de un elemento podría ser seleccionado; en algunas situaciones, ninguno de ellos lo será. Cada uno de estos resultados determinará un tipo de modificación característica de K , sobre la que volveremos más adelante. Para llevar a cabo dicha evaluación, Levi calcula la utilidad epistémica esperada de cada una de las hipótesis. Veamos cómo.

En primer lugar, debemos definir una función de probabilidad P_K , relativa a K , sobre los elementos de la partición última. (En lo que sigue eliminaré el subíndice K de P_K toda vez que no haya riesgo de confusión).

En segundo lugar, debemos evaluar la *utilidad* de cada uno de dichos elementos. Establecer funciones de utilidad adecuadas para los elementos de U_K no es una tarea sencilla. Reinterpretando una vieja idea de William James, Levi sugiere que en toda investigación coexisten dos objetivos o metas en pugna entre sí; dos *desiderata*.²² Una función de utilidad apropiada deberá tenerlos en cuenta a ambos. Uno de ellos consiste en evitar el error, o evitar creer cosas erróneas. Diremos que

22. James (1897), p. 17 y ss.

$T(h,x)$ es una función que arroja diferente resultado dependiendo de si el error fue evitado ($x = v$) o no ($x = f$). Sea $T(h,v) = a$ y $T(h,f) = b$, donde $a > b$. Para simplificar, estipularemos que $a = 1$ y $b = 0$.

El otro objetivo es adquirir nueva información valiosa. Sea $C(h,x)$ una función que especifica el incremento en valor informacional obtenido. Se presupondrá que este valor es el mismo ya sea que $x = v$ o que $x = f$; esto es, $C(h,v) = C(h,f) = C(h)$. No entraré aquí en los pormenores de cómo calcular C . Solamente diré, de manera un tanto imprecisa, que C mide el *contenido* de la hipótesis.

En síntesis, de acuerdo con la teoría de los dos *desiderata* de la investigación, por un lado deseamos salir de situaciones de incertidumbre y adquirir nueva información, pero por otro lado, también deseamos evitar el error. La utilidad cognitiva $V(h,x)$ de una hipótesis, entonces, debería expresar cuantitativamente el grado en el cual la hipótesis satisface ambos objetivos. Pero salta a la vista que éstos no pueden satisfacerse máximamente de manera simultánea: maximizar el deseo de evitar el error equivale a minimizar el deseo de adoptar nuevas creencias dudosas, y viceversa. De modo que una función de utilidad lo más que puede hacer es expresar el grado en que una hipótesis es capaz de satisfacer *cierto compromiso* entre ambos objetivos. Sea pues $V(h,x) = \alpha T(h,x) + (1-\alpha)C(h,x)$. El coeficiente α , o índice de cautela,²³ puede verse como una manera de representar la importancia relativa atribuida a la preocupación por evitar el error, y a la preocupación por maximizar el valor informacional de nuestro cuerpo de creencias.

Con esto ya tenemos todas las herramientas necesarias para calcular la utilidad cognitiva esperada (UE) de cada elemento de la partición última. Es claro que, para cada hipótesis, hay dos únicos estados posibles del mundo que son relevantes: o bien la hipótesis es verdadera, o bien es falsa. Entonces, la utilidad cognitiva esperada de una hipótesis será la probabilidad de que la hipótesis sea verdadera, multiplicada por la utilidad asociada a que sea verdadera, más la probabilidad de que sea falsa, multiplicada por la utilidad asociada a que sea falsa. Es decir:

23. En sentido estricto, lo que Levi llama "índice de cautela" (o de audacia, según se mire) es $(1-\alpha)/\alpha = q$. q tiene ciertas propiedades muy interesantes, sobre las que lamentablemente no me puedo explayar aquí. Véase Levi (2001), pp. 32 y ss.

$$\begin{aligned} UE(h) &= P(h) V(h,v) + P(-h) V(h,f) = \\ &= P(h) [\alpha T(h,v) + (1-\alpha)C(h)] + (1-P(h)) [\alpha T(h,f) + (1-\alpha)C(h)] = \\ &= \alpha P(h) + (1-\alpha)C(h).^{24} \end{aligned}$$

De acuerdo con esto, Levi propone una regla específica para la modificación de K . No daremos aquí los detalles de dicha regla.²⁵ Sólo diré que Levi propone rechazar todos los elementos que no cumplan con cierto nivel mínimo de utilidad cognitiva esperada. Luego, se incorporará a K la disyunción de todos los elementos $h_i \in U_K$ que no han sido rechazados, para finalmente realizar la clausura deductiva correspondiente. Así, si la disyunción de elementos no rechazados es representada por H , tendremos como resultado un nuevo estado epistémico K' , conformado por K , H , y las consecuencias lógicas de ambos. En símbolos, $K' = Cn(K \cup \{H\})$, donde Cn es el operador tarskiano de consecuencia lógica. La idea es que, si hay más de un elemento no rechazado (esto es, si H no es simplemente idéntica a algún h_i de la partición última) el agente suspende el juicio entre ellos. En caso de duda severa, H podría consistir en la disyunción de todos los elementos de U_K , en cuyo caso en realidad no estaríamos haciendo modificación alguna del estado epistémico del agente, puesto que H ya pertenecería a K (dado que K es un conjunto deductivamente cerrado, y dado que inicialmente al menos un elemento de la partición última es juzgado verdadero). Por otra parte, rechazar *todos* los elementos de U_K supondría introducir una contradicción en el conjunto de creencias. De modo que, de obtenerse este resultado, el agente se vería forzado a retirar algunos elementos de su conjunto inicial K . Pero aquí ya no estaríamos tratando con expansiones inductivas, sino con *contracciones*. En este trabajo no me ocuparé de examinar las posibles estrategias de contracción de un conjunto de creencias.

Obsérvese que el mecanismo de selección descrito en esta sección puede muy bien aplicarse a un caso de IME, en el cual la partición últi-

24. Como ya he advertido, esta presentación es tan sólo una versión simplificada de la propuesta de Levi. Entre otras cosas, éste presupone que los agentes raramente cuentan con una única función P de probabilidad subjetiva. En los casos típicos, debemos suponer que disponen de un *conjunto convexo* de tales funciones de probabilidad. Sobre este tema véase "On Indeterminate Probabilities", reimpreso en Levi (1997), cap. 6.

25. Véase Levi (2001), p. 32.

ma consiste en un conjunto dado de explicaciones rivales empíricamente adecuadas. Para elegir la mejor explicación, o disyunción de mejores explicaciones, se calcula la utilidad cognitiva esperada de cada hipótesis de la manera ya señalada, para lo cual deben tenerse en cuenta las funciones *P* y *C* del investigador, así como cierto índice de cautela. Luego de realizada la selección, la teoría presentada aquí requiere que la mejor explicación, o disyunción de mejores explicaciones, se incorporen al conjunto *K* de creencias del agente, modificándolo de la manera detallada en el párrafo anterior.

Entiendo que el mecanismo de selección descripto cumple con todo lo que podemos pedir de un mecanismo selectivo en una IME. Sin embargo, creo también que el modelo general de cambio epistémico en el que dicho mecanismo de selección está inserto no permite capturar toda la gama de situaciones que pueden ocurrir en el marco de una IME. En la próxima sección argumentaré que las ideas básicas presentadas aquí deberían adoptarse en combinación con un modelo de cambio epistémico un tanto más complejo.

4. Compromisos epistémicos contextuales

De acuerdo con una distinción usual en la bibliografía reciente, hay básicamente dos maneras opuestas de encarar la problemática de la investigación racional y la creencia. Por un lado, podemos optar por un modelo cartesiano de investigación, llamado a veces el "modelo Duda-Creencia"; por otro lado, podemos adoptar un modelo peirceano, o modelo "Creencia-Duda".²⁶ El primero se identifica con la teoría del conocimiento tradicional, y viene en todo tipo de variantes (fundacionismo, coherentismo, externalismo, internalismo, etc.). Adviértase que en este contexto el término "cartesiano" no refiere a la variedad o *especie* de modelo dada por Descartes mismo, sino a su *género*. Un rasgo común de todas las variantes de este modelo es la exigencia de identificar algún criterio para separar las creencias que están justificadas de aquellas que no lo están. Por el contrario, el segundo modelo surge de raíces pragmatistas, y mantiene que sólo los *cambios* de creencias deben justi-

26. Es habitual identificar a Peirce (1877) como el escrito fundacional del modelo Creencia-Duda. Para una descripción sencilla de ambos modelos, véase la introducción de Fuhrmann (1997).

ficarse. Es este último modelo el que está habitualmente presupuesto en la mayoría de las discusiones contemporáneas sobre revisión de creencias o teoría de la decisión cognitiva.

Podríamos tratar de capturar el corazón de una epistemología pragmatista de tipo peirceano por medio de algún principio apropiado. ¿Qué podría contar como una formulación satisfactoria de tal principio? He aquí una posible sugerencia:

EP: La racionalidad requiere que los agentes no realicen cambios de creencias a menos que haya justificaciones adecuadas para tales cambios. Conversamente, prescribe que los agentes lleven a cabo dichos cambios toda vez que haya buenas razones para ello. Además, también exige que los agentes no se ocupen de justificar sus creencias presentes.

Este es sólo un principio esquemático, pues deja abierto el problema acerca de qué cosa sea una justificación adecuada. Esto es exactamente lo que deseamos; el mero compromiso con una línea epistemológica pragmatista no tiene por qué forzarnos a adoptar un criterio particular de justificación.

Trataré de aclarar ahora el sentido de "creencia" que parece darse por sentado en muchas de las discusiones contemporáneas sobre cambio epistémico. Una posible interpretación es la de creencia como "compromiso epistémico" [*doxastic commitment*].²⁷ En la manera más intuitiva de entender esta idea, decir que un agente tiene una determinada creencia significa que el agente se ha hecho a sí mismo la promesa de actuar de cierta manera, donde "actuar" ha de ser tomado en un sentido amplio. Así, si *X* cree que *p*, y le preguntamos si ocurre que *p*, y además *X* quiere revelarnos su compromiso, *X* deberá contestar la pregunta afirmativamente; si *X* se encuentra en la disyuntiva de decidir si llevar a cabo *A* o *B*, deberá tomar su decisión bajo el supuesto de que *p*; de manera análoga, si *X* se propone deliberar sobre alguna cuestión (independientemente de si luego hará algún uso práctico del resultado de la deliberación), *X* deberá proceder bajo el supuesto de que *p* es verdadero. Como puede verse, "compromiso" es un concepto normativo, que indica lo que el agente *debe* hacer, y no lo que de hecho hace en cada caso.

Así pues, decir que un agente cree que *p* significa que, toda vez que

27. Cf. Levi (1991), cap. 2.

el agente actúe o delibere, deberá dar por supuesto que p . Por supuesto, nada impide que X cambie de idea sobre p , pero mientras ello no suceda, el agente se ha resuelto a tomar a p como verdadero para toda acción o deliberación futuras, cualesquiera que éstas sean. Por supuesto, no todo lo que X haga en el futuro requerirá tomar a p en consideración. Pero esto es irrelevante; lo importante es que cada vez que la consideración de la verdad o falsedad de p sea capaz de producir algún tipo de diferencia, el agente deberá actuar como si p fuera cierto.

En síntesis, la creencia es interpretada como una actitud que el agente deberá tener en *todo contexto posible* (para el agente) *de acción o deliberación*. Sin embargo, adviértase también que la idea de compromiso, por sí mismo, no acarrea esta connotación. Un agente podría muy bien estar comprometido a actuar de cierta manera sólo bajo ciertas circunstancias. Por lo tanto, "creencia" no es simplemente sinónimo de "compromiso epistémico", sino de "compromiso epistémico para todo contexto". De modo que, a la hora de dar una definición rigurosa, deberíamos explicitar este último punto. Para ser precisos, entonces, digamos que un agente X cree que p si y sólo si X está comprometido con la verdad de p (junto con todas sus consecuencias lógicas) para todo contexto de acción o deliberación. Entiendo que esta definición captura la idea de creencia implícita en buena parte de la bibliografía sobre revisión de creencias y teoría de la decisión cognitiva.

Con este marco teórico, volvamos ahora a considerar la elección de hipótesis, o "mejor explicación", a través de un mecanismo de decisión cognitiva. Es claro que la versión ofrecida en la sección anterior debe ser entendida dentro del marco teórico de los modelos Creencia-Duda: su mecanismo central consiste en utilizar creencias ya presentes, que no se cuestionan, para evaluar la legitimidad de incorporar otras nuevas. Efectivamente, dicha propuesta puede verse como un medio para proveer un *criterio* que el agente puede usar para saber si cierta expansión es obligatoria o prohibida (dado cierto problema, cierto conjunto de respuestas potenciales, un determinado conocimiento antecedente, una función de probabilidad subjetiva, un índice de riesgo, etc.). En otras palabras, dicha propuesta completa el principio esquemático (EP) con un criterio sustantivo, aunque parcial,²⁸ sobre la justificación de los cambios.

28. El criterio es parcial puesto que no trata específicamente el caso de las *contracciones* de un conjunto de creencias.

Hay, sin embargo, un aspecto crucial de (EP) completado con el mecanismo de selección de hipótesis de la sección anterior que encuentro demasiado restrictivo.

Repasemos una vez más la interpretación de "creencias" como "compromisos para todo contexto de acción y deliberación". Entiendo que, a la hora de tomar parte en un proceso de decisión cognitiva, los agentes no siempre tienen como meta la adopción de nuevas creencias en este sentido. En numerosas ocasiones, todo lo que un agente quiere es comprometerse epistémicamente *para cierto contexto determinado*. Consideremos el enunciado que dice que la enfermedad X resulta causada por un único factor Y . Un agente puede encontrar razonable aceptar dicho enunciado como verdadero en el contexto de una investigación sobre la génesis de otras enfermedades, y, sin embargo, a la hora de diseñar una política de salud pública, tal vez piense que la cautela que debe ejercerse es mayor, y que es más adecuado suspender el juicio, *i.e.*, actuar como si su verdad o falsedad no estuviera aún establecida. En este ejemplo, evidentemente el agente *no* ha decidido *creer* que X es causada por Y , en el sentido de creencia ofrecido más arriba. Sin embargo, ha decidido dar por supuesta dicha relación causal en un ámbito importante de su vida. Considero que, lejos de ser una excepción, esta es una situación con la que nos topamos habitualmente. En otras palabras, en numerosas ocasiones los agentes llevan a cabo decisiones cognitivas *contextuales*.

De aquí en más voy a reservar el nombre de "aceptaciones" para referirme a los compromisos que un agente adquiere para *algún* contexto (no necesariamente para todos); las "creencias", como ya dijimos, referirán a los compromisos que el agente mantiene para todo contexto. Obsérvese que, bajo esta definición, una creencia es un tipo especial de aceptación.

Esta definición entraña una importante consecuencia. Puesto que es posible para un agente aceptar un elemento para cierto contexto pero no para todos, ello implica que hay contextos para los cuales el agente tiene mayor cantidad de compromisos epistémicos que creencias. Así, el conjunto de elementos que un agente toma como "base de acción y deliberación" puede cambiar según el contexto. En nuestro ejemplo anterior sobre la causa de la enfermedad X , podemos suponer que, para llevar a cabo la decisión racional que finalmente culminó con la aceptación de un nuevo compromiso, nuestro agente no solamente usó un índice de cautela propio para dicho contexto, sino que dio por sentados

elementos con los que no se comprometería en otros ámbitos, esto es, elementos que no son en sentido estricto "creencias".

Podría argumentarse que hablar de realizar problemas de decisión cognitiva contextuales simplemente equivale a poner un nuevo nombre a la vieja idea de que los agentes a veces llevan a cabo razonamientos hipotéticos. Decir que hemos adoptado el nuevo compromiso de que p para un contexto en el cual damos a q por sentado, cuando q no es un compromiso para todo contexto, no sería sino otra manera de decir que, razonando bajo el supuesto (falso, en lo que a nosotros concierne) de que q es el caso, hemos llegado a ser capaces de afirmar p .

En verdad, no tengo mayores inconvenientes en redescubrir la existencia de múltiples contextos de decisión como contextos de razonamiento hipotético. Sin embargo, hablar de decisiones tomadas bajo supuestos hipotéticos es un poco engañoso, por cuanto parece dar a entender que dichas decisiones constituyen un mecanismo de cambio epistémico provisorio. Parece implicar que alguna vez el agente buscará, digamos, "cancelar" los supuestos; parece implicar también que las únicas decisiones cognitivas realmente importantes, o que realmente cuentan, son las que atañen a modificaciones de un conjunto de creencias propiamente dicho. Pero ninguna de estas connotaciones es apropiada. Las decisiones epistémicas contextuales no son necesariamente provisorias; no son decisiones que un agente tome temporalmente hasta tanto decida qué creencias adoptar. Por el contrario, considero que la existencia permanente de múltiples contextos de decisión es parte esencial del estado epistémico habitual de un agente. Por supuesto, cómo referirse a dichos contextos constituye una cuestión terminológica menor, siempre y cuando no se generen connotaciones indeseables. Para evitar dichas connotaciones es que prefiero evitar hablar de "decisiones cognitivas hipotéticas" o de "decisiones bajo supuestos". Hablaré en cambio de "contextos de aceptación", o "contextos de decisión cognitiva", que son expresiones más neutrales, así como de "elementos aceptados contextualmente".

En vista de la existencia de múltiples contextos, el principio (EP) formulado más arriba resulta demasiado restringido, puesto que un problema de decisión cognitiva no siempre dará como resultado un *cambio de creencias*. En la próxima sección trataré con un poco más de detalle la distinción entre creencia y aceptación que acabó de sugerir, y propondré un nuevo principio fundamental para el modelo peirceano de epistemología. El resultado será una concepción más flexible de la expansión inductiva y de la inferencia a la mejor explicación.

5. Aceptar y creer

En esta sección ofreceré el esbozo de una teoría formal basada en las ideas sugeridas en la sección 4.

A la hora de describir el estado epistémico de un agente, propongo distinguir entre un conjunto K de creencias en sentido estricto y un conjunto más amplio de elementos aceptados, al que llamaremos M . Tanto "creencias" como "aceptaciones" son representadas por oraciones de un lenguaje proposicional L . M es consistente e incluye a K como parte propia. Diremos que M es el conjunto de elementos que el agente se ha comprometido a tomar como base de acción y deliberación para *algún* contexto.

Sea A_c el conjunto de aceptaciones para el contexto c . A_c está conformado por aquellos elementos con los que el agente se ha comprometido para c (esto es, que toma como verdaderos en c). Se requerirá además que A_c sea deductivamente cerrado y consistente. K es la intersección de todos los conjuntos de aceptaciones, para todo contexto, y M es la unión de tales conjuntos. Formalmente:

- $A_c = Cn(A_c) \subseteq M$ es el conjunto de elementos aceptados como verdaderos en el contexto c .
- $K = \bigcap_c A_c$; y $M = \bigcup_c A_c$, para todo c .

Ahora bien, en muchas teorías estándar de dinámica epistémica, se supone que a la hora de expandir un conjunto de *creencias*, un agente necesita ordenar de alguna manera los elementos inciertos, esto es, los elementos compatibles con K pero que están fuera de K . Esto se consigue, por ejemplo, si suponemos que el agente cuenta con una función de probabilidad subjetiva dado K . Se recordará que en la sección 3 llamamos a esta función P_K . Es claro que todos los elementos de K , o creencias, tienen probabilidad 1 según P_K . Algunos de los elementos fuera de K , pero compatibles con K , podrían también tener probabilidad 1 según P_K , aunque ello no es necesario.

Pues bien, en el nuevo esquema que sugiero aquí, supondremos que el agente cuenta con distintas funciones de probabilidad para cada contexto. Diremos que P_{A_c} es la función de probabilidad subjetiva *relativa al conjunto de aceptaciones del contexto c* . El hecho de contar con diferentes funciones para cada contexto, evidentemente, tendrá consecuencias a la hora de llevar a cabo un ejercicio de decisión cognitiva y evaluar un posible cambio epistémico.

Así, propongo analizar el cambio epistémico de la siguiente manera: al participar en un problema de decisión cognitiva, un agente busca expandir A_c , para cierto contexto c , lo cual puede llevar a expandir también M . Más precisamente, a la hora de participar en un problema de decisión cognitiva el agente usará como base para su deliberación algún conjunto $A_c \subseteq M$. Por consiguiente, el agente utilizará alguna función de probabilidad P_{A_c} que asigne probabilidad 1 a los elementos de A_c , y que asigne algún número en el intervalo $[0,1]$ a los elementos compatibles con A_c pero que no pertenecen a A_c . El resultado de una expansión por H , para H consistente con M , será entonces un nuevo conjunto $A'_c = Cn(A_c \cup \{H\})$. En este caso, si H no estaba ya en M se habrá ampliado el conjunto total de aceptaciones del agente, pero no necesariamente su conjunto de creencias (a menos que se trate del caso especial en que $A_c = K$).

En fin, la aceptación es contextual, y el agente no sabe cómo reaccionará en otro contexto c' frente a la posibilidad de incorporar H , no solamente porque $P_{A'_c}$ puede diferir de P_{A_c} , sino porque el índice de cautela que el agente decidió usar en c puede no ser el mismo que está dispuesto a usar en c' .

La idea de que un agente puede tener más de una función de probabilidad subjetiva no es demasiado novedosa. Sin ir más lejos, en la propuesta de Levi el agente cuenta en realidad con un *conjunto convexo* de tales funciones.²⁹ Ello intenta recoger la intuición de que los agentes raramente asignan valores numéricos precisos a sus incertidumbres. Supongamos que Juan tiene un alto grado de confianza en que mañana lloverá, aunque no tiene certeza total. Nos parecería extraño, sin embargo, si Juan nos dijera que tiene una confianza de grado 0,957 en que mañana habrá lluvias. La confianza de los agentes más bien se refleja en *intervalos* numéricos; Juan podría afirmar, por ejemplo, que su nivel de confianza en que mañana lloverá se ubica entre 0,8 y 1. Pues bien, Levi muestra que esto equivale a reconocer que Juan cuenta en verdad con un *conjunto* de funciones de probabilidad, tal que cada una de ellas atribuye un número preciso entre 0,8 y 1 a la proposición de que habrá lluvia mañana. Sin embargo, en el esquema de Levi todas estas funciones son relativas al conjunto K . Es decir que la "base de acción y delibera-

29. Un conjunto Δ no vacío de funciones de probabilidad subjetiva es *convexo* si y sólo si: si $P, P' \in \Delta$, entonces todo P'' que resulte de un promedio ponderado entre P y P' también pertenece a Δ . Véase Levi (1997), cap. 6, p. 127.

ción" es siempre el conjunto K . Yo propongo en cambio permitir que dicha base cambie de acuerdo a los contextos.

En verdad, adaptando la idea de Levi, bien podríamos suponer que el agente cuenta con un conjunto convexo de funciones (en lugar de una única función) *para cada contexto*. Si realizamos esta suposición adicional, podemos obtener maneras alternativas de elucidar la idea de aceptación contextual. Por ejemplo, podríamos tomar *como punto de partida* el conjunto de probabilidades subjetivas del agente, y con base en ellas determinar cuáles son los contextos posibles para el agente, y cuáles los elementos aceptados para cada contexto. Así, cada contexto podría muy bien ser caracterizado por cierto conjunto de probabilidades. Veamos cómo:

Sea Δ un conjunto convexo³⁰ de funciones de probabilidad subjetiva P_n sobre L , y sea T el conjunto de todas las teorías T_i (conjuntos deductivamente cerrados) de L . Entonces, puede definirse una función $f: \Delta \rightarrow T$ tal que $f(P_n) = T_i$ si y sólo si para todo elemento $a \in T_i$, $P_n(a) = 1$, y para toda otra teoría T_j , si $P_n(a) = 1$ para todo $a \in T_j$, entonces $T_j \subseteq T_i$. Esto es, cada elemento P_n de Δ selecciona algún miembro de T , a saber, la mayor teoría T para la cual $P_n(a) = 1$, para todo a en T . Típicamente, algunos miembros de Δ seleccionarán el mismo elemento en T (y algunas teorías en T no serán seleccionadas en absoluto, por supuesto). Un conjunto de aceptaciones contextuales será entonces una teoría T de T que es seleccionada por algún miembro de Δ . Un contexto c puede caracterizarse mediante el subconjunto $\Delta' \subseteq \Delta$ compuesto por todos aquellos elementos P_n que seleccionan la teoría $T = A_c$. Dicho de otra manera, cada contexto puede identificarse con el conjunto de funciones de probabilidad subjetivas del agente que determinan unívocamente cierto conjunto de aceptaciones (a saber, el conjunto de aceptaciones para dicho contexto).

Luego podemos definir K como la mayor teoría incluida en todo contexto. Es claro que entonces toda función P_n asigna probabilidad 1 a todo elemento de K , como debe ser. Por otra parte, no es necesario que haya una función P_n que atribuya probabilidad 1 a todo elemento de M , si bien todo elemento de M recibe probabilidad 1 por parte de alguna (no necesariamente la misma) función P_n .

30. Para una explicación de la propiedad de convexidad, véase la nota anterior.

En este escenario, registrar un cambio en un conjunto de elementos aceptados es en verdad una manera diferente de expresar que ha ocurrido un cambio en el conjunto de probabilidades subjetivas del agente, esto es, que ha tenido lugar un cambio en Δ . Dicho cambio, sin embargo, no necesariamente será el resultado de un cambio en K , esto es, de un cambio en sus *creencias* propiamente dichas.

La posibilidad de esta última elucidación es de alguna importancia, según creo. Por un lado, permite simplificar los elementos primitivos del sistema: todo lo que necesitamos es suponer que un agente cuenta con un conjunto de funciones de probabilidad. Por otro lado, disipa el posible recelo que la idea de contexto puede generar, puesto que muestra que hablar en términos de aceptaciones contextuales no nos compromete con entidades dudosas, ni está reñido con la posibilidad de ofrecer un análisis riguroso.

Si el análisis anterior es plausible, entonces debemos considerar un nuevo principio fundamental de la epistemología peirceana. A saber:

EP': Los agentes pueden modificar su conjunto de elementos aceptados sólo si existe justificación adecuada para dicho cambio. Conversamente, los agentes deberán modificar tal conjunto toda vez que haya razones de peso para ello. Además, los agentes no deben preocuparse por justificar elementos previamente aceptados.

Obsérvese que obtenemos a (EP) como caso límite de (EP').

¿Qué podemos concluir de toda esta discusión respecto de nuestro problema original, esto es, el problema de elucidar la idea de Inferencia a la Mejor Explicación? En la sección 2 propuse entender una IME como un proceso que culmina con una modificación del estado epistémico de un agente. Creo que ahora es claro cómo ha de interpretarse dicha modificación. Típicamente, una IME se formula para resolver un problema específico, y en el marco de un contexto específico, lo que supone un conjunto particular de funciones de probabilidad subjetivas, y, como consecuencia, un conjunto particular de aceptaciones. Además, en general el índice de cautela que el agente está dispuesto a usar no será trasladable a otros contextos, aún si se tratara de evaluar, en un contexto diferente, el mismo conjunto de hipótesis, esto es, la misma partición última. Por todo ello, el resultado de una IME será una nueva aceptación del agente, aunque no necesariamente una nueva creencia.

6. Inferencia a la mejor explicación y realismo científico

La mayoría de los estudios filosóficos sobre IME aparecidos recientemente realiza algún tipo de evaluación sobre la relación entre IME y la posición epistemológica conocida como "realismo científico". De hecho, en tiempos recientes raramente encontramos publicaciones de alguna de las dos temáticas sin referencia a la otra. El presente estudio no estaría completo, pues, sin al menos una breve mención sobre un posible vínculo con el realismo. Por otra parte, algunos autores han argumentado que la relación entre IME y realismo brinda un motivo de sospecha sobre la legitimidad de las IME, de modo que dicho vínculo nos presenta en verdad con un problema al que hay que dar alguna solución.³¹ En los párrafos que siguen espero mostrar que la elucidación propuesta en este trabajo permite evaluar el supuesto compromiso realista de las IME desde una nueva perspectiva.

Ante todo, es conveniente introducir algunas aclaraciones. La posible conexión entre IME y realismo puede aludir a dos cuestiones muy diferentes que no deben confundirse entre sí. Lamentablemente, rara vez han sido enunciadas con claridad como cuestiones independientes. Una de ellas es si la idea misma de IME se compromete con una posición realista; otro problema es si es posible defender el realismo científico usando una estructura argumentativa que responda al esquema de una IME. Como ejemplos de autores involucrados en la primera polémica, podemos citar a Bas van Fraassen (en calidad de defensor de la estrecha conexión entre IME y realismo)³² y a Arthur Fine (en el papel

31. Hay otro problema importante asociado con las IME; a saber, se ha argumentado que las IME constituirían un patrón incoherente de cambio epistémico. Esta objeción se debe esencialmente a van Fraassen (1989, p. 166 y ss.), y hace uso de un argumento más general que se conoce como "Dutch Book diacrónico". Dicho argumento aparece por primera vez en Teller (1973), quien recoge una idea original de David Lewis. Lamentablemente no tengo espacio en este artículo para desarrollar el problema con el detalle que merece, de modo que quedará para otra oportunidad. Una estrategia posible para desactivar la objeción consiste en intentar invalidar la legitimidad de los *Dutch Books* diacrónicos en general. Existe una muy amplia bibliografía sobre el tema. Como muestra, véanse Douven (1999), Machina (1991), Maher (1993), cap. 5, o Levi (1987).

32. Van Fraassen (1989). Como es bien conocido, van Fraassen mantiene una perspectiva crítica hacia el realismo científico y, por consiguiente, hacia las IME.

de escéptico respecto del compromiso realista de una IME).³³ Como ejemplos de la segunda polémica, encontramos a Richard Boyd³⁴ (defendiendo la legitimidad de una estructura de IME para defender al realismo) y a Peter Lipton³⁵ (en el papel de escéptico respecto de dicha legitimidad). Para asegurarnos que ambas discusiones son independientes, basta notar que, al menos en principio, resultaría legítimo apelar a una IME para apuntalar al realismo *si y sólo si* mantenemos la neutralidad epistémica de las IME, so pena de circularidad. Aquí me ocuparé brevemente de la primera polémica, pero no diré nada acerca de la segunda.

Consideremos, pues, la posibilidad de que la IME se comprometa intrínsecamente con el realismo científico. Adoptar un compromiso realista no es necesariamente un problema, por supuesto. Sin embargo, si la conexión es tan fuerte como algunos autores parecen creer, la suerte de toda IME está atada a la del realismo; todo aquel que encuentre a este último insatisfactorio rechazará igualmente las IME.

Esta es precisamente la posición de van Fraassen. He aquí una reconstrucción de uno de los argumentos presentados por van Fraassen en (1989):

- (1) El realismo científico es la posición que se compromete con la verdad de las mejores explicaciones disponibles, o, al menos, con su verosimilitud.
- (2) Una IME es esencialmente una regla de cambio epistémico según la cual las mejores explicaciones disponibles deben incorporarse al conjunto de creencias de un agente, y por ende deben adoptarse como verdaderas (para todo contexto).

33. Fine ha expresado en diversas ocasiones que el resultado de una IME bien podría ser una hipótesis que "salve los fenómenos", al estilo instrumentalista (cf. por ejemplo Fine (1984)). Fine, sin embargo, no distingue ambas polémicas del modo en que las hemos presentado aquí, y también podría ser ubicada como escéptico en la segunda polémica que menciono en el trabajo.

34. Por ejemplo en Boyd (1983).

35. Véase por ejemplo Lipton (1991), cap. 9. Cabe observar que Lipton sí cree que hay argumentos para defender al realismo científico. Pero concede que un antirrealista no podría ser convencido por medio de un argumento que apele a la idea de que la verdad de las teorías exitosas constituye la *mejor explicación de su éxito* (pp. 160-168).

- (3) Pero no hay garantías para afirmar que las mejores explicaciones disponibles sean verdaderas, ni para afirmar que puedan estar remotamente cerca de la verdad en algún sentido interesante: el realismo científico es una posición infundada.
- (4) Por ende, toda regla que recomiende *creer* las mejores explicaciones disponibles es una mala regla de cambio epistémico.
- (5) La IME es una mala regla de cambio epistémico.³⁶

Pero de acuerdo con el análisis propuesto en este artículo, la premisa (2) es falsa, de modo que el argumento queda desactivado, independientemente de que aceptemos o rechacemos la premisa (3). El vínculo entre IME y realismo sólo se sostiene bajo el supuesto de que el producto de un problema de decisión cognitiva es siempre la recomendación de agregar una nueva creencia al cuerpo epistémico del agente. Pero se ha argumentado que un problema de decisión cognitiva no necesariamente culmina con un nuevo conjunto de creencias; de hecho, en los casos típicos tan sólo se obtiene una nueva aceptación contextual. En síntesis, el mecanismo de una IME no fuerza a un agente a adoptar una nueva *creencia*, si bien tampoco lo prohíbe. El uso de IME como una regla para el cambio epistémico no nos compromete con una epistemología realista.

7. Conclusión

El concepto de inferencia a la mejor explicación se relaciona estrechamente con muchos de los temas centrales de la epistemología contemporánea, lo cual lo transforma en un objeto de estudio particularmente interesante. Comencé por observar que en numerosas ocasiones lo encontramos confundido con la idea de abducción. Argumenté aquí que la identificación entre ambos conceptos debe ser evitada. Sugerí en cambio reservar el nombre de abducción para una de las dos etapas que componen el proceso de una IME.

En la primera etapa, o etapa abductiva, se generan las explicaciones potenciales que pueden dar respuesta a algún interrogante del agente. Quedará para otra oportunidad un análisis más pormenorizado de este

36. Más adelante van Fraassen considera una regla de IME probabilística, e intenta desacreditarla mediante el llamado argumento del *Dutch Book* diacrónico. Véase la nota 31 de este trabajo.

ámbito. En la segunda etapa, o etapa de selección de hipótesis, se lleva a cabo la elección de la mejor (o mejores) hipótesis disponibles. Propuse llevar a cabo dicha selección con ayuda de algunos instrumentos de teoría de la decisión cognitiva. Para ello me basé en la propuesta de expansión inductiva de Isaac Levi. Argumenté que la teoría de Levi presupone cierto modo peculiar de entender los lineamientos básicos de una epistemología pragmatista de tipo peirceano, y argumenté también que dicha interpretación de la epistemología peirceana es demasiado restringida. Más precisamente, argumenté que un principio básico de epistemología pragmatista no debe ocuparse de regular el cambio racional de *creencias*, sino que debe limitarse a reglamentar el cambio racional de elementos aceptados, o *aceptaciones*.

La razón es la siguiente. Comencé por defender la conveniencia de elucidar la idea pragmatista de creencia en términos de un "compromiso epistémico para todo contexto de acción o deliberación". Ahora bien, a la hora de llevar a cabo una investigación, nos encontramos con que los agentes muchas veces toman en consideración elementos esencialmente contextuales, que raramente pueden ser generalizados para todo contexto. Es decir que los agentes típicamente evalúan las ventajas o desventajas de comprometerse con cierta hipótesis para el contexto particular en el cual se lleva a cabo la investigación. De ello resulta que, en general, una explicación potencial podrá ser aceptada o rechazada *para el contexto en cuestión*.

De modo que propuse una manera más amplia de entender la epistemología peirceana y, correspondientemente, una versión un tanto más compleja de cómo procede un mecanismo de selección de hipótesis o explicaciones potenciales. Para ello, esboqué un modelo epistémico que toma como central la idea de conjuntos de elementos aceptados para diferentes contextos. *K*, o el conjunto de creencias propiamente dichas del agente, es la intersección de la totalidad de tales conjuntos. Dado que un proceso de decisión cognitiva es típicamente relativo a cierto contexto *c*, el resultado de llevar a cabo una inferencia a la mejor explicación es un nuevo conjunto de aceptaciones para algún *c*. También sugerí que cada contexto puede identificarse con un conjunto de funciones de probabilidad subjetivas del agente, las que determinan unívocamente cierto conjunto de aceptaciones (a saber, el conjunto de aceptaciones para dicho contexto).

Argumenté, finalmente, que mi propuesta permite revelar que una inferencia a la mejor explicación es en principio un mecanismo de cambio epistémico neutral respecto del realismo científico.

Por supuesto, aún queda mucho por hacer. Por ejemplo, no me ocupé de dar cuenta de las posibles contracciones del estado epistémico de un agente. Por otra parte, el presente trabajo podría ser suplementado con un análisis de cuándo una hipótesis es en verdad *explicativa*, lo cual correspondería, de acuerdo con el esquema sugerido, a un estudio de la etapa abductiva de las IME. Confío, sin embargo, en que este trabajo pueda verse como un primer paso hacia la construcción de una teoría general de cambio epistémico, que sea además capaz de tratar satisfactoriamente el problema de la elección entre hipótesis rivales empíricamente adecuadas.

Columbia University

BIBLIOGRAFÍA

- BOYD, Richard (1983), "On the Current Status of the Issue of Scientific Realism", *Erkenntnis*, 19, pp. 45-90.
- COHEN, Jonathan (1992), *An Essay on Belief and Acceptance*, Oxford, Clarendon Press.
- DANCY, Jonathan, y SOSA, Ernest (1993), *A Companion to Epistemology*, Oxford, Blackwell.
- DOUVEN, Igor (1999), "Inference to the Best Explanation Made Coherent", *Philosophy of Science*, 66 (Proceedings), pp. S424-S435.
- FLACHS, Peter y KAKAS, Anthony (2000), *Abduction and Induction: Essays on their Relation and Integration*, Dordrecht, Kluwer.
- FINE, Arthur (1984), "The Natural Ontological Attitude", reimpreso en Papineau, David (ed.) *The Philosophy of Science*, Oxford, 1996, pp. 21-44.
- FUHRMANN, André (1997), *An Essay on Contraction*, Stanford, California, CSLI Publications.
- GÄRDENFORS, Peter (1988), *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, Cambridge, Mass., The MIT Press.
- HANSON, Norwood R. (1958), *Patterns of Discovery*, Cambridge, Cambridge University Press.
- HANSSON, Sven Ove (1999), *A Textbook on Belief Dynamics: Theory Change and Database Updating*, Dordrecht/ Boston/ London, Kluwer.
- HARMAN, Gilbert (1965), "The Inference to the Best Explanation", *The Philosophical Review*, vol. 74, I, pp. 88-95.

- HARMAN, Gilbert (1968), "Enumerative Induction as Inference to the Best Explanation", *The Journal of Philosophy*, 65(18), pp. 529-533.
- JAMES, William (1897), *The Will to Believe and Other Essays in Popular Philosophy*. Reimpreso en New York, Dover, 1956.
- JOSEPHSON, John y JOSEPHSON Susan (1994), *Abductive Inference: Computation, Philosophy, Technology*, Cambridge, Cambridge University Press.
- MACHINA, Mark (1991), "Dynamic Consistency and Non-Expected Utility", en: Bacharach, Michael y Hurley, Susan, *Foundations of Decision Theory: Issues and Advances*, Oxford, Blackwell.
- MAHER, Peter (1993), *Betting on Theories*, Cambridge, Cambridge University Press.
- LAUDAN, Larry (1981), "Why was the Logic of Discovery Abandoned?", en: *Science and Hypothesis: Historical Essays on Scientific Methodology*, Dordrecht, D. Reidel, cap. 11, pp. 181-191.
- LEVI, Isaac (1984), *Decisions and Revisions*, Cambridge, Cambridge University Press.
- LEVI, Isaac (1987) "The Demons of Decision", *The Monist*, 1987, pp. 193-211.
- LEVI, Isaac, (1991), *The Fixation of Belief and Its Undoing: Changing Beliefs Through Inquiry*, Cambridge, Cambridge University Press.
- LEVI, Isaac (1997), *The Covenant of Reason*, Cambridge, Cambridge University Press.
- LEVI, Isaac (2001), "Inductive Expansions and Non-monotonic Reasoning", en: Rott, Hans y Williams, Mary Ann, *Belief Revision*, Dordrecht, Kluwer.
- LIPTON, Peter (1991), *Inference to the Best Explanation*, London, Routledge.
- PAGNUCCO, Maurice (1996), *The Role of Abductive Reasoning within the Process of Belief Revision*, Tesis de Doctorado, Basser Department of Computer Science, University of Sydney, Australia.
- PEIRCE, Charles Sanders (1877), "The Fixation of Belief", reimpreso en: Houser, N. y Kloesel, Ch., *The Essential Peirce*, Vol. I, Indiana University Press, cap. 7, 1992, pp. 109-123.
- PEIRCE, Charles Sanders (1878), "Deduction, Induction and Hypothesis", reimpreso en: Houser, N. y Kloesel, Ch., *The Essential Peirce*, vol. I, Indiana University Press, cap. 12, 1992, pp. 186-199.
- PEIRCE, Charles Sanders (1908), "A Neglected Argument for the Reality of God", reimpreso en: *Selected Writings*, New York, Dover, 1966, pp. 358-379.

- PEIRCE, Charles S. (1958), *Collected Papers of Charles Sanders Peirce*, Cambridge, Mass., Harvard University Press. Editado por C. Harsthorne, P. Weiss. y A. Burks.
- SAVAGE, Leonard J. (1954), *The Foundations of Statistics*, 2ª edición, New York, Dover, 1972.
- TELLER, Paul (1973), "Conditionalization and Observation", *Synthese*, 26, pp. 218-258.
- THAGARD, Paul (1978), "The Best Explanation: Criteria for Theory Choice", *Journal of Philosophy*, vol. 75, 2, pp. 76-92.
- VAN FRAASSEN, Bas (1989), *Laws and Symmetry*, Oxford, Clarendon Press.

ABSTRACT

My aim in this paper is to offer an elucidation of the idea of Inference to the Best Explanation (IBE). I suggest to understand IBE as a complex process involving two steps: an abductive stage, where the agent chooses among the pool of elements previously identified as candidates for explanation. I give an account of the selective stage in terms of the so-called cognitive decision theory. In addition, I distinguish between believing a hypothesis and accepting it for a particular context of research, and argue that typical cases of IBE are *contextual* in this sense – hence an IBE typically yields a new contextual acceptance, but not necessarily a new belief. Finally, I show that the use of IBE, so understood, does not commit us to scientific realism.

REVUE INTERNATIONALE DE PHILOSOPHIE

Editor: Michel MEYER

Université de Bruxelles, C.P. 188
50, avenue F.D. Roosevelt - 1050 Bussels, Belgium

Each number is devoted to a particular movement, a particular philosopher, or a particular problem.

We publish 4 issues annually. Articles are written in English, French, German, or Italian.

Our last issue was devoted to *Ludwig Wittgenstein*.

Articles by: Jean-Pierre Cometti: *Présentation* – Garry Hagberg: *The Self, speaking: Wittgenstein, Introspective Utterances, and the Arts of Self-Representation* – Simo Säätelä: *"Perhaps the most important thing in connection with aesthetics": Wittgenstein o "aesthetic reactions"* – Jean-Pierre Cometti: *Merleau-Ponty, Wittgenstein, and the question of expression* – Richard Shusterman: *Wittgenstein's Somaesthetics: Body Feeling in Philosophy of Mind, Art, and Ethics* – Jean-Maurice Monnoyer: *Voir comme autrement* – François Latraverse: *Signe, proposition, situation: éléments pour une lecture du Tractatus logico-philosophicus* – Comptes Rendus: J.-P. Cometti et K. Mulligan (éd.), *La philosophie autrichienne de Bolzano à Musil, Histoire et actualité* (Roger Pouivet) – Daniel Giovannangeli, *Le retard de la conscience, Husserl, Sartre, Derrida*, (Grégory Cormann) – Sophie Jama, *La Nuit de songes de René Descartes* (Philippe Ducat) – Sylvain Auroux, *La Raison, le langage et les normes*, (Jean-Michel Fortis) – Anouk Barberousse, *La physique face à la probabilité* (Angèle Kremer-Marietti) – Cynthia Fleury, *Métaphysique de l'imagination* (Frédéric Laville).

Coming number: Bourdieu et la philosophie – L'Esthétique de Hegel – Ryle – Le Corps – Dilthey

Revista Latinoamericana de Filosofía, Vol. XXVIII N° 2 (Primavera 2002)

EL COMPORTAMIENTO COLECTIVO Y LA RELACIÓN ENTRE RESPONSABILIDAD Y CAUSACIÓN*

Carolina Sartorio

1. Introducción

En este trabajo señalo una consecuencia de ciertos tipos de comportamiento colectivo. Un "comportamiento colectivo" es el comportamiento de un grupo de agentes morales, en donde el concepto de grupo debe entenderse, como sugeriré más adelante, de manera más o menos laxa. Así como el comportamiento individual, el comportamiento colectivo puede tener consecuencias morales de diversos tipos, como injusticias y daños a otros agentes o grupos de agentes. En este trabajo sugiero que cierto tipo de comportamiento colectivo tiene consecuencias importantes para la relación entre la responsabilidad moral y la causación. En particular, argumento que la reflexión acerca de un tipo específico de comportamiento colectivo nos permite ver que la posición intuitiva sobre la relación entre responsabilidad moral y causación está equivocada. Por "responsabilidad moral" entenderé, a menos que especifique lo contrario, responsabilidad moral *individual*. En otras palabras, en este trabajo argumento que ciertos comportamientos *colectivos* refutan la posición intuitiva sobre cómo están relacionadas la responsabilidad moral *individual* y la causación.

* Presenté una versión de este trabajo en una serie de conferencias en el Instituto de Filosofía de la Universidad de Buenos Aires organizadas por Eduardo Barrio en julio de 2002. Agradezco a los asistentes a la conferencia por sus preguntas, comentarios y objeciones. Varias personas contribuyeron de maneras diversas a la evolución de este trabajo. Entre ellas quiero mencionar a Alejandro Cassini, Eduardo Rivera López y, especialmente, Juan Comesaña, Ned Hall, Judith Thomson y Stephen Yablo. Este trabajo fue realizado gracias al apoyo de la Fundación Antorchas.

2. Un contraejemplo a la posición intuitiva sobre la relación entre responsabilidad y causación

¿Cuál es la posición intuitiva acerca de la relación entre responsabilidad moral y causación? Supongamos que nos preguntamos por la responsabilidad moral de un agente determinado por hechos en el mundo, externos a él.¹ ¿Cuándo es un agente moralmente responsable por un hecho de este tipo?

Víctima yace en el suelo sin vida con una herida de bala. Tirador y Asesino a Sueldo tienen en sus manos cada uno un revólver humeante. Cuando se hace la autopsia de Víctima, se determina que la bala que mató a Víctima provino del revólver de Asesino a Sueldo, no de Tirador. ¿Es Tirador moralmente responsable por la muerte de Víctima? No. Tirador es moralmente responsable por haber intentado matar a Víctima, pero no por su muerte. Intuitivamente, esto es así porque, si bien la muerte ocurrió, Tirador no la causó. Más precisamente, el haber disparado su revólver, el único comportamiento de Tirador que podría haber causado la muerte de Víctima, de hecho no la causó. Esto es, la posición intuitiva sobre la relación entre responsabilidad y causación es que responsabilidad *implica* causación (al menos para hechos en el mundo, externos al agente): si Tirador no causó la muerte de Víctima, entonces parece seguirse que no es responsable por la muerte. ¿Cuándo diríamos que Tirador causó la muerte de Víctima? Cuando alguna acción u omisión de Tirador la causó. En otras palabras, la posición intuitiva sobre la relación entre responsabilidad y causación está expresada por la siguiente tesis:

MORAL IMPLICA CAUSAL: Si un agente es moralmente responsable por un hecho en el mundo, entonces alguna acción u omisión del agente causó ese hecho en el mundo.

1. Estoy suponiendo que tiene sentido hablar de responsabilidad moral por cosas que no dependen pura y exclusivamente de nosotros. Ver Nagel (1979), cap. 3, Williams (1981), cap. 2, y Thomson (1989). A menos que aceptemos que podemos ser moralmente responsables por cosas que no dependen pura y exclusivamente de nosotros, se seguiría que no podemos ser responsables por prácticamente nada. Para una defensa de esta posición, ver Walker (1991).

En lo que sigue argumento que cierto tipo de comportamiento colectivo es un contraejemplo a MORAL IMPLICA CAUSAL.

Consideremos el siguiente caso:

NEGLIGENCIA CONJUNTA: Homero y Mario son empleados en una planta nuclear. Hay una emergencia, una alarma suena y tanto Homero como Mario saben que cada uno de ellos debe apretar inmediatamente un botón determinado o una explosión ocurrirá: la explosión se evitará si y sólo si los dos botones se aprietan en el momento adecuado. Tanto Homero como Mario pueden fácilmente apretar sus respectivos botones. Sin embargo, están en distintas habitaciones y no tienen manera de saber qué hará el otro. Tanto Homero como Mario deciden no apretar sus botones y dejar que la explosión ocurra.

Argumentaré que NEGLIGENCIA CONJUNTA es un contraejemplo a MORAL IMPLICA CAUSAL. Es decir, argumentaré lo siguiente:

- 1) Homero es responsable por la explosión en NEGLIGENCIA CONJUNTA.
- 2) Ninguna acción u omisión de Homero causó la explosión en NEGLIGENCIA CONJUNTA.
- 3) Por lo tanto, MORAL IMPLICA CAUSAL es falsa.

En las dos secciones siguientes argumento a favor de mis premisas, 1) y 2).

3. La primera premisa: responsabilidad colectiva y responsabilidad individual

En esta sección argumento a favor de la primera premisa en mi argumento:

- 1) Homero es responsable por la explosión en NEGLIGENCIA CONJUNTA.

Recordemos que por "responsable" quiero decir "responsable individualmente". En lo que sigue argumento que Homero es responsable por la explosión a partir de la responsabilidad *colectiva* de Homero y Mario. Hay quienes creen que la responsabilidad colectiva de un grupo de agentes no puede analizarse de manera completa en términos de la responsabilidad individual de los miembros del grupo. En particular,

hay quienes creen que un grupo puede ser responsable por algo sin que ningún integrante del grupo sea responsable. Sin embargo, me propongo argumentar que, aun cuando esto sea cierto, NEGLIGENCIA CONJUNTA es un caso en el cual la responsabilidad del grupo claramente se distribuye entre los integrantes del grupo. Por consiguiente, dado que el grupo es responsable, Homero es responsable.

Antes que nada, veamos por qué el grupo formado por Homero y Mario es responsable por la explosión. Del mismo modo que la responsabilidad individual de agentes morales puede ser por acción u omisión, la responsabilidad colectiva de agentes morales en tanto forman un grupo determinado puede ser por acción u omisión. Un grupo es responsable por algo *por acción* cuando es responsable en virtud de alguna acción conjunta del grupo. Por ejemplo, un ejército puede ser responsable por la muerte de inocentes en virtud de haber atacado una cierta ciudad. Por el contrario, un grupo es responsable por algo *por omisión* en virtud de alguna omisión conjunta del grupo. Una omisión conjunta puede ser planeada con anticipación entre los miembros del grupo, como cuando un equipo de médicos decide no operar a un paciente después de haber discutido el asunto y haber llegado a la conclusión de que la operación sería muy costosa y el paciente no puede pagarla, o puede ser algo totalmente espontáneo, como cuando un grupo azaroso de personas permanece inmóvil ante el ataque de una pandilla a una anciana en la calle en pleno día. La responsabilidad colectiva de Homero y Mario en NEGLIGENCIA CONJUNTA es de este último tipo: Homero y Mario son colectivamente responsables por la explosión en virtud de su omisión conjunta no planeada. Parte de su trabajo en tanto empleados de la planta nuclear consistía en prevenir accidentes del tipo que no previnieron y que habría sido prevenido si cada uno de ellos hubiera apretado su botón. Es claro, entonces, que Homero y Mario son colectivamente responsables por la explosión.

Como dije, mi intención es sugerir que del hecho de que Homero y Mario son colectivamente responsables por la explosión se sigue que Homero es responsable por la explosión. Para mostrar que Homero es responsable alcanza con mostrar que *alguno* de los dos miembros del grupo, Homero o Mario, es responsable, es decir, que la responsabilidad del grupo implica la responsabilidad individual de alguno de los integrantes del grupo. Esto es así porque la situación es simétrica por estipulación: toda razón para pensar que Homero es responsable es también una razón para pensar que Mario es responsable, y viceversa.

La posición intuitiva sobre la relación entre responsabilidad colectiva y responsabilidad individual parece ser que, siempre que un grupo es responsable, algún integrante del grupo es individualmente responsable. Sin embargo, como ya dije, hay quienes creen que a veces hay responsabilidad colectiva sin responsabilidad individual de *ninguno* de los miembros del grupo. Argumentaré, entonces, que, aun cuando esto sea cierto, el caso particular que nos ocupa, NEGLIGENCIA COLECTIVA, es un caso en donde la responsabilidad colectiva está acompañada de responsabilidad individual.

Los casos en donde, según algunos filósofos, hay responsabilidad colectiva sin responsabilidad individual son del siguiente tipo. Un grupo ejercita una práctica injusta para ciertos individuos o ciertos sectores de la población. Sin embargo, todos los miembros actuales del grupo fueron criados desde su nacimiento de manera tal que ven a la práctica como algo natural, no como algo injusto. En esos casos, algunos filósofos sugieren que el grupo es responsable por la injusticia resultante pero ninguno de los miembros del grupo es responsable: sería demasiado pedir que alguno de ellos pudiera ver con claridad la injusticia de la práctica a la que han sido acostumbrados desde su nacimiento.

Por ejemplo, David Cooper propuso el siguiente caso.² Imaginemos un pueblo chico del oeste de Estados Unidos en el siglo XIX. Hace varios años los habitantes del pueblo formaron un comité cuyo objetivo es proteger los intereses de sus ciudadanos frente a los de los habitantes de otros pueblos. Entre otras cosas, el comité instituyó una práctica que contempla echar sin más del pueblo a cualquier visitante que algún ciudadano del pueblo considere amenazante. La práctica es injusta, pero está establecida desde hace años y los habitantes del pueblo la consideran natural. Es más, podemos imaginar que cualquier habitante que se rebelara en contra de la práctica podría enfrentar el rechazo de sus pares y de sus familiares por no estar dispuesto a ofrecerles la protección que ellos requieren. En este caso, cree Cooper, el pueblo es responsable por las injusticias a las que esta práctica da lugar, pero ninguno de sus integrantes es individualmente responsable.

¿Tiene razón Cooper? ¿Es éste un caso de responsabilidad colectiva sin responsabilidad individual? No me parece tan claro, porque no me parece tan claro que sea un caso de responsabilidad colectiva. Es claro

2. Cooper (1972), pp. 88-89.

que existe una injusticia, y que la práctica instituida en el pueblo es la que causó esa injusticia, pero de ahí no se sigue que el pueblo es moralmente responsable. Al menos en el caso de la responsabilidad individual, muchas veces somos causalmente responsables por cosas por las que no somos moralmente responsables. En otras palabras, causar algo no implica ser moralmente responsable por ello (si bien la conversa es generalmente considerada cierta, y es la tesis que me propongo atacar aquí). Si alguien conecta mi cuerpo a un detonador, y yo sin saberlo activo el detonador al moverme, lo cual causa una explosión, soy una causa de la explosión pero no soy moralmente responsable por ella. Si causación no implica responsabilidad individual, probablemente tampoco implique responsabilidad colectiva. En otras palabras, no veo por qué tendríamos que sentirnos inclinados a aceptar que el pueblo que Cooper imagina es *moralmente* responsable por las injusticias causadas.³

Dados mis objetivos en este trabajo, sin embargo, no necesito argumentar que *no puede* haber responsabilidad colectiva sin responsabilidad individual. Me alcanza con mostrar que NEGLIGENCIA COLECTIVA no es el tipo de caso en donde puede haber responsabilidad colectiva sin responsabilidad individual. Y esto es bastante fácil de ver. Una característica fundamental del ejemplo de Cooper es que el comportamiento de cada uno de los integrantes del grupo es *excusable*. No podemos hacer responsables a los habitantes del pueblo individualmente porque cada uno de ellos tiene una buena excusa (una excusa *exculpatoria*) para comportarse como se comporta. En particular, cada uno de ellos percibe la práctica injusta como algo totalmente natural y por lo tanto no ve la injusticia en ella. Además, aunque le surgieran dudas, si intentara rebelarse en contra de la práctica, sus pares y sus familiares probablemente le tendrían rencor; por lo tanto, rebelarse en contra de la práctica es supererogatorio.

Pero, claramente, NEGLIGENCIA COLECTIVA *no* es un caso de este tipo. En NEGLIGENCIA COLECTIVA, ni Homero ni Mario tienen una buena excusa para comportarse como se comportan. Tanto Homero como Mario sabían lo que hacían al no apretar sus botones, y nada malo les habría ocurrido si hubieran apretado sus botones. Por lo tanto, no hay ninguna razón para pensar que (al menos en un caso de este ti-

3. Benjamin critica la posición de Cooper de esta manera en Benjamin (1976), pp. 103-104.

po) la responsabilidad del grupo no se distribuye entre los miembros del grupo. Dado que, además, en este caso, toda razón para considerar responsable a uno de los integrantes es también una razón para considerar responsable al otro, se sigue que, en particular, Homero es responsable. En otras palabras, 1) es verdadera.

4. La segunda premisa: causación y sobredeterminación

En esta sección argumento a favor de la segunda premisa en mi argumento:

2) Ninguna acción u omisión de Homero causó la explosión en NEGLIGENCIA CONJUNTA.

Es claro que ninguna otra acción u omisión de Homero *fuera de* la omisión consistente en no haber apretado su botón causó la explosión. Lo único que necesito mostrar, entonces, es que la omisión de Homero consistente en no haber apretado su botón no causó la explosión. Si logro mostrar esto, se sigue que 2) es verdadera.

Voy a empezar por mostrar que un argumento particular (y que puede fácilmente ser considerado persuasivo) en contra de mi posición es falaz. Esto me permitirá descartar una posible motivación para rechazar 2), y abrir el camino a mi argumento positivo a favor de 2), que seguirá a continuación.

Alguien podría pensar que 2) es falsa (esto es, que la omisión de Homero causó la explosión) por la siguiente razón: es claro que el hecho de que los botones no fueron (simultáneamente) apretados causó la explosión, y de ahí se sigue que, en particular, el hecho de que Homero no apretó su botón causó la explosión. Sin embargo, este argumento es falaz. Estoy de acuerdo en que el hecho de que los dos botones no fueron simultáneamente apretados causó la explosión. Después de todo, si los dos botones *hubieran* sido simultáneamente apretados, la explosión *no* habría ocurrido. Pero de ahí no se sigue que la omisión de Homero causó la explosión (nótese, en particular, que *no* es cierto que, si Homero hubiera apretado su botón, la explosión no habría ocurrido: dado que Mario tampoco apretó su botón, la explosión habría ocurrido igual).⁴

4. No pretendo sugerir que ciertas conexiones contrafácticas son ni suficien-

¿Por qué alguien podría pensar que habría de seguirse que la omisión de Homero causó la explosión? Pensemos cuál es la relación entre el hecho de que los dos botones no fueron simultáneamente apretados y la omisión de Homero. La relación es la siguiente: el hecho de que Homero no apretó su botón *implica* el hecho de que los dos botones no fueron simultáneamente apretados. (Dado que Homero no apretó su botón, no importa qué es lo que ocurrió con el otro botón, se sigue inmediatamente que los dos botones no fueron simultáneamente apretados.) De modo que lo que podría estar pensando quien sostuviera que 2) es falsa es lo siguiente:

Principio de implicación y causación: Si C causa E, y D implica C,⁵ entonces D causa E.

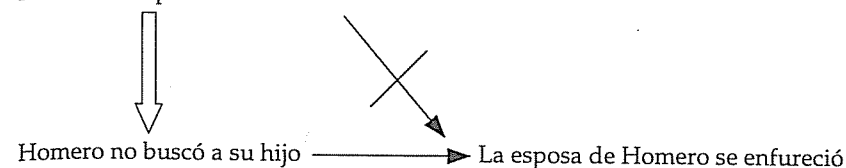
Si este principio fuera cierto, dado que el hecho de que los dos botones no fueron apretados causó la explosión, y dado que la omisión de Homero implica el hecho de que los dos botones no fueron apretados, se seguiría que la omisión de Homero causó la explosión.

Sin embargo, este principio es falso. Supongamos que Homero prometió pasar a buscar a su hijo por la escuela a las 5 de la tarde. A la salida del trabajo, Homero se olvida de su promesa y camina de vuelta a su casa. En el camino, cuando son exactamente las 5 de la tarde, compra rosas para su esposa. La esposa de Homero se enfurece al verlo volver solo a su casa. ¿Qué causó que la esposa de Homero se enfureciera? Claramente, fue el hecho de que Homero no buscó a su hijo por la escuela, como había prometido. La causa de que su esposa se enfureciera no fue que (en lugar de buscar al hijo por la escuela) Homero compró rosas para su mujer. Sin embargo, el hecho de que Homero compró rosas para su esposa a las 5 de la tarde implica que no buscó a su hijo por la escuela a las 5 de la tarde: si estaba comprando rosas, no podría haber estado buscando a su hijo al mismo tiempo. Podemos representar la situación en el siguiente diagrama, en donde la flecha doble es la relación de implicación, y la flecha simple es la relación de causación:

tes ni necesarias para la existencia de causación. Al igual que mucha gente, tomo la existencia (inexistencia) de relaciones contrafácticas como una razón *prima facie* para pensar que hay (no hay) causación.

5. "D implica C" (en donde D y C son hechos) puede entenderse como "En todo mundo posible en donde D ocurre, C también ocurre".

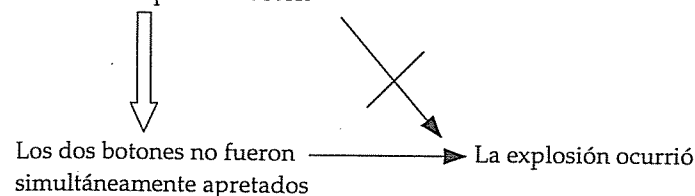
Homero compró rosas



Este ejemplo muestra que el Principio de implicación y causación es falso: es posible que C cause E, D implique C, pero D *no* cause E.

Mi sugerencia es que esto es precisamente lo que pasa en NEGLIGENCIA CONJUNTA: el hecho de que los dos botones no fueron apretados causó la explosión, la omisión de Homero implica que los dos botones no fueron apretados, pero la omisión de Homero no causó la explosión. En un diagrama, mi propuesta es que la situación en NEGLIGENCIA CONJUNTA es la siguiente:

Homero no apretó su botón



Hasta aquí he refutado un argumento en contra de mi premisa 2). Ahora paso a dar razones *positivas* a favor de 2). ¿Qué razones positivas hay para pensar que la omisión de Homero no causó la explosión? En lo que sigue argumento que hay por lo menos las siguientes dos razones: a) si dijéramos que la omisión de Homero causó la explosión, se seguiría que hay mucha más *causación* que la que parece haber, y b) se seguiría que hay mucha más *sobredeterminación* que la que parece haber.

Veamos a) primero. Supongamos que decimos que la omisión de Homero causó la explosión. Notemos nuevamente que, aunque Homero hubiera apretado su botón, la explosión habría ocurrido, ya que los dos botones debían ser apretados para prevenir la explosión. Lo que es cierto es lo siguiente: si Homero hubiera apretado su botón, y si Mario también hubiera apretado el suyo, entonces la explosión no habría ocurrido. Ahora bien, el hecho de que otro ser humano, Mario, se encontraba a cargo del otro botón claramente no puede hacer ninguna diferencia a

los poderes causales de Homero. Los únicos hechos que involucran al otro botón y que pueden afectar los poderes causales de Homero son hechos acerca de si el otro botón fue apretado y hechos acerca de cómo están relacionados los botones con la explosión. Que el otro botón estuviera siendo controlado por un ser consciente o por un ser no consciente no puede hacer diferencia alguna a los poderes causales de Homero.

Pero entonces imaginemos que Mario no se encontraba en la otra habitación sino que, en su lugar, había un dispositivo automático listo para accionarse en el momento adecuado. Sin embargo, en el instante preciso en que el botón va a accionarse, una aplanadora fuera de control destruye por completo la habitación en donde se encontraba el botón. Homero, por su parte, no aprieta su botón, y la explosión ocurre. ¿Es la omisión de Homero una causa de la explosión en este caso? Parece que no. Pero, si dijéramos que la omisión de Homero es una causa de la explosión en NEGLIGENCIA CONJUNTA, también deberíamos decir que lo es aquí. Después de todo, si Homero hubiera apretado su botón, *y si el otro botón no hubiera sido destruido*, la explosión no habría ocurrido.⁶

O consideremos el siguiente caso. Un doctor examina a un paciente y descubre que el paciente tiene un tumor. Si el tumor no se opera inmediatamente, el paciente vivirá sólo un par de días. Tumores de ese tipo son operables con la tecnología adecuada. Sin embargo, el hospital en donde está internado el doctor no dispone de la tecnología necesaria. Además, el paciente es alérgico a la anestesia, con lo cual cualquier operación lo mataría también al cabo de un par de días. El doctor decide no abrir al paciente y el paciente muere. ¿Es la omisión del doctor una causa de la muerte del paciente? Claramente no: abrir al paciente no hubiera servido de nada.⁷ Pero, si rechazáramos 2) y dijéramos que

6. Alguien podría pensar que hay una diferencia importante entre NEGLIGENCIA CONJUNTA y este caso, a saber: la presencia de un ser humano en NEGLIGENCIA CONJUNTA da lugar a una cierta *indeterminación* que está ausente en el caso de la aplanadora. Sin embargo, esta diferencia no parece ser pertinente. Porque podemos imaginar que la aplanadora estaba controlada por un mecanismo indeterminista. Aun así, lo único que parece importar es que lo que pasó *de hecho* es que la aplanadora destruyó el otro botón, no qué tan probable era que lo hiciera. Dado que *de hecho* el botón fue destruido en el momento preciso, la omisión de Homero no parece ser una causa de la explosión.

7. ¿Sabía el doctor o no de la alergia del paciente y del hecho de que el hos-

la omisión de Homero causó la explosión en NEGLIGENCIA CONJUNTA, también deberíamos decir que la omisión del doctor causó la muerte del paciente. Después de todo, si el doctor hubiera abierto al paciente, *y si el hospital hubiera tenido la tecnología adecuada, y si el paciente no hubiera sido alérgico a la anestesia*, el paciente se habría salvado.

Y así podemos obtener resultados cada vez más y más implausibles. Supongamos que, en vez de *dos* botones, *cient* botones deben ser apretados para prevenir la explosión. Homero está a cargo de uno, y los noventa y nueve restantes están programados para accionarse automáticamente. Sin embargo, en el momento en que van a ser apretados, un meteorito cae en la planta y destruye los noventa y nueve botones. Homero y su botón se salvan, Homero no aprieta su botón, e inmediatamente después ocurre la explosión. Es claro que la omisión de Homero no causó la explosión. Pero deberíamos decir que sí la causó si dijéramos que la omisión de Homero causó la explosión en NEGLIGENCIA CONJUNTA. Después de todo, si Homero hubiera apretado su botón, *y si los noventa y nueve botones restantes no hubieran sido destruidos*, la explosión no habría ocurrido.

Sugerí que debemos aceptar mi premisa 2) o de otro modo nos comprometeríamos con la existencia de mucha más causación que la que estamos preparados a aceptar. Pasemos ahora a b). En lo que sigue argumento que, si no aceptamos 2), no sólo nos comprometeríamos con mucha más causación que la que estamos preparados a aceptar, sino también con mucha más *sobredeterminación* que la que estamos preparados a aceptar. En primer lugar, introduzco el concepto de sobredeterminación. Luego discuto bajo qué condiciones y en qué medida podemos estar dispuestos a aceptar la existencia de sobredeterminación. Finalmente argumento que, si rechazáramos mi premisa 2), nos comprometeríamos con mucha más sobredeterminación que la que estamos dispuestos a aceptar.

La sobredeterminación es un tipo muy especial de causación. Hay quienes rechazan por completo la posibilidad de sobredeterminación. Pero, de manera más o menos general, se considera que la sobredetermina-

pital no tenía la tecnología necesaria? Estoy dejándolo abierto intencionalmente. Dado que la relación de causación es una relación metafísica, no epistémica, las creencias del doctor no son pertinentes a sus poderes causales. Si la omisión del doctor es una causa de la muerte del paciente, lo es independientemente de cuáles sean las creencias del doctor.

ción es un fenómeno *poco común* (si bien posible).⁸ Para los que aceptan la posibilidad de sobredeterminación, un ejemplo paradigmático de sobredeterminación es el siguiente: Homero y Mario juegan a tirar piedras a una botella. Homero tira una piedra y Mario tira otra piedra, las dos piedras impactan la botella *simultáneamente* y la botella se rompe. Además, cualquiera de las piedras por sí sola hubiera sido suficiente para que la botella se rompiera. En ese caso, las dos piedras *sobredeterminan* el hecho de que la botella se rompiera. Podemos trazar dos cadenas causales que conducen a la ruptura de la botella: la cadena iniciada por Homero al tirar su piedra y la cadena iniciada por Mario al tirar la suya. Cualquiera de esas cadenas causales hubiera alcanzado para explicar por qué se rompió la botella. Además, ninguna de las piedras desvió a la otra, y las dos piedras tocaron la botella a exactamente el mismo tiempo (tenemos que suponer esto; de lo contrario la piedra que llegó antes sería la causa de que la botella se rompiera, mientras que la que llegó después no).

¿Por qué tiende a creerse que la sobredeterminación es un fenómeno poco común? Una razón es que, cuando uno trata de imaginar un caso de sobredeterminación, se da cuenta de que varias condiciones improbables tienen que darse para que sea posible (en nuestro ejemplo, las dos piedras tienen que tocar la botella en exactamente el mismo instante). Otra razón, más teórica, es que la sobredeterminación da lugar a *redundancia* en la explicación de los fenómenos que ocurren en el mundo. Cuando queremos explicar por qué la botella se rompió, hay dos explicaciones alternativas y autosuficientes de este hecho. No hay ninguna razón por la cual debiéramos elegir una explicación en lugar de la otra. Pero, al mismo tiempo, parece que, dada cualquiera de las explicaciones por sí sola, la otra está de más. Es decir, por un lado, no hay ninguna razón fundada para preferir una explicación a la otra pero, por otro lado, las dos explicaciones parecen competir por su poder explicativo. En consecuencia, si bien podemos llegar a aceptar la existencia de casos contados y únicos de este tipo, no podemos aceptar que sea algo que ocurre comunmente.

En particular, Jaegwon Kim ha argumentado en contra de ciertas

8. Bunzl argumenta que no hay sobredeterminación genuina en Bunzl (1979). Por su parte, David Lewis acepta la posibilidad de sobredeterminación en Lewis (1986), pp. 208-212, pero considera que es un fenómeno lo suficientemente raro como para que nuestras intuiciones causales se vuelvan borrosas.

teorías acerca de la naturaleza de la mente sobre la base de que la sobredeterminación no es un fenómeno común.⁹ Kim muestra que, si ciertas teorías acerca de la mente fueran ciertas, entonces se seguiría que hay mucha sobredeterminación. (En particular, Kim muestra que se seguiría que, siempre que un hecho mental causa algo, el efecto del hecho mental está sobredeterminado por el hecho mental y el hecho físico en el cual el hecho mental sobreviene.) Pero, sigue el argumento, es implausible pensar que puede haber tanta sobredeterminación. Por lo tanto, aquellas teorías sobre la naturaleza de la mente deben ser rechazadas.

Mi argumento a favor de 2) es similar. Es el siguiente: rechazar 2) tiene la consecuencia de que habría mucha sobredeterminación. Pero es implausible pensar que pueda haber tanta sobredeterminación. Por lo tanto, debemos aceptar 2). En lo que sigue muestro por qué rechazar 2) tiene la consecuencia de que habría mucha sobredeterminación.

Si rechazamos 2), decimos que la omisión de Homero causó la explosión en NEGLIGENCIA CONJUNTA. Pero si la omisión de Homero la causó, entonces también la causó la omisión de Mario. Esto es así porque la situación es completamente simétrica en lo que respecta a Homero y a Mario: ambos omitieron apretar los botones que, de ser apretados conjuntamente, habrían prevenido la explosión. De modo que, si rechazáramos 2), tendríamos que decir que tanto la omisión de Homero como la omisión de Mario fueron causas de la explosión. Ahora bien, dado que la explosión sólo hubiera sido prevenida si los *dos* botones hubieran sido apretados simultáneamente, cada una de estas omisiones (la omisión de Homero y la omisión de Mario) habría sido suficiente por sí sola para que la explosión ocurriera. La omisión de Homero sin la omisión de Mario (es decir, aun cuando Mario hubiera apretado su botón) habría alcanzado para que la explosión ocurriera. Y, paralelamente, la omisión de Mario sin la omisión de Homero (es decir, aun cuando Homero hubiera apretado su botón) habría alcanzado para que la explosión ocurriera.

De modo que, si rechazáramos 2), deberíamos decir que la explosión tuvo dos causas cada una de las cuales era por sí sola suficiente para que ocurriera la explosión. En otras palabras, deberíamos decir que la omisión de Homero y la omisión de Mario son causas que *sobredetermini-*

9. Kim (1998), pp. 44-45. El argumento de Kim ataca a teorías según las cuales los hechos mentales son causalmente eficaces pero no son reducibles a hechos físicos, si bien sobrevienen sobre hechos físicos.

nan la explosión. Este resultado no sería tan grave si fuera un resultado aislado, pero el problema es que puede generalizarse fácilmente, como paso a mostrar ahora.

Recordemos el ejemplo de Homero y los noventa y nueve botones restantes: Homero no aprieta su botón, pero los noventa y nueve botones restantes son destruidos por un meteorito en el momento en que iban a ser automáticamente apretados. Dije que rechazar 2) tiene la consecuencia de que la omisión de Homero es una causa de la explosión, aun cuando los noventa y nueve botones restantes no fueron apretados y, por lo tanto, aun cuando el que Homero hubiera apretado su botón no hubiera ayudado a prevenir la explosión. Ahora bien, si la omisión de Homero fue una causa de la explosión en estas circunstancias, también lo fue el hecho de que el botón número dos no fue apretado en el momento adecuado, el hecho de que el botón número tres no fue apretado en el momento adecuado, y así con cada uno de los noventa y nueve botones. Pero cada uno de estos hechos hubiera alcanzado por sí solo para que la explosión ocurriera. De modo que, si rechazáramos 2), deberíamos decir que la explosión en el caso de los cien botones está sobredeterminada, y está sobredeterminada por *cien* causas diferentes. Pero esto es altamente implausible.¹⁰

Argumenté que rechazar 2) nos llevaría a decir que hay mucha más causación y, en particular, sobredeterminación, que la que es plausible pensar que hay. Mi conclusión es que debemos aceptar 2): la omisión de Homero *no causó* la explosión en NEGLIGENCIA CONJUNTA.

5. Conclusión

En resumen, propuse el argumento siguiente:

10. Quiero enfatizar que lo que es implausible acerca de este resultado no es que haya cien causas, sino que haya cien causas que son *sobredeterminantes*. Muchos fenómenos tienen múltiples causas, pero son causas que sólo son *en conjunto* suficientes para que ocurra el efecto. Por ejemplo, un incendio puede tener todas las siguientes causas: alguien tiró una colilla de cigarrillo todavía encendida al suelo, había gasolina derramada en el suelo, había oxígeno suficiente en el ambiente, no había un matafuego en toda la casa, los bomberos tardaron demasiado en llegar, etc. Todas estas condiciones son causas del incendio pero no lo sobredeterminan: cada una de ellas sin las restantes no hubiera alcanzado para que el incendio ocurriera.

- 1) Homero es responsable por la explosión en NEGLIGENCIA CONJUNTA.
- 2) Ninguna acción u omisión de Homero causó la explosión en NEGLIGENCIA CONJUNTA.
- 3) Por lo tanto, MORAL IMPLICA CAUSAL es falsa.

MORAL IMPLICA CAUSAL, recordemos, refleja la posición intuitiva sobre la relación entre responsabilidad y causación, según la cual no se puede ser responsable por un hecho en el mundo sin causarlo. El argumento que propuse muestra que la posición intuitiva sobre la relación entre responsabilidad y causación está equivocada: contrariamente a lo que la intuición indica, *es posible ser responsable por algo sin causarlo*.

Mi argumento se basó en un tipo especial de comportamiento colectivo (y, en particular, de comportamiento colectivo por omisión). ¿Qué conclusiones debemos sacar acerca de la importancia del comportamiento colectivo para nociones como la responsabilidad moral? Mi trabajo se aparta de la mayoría de la bibliografía existente sobre comportamiento colectivo y responsabilidad moral en que no pretende esclarecer la disputa sobre si la responsabilidad colectiva es reducible a la responsabilidad individual. (Algo dije sobre este tema, en la sección 3 de este trabajo, pero no con el objetivo de esclarecer la relación entre responsabilidad colectiva y responsabilidad individual, sino con el objetivo de mostrar que hay responsabilidad individual en NEGLIGENCIA CONJUNTA, el caso sobre el que trata mi argumento.) Lo que intenté mostrar con mi argumento es que el comportamiento colectivo es importante, no para el proyecto de esclarecer ciertos conceptos morales (el concepto de responsabilidad moral en general, y el de responsabilidad colectiva e individual en particular), sino para el proyecto de esclarecer la relación entre un concepto moral (responsabilidad) y un concepto metafísico (causación).

Mi conclusión final es la siguiente. Sólo podemos ser responsables por hechos en el mundo si estamos conectados de alguna manera con esos hechos; mi argumento basado en ciertos tipos de comportamiento colectivo no niega esto. Lo que mi argumento niega es que la *única* manera de estar conectados con hechos en el mundo (hechos por los que podemos ser responsables) es a través de los poderes causales de nuestras acciones y omisiones individuales. Lo que mi argumento muestra, en su lugar, es que podemos estar conectados con hechos en el mundo (hechos por los

que podemos ser responsables) a través de ciertos *comportamientos colectivos* de los cuales nuestras acciones y omisiones son parte.

Massachusetts Institute of Technology

BIBLIOGRAFÍA

- BENJAMIN, Martin (1976), "Can Moral Responsibility be Collective and Nondistributive?", *Social Theory and Practice* 4.
- BUNZL, Martin (1979), "Causal Overdetermination", *Journal of Philosophy*, 134-150.
- COOPER, David (1972), "Responsibility and the System", en Peter A. French, ed., *Individual and Collective Responsibility*, Cambridge, MA, Schenkman.
- KIM, Jaegwon (1998), *Mind in a Physical World*, Cambridge, The M.I.T. Press.
- LEWIS, David (1986), "Causation", en *Philosophical Papers*, Oxford, Oxford University Press.
- NAGEL, Thomas (1979), *Mortal Questions*, Cambridge, Cambridge University Press.
- THOMSON, Judith Jarvis (1989), "Morality and Bad Luck", *Metaphilosophy* 20, 203-221.
- WALKER, Margaret (1991), "Moral Luck and the Virtues of Impure Agency", *Metaphilosophy* 22, 14-27.
- WILLIAMS, Bernard (1981), *Moral Luck*, Cambridge, Cambridge University Press.

ABSTRACT

In this paper I argue that the existence of certain types of collective behavior has important consequences for the relation between moral responsibility and causation. In particular, I argue that those types of collective behavior undermine the intuitive view about how individual moral responsibility is tied to causation. My argument contains a discussion of the conditions under which the responsibility of a group distributes among the members of the group, as well as a discussion of the conditions under which an agent causes, and causally overdetermines, an outcome.

EL TEDIO Y EL TERROR: LA CRISIS DE 1929 EN EL CURSO DE HEIDEGGER LOS CONCEPTOS FUNDAMENTALES DE LA METAFÍSICA

Luis Alejandro Rossi

Introducción

En el curso del semestre de invierno de 1929, *Los conceptos fundamentales de la metafísica*¹ Heidegger se plantea explícitamente los problemas de la filosofía de la cultura y, presentándolo como la nota característica del ser-ahí contemporáneo, realiza un juicio general sobre la situación del momento, a diferencia de otras obras donde las referencias epocales no se presentan de forma manifiesta. El hecho de que Heidegger se ocupe de esta problemática puede parecer paradójico, pues apenas dos años antes, en *Ser y Tiempo*, rechazaba las pretensiones teóricas de la filosofía de la cultura.² Cabe preguntarse por qué Heidegger accede a discutir los diversos diagnósticos de la época, ya que no sólo intenta refutarlos, sino que, además, les niega cualquier tipo de rigor filosófico. Creemos que para comprender las razones de este giro se debe atender primeramente a las fechas en que se desarrolló el curso. El semestre de invierno es el que corresponde a la segunda mitad de 1929, es decir, es coincidente con la crisis económica mundial que se desencadena a fines de octubre de 1929.³ En segundo lugar hay una razón biográfica. Otto

1. Martin Heidegger, *Die Grundbegriffe der Metaphysik. Welt – Endlichkeit – Einsamkeit*, (Martin Heideggers Gesamtausgabe Band 29/30), Frankfurt, Vittorio Klostermann, 1992 (2ª ed.), en adelante citaremos GM y el número de página.

2. Al respecto declaraba Heidegger en *Ser y Tiempo*: «[...] la exégesis tiene una mira puramente ontológica y está muy lejos de ser una crítica del "ser ahí" cotidiano bajo un punto de vista moralizador y de tener aspiraciones propias de una "filosofía de la cultura"», M. Heidegger, *El ser y el tiempo*, trad de José Gaos, Buenos Aires, F.C.E., 1981, p. 186 (*Sein und Zeit*, Frankfurt, Vittorio Klostermann 1977, p. 222).

3. Recordemos que ella tuvo una rápida repercusión en Alemania, que era el

Pöggeler sostiene que los efectos de la crisis económica y su reflejo en la política alemana en el invierno de 1929/1930 coincidieron con una crisis de las convicciones metafísicas y religiosas del propio Heidegger. Ello lo llevó a enlazar las preguntas acerca de la posibilidad de la filosofía con la del papel de la universidad en la sociedad actual.⁴ En tercer lugar debe atenderse a la expansión y la popularidad que habían alcanzado los diagnósticos de la crisis civilizatoria que se venían repitiendo en Alemania desde la derrota en la Primera Guerra Mundial. Por último, el auge de interpretaciones de los tiempos como las corrientes en la filosofía de la cultura son, a juicio de Heidegger, un síntoma del temple de ánimo predominante en el momento y busca sacarlo a la luz, de forma que quede manifiesto en su plenitud aquello que la filosofía de la cultura es incapaz de comprender y, por ello, oculta. Demostrar que su filosofía puede responder a las cuestiones que subyacen a esta discusión a través del replanteo de los problemas metafísicos, en especial a partir de la cuestión de los temples de ánimo, interesa vivamente a Heidegger y un propósito del curso es descartar la idea misma de que se pueda dar un «diagnóstico» de la época (y la «prognosis» correspondiente) para proponerse, por el contrario, volver manifiesto lo que juzga como el fundamento metafísico de la situación. Heidegger rechaza la posibilidad de «diagnosticar» la época; este rechazo nos muestra, a partir de este punto de su trayectoria intelectual, una acentuación de los rasgos ontohistóricos de su metafísica. La cuestión de la historia (aquello que más tarde tematizará como «historia del ser») es abordada cada vez menos desde la teoría de la historicidad de *Ser y Tiempo*, ya que el marco de la analítica existencial radicaliza la «singularidad» del ser-

país europeo con mayor porcentaje de inversiones norteamericanas. El alza de la desocupación fue fulminante y a partir de este momento el nazismo deja de ser un partido local bávaro para convertirse en un partido de orden nacional. En mayo de 1928 tenía sólo doce bancas en el *Reichstag*, luego de las elecciones de septiembre de 1930 su bancada tendrá 107 diputados y se convierte en la segunda fuerza parlamentaria.

4. Según este intérprete, esa línea de pesquisa culminará en el discurso del rectorado en 1933, en el que ambas cuestiones aparecen tratadas. Cfr. Otto Pöggeler, «Heideggers politisches Selbstverständnis», en Annemarie Gethmann-Siefert und Otto Pöggeler (Hrsg.), *Heidegger und die praktische Philosophie*, Frankfurt, Suhrkamp, 1989, 2ª ed., p. 22.

ahí. Esto significa que si nos atenemos a su desarrollo cronológico, la «vuelta» (*Kehre*) del pensamiento de Heidegger no debe explicarse a partir de la pregunta por la verdad del ser (como se hace usualmente, tomando como referencia el término del proceso), sino que, a la inversa, el propósito heideggeriano de dar cuenta del temple de ánimo predominante en el momento contemporáneo lo lleva a abandonar paulatinamente el marco de la ontología fundamental y su reflexión se dirige hacia un preguntar por los temples como formas de apertura del mundo que «acaecen», es decir, se suceden sin que pueda encontrarse una lógica que explique la sucesión de uno a otro. No obstante, cabe aclarar que este proceso comienza a desenvolverse a partir de esta obra, sin que en ella se planteen todas estas cuestiones explícitamente.

Heidegger considera que la reflexión a partir de los temples de ánimo puede evitar la deficiencia principal de los «diagnósticos» de la filosofía de la cultura: la elevación unilateral de algún tópico cultural del momento a explicación del conjunto de la realidad. Eludiría ese riesgo al partir de los temples de ánimo, ya que, de acuerdo a su metafísica, ellos se fundamentan en el suelo mismo de la facticidad del ser-ahí y posibilitan su historicidad. La pregunta por el temple de ánimo predominante le permitirá traspasar los límites de la ontología del ser-ahí, posibilitándole desplazarse hacia la dimensión colectiva del ser-ahí en forma más explícita que en *Ser y Tiempo*. Como es común en su argumentación, su estrategia es deconstructiva, esto es, examinará diversos ejemplos de filosofías y caracterizaciones de la época contemporánea para descubrir en ellos el fundamento común que los hace posibles. La respuesta de Heidegger es que ese fundamento común es la filosofía de Nietzsche. La postulación de la «necesidad» de confrontar la filosofía actual con la obra de Nietzsche, que Heidegger señala en este curso por primera vez, es el inicio del camino en el que abandonará progresivamente la problemática fenomenológico-existencial (esto es, centrada en temas de las filosofías de Husserl y Kierkegaard). Esa tarea de confrontación con Nietzsche (y también con Hölderlin) revestirá especial importancia para la trayectoria intelectual de Heidegger en las décadas posteriores.

Teniendo en cuenta esta paulatina transformación tanto del contexto como de la propia problemática de Heidegger, nos interesa resaltar el carácter peculiar de este curso y mostrar que la problemática política en sus textos sigue un derrotero que si bien no puede considerarse rectilíneo, sí opera por acumulación. El juicio de Heidegger sobre su época

no es una digresión en un curso que se ocupa de otros problemas, sino que es contrapuesto a los diagnósticos que critica y, por tanto, forma parte esencial de sus conclusiones. Por todo ello es que sus reflexiones sobre la situación del ser-ahí contemporáneo están lejos de ser comentarios azarosos ligados a una inquietud pasajera, sino que profundizan la *Zeitkritik* ya presentada en *Ser y Tiempo* y modifican algunas de sus ideas principales, de forma tal que van ganando disponibilidad política. Con este término queremos designar el resultado del tratamiento de los fenómenos de la época por parte de su filosofía. Ello no produce meras digresiones en medio de obras «apolíticas», sino que implica un compromiso mayor de ella con su tiempo. Asimismo, le permite abrazar un compromiso político explícito a partir de 1933. Sería absurdo suponer que el juicio que Heidegger realiza en 1929 es nacionalista *per se*. Sin embargo, no puede pasarse por alto que ese mismo juicio contiene elementos que serán utilizados nuevamente en el discurso rectoral y que posibilitarán el compromiso de Heidegger con el nazismo. En 1929 todavía no hay una toma de posición política definida, pero sí un planteo totalizador de la situación vivida cuya respuesta se dará a partir del compromiso político rotundo de Heidegger.⁵ En suma, queremos indagar en este curso las líneas principales de la reflexión heideggeriana sobre la cultura contemporánea, en la que despuntan las cuestiones que los escritos posteriores intentarán responder.

El filosofar y el terror

El curso está dividido en dos partes, la primera, en la que se ocupa de la teoría de los templos de ánimo, analizando en detalle uno de ellos, el tedio profundo, y la segunda, que trata del concepto de «mundo». Tanto la primera como la segunda parte reelaboran temas de *Ser y Tiempo*. A través del análisis del tedio, Heidegger retoma las cuestiones referidas a la temporalidad y en la segunda parte el concepto de «mundo» de aquella obra es ampliado por comparación con la situación del

5. Al respecto, cfr. Winfried Franzen, «Die Sehnsucht nach Härte und Schwere. Über ein zum NS-Engagement disponierendes Motiv in Heideggers Vorlesung »Die Grundbegriffe der Metaphysik« von 1929/1930», en Annemarie Gethmann-Siefert y Otto Pöggeler, *Heidegger und die praktische Philosophie*, Frankfurt, Suhrkamp, 1989, 2ª ed., pp. 78-92.

animal. De este modo, el curso invierte la estructura de *Ser y Tiempo*. Su título, *Los conceptos fundamentales de la metafísica: mundo, finitud, soledad*,⁶ revela que la intención que mueve a Heidegger es establecer nuevamente cuáles son las cuestiones propias de la metafísica, descartando aquéllas de las cuales ella se ha ocupado hasta ahora, así como negar cualquier validez a concepciones para esta disciplina que tomen a la ciencia como modelo. La necesidad de plantear de modo radical la pregunta por el concepto «mundo» se debe a que la filosofía ha permanecido ciega frente a él, ella se ha, por así decirlo, «olvidado del mundo», no viendo más que los entes. Todo ello, sin embargo, no significa una novedad con respecto a *Ser y Tiempo*. Con todo, Heidegger se mueve en dirección a una radicalización de las conclusiones de esa obra, ya que la pregunta por los tres conceptos mencionados implica un replanteo completo de la metafísica como disciplina, pues el sentido que Heidegger les atribuye representa una inversión de las tres ideas kantianas: en efecto, podemos afirmar que hay una cierta correspondencia con ellas, ya que las ideas de Dios, alma y mundo, son reemplazadas por los conceptos «finitud», «individuación» y «mundo», el cual es entendido a partir de su exposición en *Ser y Tiempo*.

Nos ocuparemos primero de las cualidades que Heidegger atribuye a la actividad filosófica. Este tema, que puede parecer menor con respecto a los objetivos del curso, nos parece relevante porque la caracterización heideggeriana de la filosofía convierte a ésta en una disciplina que no puede ser considerada ni como contemplativa ni como práctica. Todo ello ya había sido dicho en *Ser y Tiempo*, sin embargo, esa concepción ahora es puesta al servicio de una radicalización de la crítica de la cultura contemporánea en forma más visible que en aquella obra. Su impugnación de la idea de filosofía como ciencia estricta significa el rechazo de la teoría como actitud fundante del ser-en-el-mundo del ser-ahí, por tanto, las cualidades propias del filosofar revelan la disposición básica que debe poseer aquél que emprende esa reflexión. Debemos acentuar el término «disposición», pues el rechazo de Heidegger al pri-

6. Según informa el editor del texto del curso, F.-W. von Hermann, asistentes al curso afirman que el subtítulo que figuraba en la pizarra, escrito por el propio Heidegger, era «Mundo, finitud, individuación», pero en la publicación se decidió mantener el término «soledad» pues era el que figuraba en el manuscrito, cfr., *Die Grundbegriffe der Metaphysik, op.cit.*, p. 537.

mado de la actitud teórica conlleva, en consecuencia, la preeminencia de la facticidad y del «encontrarse» como horizontes inescindibles de todo pensar. Todo ello resulta en la necesidad de la elucidación del temple de ánimo al que la filosofía obedece en tanto actividad específica y diferenciada.⁷ Recordemos que Heidegger atribuye a la ontología fundamental ser un discurso previo a cualquier división en «teoría» y «praxis». En este curso se repite la cuestión, pero de un modo más ambiguo que en su obra principal. Por un lado, para resolver este carácter peculiar del discurso filosófico, que debe ser previo a cualquier osificación del fluir de la facticidad en la teoría, Heidegger afirma que el carácter de los conceptos filosóficos es el de «indicaciones formales», por otro le atribuye un carácter último y extremo que lo vuelve «peligroso», esto es, donde no está presente esta peligrosidad, no hay filosofía.

La caracterización de los conceptos filosóficos como «indicaciones formales» era tratada por Heidegger desde los primeros años veinte. El sentido de «indicación formal» está emparentado con lo que en *Ser y Tiempo* denominaba como la «peculiar vacuidad del discurso ontológico». Con «indicación formal», Heidegger quiere expresar que los conceptos filosóficos no deben ser comprendidos como designaciones que hacen referencia a características o propiedades de algo «meramente presente» (*vorhanden*), sino que en tanto que indican, señalan un camino. El problema es determinar si la ontología fundamental posee un carácter prescriptivo oculto por detrás de la mera enunciación formal. A diferencia de *Ser y Tiempo*, donde no desarrolla el tema, en *Los conceptos fundamentales de la metafísica*, Heidegger lo explica, pero haciendo hincapié fundamentalmente en lo que implica la «indicación». A su juicio, que los conceptos filosóficos sean indicaciones formales obliga a quien los comprende a liberarse de las concepciones que tiene cotidianamente, de otro modo la comprensión no tiene lugar. Por ello podemos afirmar que la comprensión de los conceptos filosóficos es, en este sentido, una especie de conversión religiosa, ya que implica una transformación del propio ser:

7. Heidegger rechaza considerar a la filosofía una praxis, por eso, en lugar de «actividad», «disposición» expresa mejor el carácter fundamental de la filosofía. Heidegger, utilizando la expresión «el filosofar» (*das Philosophieren*) sustantiviza el verbo para poder salir de la dicotomía teoría-praxis.

En cada uno de esos conceptos —muerte, resolución, historia, existencia— yace la demanda de esta transformación, y por cierto no como una aplicación posterior, llamada ética, de lo comprendido, sino como una apertura previa de la dimensión de lo comprensible. (GM: 428-429)

Heidegger demanda aquí explícitamente aquello que en *Ser y Tiempo* subtiende al texto: la filosofía implica una apertura que rompe con la cotidianidad, por tanto, la comprensión exigida por la filosofía nunca podría ser equiparada con la mera teoría. Todos los conceptos que en *Ser y Tiempo* designa como «existenciales» son indicativos-formales. El carácter indicativo-formal reúne sobre sí ambos aspectos, teórico y práctico, tal como se deja ver en la expresión «apertura previa de la dimensión de lo comprensible». Este horizonte de apertura previa es su objeto de indagación. En tanto es previo a la dimensión de lo comprensible, es aquello sin lo cual lo comprensible no puede ser pensado. A diferencia de lo que ocurrirá en los textos posteriores, Heidegger insiste en el hecho de que la tematización de la apertura previa no puede estar desligada de la «resolución», de allí la asociación entre los cuatro términos que enumera en la cita. El carácter indicativo-formal implica que sacar a la luz esa apertura previa demanda una transformación del ser-ahí de acuerdo a su verdad, a la verdad que está en él en tanto ser-ahí y no como si se adoptara una concepción del ser humano entendido como «animal racional» u otras semejantes, en las que una propiedad pasa a convertirse en diferencia específica.

Pero además de encontrar en la filosofía este carácter indicativo-formal, Heidegger va más allá: afirma que la filosofía se caracteriza por su peligrosidad porque es última y extrema. Insiste en esta caracterización y la contrapone a la idea de filosofía como saber absoluto: ésta, afirma, tiende rápidamente a olvidar esta ambigüedad de la propia filosofía y por ello no permite la aparición de la instancia fundamental a partir de la cual se inicia el filosofar: los templos de ánimo que «abren» el mundo. Teniendo en cuenta, entonces, el carácter indicativo-formal de la filosofía y la transformación del ser-ahí a la que está ligada su comprensión, se vuelve más clara la afirmación de Heidegger que la verdad de la filosofía es la del ser-ahí. Esta misma concepción de la verdad era expuesta en el § 44 de *Ser y Tiempo*. Sin embargo, en esa obra, no aparecen ligadas, como sí ocurrirá posteriormente, la verdad y la libertad. En *Los conceptos fundamentales...*, Heidegger comienza a delinear la concepción de la esencia de la verdad como radicada en la libertad, aun cuando la

verdad como verdad del ser-ahí es entendida todavía en el sentido de la singularización de la que hablaba *Ser y Tiempo*. La peligrosidad de la filosofía, entonces, se debería al hecho de que sacude las concepciones predominantes en la cotidianidad: Heidegger afirma que la filosofía está en «lucha» con la cotidianidad. El filosofar debe provocar el «terror» (*Schrecken*):

En [el terror] se manifiesta precisamente algo esencial de todo comprender filosófico, que el concepto filosófico es un *ataque* al ser humano [...] El atacante no es el ser humano, el dudoso sujeto de la cotidianidad y de la felicidad del conocimiento, sino *el ser-ahí en el ser humano, que en el filosofar dirige el ataque al hombre*. (GM: 31, énfasis del original)

Heidegger se esfuerza por diferenciar su concepto de la filosofía de las concepciones del llamado «pesimismo cultural», dominante en el clima de ideas de la República de Weimar. Rechaza enfáticamente que pueda interpretarse su idea de la filosofía como una actividad desesperanzada, que se agota en sí misma, sombría, «pesimista». Observa que este pesimismo se ha convertido en la coincidencia de numerosos escritores y autores de las más variadas cosmovisiones (incluso reconoce que su propia obra podría ser interpretada de ese modo). Pero esta concepción «pesimista», no sólo no es nueva, afirma, sino que se deriva directamente de las convicciones del hombre promedio, «normal». En este punto se comprende mejor el carácter «peligroso» de la filosofía: nuestro autor afirma que se está frente a una disyuntiva, o se convierte al hombre «normal», al pequeño burgués con toda su mediocridad, en el juez de lo que es la filosofía o aquéllos que filosofan se resuelven a que, gracias al filosofar, el ser-ahí se convierta en lo que puede ser. La «resolución» expuesta en su obra principal encuentra así una ilustración que, aunque pueda ser entendida como «práctica», es, en el marco de la argumentación heideggeriana, la precondition de lo histórico. La ruptura con las convicciones predominantes no se refiere exclusivamente al conformismo generalizado que en *Ser y Tiempo* describía bajo el concepto del «uno», sino que debe dirigirse, en primer lugar, hacia las ideas que se presentan como «críticas» de lo existente. Frente a estas formas de «crítica» que son en realidad otra forma de conformismo, Heidegger insiste en la transformación que el filosofar despierta en los que tienen el valor de llevarlo a cabo. Citando a Platón, Heidegger afirma que

El ser humano que no filosofa, incluso el científico, ciertamente existen, pero duermen, sólo el ser-ahí filosofante está despierto y es algo completamente diferente, incomparablemente autónomo (GM: 34)

La filosofía se sitúa así como el punto de máxima divergencia con respecto a cualquier otro saber existente, el cuestionamiento más radical posible para el ser humano. A lo largo de este curso, Heidegger recurre permanentemente a la metáfora de que se debe «despertar» aquello que está dormido en nosotros. «Despertar» es lo contrario de una toma de conciencia, si por tal entendemos distanciarnos de lo observado para poder objetivarlo. «Despertar» un temple de ánimo significa comprender cómo vivimos en él, cómo se convierte en la condición sin la cual aquello que realizamos no puede explicarse. En este punto Heidegger vuelve a separarse de Husserl: rechaza explícitamente que sea posible aislar un temple de ánimo para considerarlo como un objeto en el flujo de la conciencia. Los temples de ánimo no son un estrato de conciencia entre otros. La filosofía debe preservar y mantener la inmediatez del ser-ahí cotidiano, por ello «despertar» los temples de ánimo requiere dejar libre nuestra perspectiva cotidiana de toda previa teoría de la conciencia.⁸ La autonomía del ser-ahí filosofante es incomparable porque él es el único que se «despierta» de las ilusiones de la objetividad y comprende la apertura previa que le es necesaria a aquélla. La peligrosidad de la filosofía consiste en este «despertar», que libera al ser-ahí al modo en que los presos en la caverna platónica eran «despertados» por la palabra del filósofo y gracias a ella podían descubrir su verdadera condición. No obstante, ello no llega a explicar suficientemente la «peligrosidad» de la filosofía. La frecuencia con que el tópico se repite en la obra no deja de ser llamativa; con todo, el resultado buscado es atribuir a la filosofía un carácter «antiburgués» (entendiendo por «burgués» un tópico cultural y no una clasificación sociológica) y discordante con respecto al mundo de las valoraciones cotidianas. La insistencia de Heidegger acerca del tema es una forma de presentar a la filosofía como enfrentada al mundo de la cotidianidad, esto es, del «uno». Frente a ambas formas de pensamiento objetivador, la ciencia—su forma más alta—y la comprensión vulgar, la filosofía «despierta» al peligro, al «terror», ya que significa asumir plenamente el poder-ser del

8. GM, p. 137.

ser-ahí, frente al «hombre promedio» y a la seguridad que acompaña su mundo. Es claro, además, que si la filosofía debe despertar al terror, es porque éste de algún modo se ha desvanecido de la existencia cotidiana. Esta afirmación parece traslucir algún tipo de nostalgia por el terror, lo que debe ser elucidado. Winfried Franzen, en el artículo antes citado, afirma que la crítica de Heidegger usualmente es malinterpretada como crítica a la superficialidad de la vida moderna, por el contrario, afirma, ella va más allá y se dirige contra «la normalidad de la vida normal»,⁹ de allí su carácter «antiburgués», ya que lo que se busca ante todo es invalidar la existencia *cultural* del burgués, su búsqueda de seguridad, sus valores universalistas y sus garantías. El enigma es determinar qué significa la ausencia del terror en el presente, cuál es su fundamento y por qué la seguridad a la que aspira la vida burguesa sería un problema al cual la filosofía debería replicar, todo lo cual nos lleva directamente a la reflexión de Heidegger sobre su época.

Los diagnósticos de la época en el pensamiento contemporáneo

En *Ser y Tiempo* la cuestión de la singularización del ser-ahí era radicalizada a través del ser-relativamente-a-la-muerte, el cual, al igual que el ser-ahí, es en-cada-caso-mío. En *Los conceptos fundamentales de la metafísica*, la cuestión de la singularización sigue estando presente, pero el planteamiento de la pregunta se desplaza paulatinamente desde el ser-ahí al «nosotros». Se percibe el carácter de transición del texto por el hecho de que Heidegger, como vimos que ocurría con el concepto de «indicación formal», todavía atribuye una importancia central a la transformación del ser-ahí a partir de la comprensión. Asimismo, el título original del curso mentaba el problema de la individuación, esto es, la cuestión de la individualidad y singularidad del ser-ahí. Sin embargo, tal como ocurría en *Ser y Tiempo* con la teoría de la historicidad, cuando Heidegger se ocupa de lo público, aborda la cuestión de la identidad a partir del colectivo «nosotros», que adquiere paulatinamente una importancia cada vez mayor. Determinar «quiénes somos nosotros» lleva a la pregunta acerca de cómo delimitar la situación del «nosotros». A diferencia de *Ser y Tiempo*, donde se postulaba el carácter

9. W. Franzen, *op. cit.*, p. 85.

inexorablemente situado de todo ser-ahí y, en consecuencia, se indagan las estructuras del ser-en-el-mundo del ser-ahí como tal, en este curso Heidegger avanza todavía más en la concreción de la pregunta, pues ya no inquires por la situabilidad del ser-ahí en sí, sino por su situación concreta: se pregunta por «nuestra» situación en este momento histórico determinado. Para ello recurre a los diagnósticos de cuatro célebres ensayistas del momento: Oswald Spengler, Ludwig Klages, Max Scheler y Leopold Ziegler. Heidegger no se detiene a realizar un análisis detallado de cada uno de ellos, sino que se pregunta cuál es el suelo común a esos cuatro diagnósticos. El valor real que les otorga queda puesto de manifiesto, cuando una vez expuestos, afirma:

Esto es sólo una referencia sumaria y formularia de lo que hoy se conoce, de lo cual se habla y de lo que en parte ya ha sido nuevamente olvidado, interpretaciones que en parte son tomadas de segunda o tercera mano y organizadas en una imagen de conjunto, imágenes que eventualmente penetran en el periodismo más alto de nuestra época y —si se puede hablar así— crean el ambiente espiritual en el que nos movemos. (GM: 106)

Por tanto, lo que se pone de manifiesto en las interpretaciones expuestas sumariamente es lo que Heidegger entiende que es el catálogo de los lugares comunes del momento. Toda discusión acerca de la situación histórica concreta debe desembarazarse de ellos filosóficamente, esto es, establecer cuál es el principio al que recurren y vulgarizan y, una vez hecho esto, pasar a la discusión con el origen mismo de esos diagnósticos. Heidegger repite que los planteos de la filosofía de la cultura «no nos alcanzan», es decir, no dan cuenta de nuestra situación. La perspectiva a partir de la cual las cuatro interpretaciones mencionadas analizan la época es la oposición entre «vida» y «espíritu». Esta oposición, sostiene, no expresa componentes o partes del ser humano, sino que ambos significan orientaciones fundamentales de él. En el marco de los autores analizados por Heidegger, dicha oposición presenta a la «vida» como lo instintivo, como el fluir indetenible e inasible y al «espíritu» como la razón que clasifica, rigidiza y ahoga lentamente la espontaneidad vital. Las cuatro interpretaciones difieren en cuanto a la prognosis que realizan, sin embargo, —con excepción de Scheler— coinciden en señalar la necesidad de reafirmar lo vital frente a las pretensiones del espíritu, que actúa como fuerza racionalizadora.

Heidegger interpreta esta pareja de opuestos como una traducción

popular de la oposición nietzscheana entre lo dionisiaco y lo apolíneo. No obstante, no está interesado en determinar si esos diagnósticos son verdaderos, o si interpretan a Nietzsche de manera adecuada. Su objetivo es mostrar el carácter estereotipado de la comprensión corriente de Nietzsche y para ello remite a algunos fragmentos póstumos, en los que este filósofo explica el significado de «apolíneo» y «dionisiaco». Heidegger observa que, por una parte, el mismo Nietzsche es consciente de la importancia de esta oposición para su propio pensamiento, y, por otra, que a lo largo de su trayectoria irá modificando su significado. De aquí que a Heidegger le interese atenerse a la forma que ella alcanza en los últimos años del filósofo. Esta oposición, sugiere, mantiene en Nietzsche una forma abierta que se pierde y se endurece en los diagnósticos mencionados. Mientras que en éstos, «vida» tiene un significado positivo y «espíritu» uno negativo, «dionisiaco» y «apolíneo» se mantienen para Heidegger como instancias abiertas de significado variable. A pesar de que reconoce que el mismo Nietzsche concibe su propia tarea como «dionisiaca», es decir, como una justificación de la vida, lo cual acercaría la oposición originaria a la de las valoraciones implícitas en la dicotomía «vida/espíritu», no pierde la ocasión de señalar además que Nietzsche opone «Dionisio» al «Crucificado». Por ello, a Heidegger le interesa enfatizar que, si bien en Nietzsche la antítesis se presenta partiendo de lo dionisiaco como justificación de la vida en todos sus aspectos –incluso en los más aterradores– y lo apolíneo como la demanda de lo que simplifica y busca claridad, la plenitud de poder y la moderación, el espíritu griego, no obstante, no se agota en la contraposición de ambos, sino que es la formación de lo apolíneo a partir de un sustrato cultural dionisiaco. Lo apolíneo es una conquista de los griegos luchando contra su propia naturaleza «asiática». La comprensión heideggeriana de la dicotomía nos la presenta contenida en una unidad más amplia: el espíritu griego, el cual es el producto, la conquista alcanzada por el desarrollo de ese antagonismo. La esclerosis de esta oposición en la filosofía de la cultura se descubre en el hecho de que ella ya sólo encuentra dos principios en lucha, uno positivo, el otro negativo, sin poder percibir que la relación entre ambos es más que una mera oposición. Esta pérdida de perspectiva con respecto al pensamiento nietzscheano lleva a que tampoco se perciba la última forma de la oposición, cuando a «Dionisio» se le opone el «Crucificado», ya que el Dios en la cruz es, a juicio de Nietzsche, equivalente a maldecir la vida, en lugar de justificarla. Heidegger remite al dinamismo de la oposición

nietzscheana originaria porque encuentra en ella la posibilidad de pensar la época a partir de las intuiciones implícitas en esos conceptos. El problema, por tanto, es establecer un esbozo de la situación contemporánea a partir de las sendas abiertas por Nietzsche y que, a su vez, evite perderse en el puro presente, como ocurre con la filosofía de la cultura, lo cual Heidegger realizará a partir de la afirmación de que el misterio y el terror se han desvanecido en la existencia contemporánea.

La angustia y el tedio profundo

En la primera parte de este curso Heidegger lleva a cabo el desarrollo más minucioso de su teoría de los templos de ánimo, en parte expuesta en *Ser y Tiempo*. Con todo, hay un cambio notable con respecto a esta obra. Allí Heidegger afirmaba que entre todos los templos de ánimo, la angustia era un templo «señalado», ya que revelaba el carácter de «arrojado» del ser-ahí y permitía el acceso a su ser-relativamente-a-la-muerte, lo que implica que su carácter «señalado» no está ligado a ninguna configuración epocal determinada, sino a la estructura misma de su ser-en-el-mundo. En *Los conceptos fundamentales...*, la angustia pierde este carácter exclusivo frente al tedio profundo. La importancia del cambio reside en que éste no se dirige a mostrar únicamente que el tedio podría compartir con la angustia su carácter de templo de ánimo señalado, sino en que Heidegger busca descubrir, a través del examen del tedio profundo, el carácter del ser-ahí contemporáneo. En «¿Qué es metafísica?», su conferencia inaugural en Friburgo, leída en julio de 1929 poco antes del inicio de este curso, remite a la experiencia del tedio profundo como reveladora de la nada, pese a ello sigue tratando el problema a través del primado de la angustia con respecto a los otros templos de ánimo. Tal como se presenta en esa conferencia, la argumentación heideggeriana podría hacer pensar en una extensión de los templos de ánimo relevantes para el carácter descubridor de la ontología, sin embargo, en *Los conceptos fundamentales de la metafísica* el interés principal de Heidegger en el tedio profundo está en su carácter descubridor con respecto a la propia situación del ser-ahí en la época contemporánea. Como la argumentación se realiza a partir de los mismos conceptos de la analítica existencial de *Ser y Tiempo*, la idea de que se trata de una extensión de la temática original parece plausible, pero en realidad ella es diferente ya que, si bien el tedio profundo se convierte en un acceso a la temporalidad, en la obra resulta ser tan o más revela-

dor acerca de la circunstancia que afecta al ser-ahí que sobre la estructura del ser-en-el-mundo en sí. La angustia y el tedio profundo tienen en común, en primer lugar, como es evidente, su carácter de temples de ánimo, lo que significa que sólo se dan como formas del «encontrarse» del ser-ahí. Asimismo distancian al ser-ahí de la cotidianidad, producen indiferencia donde antes había interés y ocupación con lo que rodea (el «estar caído» del ser-ahí). La «resolución» es posible sólo a partir de esta indiferencia y distancia y con ella la historicidad y el gestarse-con de un pueblo en comunidad. Sin embargo, también son perceptibles las diferencias entre ambos. El tedio profundo es un «encontrarse» en el que sólo hay indiferencia pues no hay proyecto, no hay comprensión ni tampoco hay sentido, es la pura facticidad del «encontrarse». En *Los conceptos fundamentales...* Heidegger afirma que debemos dejar de lado las formas superficiales del tedio para «despertar», para permitir que salga a la luz el tedio profundo. Lo mismo afirmaba acerca de la angustia en *Ser y Tiempo* y en «¿Qué es metafísica?» Pero el hecho de no retroceder ante la inminencia de ambos temples no significa que tengan una misma dirección. Claramente la angustia es en esas obras un temple de ánimo del ser-ahí, no una disposición de la época. El tedio profundo, al igual que la angustia, es evitado por los hombres, que se refugian en formas más superficiales de él, pero es claro, simultáneamente, que según Heidegger el tedio profundo es una característica del ser-ahí contemporáneo, es decir, de los tiempos que corren. La angustia aparece en *Ser y Tiempo* ligada a una característica «estructural» del ser humano como es la muerte. El tedio también permite acceder a la temporalidad y a la radical finitud del ser-ahí, pero a diferencia de lo que ocurría con la angustia, el tedio aparece como un temple de ánimo destacado por encontrar un reflejo en la misma situación contemporánea. En «¿Qué es metafísica?» Heidegger ironizaba acerca de la posibilidad de que todos vivieran angustiados, explicando el carácter fugaz e intransferible de la angustia y negándole así un rasgo colectivo.¹⁰ Por el contrario, el tedio profundo, aunque se lo confunda con otras formas más superficiales, parecería ser un sustrato que se manifiesta en la época misma. Aún más, los temples de ánimo no son estados subjetivos, no ocurren en un sujeto, ni son una «propiedad» suya. El ser-ahí se encuentra bajo un

10. Martin Heidegger, «Was ist Metaphysik?», en *Wegmarken*, Frankfurt, Vittorio Klostermann, 1976 p. 115.

temple determinado hasta que ese temple lo abandona, desplazado por otro. Esta naturaleza fundamental del «encontrarse» no puede explicarse bajo la relación de causalidad, ya que la sucesión de los temples entre sí no sigue ninguna regla. Elucidar el tedio profundo como nota característica del ser-ahí contemporáneo significa sacar a la luz la condición metafísica en la que vive. Si los temples de ánimo son caracterizados por Heidegger como esencialmente descubridores, lo que el tedio profundo descubre es la carencia de misterio del ser-ahí.

Al igual que el filósofo griego consideraba al asombro como la disposición de ánimo necesaria para adentrarse en la filosofía, en *Ser y Tiempo* la angustia lo es para el conocimiento del ser-ahí. Asimismo, en esta obra los temples de ánimo son elucidados a partir de la dicotomía «propiedad/impropiedad». La angustia abre el mundo en relación con la resolución, la cual posibilita el ser-propio. Los temples permiten comprender por qué el ser-en-el-mundo del ser-ahí nunca puede ser entendido como una relación *teórica*, sino como un «encontrarse» en el mundo que se abre. En su obra principal Heidegger no se plantea el problema de un temple de ánimo que fuera predominante en el momento contemporáneo, sino que los temples son comprendidos como uno de los elementos estructurales de la ontología del ser-ahí. Hay una cierta ambigüedad en la presentación del concepto en *Ser y Tiempo*, ya que a pesar de entender a los temples de ánimo como estados no subjetivos, sino que colorean la forma en que el ser-ahí abre el mundo, Heidegger no llega a afirmar que los temples sean públicos, sólo los explica de forma que se descarte cualquier cartesianismo a la hora de comprenderlos, es decir, que no se los entienda como estados de la conciencia. A partir de la obra que estamos analizando, transformará gradualmente su concepción original pues, si bien no abandona la noción sostenida en *Ser y Tiempo*, la búsqueda de un temple predominante en la época excede los marcos planteados por la ontología fundamental. Conceptualmente ello es posible a través de la relación entre el concepto de «temple» con el de ser-con. En los § 29 y 30 de esa obra, Heidegger no establece una relación explícita de los temples con el ser-con. Contraargumentar que el ser-con forma parte de la estructura del ser-en-el-mundo y que, por tanto, la relación está implícita en la elucidación que se hace de ella, pasaría por alto que Heidegger se detiene especialmente en la relación del «encontrarse» con el «estar arrojado» del ser-ahí, el cual implica el nudo *factum* de que el ser-ahí tiene que ser y desconoce tanto el «de dónde» como el «hacia dónde». Este énfasis en la relación

de los temples con el «estar arrojado» presenta al ser-ahí desde las condiciones que lo aíslan y lo separan de sus semejantes, no con las que se presentan como telón de fondo de una experiencia compartida. El ser-con también forma parte de la facticidad del ser-ahí, sin embargo, Heidegger no relaciona a los temples de ánimo con el ser-con-otro del ser-ahí. Por el contrario, en *Los conceptos fundamentales...*, el concepto de temple de ánimo está ligado de manera esencial al ser-con. Al descartar la posibilidad de considerar al temple de ánimo como un fenómeno meramente psicológico, Heidegger afirma que «el temple de ánimo no es un ente que aparece en el alma como experiencia, sino el cómo de nuestro ser-ahí-con-otro». Asimismo agrega: «es claro que los temples de ánimo no son algo que sólo es meramente presente (*vorhanden*), sino que ellos mismos son precisamente una forma fundamental y un modo fundamental del ser, y por cierto del ser-ahí (*Da-seins*), y ello siempre implica inmediatamente el ser-con-otro» (GM: 100-101).

Heidegger no se pregunta acerca de la época en forma directa. Su interrogación inquiriere por las características del ser-ahí contemporáneo y los rasgos que delinea son atribuidos a la época estableciéndolos como características del hombre contemporáneo. Por esta razón, la marca de la época que persigue Heidegger puede ser identificada con un *Zeitgeist* y, al convertirse los temples de ánimo en una traza del momento, reciben algo del significado de aquella noción hegeliana. Pero la dirección de la pregunta de Heidegger no coincide completamente con la concepción hegeliana: lo relevante del tedio profundo como temple de ánimo predominante en la época es que se oculta a quienes viven en él. El par oculto/desoculto es el eje sobre el cual Heidegger comprende el problema. El «estado de oculto» del temple es su característica más relevante, por ello sostiene que cuanto más se oculta un temple, más poderoso es. Si recordamos el párrafo citado en la página 261 queda más clara la diferencia entre el carácter público de un temple y el «ambiente espiritual» o «clima de ideas» de una época. Esta última noción remite a las «habladurías» de las que el «uno» se ocupaba, pero ellas no constituyen un temple, sino que él permanece oculto, esto es, ellas mismas lo ocultan. El carácter público del temple domina en la medida en que colorea las manifestaciones de una cultura sin dejarse ver por ello, sino que sólo una hermenéutica adecuada puede sacarlo a la luz y no confundirlo con otras formas más superficiales. En *Ser y Tiempo* la angustia no tiñe el modo de ser del ser-ahí contemporáneo, sino que ella permanece oculta como una posibilidad para los que no retroceden ante ella. En

Los conceptos fundamentales de la metafísica el tedio sí se presenta como el origen oculto de la situación en que «se dan las cosas» para el ser-ahí. De allí que en esta obra Heidegger aparece más interesado en indagar este carácter público de los temples, especialmente el del tedio, que en remitirlos a la estructura del ser-en-el-mundo. No obstante, ello no destruye la relación de los temples con el par «propiedad»/«impropiedad». La existencia impropia se deja llevar por las cosas, por ello no reconoce el tedio profundo, ni antes la angustia y las confunde con otras formas más superficiales del tedio y con el miedo, respectivamente. La angustia y el tedio profundo abren el mundo, porque permiten al ser-ahí alcanzar su verdadera situación, pero el lugar de cada uno de ellos en la estructura conceptual heideggeriana es diferente. Los desarrollos acerca de los temples de ánimo de *Los conceptos fundamentales...* son más ricos y detallados que los de *Ser y Tiempo*, con todo, la angustia no tiene la misma relación con la historia que el tedio profundo. Ella se presenta como aquel temple que vuelve posible al ser-ahí la comprensión de la posibilidad de su imposibilidad, mientras que la pérdida de sentido general que el ser-ahí descubre a partir del tedio profundo está conectada en forma mucho más directa con la facticidad histórica. Puede aducirse que la historicidad del ser-ahí no es posible sin la angustia, y que, por tanto, el gestarse-con de un pueblo, de acuerdo a como lo describía Heidegger en *Ser y Tiempo*, supone la angustia como temple de ánimo predominante. Sin embargo, no se plantea en la obra una relación entre los temples de ánimo y la historicidad. Los temples y el «encontrarse» sólo son analizados en una dimensión exclusivamente referida al ser-ahí, es decir, individual (o en términos de Heidegger, singular). Klaus Held sostiene que los temples de ánimo constituyen el punto de partida de la «resolución» y, por tanto, de la historicidad. Se puede aducir la siguiente afirmación en ese sentido: «[la] libertad del ser-ahí sólo es en la *propia liberación* (*Sichbefreien*) del ser-ahí. Sin embargo, la liberación del ser-ahí sólo acaece si cada vez él *se resuelve con respecto a sí mismo*, esto es, se resuelve por sí mismo como ser-ahí». ¹¹ En otros términos, la libertad para Heidegger sólo puede existir en actos concretos de liberación. El ser-ahí

11. Cfr. Klaus Held, «Fundamental Moods and Heidegger's Critique of Contemporary Culture», en John Sallis (ed.), *Reading Heidegger. Commemorations*, Bloomington and Indianapolis, Indiana University Press, 1993., p. 223 (énfasis del original).

sólo puede ser libre si acepta la carga de *su* ser-ahí y para ello son necesarios tanto la resolución como el temple que la posibilita, de allí la interpretación del temple de ánimo como condición de la historicidad. Este planteo, así formulado, no difiere del de *Ser y Tiempo*. Esta interpretación, no obstante, implica una comprensión *política* de la noción de temple de ánimo. El § 74 de esa obra, donde se expone la constitución de la historicidad, presenta su desenlace *político* ya que la realización plena de la historicidad se da, en *Ser y Tiempo*, como el gestarse-con de un pueblo. Siguiendo esta línea, debemos interpretar la elucidación de los temples de ánimo en *Los conceptos fundamentales...* como un paso más en ese mismo sentido político, ya que la noción de temple de ánimo explica en esta obra el carácter predominante de la época unido a una transformación de la existencia que juzga inminente.

El tema de la transformación de la existencia había sido expuesto en *Ser y Tiempo* bajo la rúbrica del «recursar la tradición» por parte del ser-ahí, sin embargo, en *Los conceptos fundamentales...*, al retomar este tema Heidegger lo carga con una premura de la que carece en *Ser y Tiempo*. Por ello no puede comprenderse su elucidación de los temples de ánimo si se pierde de vista que ella es realizada en función de esta transformación de la existencia que considera como una necesidad urgente y también de realización inminente. Se percibe mejor la forma en que la problemática de la «transformación de la existencia» va adquiriendo un carácter político si se atiende a que en *Ser y Tiempo*, Heidegger presenta al ser-ahí instalado en su poder-ser más propio a través de la singularización,¹² esto es, su aislamiento, si bien la resolución del ser-ahí tiene que conformarse colectivamente en la historicidad; mientras que en el curso que estamos analizando, a qué se debe resolver hoy, en esta situación, el ser-ahí, aparece incierto. Ello no significa que la noción de «propiedad» haya desaparecido, más bien podría afirmarse lo contrario, ella sigue presente y orienta los análisis; pero, a diferencia de *Ser y Tiempo*, el problema se presenta ligado de manera inmediata a resolverse con respecto a una situación contemporánea que se presenta como apremiante. A medida que en la filosofía de Heidegger los motivos filosóficos ligados al protestantismo (como el del «estado de caído» del ser-ahí y el de su singularización o individuación) van perdiendo vigor, las

12. *Sein und Zeit*, p. 445.

cuestiones que los vehiculizaban son reinterpretadas paulatinamente de forma que adquieren un sentido cada vez más político, ya que la tendencia al aislamiento con respecto a la sociedad del ser-ahí en su «singularización» desaparece en función de una interrogación que ya no se realiza en los términos individuales de la filosofía de la existencia, sino bajo el carácter perentorio de la necesidad de una resolución, pero de carácter colectivo.

Durante el curso se repite la referencia de Heidegger al problema de «la transformación de la existencia». Ya lo señalamos al tratar el carácter de «indicaciones formales» que Heidegger atribuye a los conceptos filosóficos y se reitera en su crítica a Husserl:

No es un capricho o una excentricidad de la filosofía que nosotros hoy no hablemos más de vivencias o de vivencias de la conciencia o de conciencia, sino que *estamos obligados a otro lenguaje a causa de una transformación de la existencia*. Más precisamente, esta transformación acaece con este otro lenguaje (GM: 298, énfasis del original).

La temática de «la transformación de la existencia» se modificará a partir de este curso. Desde el planteo originario de *Ser y Tiempo*, donde se trata de recuperar al ser-ahí «caído» en la cotidianidad hacia su poder-ser más propio, el tema quedará enlazado a partir de este curso con la pregunta sobre la identidad, la cual es reformulada desde la interrogación puramente singular de *Ser y Tiempo* a la pregunta por «quiénes somos nosotros». Bajo esta forma la filosofía heideggeriana encontrará el modo de ocuparse de las dimensiones colectivas y públicas. En los textos posteriores la pregunta sobre la transformación de la existencia desaparecerá y será reemplazada por el tema del «nuevo inicio», al que tanta importancia dará en sus cursos sobre Hölderlin. De ese modo, las resonancias de la filosofía de la existencia, todavía fuertes en 1929/30, se irán apagando lentamente y «la transformación de la existencia» dejará de ser una cuestión referida a las posibilidades más propias del ser-ahí para convertirse en el problema del «nuevo inicio» y del cambio epocal esperado por Heidegger durante los años treinta, años que pueden ser considerados como la espera del «adviento» del nuevo inicio. Además de la experiencia de un temple de ánimo «contemporáneo», la otra noción que posibilita esa metamorfosis (la cual es, a grandes rasgos, la *Kehre*) es la de «instante». En *Los conceptos fundamentales...* repite la noción de «instante» de *Ser y Tiempo* pero ahora lo define como la re-

solución del ser-ahí por sí mismo. Más adelante, Heidegger retomará y reformulará la noción de instante, desligándola de la ontología fundamental, bajo la forma del *kairós*. La idea de que el instante es el tiempo indicado, el momento culminante para una realización, resuena tanto en *Ser y Tiempo* como en *Los conceptos fundamentales de la metafísica*. A partir del «instante» el ser-ahí ya no es algo meramente presente junto a otras cosas, sino que está emplazado en el ente a través de la manifestación del horizonte temporal pleno. Todavía en 1929, Heidegger muestra tal entusiasmo con esta noción (diferenciándola del tratamiento original kierkegaardiano) que poco más adelante llega a afirmar que con la comprensión del instante comenzó una nueva época en la filosofía por primera vez desde la antigüedad (GM: 225).

En *Los conceptos fundamentales de la metafísica* Heidegger convierte la pregunta originaria por los temples de ánimo, tal cual era formulada en *Ser y Tiempo*, en una pregunta sobre la época a través de dos preguntas:

Nuestra pregunta —¿se ha vuelto el ser humano hoy tedioso para sí mismo?—, sólo puede significar: ¿Es finalmente tedioso para el ser-ahí el ser humano actual como tal? Preguntamos por un tedio profundo, por uno, esto es, por uno determinado, esto es, por un tal de nuestro ser-ahí, no por el tedio profundo en general y universalmente. (GM: 242, énfasis del original)

El significado de la pregunta concierne a *nuestro ser-ahí, nuestra situación*. El tedio profundo sólo puede entenderse a partir de una situación particular. A diferencia de la angustia, que revelaba el carácter finito del ser-ahí como tal, el tedio profundo sólo puede determinarse desde la particularidad y, desde allí, comprender sus formas más superficiales. Heidegger lo expone así:

Si antes preguntamos así [de forma general] y diera la impresión de que en la anterior interpretación del tedio tuviéramos una receta objetiva del tedio frente a nosotros, eso, como ahora se muestra, era una confusión. Comprendemos el tedio profundo como tal sólo a partir de uno determinado, esto es, un tedio esencial, y por consiguiente de allí recibe su dirección y su luz toda interpretación de las formas superficiales. El conocimiento esencial es posible sólo desde y en la originariedad de un preguntar. (GM: 242, énfasis del original)

Debe atenderse a que Heidegger entiende por tedio esencial uno determinado. Su metodología, por tanto, no está repitiendo la metodolo-

gía aristotélica de que se debe conocer en primer lugar lo primero para nosotros, esto es, lo más fácil de conocer y luego, lo primero en sí, lo más general. Por el contrario, en su elucidación del tedio Heidegger está dejándolo de lado un ascenso en la abstracción semejante. De allí su afirmación de que el preguntar originario sólo puede dirigirse a una situación concreta. No puede haber un tratamiento del tedio que abarque al ser humano de forma ahistórica. La concreción de la situación significa la situación de *nuestra* existencia, de *nuestro* ser-ahí, es decir, la concerniente a su contemporaneidad. La pregunta nunca se dirigió, en consecuencia, a elucidar estructuras del ser-ahí intemporalmente, sino a la situación particular que Heidegger y su auditorio perciben como propia. De allí su exposición de los diagnósticos de la época al inicio del curso. El preguntar esencial responde a lo mismo que intentan responder aquéllos, pero se diferencia de ellos precisamente por lo que considera su esencialidad, esto es, en la capacidad de elucidación que le brindan los conceptos de la ontología fundamental puestos al servicio del esclarecimiento del ser-ahí coetáneo, esto es, del carácter de la época. De un salto hemos entrado en el ámbito de lo preguntado, precisamente gracias al poder descubridor de los temples de ánimo.

La cuestión concerniente al tedio profundo en el ser-ahí contemporáneo es abordada por Heidegger de un modo que recuerda a «¿Qué es metafísica?». Allí se afirma que el tedio profundo nos revela el ente en su totalidad. En *Los conceptos fundamentales...*, se sostiene algo similar, el tedio profundo nos entrega al rehusarse del ente en su totalidad, este rehusarse es, entonces, un total vacío. Sin embargo, en la conferencia inaugural Heidegger remitía a este carácter de revelación de la totalidad por parte de los temples para explicar la apertura que se revela mediante la angustia. En el curso, en cambio, se insiste en que el vacío que revela el tedio no debe entenderse como una pura nada, sino como una carencia, una necesidad. Éste es el punto central de la argumentación heideggeriana: se pregunta si estamos afectados por alguna necesidad, a lo que se responde con una clara alusión a la situación social resultante de la crisis económica

[...] por todos lados hay estremecimientos, crisis, catástrofes, necesidades: la miseria social actual, la confusión política, la impotencia de la ciencia, el socavamiento del arte, la falta de enraizamiento (*Bodenlosigkeit*) de la filosofía, la falta de energía de la religión. Ciertamente, hay necesidades por todos lados (GM: 243)

Las necesidades se multiplican en la sociedad y Heidegger advierte los esfuerzos que se realizan para controlarlas y para satisfacerlas; aún más «[...] todos y todo está organizado contra las necesidades y toda organización tiene su programa.» El problema que se deriva de ello según Heidegger es que esta ambición social de controlar cualquier tipo de necesidad a fin de satisfacerla impide el surgimiento de una necesidad en su totalidad. «Totalidad» significa aquí lo contrario de una mera suma de individuos: que aparezca una necesidad en su totalidad significa que el conjunto de la sociedad esté determinado por ella, al modo de un horizonte al cual ella está sujeta. La multiplicación de atenciones a las diversas necesidades que Heidegger describe alude al desarrollo del Estado Social y la asistencia que éste brinda se asemeja a la descripción que hace en *Ser y Tiempo* del «procurar por» no liberador, el cual es rechazado por Heidegger señalando que es una forma de dirigirse al otro que no le permite disponer de su propio ser-ahí. Frente al convencimiento colectivo de que las necesidades deben ser satisfechas y que ello puede hacerse mediante el cálculo, Heidegger reclama que no es la necesidad lo que nos oprime, sino que por el contrario nos oprime la ausencia de una opresión esencial de nuestro ser-ahí en su totalidad.

Preguntamos por un vacío total (*Leere in Ganzen*), por una necesidad que ciertamente no puede tener el carácter de las necesidades enumeradas. [...] no es esto la necesidad, que esta o aquella necesidad nos oprima de algún modo, sino que lo más oculto y profundamente opresivo es más bien la ausencia de una opresión esencial de nuestro ser-ahí en totalidad (GM: 244, énfasis del original)

Ésta es la marca de la época según Heidegger: la ausencia de una opresión de nuestro ser-ahí en su conjunto equivale al vacío total. Esta disolución que detecta no se refiere a las consecuencias de la crisis económica, sino que apunta a la desaparición del sentido de comunidad, lo que impide la «[...] unidad enraizada de una acción esencial». Heidegger afirma que todos son sirvientes de programas, pero nadie es el «custodio» (*Verwalter*) de la «interna grandeza» del ser-ahí. ¿Puede sostenerse que su descripción está motivada por el rigor de la crisis en la sociedad alemana? No debe descartarse que su magnitud provocara que se planteara problemas que podrían parecer ajenos a su filosofía y que alejándolo de la metafísica pura lo abismaran en la coyuntura. Recordemos que Otto Pöggeler sostenía la coincidencia de una crisis per-

sonal en Heidegger junto con la crisis de 1929. Sin embargo, su crítica va más allá de la circunstancia específica de la crisis ya que retoma convicciones profundas que ya se habían presentado en *Ser y Tiempo*, especialmente en los párrafos dedicados a la «publicidad» (27, 35, 36 y 37). A diferencia de aquella obra, no obstante, la temática de la falta de una opresión esencial es nueva e indica que para Heidegger la crisis no ha hecho más que profundizarse, ya que esa falta denuncia un ahondamiento en la autosatisfacción indulgente que permite al hombre contemporáneo no hacerse cargo de su ser-ahí. La crisis mentada por Heidegger se manifiesta en el plano socioeconómico, entre otros, pero su naturaleza es diferente, pues consiste en la confusión de las necesidades sociales con la falta de una necesidad esencial, ésa es la carencia que marca la época. Ella sólo es posible en el marco de la disolución de la comunidad y de todo sentido de pertenencia, como Heidegger se encarga de señalar refiriéndose repetidas veces a la falta de enraizamiento, tanto de la filosofía como de la acción común. Ello permite comprender la tarea de «despertar al terror» que Heidegger atribuía a la filosofía. El ser-ahí contemporáneo satisface sus necesidades mediante una manipulación técnica, lo que ha significado la desaparición de todo misterio, por tanto, él y la sociedad carecen de «misterio», así como del terror que todo misterio convoca. Aunque Heidegger todavía no utilice el término, podría sostenerse que la sociedad que describe es, a su juicio, una sociedad nihilista, que cree que ha disuelto todo enigma en prescripciones técnicas y se muestra satisfecha de ello. Sin duda Heidegger siente nostalgia por el terror, esto es, por el terror que lo sagrado convoca en el hombre. La falta de una necesidad esencial es la característica de una sociedad tecnificada en todos sus aspectos, sólo el terror que ha desaparecido puede restaurar en el ser-ahí la conciencia de una necesidad que convoca a la grandeza y que no puede ser explicada en los términos técnicos de los programas ni de la ingeniería social. De allí la crítica de Heidegger a la política contemporánea:

Pues en todo este organizar, construir programas y ensayar hay en última instancia una universal y satisfecha dicha en la falta de peligro. Esta dicha en el fundamento de nuestro ser-ahí, a pesar de todas las variadas necesidades, hace que creamos que ya no se necesite más ser fuertes en el fundamento de nuestra esencia (GM: 245)

La inflexión nietzscheana de la crítica de Heidegger se vuelve más

clara, ya que la crítica se dirige contra los esfuerzos igualitarios operantes en la política europea desde fines del siglo XIX y los enfoca a partir de la debilidad que resultaría de ellos, perspectiva desde la que Nietzsche trató estas cuestiones. Entre los tópicos literarios y políticos de esa época la crítica de la existencia burguesa, a la cual se la identifica con la mezquina búsqueda de seguridad y autocomplacencia, ocupará un lugar privilegiado, tanto a la izquierda como a la derecha. Esta búsqueda de seguridad se ha vuelto el fundamento del ser-ahí contemporáneo, como agrega Heidegger ironizando sobre ella y contraponiéndola a las crecientes demandas que ella misma genera debido a la crisis. El resultado de esta búsqueda universal de seguridad es que se acepte que «no se necesite más ser fuertes». Heidegger establece de este modo con bastante claridad un conjunto de oposiciones: seguridad/peligrosidad, programar/fortalecer, necesidades «sociales»/falta de una opresión esencial. El segundo término de las oposiciones expresa lo que Heidegger encuentra ausente en la sociedad contemporánea y que ella no hace más que profundizar en ese camino de debilitamiento, lo que queda manifiesto en la siguiente cita:

Nos afanamos sólo por capacidades que pueden ser inculcadas. El presente está colmado de cuestiones pedagógicas. Sin embargo, nunca se podrá reemplazar el vigor y el poder a través de la acumulación de capacidades, sino que si algo se alcanza a través de esto es su ahogo (GM: 245)

Si se unen ambas citas, es claro que la universal vocación técnico-pedagógica sofoca lo que Heidegger denomina «el fundamento de nuestra esencia», del cual espera su fortalecimiento, contra los aires de los tiempos. Estos fenómenos serán las notas características de lo que Heidegger designará unos años después como «nihilismo». La desaparición de todo misterio es sinónimo de nihilismo porque se disuelve todo sentido que vaya más allá de la mera acción técnica. El tedio profundo es el temple de ánimo del ser-ahí contemporáneo porque es el que corresponde a una situación en la que el sentido se ha disuelto. Este temple revela el vacío al que el ser-ahí está sometido, la carencia de una opresión esencial y la falta de una necesidad que no pueda ser abordada en un marco exclusivamente técnico. A juicio de Heidegger no hay nada de paradójico en que la falta de una opresión esencial no signifique la liberación del ser-ahí.

Conclusión: lo político y la muerte de Dios

La forma en que estos temas aparecen vinculados en el curso revela que la fenomenología del tedio profundo es el camino elegido por Heidegger para abordar el problema nietzscheano de la «muerte de Dios». El tedio profundo constituye el rasgo que caracteriza al ser-ahí contemporáneo pues implica la pérdida general de sentido del mundo, y esta pérdida se manifiesta de forma más visible en la desaparición del misterio y del terror. En el aforismo 125 de *La Gaya Ciencia*, «El hombre loco», Nietzsche describe la muerte de Dios como un acontecimiento histórico. No se trata de que Dios nunca haya existido, sino que, por el contrario, durante siglos su presencia conformó el mundo espiritual de la humanidad europea y el coronamiento de esa cultura ha sido la desdivinización del mundo mediante la ciencia. La magnitud de este acontecimiento todavía no ha sido comprendida, afirma Nietzsche, por eso los hombres, habiendo matado a Dios, no son capaces de comprender las preguntas del «hombre loco», el que anuncia la muerte de Dios:

¿Hacia adónde nos movemos nosotros, apartándonos de todos los soles? ¿No nos precipitamos continuamente?, ¿hacia atrás, adelante, a un lado y a todas partes? ¿Existe todavía para nosotros un arriba y un abajo?, ¿no vamos errantes como a través de una nada infinita? ¿no nos absorbe el espacio vacío?, ¿no hace más frío? ¿No viene la noche para siempre, más y más noche?¹³

El tedio profundo elucidado por Heidegger retoma esta noche anunciada por Nietzsche. La falta de una opresión esencial en el ser-ahí contemporáneo es análoga a la actitud de los hombres que escuchan al «hombre loco» nietzscheano, quienes se extrañan ante su anuncio y no tienen nada para decir. Del mismo modo, el ser-ahí contemporáneo carece para Heidegger de esa opresión ejercida por lo sagrado, por eso lo aquejan necesidades de la más variada índole, pero ninguna puede surgir como una necesidad en su totalidad porque unas y otras se suceden sin que ninguna tenga una importancia real. Tras la muerte de Dios la filosofía debe convocar al misterio porque éste se ha esfumado del

13. Friedrich Nietzsche, *Die fröhliche Wissenschaft*, trad. esp. De L. Jiménez Moreno, *El gay saber*, Madrid, Espasa, 1986, p. 155.

mundo de los hombres. Esto se presenta como una opresión porque el ser-ahí no reacciona frente a la desaparición del misterio, sino que se afana en la mayor organización posible de las necesidades, profundizando el nihilismo. Ello ya había sido tratado por Nietzsche, cuando asocia la «muerte de Dios» con la búsqueda de seguridad:

Donde la vida social tiene un carácter *menos violento*, las decisiones últimas (sobre las llamadas cuestiones eternas) pierden importancia. Piénsese que hoy un hombre tiene raramente que ver con ellas.¹⁴

En este curso la cuestión de la «muerte de Dios» queda abierta, exponiendo Heidegger su necesidad y la necesidad de que el ser-ahí se resuelva a la demanda más extrema, esto es, a romper con la ocupación exclusiva en las necesidades para quedar abierto a aquella necesidad esencial acallada, pero todavía su respuesta aparece ligada a los lineamientos de la filosofía existencial, esto es, el ser-ahí debe resolverse a su necesidad más profunda, a su libertad, o, como también afirma, el hombre debe resolverse a lo que el ser-ahí en él le exige, todo lo cual no se aleja de la forma en que la «resolución» era tratada en *Ser y Tiempo*. El tema será retomado en el discurso rectoral en 1933 y a partir de ese momento quedará unido a la necesidad de reemplazar ese vacío dejado por la divinidad. A modo de conclusión, sostenemos que el problema de la «muerte de Dios» es uno de los núcleos principales de lo político en Heidegger, ya que su perspectiva no es restauradora. La muerte de Dios es un resultado inevitable de la cultura europea, el problema, en consecuencia, es estar a la altura del acontecimiento, tal como «el hombre loco» apostrofaba a su auditorio, negándoles esa altura. Asumir esa condición de lo político implica para Heidegger, que otra instancia tiene que volver a convocar el misterio y el terror ligados antes a Dios; es decir, la muerte de Dios no significa la imposibilidad de lo sagrado entre los hombres. A partir del desenvolvimiento de la *Kehre* en el pensamiento heideggeriano, esta temática quedará ligada a la del «nuevo inicio» y la instancia a partir de la cual Heidegger convocará a lo sagrado en tiempos de los dioses huidos será la patria a través de la palabra fun-

14. Friedrich Nietzsche, *IV Consideración Intempestiva*, § 3 p. 350, citado por Gianni Vattimo, *Introducción a Nietzsche*, Barcelona, Ediciones Península, 1987, p. 76.

dad del poeta. Pero, a diferencia de lo que algunos intérpretes sostienen, «patria» no significará exclusivamente «tierra natal»,¹⁵ sino más bien, siguiendo a Hölderlin, *Germania*.

ABSTRACT

This article examines the way in which Heidegger explains what he considers to be the philosophical and cultural foundations of the 1929 crisis. For this purpose, he makes use of the concepts of fundamental ontology exposed in *Being and Time*, but giving them a political significance absent in that work, specially in his elucidation of profound boredom as a characteristic of contemporary being-there. We try to show that heideggerian analysis make possible the «political availability» of his philosophy, i.e., the beginning of a reflection that is not yet nationalist in itself, but with his radicalisation in later works, accepting Nietzsche's indictment about «the death of God», it will end in the rectorial discourse and his affiliation to the nazi party.

15. Como afirma Otto Pöggeler en su obra *Martin Heideggers Denkweg* (1963), trad. al español de F. Duque Pajuelo, *El camino del pensar de Martin Heidegger*, Madrid, Alianza, 1986, cap. VIII, "El otro inicio".

crítica

REVISTA HISPANOAMERICANA DE FILOSOFÍA

Vol. 34 / No. 100 / Abril 2002 / ISSN 0011-1503

Artículos

AMÁN ROSALES RODRÍGUEZ, *El determinismo tecnológico y la "dialéctica de la historia"*

GERT HELGESSON, *What Is Implicit?*

IGNACIO ÁVILA CAÑAMARES, *El nuevo enigma de la inducción y los términos de clase natural*

LEE MCINTYRE, *Supervenience and Explanatory Exclusion*

Notas bibliográficas

MARÍA LILIANA LUKAC DE STIER, *El fundamento antropológico de la filosofía política y moral en Thomas Hobbes* [Fabiola Rivera Castro]

Publicaciones recientes

CRÍTICA. Revista Hispanoamericana de Filosofía is published in April, August and December. All correspondence should be addressed to CRÍTICA, Apartado 70-447, Coyoacán, 04510-México, D.F. México
E-mail: critica@filosoficas.unam.mx

DOS ARGUMENTOS EN CONTRA DEL LENGUAJE DE PENSAMIENTO

Silvio Pinto

1 Introducción

La analogía de la mente con una computadora favorecida por varios filósofos y psicólogos cognitivos presupone la tesis de que los procesos mentales (secuencias de estados mentales) son procesos computacionales de manipulación de símbolos y como tales requieren un lenguaje para su operación (i.e. un sistema simbólico articulado capaz de expresar los contenidos de todos los estados mentales que toman parte en estos procesos). Este sistema simbólico, que Fodor llama el lenguaje de pensamiento, tiene dos características fundamentales: es innato y privado. Lo primero quiere decir que no aprendemos a manejar este lenguaje y el segundo que solamente el sujeto cuyos procesos cognitivos están siendo investigados puede entender las palabras y oraciones de su lenguaje de pensamiento. Estas dos características del lenguaje de pensamiento están obviamente ligadas: si tal lenguaje no fuera privado habría la posibilidad de aprenderlo y si no fuera innato entonces su posible aprendizaje exigiría la existencia de al menos dos hablantes del mismo lenguaje, con lo que se imposibilitaría su privacidad. El hecho de que el innatismo y el carácter privado del lenguaje de pensamiento estén tan estrechamente unidos permite que los argumentos presentados por Fodor a favor del innatismo del lenguaje de pensamiento sean también argumentos a favor de su privacidad.

Sin embargo y a pesar de que dichos argumentos son aparentemente persuasivos, uno podría preferir, por economía explicativa, una teoría de la mente que no necesitara postular la existencia de un código interno además de los códigos públicos que todo hablante utiliza para comunicarse. Creo, no obstante, que hay una manera más directa de refutar la tesis de la existencia de un lenguaje de pensamiento, a saber: si nos convencemos de la imposibilidad de un lenguaje privado por medio del argumento ofrecido por Wittgenstein en las *Investigaciones Filo-*

sóficas.¹ Esto es lo que pretendo hacer en este trabajo: a) refutar la tesis del lenguaje de pensamiento mostrando que el argumento wittgensteiniano es correcto y que además es incompatible con dicha tesis; y b) mostrar la dispensabilidad del lenguaje de pensamiento, sugiriendo maneras alternativas de explicar los hechos que Fodor cree no pueden ser explicados sin postular la existencia del lenguaje de pensamiento. Todo esto tendrá como consecuencia que nos veremos forzados a privilegiar modelos no-computacionales de la mente. Quizás los modelos conexionistas sean un primer gran paso en esta dirección.

La manera como está dividido el texto es la siguiente. La sección 2 trata de la tesis del lenguaje del pensamiento y de los argumentos ofrecidos por Fodor en su favor. La sección 3 presenta el argumento en contra del lenguaje privado, privilegiando una interpretación de dicho argumento que lo torna incompatible con la tesis del lenguaje de pensamiento. La sección 4 muestra la incompatibilidad del argumento wittgensteiniano con la tesis de Fodor y también que este argumento es correcto —es decir, en esta sección se lleva a cabo la tarea (a) mencionada arriba. Finalmente, la tarea (b) se cumple en la sección 5.

2. El Lenguaje de Pensamiento

La tesis fundamental de la llamada teoría computacional de la mente es que los procesos mentales consisten de operaciones sintácticas sobre representaciones mentales. Esto implica que el conjunto de tales representaciones debe tener las características de un lenguaje (i.e. estas representaciones deben poseer al menos propiedades sintácticas y semánticas). La idea de que necesitamos de una teoría de los procesos mentales de los organismos a los cuales tiene sentido atribuir la propiedad de poseer una mente responde al deseo de entender el mecanismo por medio del cual estos organismos elaboran respuestas tan sofisticadas como acciones lingüísticas y no-lingüísticas a partir de los estímulos que reciben. Uno de los principales problemas para cualquier teoría acerca de tal mecanismo es el de explicar como los procesos mentales pueden exhibir al mismo tiempo propiedades causales y propiedades semánticas. Ejemplos de las primeras son: estímulos de un cierto tipo (un objeto moviéndose en el campo visual de un organismo perceptor)

1. Wittgenstein 1953. De aquí en adelante simplemente *Investigaciones*.

causan una percepción, que por su vez provoca la formación de una creencia, que junto con otros estados mentales del organismo desencadenan una cierta respuesta conductual; o determinado estímulo lingüístico causa un cierto estado de comprensión lingüística en el oyente, que responde con determinadas emisiones verbales. Ejemplos de las propiedades semánticas de los procesos mentales son: alguien cree que su mamá está por llegar, y cree haber escuchado una voz en la habitación vecina, y además cree haber reconocido la voz como la de su mamá, por lo que infiere que su mamá está en la casa; o alguien quiere llegar lo más rápido posible a un determinado lugar, y cree que la mejor manera para lograrlo es tomar un avión, y cree que para tomar el avión es necesario comprar un boleto, y por lo tanto compra el boleto. La corrección de estos procesos inferenciales depende claramente del contenido de los estados mentales y del contenido de las acciones que constituyen tales procesos.

Para explicar la existencia de relaciones causales y al mismo tiempo semánticas entre los estados mentales constituyentes de diversos procesos mentales, los defensores del computacionalismo sobre la mente conciben un estado mental como una relación entre un sujeto y una representación simbólica y apelan a la idea de Turing de que las propiedades sintácticas de los símbolos determinan su papel causal. Y como hay una función que lleva de las propiedades sintácticas inferenciales de los procesos en cuestión a sus correspondientes propiedades semánticas (debido a la completud de la lógica de primer orden), se podría afirmar que dentro del computacionalismo las propiedades semánticas inferenciales de los procesos mentales determinan las conexiones causales entre los estados mentales que los constituyen.² La posibilidad de una ex-

2. Obviamente que no hay correspondencia uno a uno entre los símbolos del mentalés (Fodor a veces se refiere al lenguaje de pensamiento con esta expresión) y sus respectivos contenidos. Esto significa que no es posible identificar los contenidos de tales símbolos únicamente por el poder causal procedente de la forma sintáctica de estos símbolos ya que el mismo contenido se puede expresar en un mismo lenguaje de diversas maneras. Un ejemplo sería el de la afirmación de un contenido y su doble negación. Esto no afectaría el modelo computacionalista si se lo entiende, no como afirmando la existencia de una correspondencia uno a uno entre las propiedades semánticas de los símbolos del mentalés y sus propiedades causales, sino como comprometiéndose con la tesis

plicación del doble carácter de los procesos mentales (sus dimensiones semántica y causal) es uno de los argumentos empleados por Fodor a favor de la tesis del lenguaje de pensamiento.

Un segundo argumento tiene que ver con la supuesta imposibilidad de aprender un lenguaje a no ser que uno ya maneje otro lenguaje. El argumento contiene las siguientes 4 premisas:

- P₁ aprender un lenguaje consiste en aprender el significado de los predicados de este lenguaje;
- P₂ aprender lo que tales predicados significan incluye aprender a determinar sus respectivas extensiones;
- P₃ aprender a determinar la extensión de un predicado *P* consiste en aprender las reglas *R* que establecen las condiciones de verdad de las oraciones que contienen este predicado;
- P₄ pero, no es posible llegar a saber que *P* está gobernado por la regla *R* a no ser que uno ya posea un lenguaje en que *P* y *R* puedan ser representados;
- C por lo tanto, no es posible aprender un lenguaje a no ser que uno ya posea un lenguaje.³

No va a importar para la validez del argumento si *R* expresa las condiciones de verdad o las condiciones de verificabilidad de las oraciones de *L* o cualesquiera otras condiciones. Lo crucial es la tesis de que aprender un lenguaje *L* es formar y confirmar hipótesis sobre las regu-

de que las propiedades semánticas inferenciales de estos símbolos determinan sus propiedades causales. Visto de esta última manera, el computacionalismo estaría buscando una explicación para la mencionada tesis.

3. Fodor 1975, pp. 63-64. Habría obviamente que extender lo que se dijo para los predicados del lenguaje *L* a los otros términos de *L*. La idea por detrás de la propuesta es que el significado de un término determina su referencia para aquellos términos que se refieren a algo. Para los que no se refieren a nada (como las constantes lógicas) no habrá axiomas de una definición de verdad para *L* que indican su referencia y sí axiomas que expresan su correcto uso. Con relación a aquellos términos no abarcados por la definición de verdad para *L* tendríamos que modificar la caracterización arriba de lo que estaría implicado en aprenderlos pero es plausible pensar que alguna regla de uso correcto (aunque no sea una regla de verdad o de referencia) gobierna su empleo. Esto es suficiente para los propósitos del argumento.

laridades asociadas con el empleo de los términos de *L*. Ejemplos de tales regularidades son los axiomas que establecen la extensión de los predicados de *L*.

Uno no debe pensar que al aceptar esta tesis sobre el aprendizaje del lenguaje se está automáticamente comprometiendo con un enfoque empirista, ya que también es necesario explicar el carácter normativo que las reglas de uso correcto de los términos de *L* adquieren para el hablante. Tampoco se debe inferir de la tesis mencionada y de la concepción sobre el funcionamiento de *L* que subyace al argumento que tales reglas son algoritmos deterministas, es decir: que para toda y cualquier aplicación de *P* a un objeto (digamos) su respectiva regla de verdad *R* determine un valor de verdad para la oración resultante. Tenemos que admitir que las reglas para el uso de palabras dejan lagunas —es decir, ni siempre nos dan un valor (por ejemplo, un valor de verdad) para todos los argumentos que pueden tomar (por ejemplo, una oración y un determinado contexto de su empleo).

El punto controvertido del argumento está relacionado con la premisa P₄. Otra manera de expresarlo es la siguiente: no es posible formular y confirmar hipótesis cuyos conceptos constituyentes no están ya disponibles para su expresión. Por lo tanto, si el modelo de aprendizaje del lenguaje como un proceso de formulación y confirmación de hipótesis es correcto, entonces no se puede aprender un lenguaje si ya no se posee un lenguaje. Y para evitar el regreso al infinito que esta conclusión nos podría llevar hay que postular la existencia de un lenguaje que no es aprendido, un lenguaje en el cual se realizan los procesos cognitivos del organismo. Voy a dejar la discusión de este punto para la sección 5 en donde concentro mis críticas a los argumentos de Fodor.

Un tercer argumento a favor del lenguaje de pensamiento apela a la posibilidad de explicar la productividad y la sistematicidad del pensamiento humano.⁴ Afirmar que el pensamiento humano es productivo equivale a decir que podemos tener una cantidad potencialmente infinita de pensamientos o actitudes proposicionales con distintos contenidos. Según Fodor, la mejor y quizá la única explicación de tal capacidad es que los estados mentales intencionales (aquellos a los cuales se asocia un contenido) están estructurados como un lenguaje, es decir se compo-

4. Dos de los lugares en que él aparece en la obra de Fodor son Fodor 1985 y Fodor 1987.

nen de un número finito de constituyentes básicos que pueden hacer parte del contenido de infinitos otros estados mentales. La sistematicidad del pensamiento se traduce en la idea de que nuestra capacidad para tener cualquier pensamiento está indisolublemente ligada a la capacidad para tener varios otros pensamientos semejantes al primero en su forma. Por ejemplo, si uno es capaz de pensar que el gato está sobre la alfombra debe también ser capaz de pensar que el perro está sobre la alfombra o que el niño está sobre la alfombra o que el gato está gordo o que el gato es siamés, etc.. La mejor hipótesis para explicar la sistematicidad del pensamiento, insistiría Fodor, es que el pensamiento humano está estructurado como un lenguaje.

Pero, ¿es realmente indispensable postular la existencia del mentalés para que se pueda dar cuenta de la productividad y de la sistematicidad del pensamiento? Como en el caso de la corrección del argumento anterior, voy a posponer la respuesta a esta pregunta hasta la sección 5.

Sin embargo, debo señalar que los 3 argumentos arriba presentados ya no tienen la misma importancia para Fodor. En textos más recientes,⁵ Fodor ha rechazado, por ejemplo, la idea de que ya nacemos con los conceptos primitivos del mentalés; hay un mecanismo interno responsable de la adquisición de tales conceptos, pero este mecanismo no es consciente. Por lo tanto, tampoco se puede decir que aprendimos de manera consciente a manejar dichos conceptos. También ha rechazado el modelo inductivista del aprendizaje de nuestro primer lenguaje. Esto significa que el segundo argumento a favor del mentalés ya no posee para Fodor la fuerza persuasiva que poseía en la década de 70.

Además, quiero hacer hincapié en que estos 3 argumentos no agotan la batería de razones aducidas por Fodor en defensa de la tesis del lenguaje de pensamiento. No obstante, ellos ya nos dan una idea clara de la tarea que nos espera si queremos enfrentar la referida tesis. Veamos como Wittgenstein hubiera reaccionado a la idea de un lenguaje privado.

Pero antes debo puntualizar que he dejado de lado varios aspectos de la concepción de Fodor por la simple razón de que en mi opinión su examen no es necesario para entender la tesis del lenguaje de pensamiento. Esto se debe en parte al hecho de que la tesis en cuestión es tan básica en la teoría de Fodor que resulta imprescindible para apoyar

otras de sus tesis, y en parte a que la tesis en cuestión es independiente de algunas otras de sus tesis.

Tomemos, por ejemplo, su tesis sobre el carácter científico del tipo de explicación ofrecido por la psicología popular (*folk psychology*) y, en consecuencia, su reivindicación de la relevancia de este tipo de explicación para la comprensión de la conducta humana entendida tal y como nosotros mismos la entendemos, a saber: desde la perspectiva de su descripción intencional. Fodor sostiene que el tipo de explicación que proporciona la psicología popular así como las leyes psicológicas que utiliza son tan científicas como las de cualquier otra ciencia no-básica; lo que diferencia estas ciencias no-básicas de la física básica es que sólo para las primeras, si "todo *F* es *G*" es una ley de estas entonces tiene sentido la cuestión ¿cómo puede *poseer la propiedad F* estar relacionado causalmente con *poseer la propiedad G*? Para responder a esta pregunta es necesario, según Fodor, apelar al aparato explicativo de una ciencia más básica, mientras que para la física básica no hay ulterior mecanismo más fundamental que sea capaz de iluminar las conexiones legales que ocurren en su dominio. En el caso específico de la psicología, la pregunta mencionada arriba se responde, diría Fodor, recurriendo al nivel computacional; lo que explica la existencia de las leyes causales y intencionales de la psicología popular son los procesos de manipulación de símbolos que se realizan a un nivel más profundo que el de los estados mentales intencionales. Tales procesos computacionales requieren obviamente la tesis del lenguaje de pensamiento.⁶ La reivindicación de la relevancia explicativa de la psicología popular y su carácter científico está por lo tanto conectada con la tesis del lenguaje de pensamiento; el ejemplo muestra claramente que la referida tesis es más básica dentro de la concepción de Fodor que la mencionada relevancia del tipo de explicación de la psicología popular.

Otro ejemplo es la famosa tesis fodoriana del solipsismo metodológico.⁷ Dicha tesis afirma que los estados y procesos mentales de un agente no se ven afectados por las propiedades semánticas de las representaciones asociadas a ellos, sino solamente por sus propiedades formales (de las cuales obviamente las propiedades sintácticas son un

5. Como, por ejemplo, Fodor 1998a, capítulo 6 y Fodor 1998b.

6. Fodor 1987, capítulo 1. Una versión bastante resumida de lo que se dijo en este párrafo se puede encontrar en Fodor 1994.

7. Fodor 1980.

ejemplo). Esto significa que las transiciones entre los estados mentales que caracterizan estos procesos se explican apelando simplemente a la diferencia entre las formas de las representaciones asociadas respectivamente al estado inicial y final de dicho proceso y no a los contenidos de tales representaciones, a sus referentes o cualquier otra de sus propiedades semánticas. Un ejemplo de tales procesos son los razonamientos de tipo inductivo y deductivo de que son capaces la mayoría de los humanos. Pero decir que los procesos mentales humanos responden únicamente a la forma de sus respectivas representaciones equivale a adoptar un modelo computacional acerca del mecanismo de funcionamiento de la mente humana. Y como ya sabemos, tal modelo se compromete con la tesis del lenguaje de pensamiento. Esta es la conexión del solipsismo metodológico de Fodor y la tesis de la existencia del mentalés. Sin embargo y al igual que en el ejemplo anterior, la falta de un examen detallado de la tesis del solipsismo metodológico o de su conexión con varias otras tesis defendidas por Fodor no comprometerá nuestra crítica a la tesis del lenguaje de pensamiento.

Lo mismo vale para la distinción que Fodor también defiende entre contenido amplio y estrecho y su preferencia por la segunda noción o su reivindicación del nivel representacional con contenido amplio a fin de preservar el nivel explicativo de la psicología popular. Todas estas tesis están en mayor o menor grado conectadas con la tesis que nos ocupa. No obstante, esta última es tan básica que podemos estudiarla abstractamente de la mayoría de las demás.

Pasemos entonces al argumento wittgensteiniano en contra de la propia posibilidad de un lenguaje privado.

3 El Argumento en contra del Lenguaje Privado

La tesis de un lenguaje privado está presente en la filosofía de Bertrand Russell. La existencia de un lenguaje cuyos términos tienen como referencia objetos privados es esencial para el proyecto fundacionista russelliano de construcción del mundo físico y mental a partir de los datos de los sentidos (*sense data*) y del análisis de las palabras de los lenguajes públicos en términos de palabras que se refieren garantizadamente a los datos de los sentidos. Es esencial a la noción del lenguaje privado russelliano que los referentes de sus términos sean objetos privados, es decir objetos a los cuales sólo el sujeto que maneja este lenguaje tiene acceso.

La idea de un lenguaje privado también está presente en la primera filosofía de Rudolf Carnap, así como en la de varios filósofos empiristas. En sus primeros escritos, Carnap hablaba del lenguaje más básico, el lenguaje protocolario, como un conjunto de oraciones usado por cualquier hablante para reportar directamente sus propias experiencias privadas. Según el mismo Carnap, cualquier oración de un tal lenguaje tendría sentido solamente para el propio hablante; nadie más podría entender este lenguaje.⁸ Sin embargo, el lenguaje privado carnapiano se relaciona de alguna manera con los lenguajes públicos que usamos para comunicarnos con nuestros semejantes. No nos va a importar en este momento la manera como los lenguajes privado y público se relacionan; lo relevante es el simple hecho de que para Carnap había un lenguaje privado que está en la base de cualquier lenguaje público.

La hipótesis de un lenguaje privado también es compatible con la filosofía del empirismo británico de los siglos XVII y XVIII. Los filósofos empiristas tales como Locke y Hume creían que los átomos de nuestro conocimiento son experiencias, ideas sensibles simples o impresiones con las cuales cada uno de nosotros está familiarizado interna y privadamente; además, estas ideas como contenidos se expresan naturalmente por medio de símbolos que ambos filósofos admiten como parte de un lenguaje. Como los contenidos a que se refieren dichas palabras participan en la formación de todos los contenidos o ideas complejos, se puede decir que gran parte del lenguaje que usamos para comunicarnos consta de palabras cuyo contenido es esencialmente privado. Es un misterio el que la comunicación entre hablantes se pueda lograr en tal lenguaje.

Wittgenstein dirige su argumento en contra del lenguaje privado a algo como el lenguaje que Russell,⁹ Carnap y los empiristas clásicos tenían en mente, es decir a un lenguaje de sensaciones privadas e inmediatas.¹⁰ El argumento aparece oficialmente entre los párrafos 243 y 315 de las *Investigaciones*, pero por lo menos una de sus premisas ya había sido discutida en las famosas observaciones sobre el seguimiento de reglas lingüísticas (entre los párrafos 138 y 243). Se puede formularlo de la siguiente manera:

8. Por ejemplo, en Carnap 1934.

9. Por ejemplo en Russell 1918. El lenguaje privado russelliano era un lenguaje cuyos nombres propios se referían a los *sense data*.

10. Wittgenstein 1953, § 243.

- P₁ Es de la esencia de cualquier lenguaje *L* la existencia de criterios para el uso correcto de cada una de sus expresiones;
 P₂ Pero no puede haber tales criterios para un lenguaje privado;
 C Por lo tanto, no es posible un lenguaje privado.

Antes de discutir el argumento wittgensteiniano, debo decir que mi propósito no es primariamente exegético; no me interesa aquí examinar las innumerables interpretaciones que se ha ofrecido para tal argumento. Solamente el volumen de la literatura secundaria sobre el tema exigiría un o varios artículos exclusivamente dedicados a clasificar y evaluar las diversas líneas interpretativas acerca del argumento en contra del lenguaje privado. Mi estrategia en esta sección será por lo tanto más dogmática: empiezo presentando mi interpretación del argumento, bastante influenciada sin duda por la de Jim Hopkins,¹¹ y también muy próxima a la propuesta por Saul Kripke.¹² Sobre la marcha, discuto algunos de los temas que se mostrarán más controvertidos en mi interpretación.

Hecha esta aclaración, creo que ya podemos pasar a la discusión de las premisas del argumento wittgensteiniano. Recordemos sus premisas: (P₁) para cualquier lenguaje *L* hay criterios para empleo correcto de los símbolos de *L*; (P₂) no puede haber tales criterios para un lenguaje privado.

Empecemos con P₁. El que la caracterización de una actividad como un lenguaje presuponga la existencia de criterios de corrección para el empleo de sus respectivas expresiones es algo sobre lo cual Wittgenstein insiste bastante en las observaciones sobre el seguimiento de reglas arriba mencionadas. Por ejemplo, en el párrafo 186 de las *Investigaciones* Wittgenstein se plantea de manera dramática la cuestión de cual es el criterio de corrección para la actividad de sumar 2. De nuevo en el párrafo 190 se pregunta por el criterio para la manera en que alguien comprende una fórmula algebraica que se le presenta. Esta manera de ver P₁ como siguiéndose de la discusión wittgensteiniana sobre el seguimiento de reglas lingüísticas corresponde a la lectura que hace Kripke de las *Investigaciones* en el libro citado,¹³ según la cual el argumento en contra

11. Hopkins 1975 y Hopkins 1999.

12. Kripke 1982.

13. Kripke 1982, prefacio y introducción.

del lenguaje privado es apenas un corolario de un argumento más general presente en las mencionadas observaciones sobre el seguimiento de reglas. Dicho argumento más general tiene como conclusión que no es posible seguir una regla lingüística privadamente.¹⁴ Esto porque si uno supone que es posible un tal seguimiento de reglas, entonces se genera la famosa paradoja según la cual una determinada regla lingüística privada siempre puede ser reinterpretada de manera que cualquier curso de acción esté de acuerdo o en conflicto con ella.¹⁵ Y como esto significa que para tal regla no hay criterios de corrección y de incorrección para el empleo de la misma, la idea de una regla lingüística a ser seguida privadamente es absurda. Lo importante aquí es el énfasis en que un lenguaje cualquiera requiere de criterios de corrección para el empleo de sus respectivas expresiones.

Fodor¹⁶ y otros¹⁷ han sugerido que la exigencia wittgensteiniana de que haya criterios públicos para el uso correcto o incorrecto de palabras presupone la tesis verificacionista de que la propia posibilidad de la distinción uso correcto/uso aleatorio depende de que otros hablantes o el mismo hablante estén en posesión de estos criterios.

La acusación me parece injusta ya que, en primer lugar, se trata del manejo de un lenguaje, o sea, de un fenómeno intencional en el sentido fuerte. Esto quiere decir: un fenómeno intencional en el sentido en que la actividad de hablar un lenguaje involucra no solamente condiciones

14. Wittgenstein 1953, § 202.

15. Wittgenstein 1953, § 201.

16. En Fodor 1975, pp. 70-1.

17. Por ejemplo, Barry Stroud en Stroud 1983. Pero también Jules Alfred Ayer en Ayer 1954 y Anthony Kenny en Kenny 1971. El artículo de Edward Craig (Craig 1997) es ambiguo con relación a si el argumento wittgensteiniano presupone el verificacionismo o no. Jim Hopkins en Hopkins 1974-75 y Kripke en Kripke 1982, parecen ambos sostener que algún tipo de verificacionismo está presente en el argumento en contra del lenguaje privado en la medida en que sugieren que la noción de corrección del uso lingüístico apela esencialmente a la manera como una comunidad de hablantes actúa lingüísticamente. Kripke dice explícitamente que en las *Investigaciones* "Wittgenstein propone una concepción del lenguaje basada, no en condiciones de verdad, sino en condiciones de asertibilidad o condiciones de justificación: ¿bajo que condiciones nos es permitido hacer un determinado enunciado?" (Kripke 1982, p. 74; mi propia traducción).

de corrección como también condiciones acerca de las cuales tiene sentido suponer que los seres capaces de manejar tal lenguaje son sensibles a ellas en tanto se les puede observar corrigiendo su propia conducta lingüística y la de sus semejantes.

La acusación de verificacionista en contra de P_1 me parece injusta, en segundo lugar, porque la posibilidad de la distinción entre usos lingüísticos correctos e incorrectos, tal y como la propone Wittgenstein, de hecho no presupone la existencia de otros hablantes, además del propio hablante, capaces de hacer uso de los criterios de corrección (que tornan posible tal distinción) para evaluar las conductas lingüísticas del hablante. En contra de lo que afirma Ayer,¹⁸ por ejemplo, Wittgenstein sí considera aprobatoriamente la posibilidad de un lenguaje solitario, esto es, un lenguaje usado por solamente un hablante. El pasaje relevante es el siguiente:

Un hombre puede animarse a sí mismo, darse órdenes, obedecerse, censurarse, castigarse, formularse una pregunta y responderla. Se podría también imaginar incluso hombres que hablasen sólo en monólogo. Acompañarían sus actividades hablando consigo mismos. –Un explorador que los observase y escuchase sus discursos podría conseguir traducir su lenguaje al nuestro. (Estaría por ello en situación de predecir correctamente las acciones de estas personas, pues también las oíría formar resoluciones y decisiones.) (Wittgenstein 1953e, § 243)

En mi opinión, la gran mayoría de los comentadores no ha prestado suficiente atención a este aforismo. No obstante, él presenta el caso límite de la posibilidad de un lenguaje; bautizemos el lenguaje que aparece en este pasaje lenguaje solitario para diferenciarlo del lenguaje privado. Según Wittgenstein, el lenguaje solitario es posible pues, pese a que no hay nadie observando la conducta lingüística y no lingüística del monologante, tal conducta es observable. Si alguien estuviera observando el monologante, podría quizás extraer criterios de corrección de la misma. En el lenguaje privado, por el contrario, dicha posibilidad de observación está excluida de antemano; pues en este caso simplemente no hay conducta a ser observada. Volveremos a este punto en la discusión de P_2 . Nos basta por el momento la siguiente conclusión: la posibi-

18. En el mencionado artículo.

lidad de trazar la distinción entre usos lingüísticos correctos e incorrectos no requiere la existencia de una comunidad de hablantes con relación a la cual se puede establecer criterios de corrección para las expresiones de este lenguaje; el ejemplo del lenguaje solitario demuestra esto.

Además, Wittgenstein tampoco se compromete con la idea de que el propio hablante, al usar su lenguaje, lo hace siempre de manera reflexiva, es decir: siempre evaluando de manera consciente su praxis lingüística. Veamos el siguiente pasaje en que se discute el empleo por un hablante de la expresión “yo tengo un dolor”:

“Cuando digo ‘Tengo un dolor’ estoy en cualquier caso justificado *ante mí mismo*.” –¿Qué quiere decir esto? ¿Quiere decir: “Si otro pudiera conocer lo que llamo ‘dolor’, convendría en que empleo la palabra correctamente”?

Usar una palabra sin justificación no quiere decir usarla injustamente.

No identifico, ciertamente, mi sensación mediante criterios, sino que uso la misma expresión. Pero con ello no *termina* el juego de lenguaje; con ello comienza. (Wittgenstein 1953e, §§ 289-290)

Este y otros pasajes¹⁹ muestran que el acceso de primera persona al contenido de las expresiones lingüísticas y a su uso no siempre están mediados por reglas o criterios de corrección, y un caso particularmente relevante en que esto no se da es el del empleo de las palabras que designan sensaciones. Mi conclusión del análisis de P_1 es que la objeción de que ella presupone algún tipo de verificacionismo es completamente infundada.

Pasemos entonces al examen de P_2 . Wittgenstein la justifica mostrando que cualquier intento para establecer un criterio de corrección para el uso de los términos del lenguaje privado²⁰ está condenado al

19. Por ejemplo, Wittgenstein 1953, §§ 211, 217.

20. Algunos comentadores como Kenny (Kenny 1971) y Craig (Craig 1997) juzgarán conveniente distinguir dos sentidos de privacidad: un sentido epistemológico y otro metafísico. Un ítem mental es privado en el primer sentido cuando solamente el poseedor de tal ítem puede tener acceso a él; se dice del ítem en cuestión que es incomunicable. En contrapartida, un evento mental es privado en el segundo sentido cuando solamente su poseedor puede tenerlo; dicho evento mental se dice inalienable. Por ejemplo, alguien podría afirmar que solamente yo puedo tener mi dolor, lo que sería bastante discutible ya que hay casos como lo de los gemelos siameses en que un gemelo podría en princi-

fracaso justamente porque, como no hay la posibilidad de una perspectiva externa sobre los eventos de usar un símbolo de este lenguaje supuestamente asociado a una determinada sensación y el de tener tal sensación, tampoco tiene sentido hablar de algo objetivo que correspondería a la regularidad *uso del símbolo/ocurrencia de la sensación correspondiente*. Por cierto, la concepción de objetividad que se maneja aquí requiere por lo menos la posibilidad de una perspectiva independiente (la de la tercera persona) de la perspectiva de la primera persona, pero esto no implica, como ya vimos en la discusión de P₁, una concepción comunitarista de la objetividad. Según el comunitarismo, la objetividad del uso lingüístico está constituida por la coincidencia entre la actividad lingüística del hablante y la de un observador suyo. Sin embargo, uno podría considerar la posibilidad de la observación como indispensable para la objetividad de tal uso sin por esto exigir una coincidencia entre lo que hacen el observador y un intérprete suyo. Lo que sí es necesario en ambas posturas es que tanto los actos lingüísticos del hablante como sus actos no lingüísticos sean observables.

Wittgenstein defiende su segunda premisa (P₂) usando ejemplos como el del diario para la ocurrencia sistemática de una sensación. La situación es la siguiente. Una persona registra en un diario un determinado símbolo 'S' cada vez que le ocurre una cierta sensación. La sugerencia es la de que la conexión entre 'S' y su respectiva sensación se establece por medio de una definición ostensiva privada: el hablante concentra su atención en la sensación y usa 'S'. Tal procedimiento de ostensión privada no establece un criterio de corrección para el uso de 'S', nos dice Wittgenstein, porque con él no se nos ofreció una manera independiente de comprobar la covariancia entre el uso de 'S' y el representar con tal uso la sensación mencionada; en otras palabras, no se nos ofreció en la actividad del que compone el diario un hecho observable.²¹ ¿Qué es lo que nos garantiza, por ejemplo, que la ostensión privada conecta realmente 'S' a la sensación S y no a cualquier otra?²² A la cual el defensor del lenguaje privado podría responder apelando a la

pio sentir el dolor del otro. De todas maneras, el sentido de privacidad relevante para la presente discusión es claramente el primero.

21. Wittgenstein 1953, § 265.

22. Aquí aparece la pregunta típica del escepticismo kripkeano (ver Kripke 1982).

memoria: hay un acto inicial de bautizo de S y en cada uso futuro de 'S' el hablante se recuerda de este acto inicial. Pero, insiste Wittgenstein, ¿y cual es criterio para el correcto recuerdo de la conexión entre 'S' y S?²³

El problema es que continuamos dentro del círculo de las nociones mentales sin que se haya establecido la conexión entre estas nociones y sus correspondientes representados en el mundo o entre ellas y las acciones que evidencian la atribución correcta de estas nociones a un sujeto. Por ejemplo, la relación entre el pensar un objeto (la sensación S) y tal objeto o la relación entre la comprensión de 'S' y su uso para referirse a S o mismo la relación entre la memoria de los usos pasados "correctos" de 'S' y tales usos.

El caso del diario ejemplifica un lenguaje cuyos términos poseen sentido y referencia –para usar la terminología de Frege– privados: la definición ostensiva privada no es algo que sea posible comunicar. Sin embargo, Wittgenstein también considera un lenguaje privado en el cual los sentidos de los términos son públicos mientras que sus referencias son privadas.²⁴ Llamémosle lenguaje de escarabajos, ya que estos son los objetos privados a que se refieren los términos de este lenguaje; cada hablante se refiere al usar el término 'escarabajo' a su propio escarabajo privado. Pero supuestamente la manera como se establece la relación de referencia entre la palabra 'escarabajo' y el hipotético escarabajo de cada hablante no es privada; hablante y oyente logran comunicarse por medio del uso por el primero de la palabra 'escarabajo'. Pero entonces el contenido de lo comunicado no puede tener que ver con el mismo objeto escarabajo ya que no hay ninguna garantía de que los escarabajos privados de cada hablante sean idénticos, es decir: no hay garantía de que los hablantes de este lenguaje estén hablando de la misma cosa. Aparentemente, esto es lo que estaría pasando en el caso del lenguaje privado de sensaciones; es plausible afirmar que cada hablante tiene una sensación de rojo, por ejemplo, o de dolor de muelas estrictamente distinta de los demás. ¿Cómo es posible entonces comunicarse usando un tal lenguaje si no hay manera observable de saber si cada hablante conecta correctamente el símbolo para sensación con el tener dicha sensación? La conclusión wittgensteiniana es finalmente que si el sentido es público, entonces no es posible decir que tal sentido

23. Wittgenstein 1953, § 265.

24. Esto está entre los párrafos 273 y 293 de Wittgenstein 1953.

nos lleva a una referencia privada. La palabra 'escarabajo' no puede funcionar como nombre de un objeto.

Los dos casos —el del lenguaje con sentido y referencia privados y el del lenguaje con sentido público y referencia privada—²⁵ ilustran según Wittgenstein la imposibilidad de un lenguaje privado. Pero, ¿no habría una tercera posibilidad de lenguaje privado, a saber: un lenguaje con sentido privado y referencia pública, no considerado por el autor de las *Investigaciones*? Este parece ser justamente el caso del lenguaje de pensamiento propuesto por Fodor. En la sección siguiente examinaremos esta posibilidad.

4 Wittgenstein en contra del lenguaje de pensamiento

Como hemos visto, el lenguaje privado de Fodor tiene que tener por lo menos el mismo poder expresivo que los lenguajes públicos que aprendemos. Si no fuera así, no podríamos aprender, nos dice, estos últimos. Además, no sería posible, contra lo que sostiene Fodor, traducir de un lenguaje público al código interno. Pero si los lenguajes públicos son traducibles en el lenguaje privado entonces este último debe describir la misma realidad que los primeros. Y, ¿cómo se establecen las relaciones entre las expresiones del lenguaje de pensamiento y las cosas en el mundo? En otras palabras: ¿cómo se constituyen los criterios de corrección para los términos de este lenguaje?

De lo que Fodor dice se puede concluir que el empleo del lenguaje de pensamiento no está al nivel consciente, es decir: no pertenece al mismo nivel de las actitudes proposicionales, sensaciones, dolores y los demás estados mentales. En este sentido, ni del propio organismo al cual se le está atribuyendo un código interno se puede decir que conoce los criterios de corrección de los símbolos de su propio código interno. ¿En que sentido entonces el mentalés es privado? Justamente en tanto que el sentido de sus expresiones no es público; como lo pone Fodor, en este lenguaje no hay criterios públicos para la aplicación de sus términos.²⁶ Sin embargo, hay criterios y su establecimiento está ligado,

25. Esta manera de clasificar los casos posibles de lenguaje privado (lenguaje con sentido y referencia privados y lenguaje con sentido público y referencia privada) me la sugirió la lectura del artículo de Hopkins (Hopkins 1974-75).

26. Fodor 1975, p. 69. Véase también Fodor 1998a, capítulo 6.

en la época del *LOT*, a la manera como el organismo que maneja el código interno fue formado por la naturaleza con la capacidad para responder a determinados estímulos. En esta fase, el mecanismo de selección natural es el responsable por el ajuste fino de los criterios de aplicabilidad de los símbolos del lenguaje de pensamiento.

Como se dijo al final de la sección 2, más recientemente Fodor se ha desencantado del modelo inductivista de adquisición del lenguaje y también con la explicación evolutiva para la fijación de los criterios de corrección de los símbolos del mentalés justamente porque, según él, tal explicación no puede dispensar dicho modelo inductivista.²⁷ El desencanto del modelo inductivista de adquisición conceptual también está ligado, según el mismo Fodor, con el rechazo de cualquier concepción cognitivista de la posesión de conceptos. En *LOT*, Fodor sostenía que poseer un concepto *C* consiste en conocer la condición de verdad asociada al uso del predicado *P* que expresa *C*.²⁸ Si la posesión de un concepto cualquiera se explica en términos cognitivos, entonces debe haber por lo menos en el caso de los conceptos primitivos una relación evidencial entre el contenido del conocimiento en cuestión (en nuestro caso, la condición de verdad asociada a *P*) y las instancias empíricas que lo apoyan (los casos en que un hablante ha usado *P* y se pudo observar el cumplimiento de su condición de verdad). Tal exigencia correspondería tal vez a la mejor manera de dar cuenta de la justificación epistémica del juicio de que el hablante conoce la condición de verdad asociada a *P*. Por lo tanto, si abandonamos cualquier concepción cognitivista de la posesión conceptual, la motivación para adoptar un modelo inductivista de la adquisición también desaparece; o por lo menos esto es lo que sostiene Fodor en sus últimos escritos.²⁹

En *LOT*, Fodor afirma que hay evidencia empírica de que el lenguaje de pensamiento está siendo usado de manera correcta; según él, tal hi-

27. Por ejemplo, en Fodor 1998a, capítulo 6. Según Fodor, si elegimos el mecanismo de selección natural como el responsable por la determinación de las normas de corrección para el uso de los símbolos del mentalés de todas formas tendremos que explicar la adquisición de tales normas por parte de los organismos. Y el candidato natural a dar cuenta de dicha adquisición de normas lingüísticas es el mecanismo de aprendizaje inductivo.

28. Fodor 1975, capítulo 2.

29. Por ejemplo, en Fodor 1998a, capítulo 6.

pótesis se podría sugerir como la mejor explicación de la coherencia integral de la vida mental de un organismo.³⁰ Ahora bien, si podría argumentar que la propia idea de que el organismo tiene una vida mental y maneja un lenguaje público ya es una hipótesis; ella misma se apoya en un complejo evidencial conductual; ella misma se propone como la mejor explicación para la conducta observable del organismo. Digamos que esta última hipótesis es una teoría sobre el agente y hablante; una teoría que pretende describir su conducta como motivada por la manera como representa el mundo y por sus propósitos y aspiraciones; además, una teoría que pretende ofrecer criterios para la evaluación de tal conducta como estando constituida por acciones lingüísticas y no lingüísticas. El hablante usa esta teoría para guiar su conducta; el oyente impone una teoría semejante sobre la conducta del hablante para entenderlo. Si la comunicación es exitosa, sus respectivas teorías deben ser bastante semejantes entre sí.³¹

Entre estos dos puntos de vista sobre lo que necesita de explicación hay obviamente una gran diferencia, pero no me interesa aquí discutir si lo que merece ser explicado prioritariamente es la conducta del organismo o la complejidad y coherencia de su vida mental. Lo que me parece más relevante para la presente discusión es la cuestión de si Fodor es capaz de refutar la tesis wittgensteiniana de que no hay criterios de corrección para el uso de palabras en un lenguaje privado. Bien, la relación entre nuestro uso de símbolos y las actitudes proposicionales que están correctamente asociadas a ellos es mediada en el caso del lenguaje público por nuestras prácticas o convenciones lingüísticas. Un hablante del español que profiere la oración "tengo hambre" está en general expresando su creencia de que tiene hambre. Sobre este punto creo que estarían de acuerdo Fodor y Wittgenstein.³² Pero, ¿qué podría desempeñar el papel de estas convenciones en el lenguaje de pensamiento?

Según el Fodor de la época de *LOT*, hay una relación de necesidad nomológica entre los símbolos del código interno y lo que representan,

30. Ver Fodor 1975, p. 70.

31. Todo esto es parte del enfoque davidsoniano de interpretación radical. Al respecto, ver a Davidson 1973; Davidson 1974, entre otros.

32. Ver por ejemplo Fodor 1975, pp. 71-72 y Wittgenstein 1953, §§ 198, 202, 337, 355.

y tal relación supuestamente se establece por un proceso de adaptación de las reacciones del organismo a ciertos estímulos del medio en que está inserido.³³ Como vimos, esta última era la explicación darwinista de la adquisición conceptual privilegiada por Fodor en *LOT*.

La posición de Fodor sobre la posesión conceptual y por lo tanto sobre la corrección o incorrección de nuestras representaciones ha cambiado bastante. Su teoría más reciente es que poseer una representación conceptual que se aplica a un determinado conjunto de objetos consiste en la existencia de una relación de determinación metafísica entre las cosas representadas y nuestras representaciones de ellas, que se puede expresar de la siguiente manera:

en cualquier mundo metafísicamente posible en que hay mentes como las nuestras y en que los objetos posean típicamente las mismas propiedades que poseen en este mundo, una propiedad X de tales objetos va a estar por fuerza ligada a (*locked to*) nuestro concepto X.³⁴

Esta es la concepción no-cognitivistica de Fodor acerca de la posesión conceptual. Obsérvese que la relación entre representación y mundo aquí es mucho más fuerte que la propuesta en la época de *LOT*. La causalidad o el mecanismo de selección natural podían explicar el establecimiento de la relación de necesidad nomológica entre representación y representado. Estos mismos mecanismos ya se muestran impotentes para explicar la fijación de la relación de necesidad metafísica entre un determinado y los objetos a los cuales se aplica correctamente. Entraremos en detalle a continuación.

En mi opinión, hay por lo menos tres problemas con estas dos nociones de corrección para las expresiones del lenguaje de pensamiento.³⁵ El primer problema afecta la primera, aun que probablemente no afecte la segunda. El problema es el siguiente: el hecho de que representación y representado estén nomológicamente relacionados no es capaz de dar cuenta de la relación normativa que hay entre nuestras representaciones lingüísticas y las cosas que representan. La suposición de que hay una relación de necesidad nomológica entre las representaciones y las

33. Fodor 1975, pp. 70-71.

34. Ver, por ejemplo, Fodor 1998a, pp. 137, 157.

35. La noción presente en *LOT* y la más reciente.

cosas que representan es equivalente a la idea de que representación y representado covarían causalmente; *grosso modo*, siempre que percibimos un gato formamos la representación GATO. Habría una ley que conecta los dos tipos de eventos. Pero, como Fodor mismo ha reconocido en sus escritos posteriores,³⁶ la covariancia causal no puede explicar la posibilidad de la representación falsa. Por ejemplo, es un hecho que ciertos animales muy parecidos a los gatos también causan en nosotros la representación GATO, que representa incorrectamente estos animales. Pero si GATO covaria causalmente con gatos o animales muy parecidos con los gatos entonces no podemos afirmar que la aplicación de GATO a estos últimos animales sea errónea.

El segundo problema con las nociones fodorianas de corrección de uso conceptual afecta más bien su noción más reciente de corrección, a saber: la idea de que la aplicación correcta de un concepto es aquella que concuerda con el nexo metafísico entre dicho concepto y una determinada propiedad. Pero, uno podría preguntar ¿cómo se justifica la afirmación de que hay tal relación de necesidad metafísica entre los símbolos del lenguaje de pensamiento y el mundo? Ciertamente, la existencia de tal relación no se puede justificar en base a la evidencia empírica; lo máximo que tal evidencia puede apoyar es la hipótesis de que existe una ley causal que conecta el concepto y su respectiva propiedad. Sin embargo, la conexión entre el concepto MANIJA y la propiedad *ser una manija* es constitutiva en el sentido en que *ser una manija* está constituido por las experiencias con objetos que son manijas. Si es así, entonces la relación entre tales experiencias en tanto intencionalmente descritas y el concepto MANIJA es metafísicamente necesaria, o sea: estas experiencias están necesariamente ligadas (en el sentido metafísico) a dicho concepto.

Parece que el misterio sobre como una dada propiedad en el mundo y un determinado contenido conceptual primitivo (por ejemplo, *ser una manija* y el concepto MANIJA) se conectan finalmente ha desaparecido. No obstante, la apariencia de solución del misterio es engañosa. Pues, la pregunta es: ¿qué es lo que constituye el contenido de un determinado concepto? Y la respuesta de Fodor es simplemente ésta: un concepto hereda su contenido del contenido de las experiencias que se encuentran metafísicamente ligadas a él; las experiencias de objetos *en tanto*

36. Por ejemplo, en Fodor 1987, cap. 4 y en Fodor 1998a, caps. 6 y 7.

manijas están necesariamente ligadas al concepto MANIJA. Pero entonces obviamente la pregunta sobre la constitución del contenido de un determinado concepto se transfiere automáticamente para las experiencias ligadas a él por necesidad metafísica. ¿Qué es lo que constituye el contenido de una experiencia de algo *en tanto* manija? Decir que propiedad y contenido coinciden en el caso de los conceptos y experiencias primitivos no responde a nuestra pregunta. Pues, queremos caracterizar las nociones de corrección y incorrección del uso de conceptos y otras representaciones con contenido (incluidas las experiencias) pero la identificación de una propiedad de objetos externos a nosotros con experiencias de un cierto tipo ciertamente no permite establecer la distinción apariencia/realidad que es crucial en el argumento wittgensteiniano. De acuerdo con la concepción de Fodor, no sería posible tener una experiencia equivocada de un objeto *en tanto* manija. Por la misma razón, tampoco sería posible una experiencia correcta de una cosa concebida como manija.

El tercer problema con la propuesta de Fodor afecta sus dos nociones de corrección. El problema es que los criterios de corrección que gobiernan el lenguaje de pensamiento ni son públicos ni están a la disposición del organismo que emplea tal lenguaje. Esto significa que un observador externo no puede corregir el usuario de este lenguaje y tampoco lo puede hacer el mismo usuario. Pero si es así, entonces ¿en qué sentido éstos se pueden llamar criterios de corrección para el mentalés? ¿No es esencial al uso de un lenguaje la posibilidad de aprobarse o corregirse a sí mismo o a los otros?

Fodor parece no ver ningún problema con esto. Él da el ejemplo de los lenguajes de computación para ilustrar su punto de que la noción de corrección no exige que el propio usuario del lenguaje sea capaz de determinar si su uso de los términos de este lenguaje es correcto o no; lo único que se exige es que ellos de hecho sean usados correctamente.³⁷ Pero aunque la computadora misma no sea capaz de corregir su conducta lingüística, su programador lo puede hacer y de hecho lo hará para construir una máquina que haga lo que él quiere. Si el papel del programador es desempeñado por el mecanismo de selección natural en el caso de los organismos que operan un lenguaje de pensamiento como en *LOT*, entonces sí se podría quizás recuperar alguna idea de co-

37. Fodor 1975, pp. 70-71.

rrección.³⁸ Sin embargo, esta no sería una noción de corrección tan robusta como la que se aplica a los lenguajes públicos que usamos, ya que no hay para el mentalés conciencia de los criterios de corrección de sus términos. Por lo tanto, no hay ni la posibilidad del hablante de guiar su propia conducta lingüística de acuerdo con tales criterios, ni la posibilidad de corregirse a sí mismo intencionalmente cuando él cree que se ha desviado de la norma lingüística, ni tampoco la posibilidad de que otros lo corrijan cuando lo mismo ocurre. La idea de corrección que es esencial al lenguaje natural no se puede aplicar al lenguaje de pensamiento, y esto significa que este último no es un lenguaje en el mismo sentido en que el primero lo es.

Las perspectivas de solución de este tercer problema no son mejores si se adopta la más reciente noción de corrección conceptual propuesta por Fodor. Recordemos que la nueva propuesta acerca de la posesión de un concepto es no-cognitivist; esto quiere decir que poseer un concepto no se identifica con ningún elemento cognitivo teórico o práctico. Si Fodor tiene razón, entonces es posible poseer un concepto sin conocer los criterios para la correcta aplicación del mismo. Pero esto es justamente lo que el argumento wittgensteiniano muestra que es imposible. La conclusión a que llegamos es por lo tanto la misma: el mentalés no es un lenguaje en la medida en que la noción de corrección del uso lingüístico no es caracterizable para él.³⁹

38. Esto me parece bastante discutible. Uno ciertamente podría hablar de la corrección de la conducta de un organismo en términos de su adecuación a las exigencias del medio en que vive, en el sentido en que la corrección sería premiada con una ventaja en términos de sobrevivencia de la especie y la incorrección sería castigada en los casos más extremos con el desaparecimiento de los organismos que exhibieran tales capacidades para las mencionadas conductas inadecuadas. El problema es que la diversidad de las posibles conductas lingüísticas es tan grande que es difícil imaginar como la corrección o incorrección de cada una de ellas pueda ser explicada en términos de su adecuación o inadecuación a las exigencias del medio. Un interesante propuesta en esta dirección es la que aparece en Millikan 1984.

39. Un posible objetor podría decir: "pero lo único que has mostrado es que Fodor y Wittgenstein manejan dos conceptos distintos de lenguaje: un concepto formal y otro natural de lenguaje. No obstante, no nos has convencido de que el concepto manejado por Wittgenstein es el más robusto y, por lo tanto, el que debemos adoptar." La distinción nos remite claramente a David Lewis en su

5. La Dispensabilidad del Lenguaje de Pensamiento

Paso ahora a mi segundo argumento contra el mentalés: el de que la hipótesis de un lenguaje de pensamiento es dispensable. El argumento consiste en mostrar que se puede explicar los hechos evocados en cada uno de los 3 argumentos mencionados en la sección 2 en favor del mentalés sin que sea necesario postular la existencia de un tal lenguaje. Recordemos que el primer argumento de Fodor llevaba a la conclusión de que sin un código interno no es posible explicar el hecho de que los procesos mentales exhiben propiedades causales y al mismo tiempo propiedades semánticas inferenciales. Esto me parece un non-sequitur. El contra-ejemplo quizás más conocido es el monismo anómalo de Donald Davidson. De acuerdo con Davidson, los eventos mentales son físicos (tesis monista) pero no hay leyes estrictas que relacionan las propiedades mentales a las propiedades físicas (tesis de la anomalía de lo mental). Los eventos mentales tienen eficacia causal tanto sobre otros eventos mentales (la creencia de que hoy es lunes y el recuerdo de que tengo una cita el lunes a las 6 de la tarde con el dentista causan *ceteris paribus* mi creencia de que voy a estar en el dentista a las 6 de la tarde)

"Languages and Language" (Lewis 1975). En este artículo, Lewis traza una distinción entre una noción de lenguaje como una función que asigna significados a secuencias de palabras o sonidos (el concepto formal) y otra noción de lenguaje como un fenómeno social por medio del cual unas personas profieren ciertas secuencias de sonidos o escriben ciertas secuencias de marcas y otras personas responden a tales proferencias o inscripciones con pensamientos o con acciones (el concepto natural). El propio Lewis afirma que estos dos conceptos de lenguaje son incompletos y que la noción integrada de lenguaje requiere la unión de dichos conceptos a través de la idea de convención lingüística. Este último concepto es lo que permite afirmar que una determina población habla un determinado lenguaje (una determinada función de secuencias que va de palabras a significados). Si Lewis tiene razón, entonces el lenguaje también es un conjunto de conductas intencionales y, por lo tanto, un fenómeno gobernado por las normas que conectan ciertos estados mentales intencionales y dichas conductas. Ahora bien, el argumento de Wittgenstein lo que muestra es que no puede haber normas del uso lingüístico para un lenguaje privado. Esto significa que el concepto de lenguaje manejado por Fodor es incompleto porque no logra dar cuenta de las normas asociadas al fenómeno esencialmente intencional que es el de hablar un lenguaje.

como sobre eventos que no tienen ninguna descripción mental (que la aparición de un perro ladrando enfrente de mí me haga creer que hay un perro feroz delante de mí).

Sin embargo, mientras que la teoría computacional explica la causalidad entre estados mentales en términos de las relaciones sintácticas inferenciales entre estos estados concebidos como símbolos, y por lo tanto, en términos de las relaciones semánticas inferenciales entre ellos, la concepción davidsoniana da cuenta de la causalidad mental apelando a los contenidos de los estados relacionados causalmente. Tales contenidos son identificados por oraciones de los lenguajes públicos que usamos para comunicarnos. Para Davidson como para el segundo Wittgenstein, sin el lenguaje público no habría mentalidad (*mentality*) del tipo de la mentalidad humana; no es necesario un lenguaje de pensamiento para individuar los estados mentales por su contenido. Por tanto, las relaciones causales entre los estados que constituyen nuestros procesos mentales están determinados, también según Davidson, por las relaciones entre los contenidos de estos estados.

Uno podría preguntar: pero ¿cómo es posible aprender un lenguaje si uno ya no poseyera un lenguaje innato? Lo que nos lleva a la crítica al segundo argumento de Fodor a favor del mentalés. Bajo la hipótesis de que aprender un lenguaje es formular hipótesis y confirmarlas obviamente no sería posible aprender un lenguaje sin ya estar en posesión de un lenguaje. Y como la idea de que todo lenguaje es aprendido nos lleva a un regreso al infinito, es necesario suponer que hay un lenguaje innato. Pero, ¿qué nos obliga a suponer que todos los casos de aprendizaje de un lenguaje son casos en que se formula y confirma una hipótesis en el lenguaje que uno ya domina? ¿No podría pasar que el aprendizaje del primer idioma ejemplificara un modelo esencialmente distinto del modelo sugerido por Fodor? Wittgenstein sugiere algo en esta dirección en las *Investigaciones*.⁴⁰ Según él, uno aprende su primer lenguaje a través de ejemplos y de práctica; el enseñante exhibe al aprendiz varias cosas que ejemplifican el mismo concepto; le pide que reconozca ciertos patrones, que continúe ciertas series, que emplee expresiones lingüísticas de una determinada manera; el aprendiz lo hace bien o mal; el instructor le dirige el aprendizaje por medio de expresiones de acuerdo o desacuerdo con lo que hace el aprendiz. Así una persona captura

40. Párrafo 208.

los conceptos de su primer idioma. Si esto es correcto, entonces no tenemos por que aceptar que la existencia del lenguaje de pensamiento es condición de posibilidad del aprendizaje del primer lenguaje público que manejamos. También como ya se dijo al final de la sección 2, el propio Fodor ya no cree en este argumento en favor del innatismo del mentalés ya que el mecanismo de aprendizaje inductivo ya no le parece la única manera de explicar nuestra adquisición de un lenguaje. Además, su posición actual es que el mentalés es adquirido aunque no sea aprendido de manera consciente.⁴¹ En este sentido más blando, aún se puede decir que el mentalés es innato.⁴²

Pasemos entonces al tercer y último argumento mencionado en la sección 2 a favor del lenguaje de pensamiento. Recordemos que este argumento tiene como conclusión que debe existir un código interno a partir de la necesidad de una explicación (talvez la única posible) de la productividad y sistematicidad del pensamiento humano. De acuerdo con Fodor, si las actitudes proposicionales son estructuradas como un lenguaje (es decir: como relaciones entre un sujeto y las oraciones del código interno), es fácil dar cuenta de la posibilidad de actitudes tan complejas cuanto se quiera y también del hecho de que las capacidades asociadas a tales actitudes no se les puede atribuir a un organismo de manera aislada, sino holísticamente.

Pero, ¿y si decimos que las actitudes proposicionales son productivas y sistemáticas porque su contenido presenta estas propiedades, y que este contenido exhibe productividad y sistematicidad porque se constituye de conceptos que son composicionales, y finalmente que tales conceptos se constituyen de capacidades que no pueden ser atribuidas de manera aislada, etc.? Fodor probablemente diría que esto no nos muestra como estas actitudes pueden hacer parte de la arquitectura de un organismo de tal manera que por sus propiedades semánticas éstas puedan tener efectividad causal en los procesos mentales. Y es cierto que de acuerdo con el modelo davidsoniano de explicación de la mentalidad humana, las actitudes proposicionales no son ellas mismas parcialmente lingüísticas; pero bajo su descripción mentalista tienen un contenido que es identificado de manera lingüística. Este lenguaje no es más que un lenguaje público, por medio del cual un observador que es

41. Sobre esto, vease Fodor 1998a, capítulo 6.

42. Como Fodor lo afirma explícitamente en Fodor 1998b.

capaz de hacer inteligible el poseedor de estas actitudes llega a expresarlas correctamente. El modelo explicativo davidsoniano me parece bastante convincente pero claramente no asocia los procesos mentales con determinados mecanismos a nivel sub-personal (por ejemplo, mecanismos computacionales como querría Fodor); éste nos responde solamente a la cuestión de la naturaleza de las actitudes y del contenido lingüístico. Sin embargo, el modelo contiene una explicación de la productividad y de sistematicidad del pensamiento. Este último exhibe estas dos propiedades porque comparte con el lenguaje público la propiedad de expresar contenidos lingüísticos.

6 Conclusión

Finalmente, me parece importante señalar que he pretendido en este artículo refutar la tesis del lenguaje de pensamiento. En este sentido, el argumento en contra del lenguaje privado fue una herramienta fundamental. Pero además intenté también motivar la tesis de la dispensabilidad del lenguaje de pensamiento a través de la crítica a los 3 argumentos presentados por Fodor a favor de dicho lenguaje. Con relación a este mi segundo objetivo, reconozco que un trabajo más abarcador requeriría un análisis de todos los argumentos a favor del mentalés. Sin embargo, esto lo tengo que dejar para una futura investigación.⁴³

Universidad Autónoma del Estado de Morelos
México

BIBLIOGRAFÍA

AYER, A. J. 1954: "Can There Be a Private Language?" in *Aristotelian Society Proceedings*. Supplementary vol. 28, pp. 63-76. Reprinted in Canfield, J. 1986: *The Private Language Argument* (London: Garland Publishing, Inc.).

43. Deseo expresar mi agradecimiento al anónimo réferi de la *Revista Latinoamericana de Filosofía* por sus excelentes comentarios a una primera versión del texto y a mi amigo y colega José Miguel Esteban por su profunda y metódica revisión del español.

- BLOCK, N. 1998: "Classical Computationalism" in Grayling, A. 1998: *Philosophy 2: further through the subject* (Oxford: Oxford University Press)
- CARNAP, R. 1934: *The Unity of Science* (London: Keagan Paul)
- CRAIG, E. 1997: "Meaning and Privacy" in Hale, B. & Wright, C. 1997
- DAVIDSON, D. 1973: "Radical Interpretation" in Davidson 1984
- DAVIDSON, D. 1974: "Belief and the Basis of Meaning" in Davidson 1984
- DAVIDSON, D. 1984: *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press)
- FODOR, J. 1975: *The Language of Thought* (Hassocks: the Harvester Press, 1976)
- FODOR, J. 1980: "Methodological solipsism considered as a research strategy in cognitive psychology" in *The Behavioral and Brain Sciences* vol. 3, pp. 63-109
- FODOR, J. 1985: "Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade-Mecum" in *Mind* vol. 94, pp. 76-100. Reprinted in Warfield, T. & Stich, S. 1994: *Mental Representation* (Oxford: Blackwell)
- FODOR, J. 1987: *Psychosemantics* (Cambridge: MIT Press)
- FODOR, J. 1994: "Fodor, Jerry A." in Guttenplan, S. 1994: *A Companion to the Philosophy of Mind* (Oxford: Blackwell)
- FODOR, J. 1998a: *Concepts* (Oxford: Clarendon Press)
- FODOR, J. 1998b: "Do We Think in Mentalese? Remarks on Some Arguments of Peter Carruthers" in Fodor, J. 1998c: *In Critical Condition* (Cambridge: the MIT press)
- HALE, B. y WRIGHT, C. 1997: *A Companion to the Philosophy of Language* (Oxford: Blackwell)
- HOPKINS, J. 1974-1975: "Wittgenstein and Physicalism" in *Proceedings of the Aristotelian Society* vol. 75, pp. 121-146
- HOPKINS, J. 1999: "Wittgenstein, Davidson and Radical Interpretation" in Hahn, F. 1999: *The Library of Living Philosophers: Donald Davidson*. (Chicago: University of Chicago Press)
- KENNY, A. 1971: "The Verification Principle and the Private Language Argument" in Jones, O.R. 1971
- KRIPKE, S. 1982: *Wittgenstein on Rules and Private Language* (Cambridge: Harvard University Press)
- JONES, O. R. 1971: *The Private Language Argument* (London: Macmillan)
- LEWIS, D. 1975: "Languages and Language" in Gunderson, K. 1975: *Minnesota Studies in the Philosophy of Science* Vol. VII, pp. 3-35. Re-

- printed in Lewis, D. 1983: *Philosophical Papers* vol. I (Oxford: Oxford University Press)
- MILLIKAN, R. 1984: *Language, Thought and other biological Categories* (Cambridge: MIT press)
- RUSSELL, B. 1918: "The Philosophy of Logical Atomism" in Russell 1956
- RUSSELL, B. 1956: *Logic and Knowledge* (Worcester: Billing & Sons Ltd.)
- STROUD, B. 1983: "Wittgenstein's "Treatment" of the Quest for "A Language which describes my inner Experiences and which only I myself can understand"" in *Proceedings of the 7th International Wittgenstein Symposium Kirchberg 1982* (Vienna: Hölder-Pichler-Tempsky)
- VILLANUEVA, E. 1979: *El Argumento del Lenguaje Privado* (México: UNAM)
- WITTGENSTEIN, L. 1953: *Philosophical Investigations* (Oxford: Basil Blackwell)
- WITTGENSTEIN, L. 1953e: *Investigaciones Filosóficas*. Trads. Alfonso García Suárez y Ulises Moulines (Barcelona: Editorial Crítica).

ABSTRACT

In several of texts of his and in particular in *The Language of Thought* (1975), Jerry Fodor has defended the hypothesis that there is a language of thought. Within the context of the computational model of the mind proposed by Fodor, this hypothesis explains the fact that mental processes consist of causal relations between mental representations (propositional attitudes) and also of semantical relations between these attitudes' propositional contents. Moreover, the mentioned hypothesis accounts for the systematicity and the potentially infinite productivity of human thought. However, the necessarily private character of such a language constitutes a serious obstacle to the acceptance of the Fodorian hypothesis at least for somebody who is convinced by Wittgenstein's famous argument against the possibility of a private language. In this paper, I propose to clarify the extent to which Wittgenstein's argument refutes Fodor's hypothesis and offers some suggestions as to how the above-mentioned facts –the causal and the semantical relations among the attitudes as well as the systematicity and productivity of thought– could be explained in the absence of a language of thought.

BUNGE, LA LÓGICA INDUCTIVA Y EL GRADO DE CONFIRMACIÓN DE LAS TEORÍAS

Alejandro Cassini

1. La definición de Bunge del grado de confirmación

La lógica inductiva, o más precisamente, la teoría de la confirmación, es frecuentemente objeto de malentendidos e interpretaciones equivocadas. Mario Bunge no ha escapado a ello. Bunge ha hecho muchas contribuciones importantes a la filosofía de la ciencia, aunque nunca ha tratado con detalle el problema de la confirmación de teorías. Las pocas páginas que en su extensa producción dedica a este tema contienen afirmaciones que no son muy afortunadas. En esta breve nota quisiera hacer algún aporte que permita clarificar qué es la lógica inductiva y, en particular, qué es el grado de confirmación de una hipótesis y cómo se calcula.

En una obra reciente de Mahner y Bunge sobre filosofía de la biología leemos lo siguiente:

"Los confirmacionistas afirman que unas pocas excepciones apenas importan, mientras que los refutacionistas sostienen que son decisivas. Los primeros piensan en el grado de confirmación de una hipótesis como en la razón del número de casos favorables al número total de datos. Y, debido a la analogía formal entre esta razón y la definición de Laplace de la probabilidad, identifican a estas dos y a su vez a ambas con el grado de verdad." (Mahner y Bunge 1997, trad esp. pp. 148-149).

Esta es una caracterización muy curiosa de la posición confirmacionista. Tal como yo entiendo el asunto, los confirmacionistas actuales no hacen ninguna de las afirmaciones que Bunge y Mahner les atribuyen. Piensan, en cambio, que las excepciones sí importan, hasta el punto de que una sola consecuencia observacional falsa de una teoría o hipótesis reduce su grado de confirmación a cero. Además, no definen el grado de confirmación de una hipótesis como la razón entre el número de

contrastaciones favorables y el número total de datos. Esta definición de Bunge y Mahner es del todo incompatible con la idea confirmacionista del grado de confirmación, como veremos enseguida. No conozco ninguna obra confirmacionista donde se ofrezca semejante definición. Bunge y Mahner, por su parte, no la ilustran con alguna cita o referencia. Los confirmacionistas efectivamente identifican el grado de confirmación de una teoría con su grado de probabilidad, pero no lo hacen sobre la base de una analogía con la definición de la probabilidad de Laplace. Por último, no identifican el grado de confirmación de una teoría con su grado de verdad.

Creo que el origen de todos estos errores se encuentra en la propia concepción de Bunge del grado de confirmación de una hipótesis. En su amplio tratado *Scientific Research*, cuya primera edición es de 1967, Bunge ofrece la siguiente definición:

Sea $\{t_i\}$ el conjunto de las consecuencias contrastables de h derivadas y contrastadas hasta el momento, y $E = \{e_i\}$ el conjunto de informes empíricos relevantes para h , o sea, comparables, directa o indirectamente, con los t_i . Supongamos que se han realizado N contrastaciones, $n \leq N$ de las cuales son favorables a h , y el resto, $N - n$, desfavorables a h . [...] El grado de confirmación empírica de h respecto de la información empírica $E = \{e_i\}$ puede definirse así:

$$C(h | E) = n / N$$

Si $n = 0$, o sea, si todos los datos empíricos obtenidos hasta el momento son discordantes con los t_i , implicados por h , entonces $C(h | E) = 0$. Si todas las contrastaciones practicadas hasta el momento concuerdan con los t_i correspondientes [...], entonces $C(h | E) = 1$. En todos los demás casos $C(h | E)$ se encontrará entre 0 y 1." (Bunge 1967, trad. esp. p. 739).¹

Aquí está claro qué entiende el propio Bunge por grado de confirmación de una hipótesis. Lo extraño es que atribuya esta definición a todos los confirmacionistas en general, cuando nadie, aparte de él mismo, parece haberla suscrito. Bunge extrae en este pasaje las consecuencias más evidentes de su definición. Tales consecuencias son incompatibles con la lógica inductiva actual, esto es, con la teoría bayesiana de la confirmación. De las tres afirmaciones que hace Bunge sólo la primera

1. Este pasaje se escribió originalmente en 1967, pero aparece sin cambios en la tercera edición corregida del libro de Bunge, publicada en el año 2000, por lo que entiendo que representa su posición actual.

es cierta desde el punto de vista bayesiano. No cabe duda de que si todos los datos obtenidos son incompatibles con las consecuencias de una teoría o hipótesis, el grado de confirmación de dicha teoría o hipótesis es igual a 0. Las dos siguientes afirmaciones de Bunge, que son consecuencias inevitables de su definición, son falsas en la lógica inductiva bayesiana. Si todas las contrastaciones practicadas hasta un momento determinado concuerdan con las predicciones deducidas de una hipótesis, el grado de confirmación de esa hipótesis será, en general, un número mayor que 0 y menor que 1 (y no exclusivamente 1, como afirma Bunge). En todos los restantes casos el grado de confirmación será exactamente 0 (pero nunca un número entre 0 y 1, como dice Bunge).

Bunge agrega que la definición de grado de confirmación de los confirmacionistas no satisface los axiomas del cálculo de probabilidades (Bunge 1967, trad. esp. pp. 739-740). Esta afirmación, en verdad, sólo se aplica a su propia definición del grado de confirmación, pero no es cierta, como veremos, respecto de la noción bayesiana de grado confirmación.

En el apartado que sigue doy un esquema elemental de la teoría bayesiana de la confirmación y pruebo los resultados que justifican las afirmaciones que acabo de hacer en el párrafo anterior. Antes de entrar en esos asuntos técnicos, me parece pertinente señalar la implausibilidad de la idea que Bunge tiene del grado de confirmación. De acuerdo con la definición de Bunge, una teoría que haya sido contrastada en 1000 ocasiones, de las cuales 100 de ellas son desfavorables, tendría un grado de confirmación igual a 0.9. Esto es, aunque haya 100 contraejemplos verificados, la teoría tendría un grado de confirmación bastante cercano a la certeza. Esta es una consecuencia lo suficientemente inaceptable como para mostrar que la definición de Bunge del grado de confirmación debe abandonarse sin más. Confirmacionistas y refutacionistas coinciden en sostener que en un caso como este la teoría en cuestión está refutada desde el punto de vista lógico y, por consiguiente, su grado de confirmación debe ser igual a 0. El sentido común, si es lícito apelar a algo tan poco determinado, indudablemente apoya esta última conclusión.

2. Un esquema de la lógica inductiva bayesiana

Antes de 1945 la lógica inductiva prácticamente no existía y el concepto mismo de confirmación apenas había sido elucidado. Ese año tanto Carnap como Hempel publicaron una serie de artículos pioneros en los cuales analizaban el concepto cualitativo de confirmación y propo-

nían métodos para medir el grado de confirmación de las hipótesis (Hempel 1945, Hempel y Oppenheim 1945, Carnap 1945). En 1950 Carnap publicó *Logical Foundations of Probability*, que se considera habitualmente como la obra fundacional de la lógica inductiva. El programa de Carnap, basado en la interpretación lógica de las probabilidades, pronto encontró limitaciones de alcance y dificultades técnicas que lo hicieron poco viable. Aunque algunas de esas dificultades se solucionaron posteriormente, cuando Carnap murió, en 1970, su programa de lógica inductiva ya había sido abandonado por la mayoría de los filósofos de la ciencia. Mientras tanto, se produjo la emergencia de la lógica inductiva bayesiana, que aplicó a la teoría de la confirmación diversos elementos de la inferencia estadística bayesiana, desarrollada, entre otros, por Jeffreys, Good, Savage y Lindley. Este es un programa de investigación que, como tal, tiene éxitos y anomalías. Constituye, no obstante, la mejor versión disponible de la lógica inductiva y, por esa razón, es el punto de vista que adoptaré en mi discusión. La obra de John Earman *Bayes or Bust*, publicada en 1992, es la que ha establecido los términos de la discusión actual sobre lógica inductiva. Aquí sólo puedo dar un esbozo de algunos aspectos positivos de la teoría bayesiana de la confirmación, en particular, de aquellos que son necesarios para justificar mi crítica a las ideas de Bunge. Para mayores detalles, especialmente respecto de los problemas del bayesianismo, que son muchos y muy profundos, remito al lector a la notable obra de Earman.

La lógica inductiva se propone encontrar una medida del grado de confirmación de una hipótesis o teoría respecto de la evidencia disponible. Esto es, busca definir una función confirmación C que asigne un número real a cada hipótesis o teoría (o, más precisamente, a cada triada ordenada $\langle H, E, K \rangle$) que represente el grado en el cual una hipótesis se encuentra confirmada por determinada evidencia. Sea H una hipótesis o conjunto de hipótesis, E un enunciado observacional que describe una determinada evidencia y K el conocimiento aceptado en un momento dado. La expresión $C(H, E \& K) = r$ afirma que el grado de confirmación de la hipótesis H respecto de la evidencia E y el conocimiento antecedente K es igual al número real r . Todas las lógicas inductivas actualmente viables sostienen que el grado de confirmación es una medida probabilística del apoyo empírico. Se basan, por consiguiente, en la siguiente identidad:

$$[1] \quad C(H, E \& K) = P(H / E \& K)$$

Esta ecuación afirma que el grado de confirmación de H respecto de E y K es igual a la probabilidad condicional de H dados E y K . Puesto que H , E y K son proposiciones, o conjuntos de proposiciones, otra manera de leer la ecuación es la siguiente: el grado de confirmación de H respecto de E y K es igual a la probabilidad de que H sea verdadera dado que E y K son verdaderas. La función confirmación C se identifica, entonces, con la función probabilidad P tal como es usualmente caracterizada en el cálculo axiomático de probabilidades, es decir, en el de Kolmogorov o en alguna otra axiomatización equivalente.

El sistema axiomático de Kolmogorov está formulado en el lenguaje de la teoría de conjuntos, pero la lógica inductiva requiere que se lo formule en el lenguaje de la lógica proposicional clásica. Ambas formulaciones son equivalentes y sus respectivos lenguajes son intertraducibles. La formulación axiomática habitual de la teoría de la probabilidad en el lenguaje proposicional es la siguiente:

Sea $\mathcal{E}$ un conjunto de proposiciones cerrado respecto de las combinaciones finitas veritativo-funcionales. La probabilidad P es una aplicación de $\mathcal{E}$ en el conjunto de los números reales ($P: \mathcal{E} \rightarrow \mathbb{R}$) que satisface los siguientes axiomas:

Axioma 1: para toda proposición $\alpha \in \mathcal{E}$,
 $0 \leq P(\alpha) \leq 1$.

Axioma 2: para toda proposición $\alpha \in \mathcal{E}$ tal que $\models \alpha$,
 $P(\alpha) = 1$.

Axioma 3: para cualesquiera proposiciones $\alpha, \beta \in \mathcal{E}$ tales que $\models \neg(\alpha \& \beta)$,
 $P(\alpha \vee \beta) = P(\alpha) + P(\beta)$.

Axioma 4: para cualesquiera proposiciones $\alpha, \beta \in \mathcal{E}$ tales que $P(\alpha) > 0$ y $P(\beta) > 0$,
 $P(\alpha \& \beta) = P(\beta) \times P(\alpha / \beta)$.²

2. El sistema axiomático de Kolmogorov no introduce la probabilidad condicional como término primitivo, sino como término definido, lo cual permite

Los axiomas 3 y 4 se pueden escribir de una manera más general para una secuencia finita o infinita enumerable de proposiciones. Se obtiene, así, la teoría de la probabilidad con aditividad finita o contable, respectivamente. Para los fines de este trabajo, sin embargo, no resulta necesario apelar a estas formulaciones generales.

En el curso de nuestra argumentación necesitaremos emplear cuatro teoremas de la teoría de la probabilidad que, por razones de espacio, no demostraremos aquí. Son los siguientes:

Teorema 1: $P(\neg \alpha) = 1 - P(\alpha)$.

Teorema 2: $P(\neg \alpha / \beta) = 1 - P(\alpha / \beta)$.

Teorema 3: $\models (\alpha \leftrightarrow \beta) \rightarrow P(\alpha) = P(\beta)$.

Teorema 4: $P(\alpha / \beta) = P(\alpha) \times P(\beta / \alpha) / P(\beta)$.

El primero es el teorema del complemento, que afirma que la probabilidad de una proposición es igual a 1 menos la probabilidad de su contradictoria. El segundo es este mismo teorema para probabilidades condicionales. El tercero es el teorema de equivalencia, que afirma que dos proposiciones lógicamente equivalentes tienen la misma probabilidad. El cuarto es la forma más simple del teorema de Bayes, que luego explicaremos más adelante.

Todos los teoremas sobre probabilidades que se deduzcan de estos axiomas son también teoremas acerca del grado de confirmación de hipótesis. Por otra parte, toda interpretación verdadera de la teoría de la probabilidad, también será una interpretación verdadera, un modelo, de la teoría de la confirmación. La lógica inductiva no está necesariamente ligada, como a veces se cree, a una determinada interpretación

prescindir del axioma 4. Sin embargo, tanto De Finetti como la mayor parte de los bayesianos prefieren adoptar este último axioma y evitar la definición de probabilidad condicional. En la axiomatización que presento sigo esta última formulación. Los sistemas resultantes son lógicamente equivalentes. El axioma 4, en efecto, es obviamente equivalente a la definición de probabilidad condicional de Kolmogorov, que, formulada en el lenguaje proposicional, es la siguiente: $P(\alpha / \beta) = \text{def } P(\alpha \& \beta) / P(\beta)$. Los tres restantes axiomas son comunes a ambos sistemas, aunque Kolmogorov los formula en el lenguaje de la teoría de conjuntos. (Sobre las ventajas de emplear la formulación axiomática que presento aquí cfr. Gillies 2000 pp. 65-66).

del cálculo de probabilidades, sino que, como la propia teoría abstracta de las probabilidades, admite, en principio, múltiples modelos.

La concepción más exitosa de la lógica inductiva ha sido hasta el momento la teoría bayesiana de la confirmación, la cual, en su versión ortodoxa, adopta la interpretación subjetiva o personal de las probabilidades elaborada por Bruno De Finetti. Sin embargo, no se sigue de allí que toda lógica inductiva basada en la identidad [1] deba comprometerse con una concepción subjetivista de las probabilidades. De hecho, no todos los bayesianos son personalistas. En las observaciones que siguen adoptaré la teoría bayesiana ortodoxa, con la advertencia de que todos los resultados acerca de la confirmación que se deduzcan exclusivamente de los axiomas 1-4 de la teoría de la probabilidad admiten interpretaciones diferentes de la de De Finetti.

En la teoría bayesiana personalista las probabilidades se interpretan como los grados de creencia de cada sujeto individual. La expresión $P(\alpha / K) = r$ debe leerse como si afirmara que r es el grado de creencia de un determinado individuo en la verdad de la proposición α dadas las creencias contenidas en K , que acepta como verdaderas. De Finetti probó en 1937 que esta interpretación satisface los axiomas del cálculo de probabilidades (con aditividad finita). Se acepta desde entonces que la interpretación en términos de grados de creencia es un modelo de la teoría matemática de la probabilidad (para mayores detalles véase Gillies 2000, Cap. 4).

La epistemología bayesiana se basa en los supuestos de que la creencia admite grados y de que los grados de creencia de un sujeto se pueden representar mediante números reales. Esta representación numérica es una idealización, que incluso puede tomarse como una ficción útil. A estos supuestos le agrega dos principios fundamentales. El primero es que un sistema racional de grados de creencia debe ser coherente, es decir, satisfacer todos los teoremas de la teoría matemática de la probabilidad. El segundo es que los grados de creencia deben actualizarse a la luz de la experiencia mediante la regla de condicionalización. La justificación de estos dos principios ha sido objeto de una enorme bibliografía que no puede comentarse aquí. La regla de condicionalización es una regla de inferencia inductiva que determina cómo cambiar la función probabilidad cuando se adquiere información proveniente de nuevas evidencias empíricas. Sea P la función probabilidad de un sujeto antes de que se conozca la evidencia E y P' la función probabilidad del mismo sujeto después de que E se ha conocido. La regla de

condicionalización estricta determina cómo se relacionan estas dos funciones de probabilidad, bajo los supuestos de que E se ha verificado (y tiene, por tanto, probabilidad 1) y de que las proposiciones a las que se aplica la función inicial P son inciertas (es decir, no tienen probabilidad 0 ni 1). La regla de condicionalización estricta se puede enunciar, entonces, de esta manera:

$$P'(\alpha / K) = P(\alpha / E \& K), \text{ dado que } P'(E / K) = 1.$$

La aplicación de esta regla de condicionalización garantiza la conservación de la coherencia de un sistema de grados de creencia.

El componente específicamente bayesiano de la lógica inductiva aparece en relación con esta regla. De acuerdo con la teoría de la probabilidad, $P(\alpha / E \& K) = P(\alpha / K) \times P(E / \alpha \& K) / P(E / K)$. Este es el famoso teorema de Bayes (el teorema 4 de nuestra lista anterior), que se deduce directamente del axioma 4 de la teoría de la probabilidad (no daremos la demostración aquí). La regla de condicionalización es, entonces, una aplicación del teorema de Bayes para cambiar la función probabilidad de un sujeto individual.

El uso de la regla de condicionalización presupone el proceso no inferencial en el cual se pasa de una probabilidad inicial $P(E / K) < 1$ a la probabilidad final $P'(E / K) = 1$. En términos menos formales, este es el proceso por el cual se verifica la evidencia E. La lógica inductiva bayesiana no se ocupa del proceso epistemológico de verificación del conocimiento observacional. Simplemente, acepta que la evidencia es cierta o está verificada. A nadie se le escapa que los datos de que se dispone en la práctica científica son falibles y rara vez se encuentran verificados. Para manejar los casos en que la evidencia es incierta, la regla de condicionalización se puede generalizar de esta manera:

$$P'(\alpha / K) = \sum_{i=1}^n P(\alpha / E_i \& K) \times P'(E_i / K).$$

Para el caso más simple, en el que la partición de los $E_i = \{E, \neg E\}$, la regla adopta la siguiente forma: $P'(\alpha / K) = P(\alpha / E \& K) \times P'(E / K) + P(\alpha / \neg E \& K) \times P'(\neg E / K)$. Esta generalización de la regla de condicionalización la propuso Richard Jeffrey en 1965 (véase Jeffrey 1983, Cap. 11) y tiene, entre otras, la virtud de que se reduce a la regla de condicionalización simple o estricta cuando $P'(E / K) = 1$. La regla de Jeffrey permite la condicionalización cuando la evidencia es incierta, es decir,

cuando $0 < P'(E / K) < 1$. Además, el cambio de la función probabilidad de $P(E / K)$ a $P'(E / K)$ resulta reversible, mientras que en el caso de la condicionalización bayesiana estricta es irreversible (puesto que $P'(E / K) = 1$).

La teoría bayesiana de la confirmación emplea el teorema de Bayes para calcular el grado de confirmación de una hipótesis o teoría. Aquí me limitaré al caso de hipótesis o teorías deterministas, aunque la lógica inductiva bayesiana también es aplicable a las teorías probabilistas o estadísticas. En el caso de hipótesis deterministas el uso de este teorema está sujeto a los siguientes requisitos: a) $\{H, K\} \models E$; b) $0 < P(H / K) < 1$; y c) $0 < P(E / K) < 1$. El primero es la condición de implicación, característica del método hipotético-deductivo, que es necesaria para la confirmación de hipótesis o teorías deterministas. El segundo y el tercer requisito afirman que respecto del conocimiento antecedente K tanto la hipótesis H como la evidencia E son inciertas, es decir, no se conoce con certeza si son verdaderas o falsas.

Con el fin de simplificar los puntos específicos que quiero mostrar, ignoraré el uso de hipótesis auxiliares y supondré que las hipótesis y teorías que consideramos tienen por sí mismas consecuencias observacionales, es decir, implican una determinada evidencia E. El teorema de Bayes, aplicado a la confirmación de una única hipótesis o conjunción de hipótesis H adopta la siguiente forma:

$$P(H / E \& K) = \frac{P(H / K) \times P(E / H \& K)}{P(E / K)}$$

Aquí $P(H / K)$ es la probabilidad previa de la hipótesis H, relativa a la evidencia disponible K, antes de que se conozca que E es verdadera. $P(H / E \& K)$ es la probabilidad posterior de H, es decir, la probabilidad de H después que se ha verificado E. $P(E / H \& K)$ es llamada *likelihood*, y representa la probabilidad de E suponiendo que H es verdadera. $P(E / K)$, por último, es la llamada *esperabilidad*, o sea, la probabilidad de que E sea verdadera independientemente de que H sea verdadera o falsa.

Se sigue del primer requisito que $P(E / H \& K) = 1$, dado que H implica lógicamente a E. Por consiguiente, en este caso, cuando H es una hipótesis universal o determinista, el teorema se simplifica y toma la forma:

$$P(H / E \& K) = \frac{P(H / K)}{P(E / K)}$$

Antes de proseguir debemos probar que efectivamente si una hipótesis implica una determinada evidencia, la probabilidad de la evidencia dada esa hipótesis es igual a 1. Este es un caso específico de un teorema más general según el cual, si una proposición implica a otra, la probabilidad de la segunda dada la primera es igual a 1. (En esta y las siguientes demostraciones asumimos que $\alpha = H$ y $\beta = E$. Además, para simplificar las fórmulas, suprimimos el conocimiento antecedente K).

Teorema 5: $(H \models E) \rightarrow P(E / H) = 1$

Prueba:

1. $(H \models E)$ [Supuesto].
2. $\models (H \rightarrow E)$ [De 1 por teorema de la deducción].
3. $\models \neg(H \& \neg E)$ [De 2 por lógica proposicional].
4. $P(\neg(H \& \neg E)) = 1$ [De 3 y Axioma 2].
5. $P(H \& \neg E) = 0$ [De 4 y Teorema 1].
6. $P(E / H) = P(E \& H) / P(H)$ [De Axioma 4].
7. $\models H \leftrightarrow ((H \& E) \vee (H \& \neg E))$ [Teorema de lógica proposicional].
8. $P(H) = P((H \& E) \vee (H \& \neg E))$ [De 7 y Teorema 3].
9. $P(H) = P(H \& E) + P(H \& \neg E)$ [De 8 y Axioma 3].
10. $P(H) = P(H \& E)$ [De 5 y 9].
11. $P(E / H) = P(H) / P(H)$ [De 6 y 10].
12. $P(E / H) = 1$ [De 11].
13. $(H \models E) \rightarrow P(E / H) = 1$ [De 1 y 12 cancelando el supuesto] •

La noción bayesiana de confirmación consiste en el incremento del grado de confirmación de una hipótesis o teoría cuando se descubre una nueva evidencia. Se la define habitualmente del siguiente modo:

Confirmación incremental: La evidencia E confirma incrementalmente a la hipótesis H respecto del conocimiento antecedente K si y sólo si, $P(H / E \& K) > P(H / K)$.

La lógica inductiva bayesiana incorpora algunas de las ideas cualitativas fundamentales del método hipotético-deductivo. La idea central del método hipotético-deductivo (en la versión confirmacionista de

Hempel, no en la versión falsacionista de Popper) es que una hipótesis se confirma cuando se verifica alguna de sus consecuencias observacionales. Esto es, si la hipótesis H implica lógicamente a la evidencia E, entonces, E confirma a H. La noción bayesiana de confirmación incremental permite justificar esta intuición. Siempre que una hipótesis implica una determinada evidencia, la verificación de esa evidencia aumenta el grado de probabilidad de la hipótesis. Podemos probarlo como un teorema:

Teorema 6: $(H \models E) \rightarrow P(H / E) > P(H)$.

Prueba:

1. $(H \models E)$ [Supuesto]
2. $(H \models E) \rightarrow P(E / H) = 1$ [Teorema 5]
3. $P(E / H) = 1$ [De 1 y 2 por lógica proposicional]
4. $P(H / E) = P(H) \times P(E / H) / P(E)$ [Teorema 4]
5. $P(H / E) = P(H) / P(E)$ [De 3 y 4]
6. $P(H / E) > P(H)$ [De 5]
7. $(H \models E) \rightarrow P(H / E) > P(H)$. [De 1 y 6 cancelando el supuesto] •

Como ya señalamos antes, este resultado se sigue bajo el supuesto de que $0 < P(H / K) < 1$ y $0 < P(E / K) < 1$. Estas son las dos condiciones que se requieren en general para aplicar de manera útil o no trivial el teorema de Bayes, dado que la condicionalización bayesiana no puede elevar las probabilidades previas iguales a 0 ni disminuir las que son iguales a 1.

La otra idea fundamental del método hipotético-deductivo es la de refutación. De acuerdo con ella, cuando una hipótesis implica al menos una consecuencia observacional falsa, dicha hipótesis está refutada desde un punto de vista lógico. La teoría bayesiana de la confirmación recoge este resultado de manera cuantitativa porque prueba que si una hipótesis implica una evidencia falsa, su grado de confirmación se reduce a 0. Podemos demostrarlo en el siguiente teorema:

Teorema 7: $(H \models E) \rightarrow P(H / \neg E) = 0$

Prueba:

1. $H \models E$ [Supuesto]
2. $(H \models E) \rightarrow P(E / H) = 1$ [Teorema 5]
3. $P(E / H) = 1$ [De 1 y 2 por lógica proposicional]
4. $P(E / H) = 1 - P(\neg E / H)$ [Teorema 2]

5. $P(\neg E / H) = 0$ [De 3 y 4]
6. $P(H / \neg E) = P(H) \times P(\neg E / H) / P(\neg E)$ [Teorema 4]
7. $P(H / \neg E) = 0$ [De 5 y 6]
8. $(H \models E) \rightarrow P(H / \neg E) = 0$ [De 1 y 7 cancelando el supuesto 1] •

La evidencia $\neg E$ refuta a la hipótesis H porque $P(H / \neg E \ \& \ K) = 0$. La refutación es el caso de disconfirmación máxima de una hipótesis por la evidencia. Una vez que una hipótesis ha sido refutada no puede adquirir ningún grado positivo de confirmación mediante condicionalización bayesiana. Ello es así porque si $P(H / E \ \& \ K) = 0$, entonces, $P(H / E \ \& \ E' \ \& \ K) = 0$ para cualquier nueva evidencia E' implicada por H . La única manera de salir de esta situación es o bien eliminar la evidencia refutadora $\neg E$, o bien modificar el conocimiento antecedente K de modo tal que no se cumpla la condición a). En cualquiera de esos dos casos es posible que el grado de confirmación de H vuelva a ser mayor que 0. Ambas estrategias implican la contracción del conjunto de proposiciones K , lo cual no puede hacerse dentro de un marco exclusivamente bayesiano.

Eliminar la evidencia refutadora $\neg E$ consiste en revisar las observaciones o experimentos que la han verificado, de manera que se encuentre algún error en esa verificación. Revisar el conocimiento antecedente K para evitar la refutación consiste en desactivar la implicación de la consecuencia E . Esto es, si $\{H, K\} \models E$, se debe encontrar un conjunto K' tal que $\{H, K'\} \models E$. De este modo, $\neg E$ deja de ser una evidencia refutadora para H . En los casos típicos en ciencia teórica, E se deduce de H, K y un conjunto de hipótesis auxiliares A . En tal situación, se buscará desactivar la implicación de E modificando alguno de los elementos de A , esto es, reemplazando A por otro conjunto de hipótesis auxiliares A' . Si ninguna de estas estrategias resulta exitosa, entonces, será necesario abandonar H , ya que ninguna evidencia ulterior podrá elevar su grado de confirmación por sobre 0.

3. Conclusión

La manera en que la lógica inductiva bayesiana concibe el grado de confirmación de una hipótesis es completamente diferente de la idea que Bunge tiene del grado de confirmación y que erróneamente atribuye a los confirmacionistas en general. El grado de confirmación de una

hipótesis o teoría no se identifica con su probabilidad sobre la base de una analogía con la definición de Laplace. En la lógica inductiva el grado de confirmación no tiene relación alguna con el cociente entre el número de contrastaciones favorables y el número de contrastaciones realizado en un momento determinado. Ese cociente no desempeña ningún papel en la manera de concebir o de calcular el grado de confirmación de una hipótesis o teoría. Lo único que la lógica inductiva puede decir al respecto es que si tal cociente es diferente de 1, el grado de confirmación, calculado mediante el algoritmo bayesiano, debe ser igual a 0.

El grado de confirmación asignado a una hipótesis es efectivamente la probabilidad de esa hipótesis relativa a la evidencia disponible. Pero esa probabilidad representa el grado de creencia o de confianza de un sujeto particular en la verdad de dicha hipótesis dado el hecho de que ese sujeto ha aceptado la evidencia en cuestión (o bien, en la versión de Jeffrey, dado que ha probabilizado la evidencia). Se trata del grado de creencia racional de cada individuo.

Finalmente, el grado de confirmación de una hipótesis no se identifica con el grado de verdad de esa hipótesis. La lógica inductiva no proporciona una teoría de la verdad parcial o de la verosimilitud. La probabilidad inductiva que se asigna a una hipótesis no puede interpretarse como la verosimilitud de esa hipótesis por el hecho de que es una probabilidad subjetiva o personal. La probabilidad subjetiva asignada a una proposición, que es siempre una probabilidad condicional, no expresa ninguna propiedad objetiva de la proposición misma, tal como sería la verosimilitud. Representa simplemente el grado de creencia de cada individuo en esa proposición, pero ese grado de creencia no puede juzgarse como la descripción de algún atributo o propiedad de la proposición. Los grados de creencia no son objetivamente verdaderos ni falsos, ni correctos o incorrectos en algún sentido del término, como reiteradamente señalara De Finetti. Conceptualmente el grado de confirmación de una proposición y la verdad parcial de esa proposición son cosas muy diferentes. Incluso aunque las probabilidades se interpreten de manera no subjetiva, como hacen los bayesianos objetivistas, no se sigue de ello que sea lícito identificar grado de confirmación y verdad parcial.

Las concepciones realistas de la ciencia, comprometidas con la idea de que las teorías empíricas son verosímiles o parcialmente verdaderas, necesitan ciertamente establecer algún vínculo entre confirmación y ve-

rosimilitud. En general, las teorías que tienen un grado de confirmación alto deberían ser más verosímiles que las que tienen un bajo grado de confirmación. Idealmente, el grado de confirmación y la verosimilitud deberían converger, de modo tal que una teoría empírica tenga un grado de confirmación igual a 1 si y sólo si, dicha teoría es verdadera. La lógica inductiva bayesiana no puede garantizar tal cosa. No hace falta decir que no sabemos cómo alcanzar este ideal. No obstante, aunque pudiéramos obtener una teoría que justifique esta relación entre grado de confirmación bayesiano y verosimilitud, ambos seguirían siendo conceptualmente diferentes. La lógica inductiva bayesiana es neutral respecto del problema de la verosimilitud y, en principio, es igualmente compatible con concepciones realistas y antirrealistas del conocimiento científico.

Conicet-Universidad de Buenos Aires

BIBLIOGRAFÍA

BUNGE, M. (1967), *Scientific Research*, Vols. I y II, Berlin-Heidelberg-New York, Springer. (Traducción española en un volumen: *La investigación científica*, Tercera edición corregida, México, Siglo XXI, 2000).

CARNAP, R. (1945), "On Inductive Logic", *Philosophy of Science*, 12, pp. 72-97.

CARNAP, R. (1950), *Logical Foundations of Probability*, Chicago, University of Chicago Press.

De Finetti, B. (1937), "La prevision: ses lois logiques, ses sources subjectives", *Annales de l'Institut Henri Poincaré*, 7, pp. 1-68. (Traducción inglesa: "Foresight: Its Logical Laws, Its Subjective Sources", en Kyburg, H. y H. Smokler (eds.), *Studies in Subjective Probability*, New York, Krieger, 1980, pp. 53-118.

Earman, J. (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge (Mass.), The MIT Press.

Gillies, D. (2000), *Philosophical Theories of Probability*, London-New York, Routledge.

Hempel, C. G., (1945), "Studies in the Logic of Confirmation", *Mind*, 54, pp. 1-26 y 97-121.

Hempel, C. G. y Oppenheim, P., (1945), "A Definition of 'Degree of Confirmation'", *Philosophy of Science*, 12, pp. 98-115.

Jeffrey, R. (1983) *The Logic of Decision*, Second edition, Chicago, University of Chicago Press. (Primera edición 1965).

Mahner, M. y M. Bunge (1997), *Foundations of Biophilosophy*, Berlin-Heidelberg-New York, Springer. (Traducción española: *Fundamentos de biofilosofía*, México, Siglo XXI, 2000).

ABSTRACT

Mario Bunge has been a persistent critic of inductive logic but, on my view, his arguments are far from being compelling. In the first part of this paper I contend that some of his criticisms are misdirected or simply erroneous. On the one hand, no inductive logician has sustained the theses that Bunge vaguely attributes to confirmationist philosophers. On the other hand, Bunge's definition of degree of confirmation is severely flawed in the light of Bayesian confirmation theory. In the second part of this paper I offer a brief outline of Bayesian inductive logic in order to support my criticism of Bunge's ideas.

Desarrollo Económico

Revista de Ciencias Sociales

COMITÉ EDITORIAL: Juan Carlos Torre (Director), Carlos Acuña, Luis Beccaria, Roberto Bouzas, Mario Damill, Juan Carlos Korol, Edith Obschatko, Juan Carlos Portantiero, Getulio E. Steinbach (Secretario de Redacción)

ISSN 0046-001X

Vol. 41

Octubre - diciembre de 2001

Nº 163

RAVI KANBUR: Política económica, distribución y pobreza: naturaleza de las discrepancias.

GARY W. COX y SCOTT MORGENSTERN: Legislaturas reactivas y presidentes proactivos en América Latina.

ERNESTO CALVO: Las transformaciones de los mercados laborales en América Latina: negociación colectiva y desigualdad salarial industrial.

GABRIEL L. NEGRETTO: Negociando los poderes del presidente: reforma y cambio constitucional en la Argentina.

GABRIELA SCHIAVONI: Economía del don y obligaciones familiares: los ocupantes agrícolas de Misiones y el debate *farmer-campesino*.

JOSÉ ANTONIO SÁNCHEZ ROMÁN: Tucumán y la industria azucarera ante la crisis de 1890.

CRÍTICA DE LIBROS

JOSÉ ANTONIO BORELLO: Cuando los economistas hablan del territorio... ¿con quién hablan?

LUCAS LÓPEZ DÁVALOS: De necesidades y contrametaforas.

Desarrollo Económico es indizada, con inclusión de resúmenes, en las siguientes publicaciones: *Current Contents* (SSCI, Institute for Scientific Information); *Journal of Economic Literature* (AEA); *Sociological Abstract* (Cambridge Scientific Abstracts); *International Bibliography of the Social Science* (British Library of Political and Economic Science y UNESCO). También en varias otras ediciones periódicas y en volúmenes especiales nacionales e internacionales, así como en diversos índices en versión electrónica.

DESARROLLO ECONÓMICO - Revista de Ciencias Sociales es una publicación trimestral editada por el Instituto de Desarrollo Económico y Social (IDES). Suscripción anual: R. Argentina, \$ 60,00; Países limítrofes, U\$S 68; Resto de América, U\$S 74; Europa, U\$S 76; Asia, África y Oceanía, U\$S 80. Ejemplar simple: U\$S 15 (recargos según destino y por envíos vía aérea). Más información en: desarrollo@ides.org.ar, o disponible en el Web site: www.ides.org.ar. Pedidos, correspondencias, etcétera, a:



Instituto de Desarrollo Económico y Social
Aráoz 2838 • C1425DGT Buenos Aires • Argentina
Teléfono: 4804-4949 • Fax: (54 11) 4804-5856
Correo electrónico: desarrollo@ides.org.ar

Revista Latinoamericana de Filosofía, Vol. XXVIII Nº 2 (Primavera 2002)

LA "MEMORIA PRODUCTIVA" ACERCA DE LA CONCEPCIÓN DE HEGEL DEL RECUERDO Y LA MEMORIA

Daniel Brauer

La cuestión de la memoria ha vuelto a estar en el centro de la discusión filosófica, ya sea motivada por la polémica en torno a la realidad del pasado desatada por el narrativismo extremo, ya sea algo antes, en el marco de una puesta al día de la fenomenología husserliana sobre el tema. En el primer caso me refiero al último libro de Paul Ricoeur: *La Mémoire, L'Histoire, L'Oubli*,¹ que ofrece además de las tesis del autor una visión de conjunto de la polémica, en el segundo caso pienso en el libro de Edward Casey: *Remembering. A Phenomenological Study*.² Pero el nombre de Hegel apenas si es mencionado. Más aún la concepción hegeliana sobre la memoria y el recuerdo difícilmente ocupe algún capítulo en un texto que recorra la historia del concepto.

El presente artículo se ocupa de un capítulo descuidado de la "Psicología" de Hegel.³ Me propongo mostrar, recurriendo a los textos de clases recientemente publicadas y a los escritos de la época de Jena, que en Hegel hay una concepción original tanto del recuerdo como de la memoria y que su teoría no ha sido desarrollada más que fragmentariamente. Lo interesante de su concepción desde la perspectiva contemporánea es el aspecto constructivo, "productivo" que Hegel atribuye tanto al recuerdo como a la memoria, así como el papel configurador que adquiere en ella el lenguaje.

Si bien Hegel ha desarrollado su teoría del conocimiento en forma sistemática en el marco de la *Enciclopedia de las Ciencias Filosóficas* (en las sucesivas versiones de sus ediciones de 1817, 1827 y 1830) principalmente

1. Paul Ricoeur: *La Mémoire, L'Histoire, L'Oubli*, Paris 2000.

2. Edward Casey: *Remembering. A Phenomenological Study*, Indiana-Bloomington 2000.

3. Salvo el trabajo de Hans-Friedrich Fulda mencionado más abajo, prácticamente no hay bibliografía sobre el tema.

en las secciones que llevan por título: *Fenomenología y Psicología*, su concepción de la facultades cognitivas no ha merecido un estudio y comentario equivalente o al menos comparable al de las otras partes del sistema.⁴ Las razones de esto son de distinta índole. Por un lado, los textos de la *Enciclopedia* (=Enz.) en general constituyen un discurso arduo y difícil de seguir. Los párrafos fueron pensados como enunciados programáticos o taxativos que serían aclarados una y otra vez en los cursos dictados por Hegel mismo, de ahí que los llamados agregados (*Zusätze*) –adjuntados por los editores siguiendo el orden de los párrafos sobre la base de apuntes de alumnos tomados en clase– resulten generalmente mucho más claros. Por otro lado, la *Fenomenología* de la *Enciclopedia*, aún cuando pretenda establecer en forma metódica el sentido de muchos temas que ya habían sido tratados desde una perspectiva genética y algo oscura en la *Fenomenología del Espíritu* de 1807, ha quedado eclipsada detrás de ésta obra genial que ha servido de fuente de inspiración de diversas teorías de seguidores y detractores. La *Psicología* en cambio, parece retomar el proyecto de la vieja *psicología racional* de la metafísica de Christian Wolff, y si bien da cuenta de las modificaciones necesarias para su rehabilitación, resulta poco atractiva luego de la demoledora crítica a que fuera sometida por Kant en su teoría de los *paralogismos* de la razón.

A esto debe agregarse el aparente carácter mecánico de la dialéctica que organiza la obra, una impresión que la dificultad y sequedad de los párrafos, destinados a servir de punto de partida del comentario oral, han contribuido a aumentar. Con todo, una lectura que dirija la mirada detrás de bambalinas descubrirá en los textos un discurso desparejo y no tan sistemático, en parte incluso conjetural.

En su tratamiento de la memoria y el recuerdo Hegel establece elementos de una concepción original que va mucho más allá de los antecedentes histórico-filosóficos de los que parte y señala hacia nuevos horizontes que siguen abiertos, a pesar de que no suele figurar en las recapitulaciones históricas de la teoría filosófica de la memoria.⁵

4. Según Herbert Schnädelbach la *Psicología* es la parte menos conocida del sistema, véase el volumen colectivo: *Hegels 'Enzyklopädie der philosophischen Wissenschaften' (1830)*, un comentario editado por Hermann Drüie, Annemarie Gethmann-Siefert, Christa Hackenesch, Walter Jaeschke, Wolfgang Neuser, y Herbert Schnädelbach, Francfort del Meno 2000, p. 13.

5. La bibliografía sobre el tema es muy escasa. Cabe mencionar en primer

Ante todo, resulta significativo pasar revista brevemente al contexto en que es ubicada su concepción de la memoria. Se trata de la segunda parte de la primera sección de la *Psicología*. Lo que Hegel llama aquí "espíritu teórico" podría ser traducido hoy por una teoría de la cognición o una "epistemología" en el sentido en que se usa la palabra en el ámbito anglosajón. El sujeto es pensado como "teórico", o sea como cognoscente a diferencia y en oposición a lo que sucede en la sección "b", cuyo tema bajo el título "el espíritu práctico", será el sujeto agente, en tanto capaz de iniciar acciones, tomar decisiones y actuar en general de modo intencional.

Hegel subdivide a su vez la sección "a", o sea "el espíritu teórico" en tres subsecciones: "a) La intuición", "b) La representación" y "g) El pensamiento". Nuestro tema concierne a la subsección beta: "representación".

Hegel trata en general a estos procesos cognitivos como una paulatina y cada vez más compleja elaboración y transformación de los datos sensibles. Pero mientras que en la parte dedicada a la "intuición" pueden percibirse con claridad huellas y variaciones del tratamiento que hace Kant de ella al comienzo de la *Crítica de la razón pura*, aún cuando esa conceptualización sea criticada, no se encuentra en los textos de Kant algo equivalente a lo que aparece en el capítulo siguiente. Me refiero a la mencionada subsección que tiene por título general "La representación", un título que no se repite además en el interior del capítulo mismo. El texto es dividido a su vez en: "1) El recuerdo", "2) La imaginación" y "3) La memoria". La "representación" no es tratada como una forma particular, sino que lo que nos re-presentamos es considerado directamente: o algo recordado, o imaginado o depositado en la memoria.

lugar: el artículo de Hans Friedrich Fulda, *Vom Gedächtnis zum Denken*, recogido en; *Psychologie und Anthropologie oder Philosophie des Geistes. Beiträge zu einer Hegel-Tagung in Marburg 1989*. editado por Franz Hespe y Burckhard Tuschling, Stuttgart 1991. Un comentario francamente superficial se encuentra en el libro de Iring Fetscher: *Hegels Lehre vom Menschen*, Stuttgart 1970, que pretende ser un comentario literal de toda la *Filosofía del Espíritu Subjetivo* de la Enz., véase p. 159 en adelante. Vittorio Hösle, en el segundo tomo de su *Hegels System*, Hamburgo 1988, véase p. 386 y sigs., señala la incoherencia de las transiciones que ofrece el texto, pero sin detenerse en el análisis de la teoría hegeliana del recuerdo y la memoria.

Lo más cercano al tratamiento hegeliano de la memoria no se encuentra por cierto en Kant sino, como señala Hans Friedrich Fulda,⁶ en el breve tratado de Aristóteles *Acerca de la memoria y la reminiscencia*, que nos ha sido legado como formando parte de los *Parva naturalia*⁷ - de un modo más directo aún que la asumida y de buena gana confesada influencia del *De anima*.⁸ Con todo, como veremos, la presencia de variaciones sobre temas kantianos no debería ser soslayada.

La tripartición que Hegel lleva a cabo de la sección: "representación" aquí en cuestión y el orden progresivo de las subsecciones resultan muy significativos. Ante todo, llama la atención el hecho que el autor (1) separe su tratamiento del recuerdo del de la memoria y (2) ponga entre ambos a la imaginación - cuando tendemos naturalmente a pensar que es más bien a la inversa, que es la imaginación la que debería presuponer la memoria para com-poner sus imágenes.

Al comienzo del parágrafo 251, con que se inicia el tratamiento de la "representación", esta es definida como una "intuición recordada" ("*ei-ne erinnerte Vorstellung*"):⁹

"La representación es la intuición recordada, el término medio entre el inmediato encontrar-se-determinado (*dem unmittelbaren Bestimm-t-sich-Finden*) de la inteligencia y la misma en su libertad, el pensamiento".

Con esto el texto parece hacer justicia a la formación de la palabra latina re-praesentatio (más que a la "*Vor-stellung*" germánica), ya que re-presentar algo, indica un volver a hacer ahora presente, a "presentar" -algo que se había dado en el pasado ("re-")- y no sólo hacerse una imagen de algo. En la "representación", tal como es expuesta aquí, está presente una huella mnémica o en todo caso, una remisión no explicitada al pasado.

Pero también es cierto que Hegel interpreta la palabra alemana equi-

6. Fulda, op. cit., p. 323 y sigs..

7. Aristóteles, *Tratados breves de historia natural*, traducción y notas de Alberto Bernabé Pajares, Madrid 1996, p. 115 y sigs.. El comentario y las notas que acompañan la traducción inglesa de Richard Sorabji resultan de gran utilidad: *Aristotle On Memory*, Worcester y Londres 1972.

8. En el § 378 Hegel considera que el texto aristotélico constituye la "única obra de interés especulativo sobre el tema", *Enz.*, op. cit., p. 11.

9. *Enz.*, op. cit., p. 257.

valente a "recuerdo" (*Erinnerung*), en un sentido literal-etimológico que difiere del de su uso coloquial. Se trataría, cómo sugiere el término alemán, de una "interiorización", "*Er-innerung*". Lo que se quiere decir con esto, si lo entiendo bien, es que recordar, no es sólo registrar algo sino al mismo tiempo una forma de apropiación del objeto o evento, y este es un punto decisivo, porque de acuerdo a esta concepción ya de entrada *el recuerdo transforma el modo de darse el objeto en la percepción*, lejos de copiarlo sin modificaciones - en el mismo acto de hacer abstracción de su materialidad y de incorporarlo al espacio y tiempo de la subjetividad.¹⁰ Con esta resignificación del término Hegel no pretende remplazar en realidad su sentido usual, sino más bien llamar la atención sobre el aspecto configurador del recuerdo.

En su concepción orgánica de los procesos cognitivos Hegel los analiza como formando parte de una progresiva serie de estructuras que se suponen unas a otras, pero que no pueden separarse en su aplicación. El acto mismo de percibir un objeto externo es al mismo tiempo ya una forma de registro de los datos en un esquema mnémico-cognitivo.

Del mismo modo que en la percepción misma interviene una estructura conceptual sin la cual, de acuerdo al precepto kantiano, ésta sería ciega -así por ej. no perciben de la misma manera el pasaje de un avión un ingeniero y un lego, puesto que el primero "ve" otras cosas que el segundo no puede "distinguir" (tipo de turbinas o de diseño aerodinámico, etc.), ya que no dispone de nociones que permitan reconocerlo -tampoco es que primero percibimos y después recordamos, como podría sugerir el sentido común e incluso el orden de los capítulos del texto, sino que percibir es visto como un proceso que consiste en transformar lo dado en recuerdo, es decir, en una forma de "representación" o si se quiere, en un lenguaje por cierto ajeno a Hegel, de codificar la simple presentación.

La "representación" es vista así como un "término medio" ("*Mitte*") -tal como surge de la cita de más arriba- entre lo que afecta al sujeto y el esquema conceptual bajo el que será subsumido. No puede dejar de reconocerse aquí detrás de la ardua terminología de Hegel la dicotomía kantiana entre lo sensible y lo inteligible, así como, aunque con modificaciones importantes, la teoría de la imaginación trascendental de Kant, que constituye junto con el texto de Aristóteles mencionado, la segun-

10. Véase *Enz.*, op. cit., § 448 y sigs, p. 249 en adelante.

da fuente de inspiración de la noción hegeliana del recuerdo y la memoria.

Se trata por cierto de una teoría esbozada en la *Crítica de la razón pura*, implícita en la célebre versión de la triple síntesis de la primera edición,¹¹ pero no desarrollada en forma sistemática por Kant mismo, de modo que para establecer el alcance y límites de su influencia ésta debe ser primero reconstruida, lo cual es una tarea aún pendiente.¹²

Lo cierto es que en el texto de Hegel ya la "representación" es vista como una "síntesis"¹³ a partir de un material sensible y algo así como un *protoconcepto*. Consecuentemente el conocimiento es descrito como un proceso gradual de depuración del elemento externo-sensible y de esta manera el "recuerdo" es considerado una re-presentación en el sentido de una inscripción interior de un dato sensible. Para Hegel la percepción es ya un acto de encodificación mnémica que ejerce una modificación en el modo de darse su objeto. De ahí el doble sentido del recuerdo: registro codificado o "engrama"¹⁴ y conversión en objeto de una percepción interior, independiente de la presencia física de lo recordado.

La sección dedicada a la "representación" comienza directamente con la noción de "recuerdo" como su forma más elemental. El tema central pasa a ser aquí la "imagen".

En el tratamiento hegeliano de la *Erinnerung* tiene lugar una curiosa simbiosis de la concepción aristotélica con la kantiana.

La puesta en relación entre memoria e imaginación aparece ya en el breve tratado aristotélico sobre la memoria y ha sido determinante para la historia del problema.¹⁵ Pero mientras que el "eikón"¹⁶ aristotélico es entendido como una "copia y recordatorio"¹⁷ del original, en Hegel la imagen misma aparece como formando parte de un proceso de "universalización". Es que la imagen ya es para él resultado de una selección, la fijación de lo que Husserl llama un *escorzo*, en la que determinados aspectos son enfatizados y otros dejados de lado: "La imagen ya no tiene la completa determinidad (*Bestimmtheit*) que tiene la intuición, ella es arbitraria o azarosa y está en general separada del lugar y del tiempo externos y del contexto inmediato en el que estaba".¹⁸

A diferencia de lo que sucede con la célebre metáfora aristotélica de la impronta o sello que dejaría la sensación para la producción del recuerdo,¹⁹ en la concepción de Hegel ella es presentada como algo "arbitrario y azaroso".

En efecto, por un lado la imagen lejos de ser "fiel" al original es ya una primera forma de abstracción en la que la cosa es transformada en un boceto que, por otro lado permite reconocer, sin embargo, sus rasgos esenciales. Así decimos acerca de un célebre cuadro de van Gogh que representa "unos" zapatos, no tales zapatos en su particularidad, más aún la particularidad es desdibujada para mostrar algo más universal. Ni siquiera la fotografía nos da una simple copia del original, ya que de lo que se trata en ella es de mostrar lo esencial o un aspecto "significativo" de lo acontecido, producto de una selección, en la mayoría de los casos entre otras imágenes posibles, llevada cabo desde el "ángulo" adecuado a fin de "hacer ver" lo que se quiere mostrar.

La imagen mimética ya es una simplificación y una reducción de la multiplicidad dada en la percepción. Pero aún así, su similitud con el original no es presentada por Hegel como su ventaja sino por el contrario, como un grado aún elemental de un progresivo proceso cognitivo de abstracción.

En la forma del recuerdo pensado como proceso de incorporación o

11. Véase *Kritik der reinen Vernunft*, ed. de Raymund Schmidt, Leipzig 1944, A 137 y sigs., p.196 y sigs...

12. Acerca del papel central de la imaginación trascendental en la construcción del conocimiento "como raíz de ambas ramas" (sensibilidad y entendimiento) ha llamado la atención Martín Heidegger en su: *Kant y el problema de la metafísica*, trad. al esp. de Gred Ibscher Roth, México 1954, véase esp. p. 119 en adelante. Pero como señala el autor, su interpretación de esta doctrina "se mueve en dirección opuesta a la del idealismo alemán", nota 196. Como quiera que fuese, la relación entre la imaginación, el recuerdo y la memoria, que es lo que nos concierne aquí, no son tratados.

13. *Enz.*, § 451, W. 10, p.257.

14. Tomo el término de la injustamente olvidada teoría de la memoria de Richard Semon por considerarlo útil, lo cual no significa asumir las tesis del autor - para una rehabilitación de la misma véase: Schachter, D. L., Eich, J. E. & Tulving, E.: *Richard Semon's Theorie of Memory*, en: *Journal of Verbal Learning and Verbal Behavior*, 17, pp. 721-743, 1978.

15. 450ª en adelante.

16. 450b20.

17. Según la trad. de A. Bernabé Pajares, p. 122, < 450b25>.

18. *Enz.* § 452, W. 10, p. 258-259.

19. 450 a30

codificación de un dato, el sujeto se apropia del objeto y lo hace parte de su interioridad: no sólo me acuerdo de algo sino que a la vez, me acuerdo de una imagen que como tal ya es parte de mi subjetividad.

Hegel habla aquí de un "pozo nocturno" e "inconsciente" (*nächtliche Schacht, bewusste Schacht*)²⁰ en el que las imágenes estarían depositadas. Se trata ahora de una interioridad pero que permanece aún ajena al poder del sujeto, su contenido y surgimiento escapan a su voluntad.

Este es el lugar en el que Hegel trata de lo que llama el "recuerdo propiamente dicho",²¹ título con el que paradójicamente se ocupa no tanto del recuerdo en sí. El tema es más bien la "representación", a la que no había dedicado hasta ahora ninguna sección.

En efecto, la "representación" es vista como una estructura cognitiva estrechamente ligada al recuerdo. Aquí el texto resulta particularmente oscuro. Lo que Hegel parece describir es un proceso que implica una duplicidad entre la imagen latente o registrada y una nueva "intuición" en la que esa imagen es reconfirmada a partir de la reiterada percepción del mismo objeto o su aparición contingente, de modo que se "reconoce" su identidad. Así, es el "reconocimiento" de una persona conocida entre cientos de miles lo que confirma lo correcto de su "representación" y hace posible a través de un proceso reiterado prescindir de la percepción al evocar la imagen en lugar de la cosa que representa.²² Este pasaje y su descripción de la formación de representaciones podría haber sido tomado directamente de un tratado empirista sobre el conocimiento.²³ Ambas concepciones tienen en común una epistemología realista. Lo que diferencia al "idealismo" Hegeliano tiene que ver más bien con su concepción de la imagen o del engrama con que el objeto es registrado y que Hegel entiende no como copia sino como formando parte de un paulatino proceso de *codificación simbólica*, de modo que la originaria semejanza es reemplazada, cada vez más por una relación "arbitraria" entre la representación y lo representado – esta progresiva

20. *Enz.* § 453, W. 10, p. 260.

21. § 454, W. 10, p. 261.

22. El ejemplo es de Hegel mismo, véase el Agregado (*Zusatz*) al parágrafo 454, op. cit. p. 261.

23. Véase: David Hume, *A Treatise of Human Nature*, edición de L. A. Selby-Bigge, Oxford 1978, Parte I, Secc. III, p. 8 y 9, Parte III, secc. V, p. 84-86.

diferencia entre ambos no es vista como un defecto del conocimiento sino como su virtud.

Mientras que Hume trata al recuerdo y a la memoria –siguiendo la tradición desde Aristóteles– junto con la imaginación, dándole a ésta última un carácter subsidiario respecto de la primera,²⁴ Hegel interpone curiosamente a la imaginación *entre* el recuerdo y la memoria.

La sección que aparece a continuación en el orden de la *Enciclopedia* está en efecto dedicada a la imaginación y comienza directamente con la noción de "imaginación reproductiva": "La inteligencia, activa en esta posesión <de imágenes mnémicas> es la *imaginación reproductiva*, el *surgir* las imágenes desde la propia interioridad del yo, que a partir de ahora se constituye en su poder."²⁵

El concepto de "imaginación reproductiva", como capacidad de evocar el contenido de una percepción – que en la tradición es atribuida a la memoria misma – proviene, como es sabido, del vocabulario filosófico de Kant. Éste define a la imaginación en general como: "la facultad de representar un objeto, aún *sin su presencia en la intuición*."²⁶ A su vez Kant distingue entre una imaginación "productiva" y una "reproductiva". Mientras que la primera –"*exhibitio originaria*"– operaría a priori y determinaría el sentido interno de acuerdo a las formas del espacio y del tiempo, o sea de modo independiente de la experiencia, la segunda –"*exhibitio derivativa*"– obtendría su material de la sensibilidad, o sea de la "intuición empírica".²⁷

Hasta aquí Hegel parece atenerse a la conceptualización kantiana, Lo que distingue a ambos sin embargo, es su visión diferente acerca del carácter a priori del espacio y del tiempo.

Dado que para Hegel la transformación de lo percibido en imagen

24. *A Treatise of Human Nature*, op. cit., Parte I, Secc. III, p. 8 y 9.

25. Op. cit., § 455, p. 262. Las palabras entre corchetes son mías, el subrayado es de Hegel.

26. *Crítica de la razón pura*, op. cit., B151, p. 166.

27. *Anthropologie in pragmatischer Hinsicht*, en: *Werkausgabe XII*, ed. Weischedel, § 25, A67-68, B69, Francfort del Meno 1977, p. 466. La distinción que requeriría una mayor elaboración que la que ofrecen los escasos textos. Kant mismo parece no atenerse estrictamente a ella cuando sostiene que (incluso) la "imaginación productiva" no está en condiciones de proveer por sí misma su propio material.

consiste precisamente en la transposición de la exterioridad espacio-temporal con que se presenta el objeto percibido al espacio y tiempo de la subjetividad, la re-producción que lleva a cabo la imaginación no es la del contenido percibido sino la de su imagen, en la que éste ya está codificado.

Otro de los aspectos que cabe destacar de la cita de más arriba, es el énfasis que pone Hegel en el sujeto como agente de su imaginación, en la medida en que éste tiene en su "poder", la facultad de suscitar las imágenes depositadas en su interioridad; tanto el recuerdo como la imaginación son considerados formas de apropiación del mundo circundante por parte de un sujeto cognoscente. Hegel habla incluso del "tesoro de sus imágenes".²⁸

Es en este marco, que Hegel –y esto sí en conformidad con Kant– trata de las "leyes de asociación de ideas", y lleva a cabo su crítica a la "psicología" humeana. Con la imaginación reproductiva tiene lugar un doble proceso de reducción: por un lado, a diferencia del mero recuerdo, que consiste en una multiplicidad de imágenes aisladas, lo que se lleva a cabo ahora es la puesta en orden, "asociación" –si bien más o menos arbitraria– de las imágenes y su condensación, de acuerdo a reglas empírico-psicológicas, por otro lado, una "subsunción" de la multiplicidad de percepciones bajo una imagen en las que ellas son reconocidas como "lo mismo", de modo que *la imagen opera como adquiriendo poco a poco cierto carácter general hasta convertirse en icónica y de esta manera en la base de un concepto abstracto*.

Por último, el orden en que los objetos percibidos aparece ante la conciencia es reemplazado por el *interés* que preside las intenciones del sujeto que ordena su evocación o "reproducción".²⁹

Los parágrafos dedicados a la transición entre la imaginación y la memoria tratan de una tercera forma de imaginación que Hegel caracteriza como "fantasía simbolizante" (*symbolisierende Phantasie*) o "productora de signos" (*zeichenmachende Phantasie*).³⁰

Nuevamente Hegel parece inspirarse aquí en la *Antropología* kantiana en cuyo parágrafo 35, Kant insinúa la presencia en nosotros de una "*facultas signatrix*" que define algo oscuramente como: "La capacidad

28. Agregado al § 455, op. cit. p. 264.

29. idem, p. 265.

30. Idem, p. 264.

del conocimiento de lo presente, como medio para vincular la representación de lo previsible con lo pasado...".³¹ Interesante para nuestro tema es la distinción que hace Kant es este texto entre el plano "intuitivo" y el "discursivo", o puramente intelectual, ubicando al símbolo por oposición al signo o "carácter" aún en el primero, mientras que en Hegel ya la imagen aparece como un término medio.

Ahora bien, aquella relación que en la imagen se presenta como cierta oposición y al mismo tiempo complementación entre ella misma y aquello de que es imagen, se profundiza en duplicidad en el caso del símbolo, que mantiene cierta relación de semejanza con lo simbolizado y alcanza su culminación con el signo, en el que la relación entre significante y la cosa significada es "arbitraria". Entre ambos Hegel coloca la noción de "significado" que pasa a ocupar el papel central.

Antes de pasar al tratamiento que Hegel hace de la memoria es conveniente detenernos brevemente en la noción de "imagen", que aparece no sólo en relación al recuerdo, sino necesariamente también a la "imaginación productiva" en su modalidad de "fantasía": "La fantasía es el término medio en el que lo universal y el ser, lo propio y el ser-encontrado, lo interior y lo exterior están completamente producidos como unidad".³²

Tal como Hegel opera con este concepto da la impresión que la imagen que hace posible el recuerdo de alguna manera es producto de la imaginación,³³ es decir, que si no poseyésemos cierta capacidad de transformar lo percibido en imagen tampoco estaríamos en condiciones de recordar- ya el engrama es una construcción cognitiva. De este modo se justificaría con mayor razón el tratamiento que hace Hegel de la memoria a continuación del de la imaginación que pone a ésta como condición de posibilidad de la primera.

En el mismo parágrafo la fantasía es descripta como "taller interior" (*innere Werkstatt*) cuyos productos pueden reconocerse en los más diversos ámbitos como resultado de una fusión de lo subjetivo y lo uni-

31. *Anthropologie*, op. cit., p. 497, acerca del papel del símbolo véase también el § 59 de la *Crítica del Juicio: Kritik der Urteilskraft*, ed. W. Weischedel, *Werkausgabe X*, p. 294 y sigs.

32. Nota al § 457 de la *Enciclopedia*, op. cit., p. 268.

33. En la misma nota Hegel menciona a la fantasía como ya estando presente en la formación de la imagen.

versal y en este sentido se identifica con el proceso de la "razón" misma, si bien aún en su carácter "formal".

Lo que sigue a continuación es la descripción de un proceso de paulatina abstracción del contenido sensible de la imagen mnémica a través de al menos dos fases. La primera es la puesta en símbolo del dato percibido: "fantasía productora de símbolos", en la que la relación entre la representación y lo representado conservan cierta relación de analogía – Hegel menciona aquí de paso a la "alegoría" como una de sus formas. La segunda fase es su conversión en signo, "fantasía productora de signos", en la que la relación entre significante y significado resulta ya "arbitraria" y debe ser aprendida.³⁴

En las *Lecciones de Estética* el concepto de símbolo –sin duda bajo la influencia de la obra de Creuzer–³⁵ aparece definido en forma mucho más clara y elaborada. Hegel diferencia y contrasta dos aspectos en él que denomina la "expresión" y su "significado".³⁶ El símbolo es considerado aquí una especie particular de signo.³⁷ Es que –así como el concepto-eje de la sección dedicada a la representación era la "imagen"– la noción central del capítulo siguiente pasa a ser la de "significado".

Pero lo interesante desde nuestra perspectiva es que ahora, para utilizar la expresión de Kant, *el orden de la intuición es reemplazado por el orden del discurso*. No es casual que Hegel sitúe en este contexto del sistema enciclopédico su tratamiento del *lenguaje*. El lugar resulta original y significativo tanto para una teoría del lenguaje mismo como para una teoría de la memoria.

Hacia el final de la nota al § 458 Hegel llama la atención acerca de su usual incómoda ubicación como "apéndice" a la Psicología o a la Lógica. La formulación que aparece al final de la nota al parágrafo contiene condensadas la tesis centrales de la sección: "Esta actividad creadora de signos puede llamarse preferentemente memoria *productiva* (la *Mnemosyne* primeramente abstracta), en cuanto que la memoria, que en la

34. Véase para todo esto los §§ 457-458 y sus respectivos Agregados, op. cit. p. 267 y sigs.

35. Véase: Friedrich. Creuzer, *Symbolik und Mythologie der alten Völker*, 1810-1812.

36. *Vorlesungen über die Ästhetik*, ed. Moldenhauer-Michel, *Werke*, tomo 13, p. 394.

37. Ibidem.

vida común a menudo es confundida y usada en el mismo sentido que el recuerdo, la representación o la imaginación, sólo tiene que ver en general con signos".³⁸

Aquí Hegel establece una delimitación del campo de la memoria que contrasta con su tratamiento tradicional: a pesar del hecho de que en el acto de recordar se manifiesta una intervención e interrelación estrecha entre el recuerdo, la imaginación y la representación, la memoria es colocada claramente en el orden de lo simbólico.

También llama la atención aquí la paradójica postulación de una "memoria *productiva*", cuando solemos asociarla con la facultad de *re*-producir ante la conciencia determinados acontecimientos, hábitos o contenidos mentales. Más aún, el criterio para juzgar una "buena" o "mala" memoria pasa por su "fidelidad" para copiar o transcribir de alguna manera sucesos pasados.

El carácter "productivo" de la memoria, tal como es concebido en este texto, tiene que ver con los aspectos: (a) *figurativos* –que como vimos, tiene lugar en la producción de las imágenes mismas o engramas–, (b) *cognitivos* –en la medida en que se trata de un re-conocimiento de la imagen como algo con sentido–, y por último, (c) *interpretativos* de la rememoración, ya que en ella se produce una reinterpretación retrospectiva en el marco del horizonte presente.

Si bien Hegel solo ha esbozado su teoría de la memoria, su concepción acerca del carácter "productivo" de la misma representa un punto de vista original y contiene interesantes sugerencias para su desarrollo, que no encajan necesariamente en el marco de una doctrina dialéctica rígida tal como suelen considerarse los textos de la *Enciclopedia*, ni siquiera puede hablarse en realidad de un tratamiento "dialéctico" del tema, éste tiene que ver en todo caso más con la disposición de los párrafos que con su contenido –aunque la necesidad de la distinción recuerdo-imaginación-memoria puede haber obedecido a las constricciones del esquema.

Ya con el orden y los mecanismos de catalogación y evocación de las vivencias, cuyos criterios deben buscarse para Hegel en el plano lingüístico-conceptual, los recuerdos aparecen siempre no como mera imagen sino como imagen de algo que posee por lo tanto un determinado *significado* por elemental u oscuro que fuera.

38. Op. cit., § 458, p. 270 y sig.

Los recuerdos son algo que entendemos o pretendemos llegar a entender, de modo que en la evocación interviene ya el pensamiento y sus categorías. Con el lenguaje –cuya formación es obra de lo que el texto califica de “instinto lógico”–³⁹ se lleva a cabo ya una interpretación de los hechos vividos que no se encuentra en el orden de lo real mismo. *Lo que la memoria re-produce no es una fotocopia de lo dado sino una compleja versión lingüístico-conceptual codificada y retrospectivamente reinterpretada de los datos de la experiencia.*

Es esta visión de una estrecha participación de la “inteligencia” y sus categorías lógico-lingüísticas en el proceso mismo de la memoria, lo que le confiere a este esbozo teórico, con todas sus deficiencias, una claridad que no se encuentra en textos contemporáneos sobre el tema, que a pesar de su valiosa y amplia base experimental, se preocupan más por establecer las bases neuronales de su buen funcionamiento.⁴⁰

Resulta por cierto cuestionable el hecho de que se niegue tácitamente en esta visión la existencia de una memoria infantil prelingüística, pero resulta bastante plausible que la adquisición del lenguaje representa al mismo tiempo una reorganización radical de la facultad de recordar.

Los párrafos 457-459 están dedicados a exponer una –desde la perspectiva de su desarrollo ulterior, por cierto rudimentaria, pero altamente sugestiva– filosofía del lenguaje. Su tratamiento excedería el marco temático de este trabajo.⁴¹ Lo que me interesa señalar aquí son algunos rasgos que conciernen a la sorprendente conexión que Hegel

39. Op. cit., § 459, p. 272.

40. Véase por ej. el por lo demás interesante libro de Daniel L. Schachter: *Searching for Memory. The Brain, the Mind and the Past*, 1996, que ofrece un panorama de las investigaciones de los últimos años.

41. Sobre esto puede consultarse el capítulo sobre “El lugar del lenguaje en la psicología de Hegel. Lenguaje y pensamiento, lenguaje e intersubjetividad” del libro de Vittorio Hösle: *Hegels System*, Hamburgo 1988, tomo 2, p. 396 y sigs., que ofrece una exposición y discusión de sus puntos principales sin por cierto ocuparse de su conexión con la memoria. Para un tratamiento más exhaustivo pueden consultarse el libro de Theodor Bodamer, *Hegels Deutung der Sprache*, Hamburgo 1969 y el de Joseph Simon, *Das Problem der Sprache bei Hegel*, Stuttgart 1966, véase también el artículo de Jacques Derrida: *El pozo y la pirámide. Introducción a la semiología de Hegel*, contenido en el volumen colectivo: *Hegel y el pensamiento moderno*, edit. por Jacques D’Hondt, México 1973.

establece entre lenguaje y memoria por un lado, y memoria y “pensamiento” (*Denken*) por el otro, en los escasos párrafos con que termina abruptamente la sección.

Por lo pronto ya en el apartado que introduce la noción de “memoria”, esta es definida de un modo extraño en relación al “nombre” –dado lo arduo de los textos la siguiente es una versión reconstructiva.

Por “nombre” debe entenderse aquí no el nombre propio, sino el vocablo o la palabra en general, que Hegel considera como “una síntesis (*Verknüpfung*) de la intuición producida por la inteligencia y de su significado (*Bedeutung*)”.⁴²

Ahora bien, lo que Hegel establece como “memoria” es un tipo particular de “recuerdo”, o por decirlo así, recuerdo de segundo orden <no de la imagen mnémica como tal sino>, aquel referido a la palabra <en tanto imagen artificial>, en la que la captación del significante se presenta indisoluble aunque arbitrariamente unida a su significado: La memoria “es el recuerdo de esta exterioridad (*Äusserlichkeit*) <de la palabra>”.

El capítulo dedicado a la memoria es particularmente breve y la versión del Agregado al párrafo 461 resulta, más clara que la del texto principal. Hegel diferencia allí tres “formas” de memoria: que califica como “retentiva de palabras” (*namenbehaltende*), “reproductiva” (*reproduktive*) y “mecánica” (*mechanische*).

Lo propio de la primera parece consistir en fijar una convención, de tal modo que la palabra termina por reemplazar a la cosa nombrada al establecerse una conexión inmediata entre la imagen acústica o escrita y su significado, que Hegel identifica directamente con su “representación”.⁴³

La memoria “reproductiva” presupone la anterior y lo propio de ella consistiría en reconocer ya en el “nombre” a la cosa misma sin necesidad de recurrir a imágenes: “Con la palabra (*Name*) león no tenemos necesidad ni de la intuición de un animal de ese tipo ni tampoco incluso de su imagen, sino que la palabra (*Name*) en cuanto la *entendemos* es la simple representación sin imagen. <Es que> es en palabras (*Namen*) que *pensamos*”.⁴⁴

42. Op. cit., § 460, p. 277.

43. Op. cit., § 462, p. 278.

44. Ibidem.

<Traduzco aquí "Name" por "palabra" y no como es usual por "nombre", porque si bien Hegel distingue entre ambos, las utiliza en una acepción propia diferente a la actual y el párrafo no daría sentido. Asociamos el vocablo "nombre" a nombre propio, mientras que Hegel lo entiende como palabra vinculada a un significado. Por el contrario por "palabra" entiende la mera imagen acústica o gráfica.>⁴⁵

Hegel ve en la "memoria reproductiva" una facultad configuradora que permite re-presentar y armar mediante palabras un estado de cosas determinado. Pensada de este modo, la memoria "reproductiva" actuaría incluso aún cuando no se refiere a hechos acaecidos en el pasado. Su contenido mnémico se encontraría más bien en la implícita asociación de signifiante y significado. Con esto Hegel se aparta por cierto tanto de la concepción vulgar como de la tradición filosófica de la memoria, cuyo carácter "re-productivo" adquiere aquí un sentido configurante más que de copia fiel.

Al igual que en Aristóteles la memoria es tratada en estos párrafos como una condición de posibilidad del "pensamiento". Hegel asocia directamente "*Gedächtnis*" ("Memoria") con "*Denken*" ("pensamiento"), aunque a diferencia del primero lo considera como un proceso sin imágenes. Los trazos mnémicos estarían implícitos en la captación del sentido⁴⁶ - es que en la Psicología hegeliana a diferencia de la teoría del alma aristotélica mediando entre memoria y pensamiento se encuentra precisamente el lenguaje.

Es esta capacidad de prescindir de la imágenes lo que hace que la memoria sea considerada por Hegel como una facultad superior a la imaginación. Más aún, la mnemotécnica - es interpretada como "un volver a rebajar la memoria a imaginación".⁴⁷

45. Véase para esto el comentario al párrafo de las *Vorlesungen über die Philosophie des Geistes*, op. cit., p. 214.

46. Véase sobre esto el mencionado artículo de Hans Friedrich Fulda: *Vom Gedächtnis zum Denken*, op. cit., p. 326 y sigs.

47. *Enzyklopädie*, op. cit., § 462, p. 279. Acerca de los "lugares de la memoria" y la mnemotécnica en general véase ante todo el breve tratado de Aristóteles mencionado más arriba, *Acerca de la memoria y de la reminiscencia*, op. cit., p. 452^a, *De anima*, 427b, *Tópica*, 163b y para su uso en la tradición retórica romana, Cicerón, *Del Orador*, II, 352-353., Hegel cita este texto en su comentario al párrafo 462 en la la versión de las *Vorlesungen*, op. cit., p. 215. Para un tratamien-

Mediante su articulación lingüística las representaciones dejan de pertenecer a la "interioridad" para adquirir una "objetividad" ("*Gegensätzlichkeit*") en la que el espíritu "se aliena" ("*sich entäußert*") a la vez que se enriquece.

Con todo, Hegel no identifica -al menos explícitamente- el ámbito del lenguaje con el de la *intersubjetividad*, al que su propia conceptualización parece naturalmente conducir. A pesar de la primacía que adquiere la palabra por sobre lo inefable, o sobre el plano de la percepción, Hegel trata al lenguaje en el marco de su Psicología, o sea desde lo que podríamos llamar hoy un individualismo metodológico, y no ve en él la presencia de los otros, o en otras palabras, en el discurso, un instrumento mediante el cual las vivencias privadas son organizadas en categorías sociales aún antes de ser comparadas con los recuerdos o experiencias de otras personas.

Por último Hegel concibe a la memoria "mecánica" curiosamente como la forma más evolucionada de esta facultad debido a que por su carácter formal conduce ya al "pensamiento": Si bien por un lado, como vimos, la mnemotécnica es despreciada por su recurrencia a las imágenes, por el otro la memoria mecánica es exaltada como una forma superior de actividad espiritual, a pesar o precisamente por el hecho de que consiste en la capacidad de repetir series de palabras o números en un orden arbitrario y azaroso.

Aquí los breves párrafos de la *Enciclopedia* resultan particularmente escuetos y oscuros. En el texto de la llamada *Filosofía Real de Jena* Hegel describe el orden que establece la memoria mecánica paradójicamente como una "necesidad arbitraria".⁴⁸ A diferencia de la mnemotécnica la ventaja de esta forma de memoria es que se mueve en un orden puramente semiótico, sin recurrir a imágenes. Mediante la repetición de series, aún sin sentido, la "inteligencia" esta en condiciones de obtener un material en el que el mundo esta ya precodificado simbólicamente y sirve de punto de partida para el establecimiento de significados generales. En el ejercicio de la repetición mecánica de una serie de contenidos simbólicos ve Hegel una forma de apropiación que establece un orden sobre el que opera el concepto.

to actual del tema resulta imprescindible el libro de Frances Yates, *The Art of Memory*, Londres 1966.

48. *Jenaer Realphilosophie*, ed. Johannes Hoffmeister, Hamburgo 1967, p. 187.

Como tratamiento de la memoria en general los textos de la *Psicología* hegeliana son sin duda poco satisfactorios. Queda totalmente fuera de su alcance la memoria corporal, la memoria de hábitos, la memoria emocional, o fenómenos tan elementales –aunque no por eso fácil de explicar– como el recuerdo de una melodía. Pero incluso los temas abordados resultan difíciles de seguir.

Hegel mismo ha llamado la atención acerca del carácter de bosquejo de su concepción de la memoria en unas líneas insólitas en el marco de su sistema: “Captar el puesto y la significación de la memoria en la sistematización de la inteligencia, del mismo modo que concebir su conexión orgánica con el pensamiento, es uno de los puntos hasta ahora completamente descuidados y de hecho de los más difíciles de la doctrina del espíritu.”⁴⁹

La concepción de una memoria “productiva” es desarrollada por Hegel paradójicamente en el marco de una epistemología realista. El carácter configurativo y no meramente mimético de las imágenes, representaciones y signos lingüísticos tiene consecuencias también para una teoría de la verdad alternativa, que Hegel no ha desarrollado,⁵⁰ que supera la metafórica de la “correspondencia con” y de la “reproducción” de una realidad a la que tendríamos acceso independientemente de un determinado marco teórico o simbólico, y que sin embargo pretende dar cuenta de ella.

Pero es precisamente este carácter incompleto y de esbozo lo que da actualidad a estos textos en los que están contenidos tanto observaciones originales como interrogantes cuya dilucidación continúa siendo una tarea de la reflexión contemporánea.

49. Op. cit., § 464, p. 283. Acerca de este texto ha llamado la atención Hans Friedrich Fulda en el artículo citado en más arriba, p. 322.

50 No me refiero con esto a la compleja concepción hegeliana de la verdad como proceso interior a la “idea”, o sea como correspondencia entre “concepto” y “realidad”. Véase sobre esto el artículo de Michael Theunissen: *Begriff und Realität. Hegels Aufhebung des metaphysischen Wahrheitsbegriff*, en el volumen colectivo: *Seminar: Dialektik in der Philosophie Hegels*, editado por Rolf-Peter Horstmann, Francfort del Meno, 1978, p. 324 y sigs.

ABSTRACT

This paper deals with Hegel's conception of memory, showing that it involves an original point of view that goes beyond its historical-philosophical background and opens new perspectives. Since the *mnemic* image is conceived as an iconic inscription, Hegel is able to talk about a creative memory, in which symbolic imagination takes part. *Remembrance* is possible not as a bare copy of what is given in perception, but as a middle term between the intuitive and the discursive level.

Asociación Filosófica Argentina (AFRA)

XII CONGRESO NACIONAL DE FILOSOFÍA

3 al 6 de diciembre de 2003

Sede: UNIVERSIDAD NACIONAL DEL COMAHUE

Ciudad de Neuquén
Pvcia. del Neuquén
República Argentina

INFORMES:

**Departamento de Filosofía
Facultad de Humanidades**

Universidad Nacional del Comahue
Dirección: Avenida Argentina 1400
(8300) Neuquén - Argentina

Tel-Fax: 54-299-4490-389
e-mail: afra@uncoma.edu.ar

DISCUSIÓN

LA LÓGICA INDUCTIVA ES INÚTIL. RESPUESTA A ALEJANDRO CASSINI

Mario Bunge

Alejandro Cassini (2002) comparte mis críticas a la lógica inductiva tradicional o ingenua, que yo analicé en varias publicaciones (p. ej. Bunge 2000). Pero me critica por haber ignorado la lógica inductiva bayesiana, centrada en el concepto de probabilidad de una proposición, que desarrollaran de Finetti, Jeffreys, Février, Good, Savage y Lindley.

El atractivo de cualquier teoría probabilista es que es exacta, puesto que emplea una teoría matemática muy rica y fecunda. Pero el que tenga contenido y sea fecunda o estéril depende de la interpretación de que sea objeto. Es sabido que el cálculo de probabilidades se usa con éxito en casi todas las ciencias fácticas, desde la física y la biología hasta la psicología y la sociología, así como en la ingeniería. En todos estos campos los que se calcula o mide es la probabilidad de estados o de cambios de estado (sucesos), no de proposiciones.

En todas las ciencias fácticas y técnicas la probabilidad de un hecho se considera una propiedad objetiva del hecho, tanto como la masa de un cuerpo o el volumen de ventas de una empresa. La propiedad consiste en la posibilidad real de un hecho aleatorio. En otras palabras, los científicos usan tácitamente la interpretación realista de la probabilidad, adoptada explícitamente por Poisson, Poincaré, Smoluchovsky, Fréchet, Bunge, Popper y otros (v. Du Pasquier 1926, Fréchet 1955, Agazzi 1988).

Ninguna teoría científica calcula o mide la probabilidad o el grado de confirmación de sus propias hipótesis o datos. Un científico se sorprendería si se le preguntase, p. ej., cuál es la probabilidad $P(h|e)$ de la hipótesis h dado el dato e que acaba de obtener. No sabría cómo asignar tal número. Y no lo sabe porque la lógica inductiva no contiene reglas para efectuar tal asignación. ¿Será por este motivo que Cassini no nos informa, por ejemplo, sobre la probabilidad de la hipótesis que afirma existencia de bosones Higgs, o la probabilidad de que una copa diaria de vino aleje el peligro de contraer la enfermedad de Alzheimer, o la

probabilidad de que Argentina salga de la crisis actual dentro del siglo?

Perdón, no debí ser tan exigente. Me conformaré con que me digan cuál es la probabilidad de viejos conocidos, tales como la segunda ley de Newton, la segunda ley de la termodinámica, las ecuaciones de Maxwell, o la ecuación de Schrödinger. Y quedaría muy agradecido a quien me explicara qué significa la expresión "la probabilidad de p " y cómo se relaciona con esta otra, que es la única que interesa en las ciencias que frecuento: "el grado de verdad de p a la luz de la evidencia empírica existente y de las teorías pertinentes a p ". ¿O será que como dijera Du Pasquier (1926: 197), la expresión de marras es "un montón de palabras carentes de sentido"?

En resumen, la teoría matemática general de las probabilidades no especifica los valores de las funciones de probabilidad (por definición de "general"). Pero esos valores aparecen en las aplicaciones de la teoría, ya que, desde el punto de vista instrumental, ésta es una herramienta para calcular probabilidades a partir de otras probabilidades.

Puesto que la lógica inductiva no contiene reglas o criterios para evaluar las probabilidades básicas, tales como $P(h)$ y $P(e)$, en forma no arbitraria, no es una aplicación *bona fide* del cálculo de probabilidades. Sirve para llenar páginas de fórmulas. Ningún científico o técnico la ha usado jamás para evaluar ideas científicas o técnicas.

Para evaluar hipótesis, los científicos y técnicos usan nociones tales como la de error accidental y el criterio χ^2 (chi al cuadrado). Por ejemplo, dicen que una hipótesis numérica de la forma " $M(a) = n$ " ha sido confirmada por el dato empírico "Medición de $M(a) = b \pm \sigma$ " es verdadera a menos del error σ , si la diferencia entre a y b , en valor absoluto, es menor que σ . Y dicen de una función o curva de interpolación que se ajusta adecuadamente a los datos si aprueba el test χ^2 (o sea, si esta cantidad es pequeña).

El cálculo de errores fue fundado por el gigantesco Gauss a comienzos del siglo XIX; y el matemático y filósofo Karl Pearson inventó el test χ^2 a fines del mismo siglo. Sin embargo, ninguno de esos conceptos forma parte de la lógica inductiva, que pretende pronunciarse sobre el grado de confirmación de teorías (no sólo de hipótesis). Tampoco figuran esos conceptos clave en los escritos de la escuela popperiana. ¿Se podrá estimar la probabilidad de que unos y otros se den por enterados, dentro del siglo actual, de esos viejos caballos de batalla de la estadística matemática?

Hay más: la lógica inductiva no sólo es inútil, sino que sufre gravísi-

mas paradojas. Algunas de éstas son tan conocidas, que figuran en mi *Diccionario de filosofía* (Bunge 1999). Otra es el popular Problema de Monty Hall. Hay tres puertas cerradas. Detrás de una ella se ha colocado un premio. Se pregunta cuál es la probabilidad de ganar el premio al abrir una de las puertas. El subjetivista responde confiado: $1/3$. El realista objeta que no tiene sentido pensar en probabilidades, porque el premio ya ha sido ubicado detrás de una de las puertas: no ha habido azar. Si es consecuente, el realista se niega a participar del juego.

Sin azar no hay probabilidades objetivas. Y si una probabilidad es subjetiva, entonces no tiene cabida en la ciencia, excepto como objeto de estudio psicológico. Más aun, si las "probabilidades" son subjetivas, o grados de creencia, entonces no satisfacen los postulados del cálculo de probabilidades. Esto lo mostraron hace tiempo el psicólogo experimental Daniel Kahneman (premio Nobel 2002) y sus colaboradores (Kahneman, Slovic y Tversky 1992). En otras palabras, contrariamente a lo que exige la lógica inductiva, los grados de creencia de una persona de carne y hueso no constituyen un sistema coherente. Cassini no cita este importante resultado de la psicología experimental, que muestra que la lógica inductiva no sólo es inútil, sino también falsa.

Termino con una confesión. También yo caí en la lógica inductiva centrada en el concepto de probabilidad, la que había aprendido de los tratados de Jeffreys, de Finetti, von Wright, y otros. Incluso escribí un artículo (Bunge 1960), lleno de fórmulas bayesianas, del que ahora me ruborizo purpúreamente. Mi entusiasmo se evaporó poco después cuando, conversando con von Wright, me contó que él y los demás consideraban la lógica inductiva como una disciplina a priori que no recurría a dato empírico alguno para estimar la probabilidad de las hipótesis. Me di cuenta entonces de que la lógica inductiva es un ejemplo de lo que Einstein llamaba despectivamente acrobacia intelectual. El mismo juicio me inspiran los trabajos de Popper y sus seguidores en los que interviene la noción de improbabilidad de las hipótesis. Todo eso sirve a lo sumo, como decía Hilbert, para conseguir ascensos. La filosofía contemporánea está llena de industrias académicas semejantes. Por esto está en crisis (Bunge 2002).

Department of Philosophy, McGill University, Montréal

BIBLIOGRAFÍA

- AGAZZI, Evandro, compil. 1988. *Probability in the Sciences*. Dordrecht-Boston: Reidel [Kluwer].
- BUNGE, Mario. 1960. Probabilidad e inducción. *Ciencia y Técnica* [órgano del Centro de Estudiantes de Ingeniería de la UBA] 129: 240.
- BUNGE, Mario. 1999. *Diccionario de filosofía*. México: Siglo XXI.
- BUNGE, Mario. 2000. *La investigación científica*, 3a. ed. México: Siglo XXI.
- BUNGE, Mario. 2002. *La filosofía en crisis*. Madrid-Buenos Aires: GEDISA.
- CASSINI, Alejandro. 2002. Bunge, la lógica inductiva y el grado de confirmación de las teorías. *Revista Latinoamericana de Filosofía* 28: 303-317.
- DU PASQUIER, L.-Gustave. 1926. *Le calcul des probabilités: Son évolution mathématique et philosophique*. París: Librairie Scientifique J. Hermann.
- FRÉCHET, Maurice. 1955. *Les mathématiques et le concret*. Paris: Presses Universitaires de France.
- KAHNEMAN, Daniel, Paul Slovic y Amos Tversky. 1992. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

COMENTARIOS BIBLIOGRÁFICOS

Víctor Palacios, *Lenguaje y pacto en Thomas Hobbes*, Buenos Aires, Trama Editorial, 2001

En la filosofía, son muchos los temas cuyo análisis nunca puede darse por agotado. Uno de ellos es el objeto de este libro de Víctor Palacios, *Lenguaje y pacto en Thomas Hobbes*. Retomando la aplicación hobbesiana del llamado "método paduano", la perspectiva desde la que se aborda el tema del pacto social no es la de sus características ni la de su función teórica, sino la de sus condiciones de posibilidad. Consciente de la dificultad metodológica que implica establecer esto último sin por ello reducir el pacto a su propia condición de posibilidad o bien de caer en una petición de principios –objeciones que analiza en el tercer capítulo–, Palacios establece que el lenguaje es condición de posibilidad del pacto.

Para ello, parte de las definiciones hobbesianas de los usos del lenguaje, uno de los cuales identifica como el más apto para realizar el pacto y que en su segundo apéndice emparenta con el concepto de "acto de habla" de Austin. Se trata de uno de los usos *públicos* del lenguaje, según el cual los hombres expresan sus propósitos y voluntad. La artificialidad de todo lenguaje, propia de la concepción de Hobbes sobre la que se argumenta en el segundo capítulo, requiere por ello de un pacto previo y es uno de los obstáculos metodológicos más importantes que Palacios sabe necesita salvar para sostener su tesis. Lo hace introduciendo una efectiva distinción entre los conceptos de comunidad *lingüística* por un lado, y comunidad *política* por el otro. La primera supone solamente el acuerdo respecto de la *comprensión* de los signos que, si bien es condición necesaria para la realización del pacto que constituye la segunda, no es suficiente para constituir la por sí solo. Con esto, queda resuelta una de las dificultades aludidas más arriba dado que así se garantiza que considerar el lenguaje como condición de posibilidad del pacto *societatis* no implica la reducción de éste a la convención que da lugar al significado compartido de los signos.

Finalmente, se consolida esta distinción asegurando la independencia del pacto lingüístico respecto del político. Palacios lleva a cabo esta tarea argumentando que los mandatos del *Leviathan* no se refieren a la comprensión ya aludida de los signos del lenguaje –es decir, que el soberano no regula la verdad de las proposiciones– sino a su *interpretación* –esto es, a su carácter sedicioso, por

ejemplo—. Esta nueva diferenciación le permite, además, superar la dificultad de incurrir en un círculo vicioso entre ambos acuerdos. Si la decisión del soberano incidiera en la verdad de los signos, entonces la comunidad lingüística supondría a la política; pero siendo el intercambio de signos independiente de las regulaciones impuestas por el *Leviathan*, entonces su anterioridad lógica queda fuera de discusión.

Lenguaje y pacto en Thomas Hobbes es, pues, un análisis claro y ordenado de la concepción del lenguaje hobbesiano puesta en relación con la problemática del pacto, cuya tesis contribuye a esclarecer un concepto central de la filosofía política moderna.

Andrea Pac
Universidad de La Matanza

Zeljko Loparic, *A semântica transcendental de Kant*, Campinas: UNICAMP, Centro de Lógica, Epistemologia e História da Ciência, 2000. (Coleção CLE, v.29). 326 pp.

Loparic presenta en este volumen una traducción corregida de la primera parte de su tesis doctoral *Scientific Problem-Solving in Kant and Mach* [1982, Universidad Católica de Lovaina]. El propósito del autor es defender la tesis de que la *Crítica de la Razón Pura* puede ser interpretada como una teoría de la resolución de los problemas de la razón teórica. Más concretamente, Loparic sugiere una lectura de la lógica trascendental de Kant como una semántica *a priori* al servicio de una teoría de resolución de problemas. Desde esta perspectiva, la tesis básica presente en la primera *Crítica* se expresa en el siguiente teorema de solubilidad: todas las cuestiones que nuestra razón nos impone debido a su naturaleza o bien son insolubles, o bien es posible ofrecer para ellas una respuesta definitiva.

Loparic entiende que, para Kant, la clase de las proposiciones solubles es equivalente a la clase de las proposiciones sintéticas posibles. A su vez, una proposición sintética es posible si, en primer lugar, todos sus conceptos no lógicos tienen referentes en un dominio de objetos sensibles y, en segundo término, si su forma lógica puede ser satisfecha en un dominio de formas sensibles. Esta tesis es desarrollada en el Capítulo 1.

Loparic reconoce en la cuestión del método una clave importante para la demostración de dicho teorema. Por ello, el autor se ocupa en el Capítulo 2 de discutir el método propio de la investigación trascendental. La tesis de Loparic es que Kant combina el análisis y la síntesis de los géometras griegos. El autor in-

terpreta que el análisis es un método heurístico de *ampliación* del conocimiento, mientras que la síntesis constituye un método de *exposición* del conocimiento ya encontrado. En este sentido, Loparic reconstruye la argumentación kantiana orientada a mostrar cómo son posibles las proposiciones sintéticas en general (no sólo las *a priori*) identificando los momentos analítico y sintético de la prueba. Según el autor, las Deducciones Metafísica y Trascendental de las categorías son instancias analíticas, mientras que la investigación del Esquematismo y del Sistema de los Principios se desarrolla sintéticamente.

Por otra parte, el estudio de la solubilidad de las cuestiones suscitadas por nuestra razón le permite a Loparic reconstruir, en Kant, una teoría de la razón humana como un dispositivo para la solución de problemas. Esta teoría es la contracara del teorema de solubilidad: en tanto los problemas a los que nos enfrenta nuestra razón son, o bien insolubles, o bien aptos para recibir una respuesta definitiva, la razón humana es entendida como el instrumento idóneo para encontrar estas respuestas. Loparic discute este punto en el Capítulo 3. Aquí ya se adelanta una tesis que será posteriormente desarrollada con más detalle en los Capítulos 7 y 8: así como la doctrina de la analítica trascendental fija el dominio del conocimiento y también las propiedades básicas de los objetos cognoscibles, la dialéctica trascendental fija sus límites. De tal modo, Loparic distingue entre el canon *doctrinal* y el canon *heurístico*. El primero, fundamentado en la analítica trascendental determina el cuadro general de aplicación correcta de las operaciones *a priori* de la facultad del juicio, al determinar conceptualmente representaciones intuitivas dadas (reflexión) o al exhibir intuitivamente conceptos dados (subsunción). Por su parte el segundo, basado en los resultados de la dialéctica trascendental, provee reglas para organizar con eficacia las operaciones gobernadas por el canon doctrinal. Así, el autor establece dos cánones kantianos de investigación científica: el canon doctrinal para la solución de problemas sobre entidades sensibles y el canon heurístico para la solución de problemas acerca de las leyes de dichas entidades.

Según Loparic, la semántica *a priori* de Kant es una teoría acerca de la interpretabilidad en términos de representaciones intuitivas, de las representaciones discursivas que componen proposiciones sintéticas. De tal modo, el autor trata sistemática e independientemente a las representaciones intuitivas (Capítulo 4), a la cuestión de la referencia y significado de los conceptos (Capítulo 5), así como al problema de la interpretación objetiva de los juicios (Capítulo 6).

En el capítulo dedicado a las representaciones intuitivas, Loparic indica que la definición kantiana de experiencia presupone percepciones ya constituidas, así como ciertos enlaces intuitivos (no conceptuales) entre las mismas. Estas percepciones y sus enlaces no lógicos constituyen lo que, siguiendo a Kant, Loparic denomina "conocimiento intuitivo". En el análisis del mismo, el autor despliega sus tesis a la luz de los cuatro pares de conceptos introducidos en la Anfibología (identidad y diferencia, concordancia y oposición, interno y exter-

no, materia y forma). Desde esta perspectiva se discuten cuestiones tales como la afinidad de las representaciones intuitivas y la construcción de *quanta* extensivos e intensivos.

Al estudiar la cuestión de la referencia y del significado de los conceptos, Loparic analiza sucesivamente los conceptos de percepción, de experiencia, los matemáticos y los conceptos puros del entendimiento. Posteriormente se ocupa de los dominios de interpretación de los conceptos posibles. Así, discute el dominio de interpretación de los conceptos empíricos, de los matemáticos y de las categorías. Dado que los esquemas trascendentales no son "objetos" en sentido estricto, existen sólo dos dominios de interpretación: el de los objetos empíricos y el de los matemáticos. De tal modo, la tesis principal de la semántica trascendental de Kant es, según Loparic, la siguiente: un concepto tiene sentido objetivo sólo si puede ser referido a representaciones intuitivas de objetos pertenecientes al dominio de los objetos empíricos o al de los matemáticos.

En el Capítulo 6 se considera la idea kantiana de una semántica *a priori* de la verdad. Según Loparic, la teoría kantiana del esquematismo trascendental es, al mismo tiempo, una semántica *a priori* de conceptos y una semántica *a priori* de juicios del entendimiento puro. Ambos aspectos constituyen en conjunto una fundamentación de la *lógica de la verdad* de Kant, que trata del contenido intuitivo *a priori* de formas proposicionales sintéticas. A la luz de esta tesis se consideran los juicios de percepción, los de experiencia y los matemáticos, así como la distinción entre verdad trascendental y verdad empírica. Por otro lado, se analiza la interpretación objetiva de los juicios de cantidad, cualidad y relación, y el principio de tercero excluido junto con la *reductio ad absurdum*. El autor finaliza el capítulo discutiendo el fenomenalismo de Kant y evaluando el constructivismo kantiano como menos riguroso y radical que el de Carnap.

En el Capítulo 7, Loparic retoma la idea de la razón humana como dispositivo para la resolución de problemas y presenta la doctrina de los principios *a priori* del entendimiento de dos maneras: en primer lugar, la presenta como una teoría trascendental de la verdad; y en segundo lugar, la presenta como el fundamento de una teoría trascendental de algoritmos y de procedimientos heurísticos. Retomando la idea de un canon doctrinal introducida en el Capítulo 3, Loparic discute los problemas cognitivos relativos a objetos distinguiendo los problemas empíricos de los matemáticos.

Considerando ahora el canon heurístico, el autor reconstruye en el Capítulo 8 la teoría kantiana de la razón como un sistema de constructos de pensamientos y máximas heurísticas. Esto le permite establecer el programa kantiano de investigación para la psicología racional, para la cosmología racional y para el estudio de totalidades empíricas absolutas. De tal modo, Loparic sugiere que la revolución kantiana en el ámbito de la metafísica consistió en probar que una interpretación objetiva directa de los principios especulativos tradicionales es imposible, y en darles el papel de reglas metodológicas generales en programas

de investigación científica. La metafísica kantiana es, según Loparic, metodología de las ciencias empíricas.

Como confirmación de esta última tesis, el autor propone en el Capítulo 9 a la teoría dinámica de la materia como un programa de investigación para la mecánica racional.

El lector hallará en *A semântica transcendental de Kant* un texto que contiene análisis de los principales capítulos de la *Crítica de la Razón Pura* a la luz de una tesis ambiciosa y prolijamente argumentada. Sin embargo, la actitud de Loparic, tal como él mismo reconoce, no es la de un comentarista kantiano sino la de un filósofo de la ciencia que se esfuerza por escapar de las dificultades propias de su oficio buscando ayuda en el pasado. De tal modo, quien se interese por cuestiones epistemológicas contemporáneas tendrá con este libro la oportunidad de ingresar en el universo de la filosofía trascendental por un sendero desde el que resulta a todas luces evidente la rigurosa actualidad del pensamiento de Kant.

Hernán Pringe
Universidad de Buenos Aires

Néstor L. Cordero (comp.), *Ontologie et Dialogue. Hommage à Pierre Aubenque*, Paris, Vrin, 2000.

El libro que reseñamos se compone de once artículos que exploran una diversa gama de temas de Filosofía Antigua. Un primer vistazo nos introduce en cuestiones metafísicas (el ser, las Formas, la sustancia), epistemológicas (verdad y falsedad, dialéctica), físicas (el tiempo) y éticas (la acción, la equidad, el *kai-rós*). Sin embargo, detrás de semejante variedad puede encontrarse una clave rigurosamente respetada —el *diálogo* en sus diferentes facetas— y a fin de cuentas un logrado homenaje al 70º aniversario de Pierre Aubenque.

El término *diálogo*, tal como lo evidencia *Ontologie et Dialogue*, es claramente multívoco. En primer lugar hallamos al *lógos* —y sus "partes"— en su dimensión ontológica, en sus relaciones con el ser. El artículo de Michel Fattal ("*Vérité et fausseté de l'ónoma et du lógos dans le Cratyle de Platon*") apunta en esa dirección, cuando estudia cómo el *ónoma* es en el *Cratilo* pasible de falsedad, y en qué medida esto permite a Platón insistir en que el objeto de investigación debe colocarse del lado de las cosas mismas más que del de los nombres. Néstor Cordero también se dedica a un diálogo platónico que escruta las relaciones entre discurso y ser —el *Sofista*— y propone comprender a este último como *dúnamis* y a la conflictiva noción de *participación* como el ser mismo de las Formas.

Incluso Pierre Aubenque, en su artículo "Sur l'ambivalence du concept aristotélicien de substance" se ocupa de algún modo de la relación entre discurso y ser, pero ahora en Aristóteles: el Director del Centre Léon Robin –que, como se sabe, consagró años de estudio a la ontología aristotélica– realiza un detallado examen de los casos en que el Estagirita mienta a la *ousía* –la primera de las categorías– como *tò tí esti* o como *tóde ti*, observando que, si bien no existe una oposición entre ambas fórmulas, Aristóteles las distingue y deja "pistas" para comprender esta dualidad. Otro texto interesante que podría subsumirse bajo el par *lógos*/ser es "Le *lógos* chez Plotin", en el que Luc Brisson explica cómo los *lógoi* –aquí "fórmulas racionales", mejor que "discursos"– que a nivel del Alma del mundo son responsables de la organización de las realidades sensibles y del mantenimiento del orden que asegura su unidad, no son sino reflejos de los *lógoi* en los que consiste el Alma hipóstasis; éstos a su vez son imágenes de las Formas inteligibles que constituyen el Intellecto, emanado de lo Uno, origen de todo lo que *es*.

Pero *diálogo* también puede entenderse como un cierto camino: el camino que se recorre para investigar una materia y alcanzar sus primeros principios. Tanto Michel Narcy como Jacques Brunschwig se ocupan de este aspecto al estudiar la dialéctica en Platón y Aristóteles. Se trata de una noción que plantea grandes conflictos a quienes la abordan: ¿la dialéctica de Aristóteles se identifica con la de Platón? ¿Hay criterios suficientes para saber cuándo el Estagirita procede dialécticamente y cuándo no?

El *diálogo* aparece, además, en la práctica. La acción –por lo menos en la filosofía aristotélica– necesita deliberación, y en este sentido el *lógos* pasa a vincularse con el *kairós*, con el tiempo práctico, con lo no calculable. Lambros Couloubaritsis ("Temps et action dans l'Éthique à Nicomaque") observa cómo en Aristóteles adquiere relevancia ese tiempo no reductible al tiempo físico (*khronos*), esa temporalidad propia de la acción. Alonso Tordesillas ("Équité et kaironomie chez Aristote") señala la importancia que reviste la equidad toda vez que la universalidad de la ley no es suficiente; en esos casos se precisa de una comprensión de las circunstancias y de una legislación oportuna: las bases de una *kaironomía* están sentadas. Jean Frère, por su parte, analiza detalladamente las apariciones de la fogosidad (*thumos*) y la cólera (*orgé*) en el *corpus* aristotélico, probando que para el Estagirita, a diferencia de Platón, tanto el *thumos* como la *orgé* son pasiones penetradas de racionalidad y capaces de conducirnos hacia ciertas virtudes como el coraje o la suavidad de carácter. Rémi Brague, finalmente, aborda la cuestión de la mundanidad de la moral en Aristóteles y en Kant: para el fundador del Liceo, el mundo es un marco para la vida moral, pero no es él mismo moral; para el pensador de Königsberg, en cambio, el mundo sí posee un status ético.

Con esto se abre otro sentido de *diálogo*: nos referimos al intercambio filosófico, en particular al diálogo permanente que se puede mantener con los Anti-

guos. El artículo de Monique Dixsaut, "Ce qu'était Aristote pour Nietzsche", apunta en esa dirección, cuando indaga cómo Nietzsche leyó a Aristóteles y en qué aspectos pueden percibirse divergencias entre ambos: el pensador alemán –explica Dixsaut– no llegó a apasionarse por los planteos aristotélicos, pero vio en el Estagirita a un hombre de la ciencia pura que contribuyó, sin ser del todo responsable, a la eliminación de la concepción trágica del mundo.

Los trabajos que figuran en el volumen reseñado se articulan claramente en torno a la cuestión del *diálogo*. Más allá de las diferentes dimensiones del *lógos* que decide tratar cada autor, en todos los casos hallamos argumentaciones sólidas y al mismo tiempo espacios concedidos a la prosecución del debate. Quizás sean éstas lecciones transmitidas por Pierre Aubenque, que dedicó una vida a la investigación, la discusión y la enseñanza –de hecho, todos los artículos de la compilación se inspiran en algunas de las tesis forjadas por él. Y en este sentido, *Ontologie et Dialogue* constituye un verdadero homenaje.

Octavio Kulesz
Universidad de Buenos Aires

Arleen Salles y María Julia Bertomeu (edit) *Bioethics: Latin American Perspectives*, (New York, Amsterdam, Rodopi, Value Inquiry Book Series, Philosophy in Latin America 2002), 110 pp.

Esta compilación de trabajos sobre filosofía aplicada brinda al lector de habla inglesa un claro aporte sobre el desarrollo de la filosofía en la región y el estado actual de las controversias más importantes.

Lo medular en estos diez trabajos es que aportan criterios universales independientes, aunque relacionados con las prácticas particulares. A pesar de sus situaciones históricas específicas, es dable que Latino América muestre una base moral compartida, por lo menos, en algunos temas de bioética.

Los artículos están agrupados en cuatro grandes temas:

- I- La autonomía y el derecho de tomar decisiones
- II- La justicia y el derecho al cuidado de la salud
- III- La experimentación en sujetos humanos, y
- IV- Temas éticos en la obtención y asignación de órganos.

En *Autonomía y Cultura: el caso de Latino América*, Arleen Salles señala que la noción de autonomía es de importancia medular para la bioética, y con un método argumentativo riguroso y lúcido analiza y rechaza las hipótesis que

cunstancias naturales o sociales tienen un derecho absoluto sobre tales ventajas. Coinciden con Dieterlen en señalar la estrecha relación entre el desarrollo económico y la salud, pero haciendo hincapié en la idea de justicia. Las autoras subrayan que la tradición que va de Kant a Rawls es la más adecuada para conceptualizar los derechos, y se detienen en la teoría de Amartya Sen que les permite –como dijimos– relacionar los conceptos de persona moral con el derecho al cuidado de la salud.

Florencia Luna, en *Investigación en los Países en Desarrollo: los Temas Éticos*, analiza los argumentos a favor y en contra de los experimentos efectuados en los países en desarrollo, y los problemas relevantes para la protección de los sujetos de investigación, entre ellos el de la corrupción estructural. Uno de los aspectos importantes del artículo es la consideración ética de las pruebas que utilizan el grupo placebo cuando ya existe un tratamiento mejor probado, señalando que violan la Declaración de Helsinki. Analiza también las particularidades degradadas del contexto de investigación en los países en desarrollo, entre los que cita algunos silenciados como el mal pago a los profesionales y el pago de los *sponsors* directamente a los profesionales, pasando por alto a la institución, lo que ve como caldo de cultivo para la corrupción.

José Roberto Goldim, en *Bioética y la Investigación en Brasil*, traza la evolución de la investigación con seres humanos y la bioética en su país durante el siglo veinte. Destaca que, por primera vez en la historia de Brasil, en 1986 la Sociedad Brasileira de Medicina Tropical, junto con la Asociación Brasileira de Antropología, propusieron un Código de Derechos de Salud de la Comunidad que abogaba por respetar y consultar a las comunidades que participaban en los proyectos de investigación, con recomendaciones muy innovadoras. El concepto de autonomía –que recorre todo el libro– tiene en este trabajo un énfasis particular sobre la autonomía de la comunidad, basada en la justicia social más que en la tradición anglosajona que coloca el énfasis en la autonomía individual.

Goldim dice que “Además de cumplir con la ley, las propuestas (de investigación) deben también estar justificadas por su contribución a la producción de nuevo conocimiento, su relevancia, y su factibilidad. Los investigadores deben conducir sus estudios tanto para el bienestar de la sociedad como un todo, como también para el de cada una y todas las personas”.

Eduardo Rivera López en *¿Qué es (exactamente) lo incorrecto en vender nuestras partes del cuerpo?* discute si la comercialización de órganos o partes del cuerpo de personas vivas es objetable desde el punto de vista ético. Y si es así, por qué lo es. El núcleo de su argumento es que lo que debe evitarse son las condiciones que tientan a la gente a vender sus órganos y no simplemente prohibir su venta. Sigue en esa línea al filósofo Radcliff-Richards quien no objeta la venta de órganos en sí misma, sino que las personas deseen venderlos, porque no encuentran una alternativa mejor en la situación imperante. Considera hipó-

crita de parte del estado prohibir por “razones morales” comportamientos que no dañan a otros y que son la consecuencia natural de ese estado de cosas.” Mientras las condiciones de pobreza persistan, dice Rivera López, la simple prohibición de ninguna manera implica una mejora moral significativa en la sociedad.

En *Comités de Ética Interdisciplinarios para la Determinación del Criterio de Distribución de Órganos*– María Graciela de Ortúzar toma como punto de partida el marco teórico que proveen John Rawls y Norman Daniels para discutir los criterios de distribución. Describe y analiza tres modelos para la toma de decisiones: el de los expertos médicos, la consulta pública basada en las preferencias sociales y el propuesto por la autora, que es el comité interdisciplinario específico. Como en el resto de los artículos que componen el libro, la autonomía tiene un lugar en el análisis. Se afirma que proveer información y educación a los pacientes acerca de sus derechos es la vía esencial para promover acciones autónomas, también en el campo de los trasplantes como vía de tratamiento. Y destaca como tarea principal de estos comités establecer un criterio público –revisado regularmente– para la selección de pacientes en la lista de espera y la asignación final de órganos.

Como los autores analizan no sólo posiciones con las que acuerdan sino también aquellas que no comparten, este libro combina un alto nivel teórico –abarcador de las distintas tradiciones en ética aplicada– con una verdadera utilidad práctica. Es sin duda útil tanto para los legos, como para aquellas instituciones y profesionales –sean o no latinoamericanos– que a diario enfrentan estos problemas.

Silvia Woods, *Maestría de Ética aplicada*
Universidad de Buenos Aires

Michael Baurmann, *El mercado de la virtud: moral y responsabilidad social en la sociedad liberal*, Barcelona, Gedisa, 1998.

De acuerdo con Michael Baurmann existen en la teoría política contemporánea dos grandes críticas que amenazan al liberalismo. La primera de estas críticas, que procede de las filas del comunitarismo, cuestiona la idea de que un comportamiento individual autointeresado por parte de los ciudadanos pueda conducir a resultados sociales y morales ventajosos. Así, frente a la prioridad que el liberalismo asigna a los derechos individuales, los comunitaristas reivindican la vinculación del individuo con valores y formas de vida conjuntamente compartidas, alegando que sólo este tipo de vinculaciones permitiría trascender

el egoísmo de los intereses puramente privados en beneficio del interés general. Desde este punto de vista, una sociedad liberal impediría que los hombres participaran en auténticas relaciones sociales y conduciría, finalmente, a la crisis moral y a la desintegración.

A esta primera crítica se suma una segunda dificultad que pone el acento sobre el hecho de que existen situaciones en las cuales la orientación racional de los individuos hacia su propio interés puede conducir a resultados que se oponen diametralmente a los intereses de todos los participantes de un grupo. Así, por ejemplo, aunque la competencia en el mercado económico es beneficiosa para todos los actores que participan de él, cada agente tiende, sin embargo, a obtener beneficios que lo sustraigan de ella. Además de la formación de monopolios, la instrumentalización del gobierno por parte de los poderes económicos y la incapacidad de las sociedades de mercado para promover virtudes que son vitales para su subsistencia son también manifestaciones de este "abismo entre la racionalidad individual y la colectiva". Y por esa razón hay filósofos y teóricos políticos que piensan que las sociedades liberales encierran una tendencia autodestructiva. Para ellos una sociedad que allane el camino para la persecución de los intereses individuales, al destruir las metas colectivas y dificultar la cooperación, termina socavando los cimientos de toda sociedad. Con estas dos críticas se abre la pregunta de si una sociedad liberal es capaz de fomentar vinculaciones y disposiciones morales entre sus ciudadanos. Y precisamente esta cuestión de la "productividad moral" de las sociedades liberales es para Baurmann el problema clave del liberalismo.

Frente a quienes proponen un rearme religioso o ideológico que fortalezca las instituciones y las virtudes de los ciudadanos, y el abandono definitivo de la mirada secular propia de la teoría liberal, Baurmann trata de probar que una sociedad liberal está en condiciones de crear y conservar la moral y el sentido comunitario de sus miembros por sus propias fuerzas. Para esto, Baurmann construye un modelo de comportamiento en el que confluyen una reelaboración de la teoría del Homo oeconomicus y el modelo weberiano de maximización racional de la utilidad, que pretende mostrar que un comportamiento virtuoso puede responder al autointerés a largo plazo de los individuos. De ese modo rompe el análisis clásico sobre la realización de los intereses individuales que la enfoca como un problema exclusivamente referido al comportamiento que parece más conveniente elegir en cada situación aislada. El hecho de que las acciones virtuosas realizadas en empresas cooperativas aumente la reputación de quien las lleva a cabo permitiéndole obtener ventajas frente a quienes demuestran carecer de virtud, es para Baurmann una clara prueba de que un comportamiento virtuoso puede ser conveniente aún para un individuo autointeresado.

Surge, empero, el problema de que la observación de las acciones de los demás sólo permite inferencias dudosas acerca de su carácter y de su auténtica

disposición a mantener un comportamiento moral. Por ese motivo, Baurmann destina todo un capítulo a explicar porqué en el contexto de las empresas cooperativas los actores pueden distinguir efectivamente a los simuladores de las personas que son "moralmente íntegras", ya que únicamente bajo este supuesto respondería al autointerés de un agente no sólo *parecer* una persona virtuosa, sino también *serlo*. Ahora bien, una de las tesis centrales del planteo de Baurmann consiste en afirmar que esas virtudes que las personas adquieren al participar en empresas cooperativas pueden ser no solamente útiles para ellas mismas, sino también para la estabilidad de un orden social considerado en su conjunto. En ese sentido, cree que si los individuos se orientan en su acción colectiva por principios de respeto interpersonal y de equidad social, contribuirán espontáneamente a la protección y a la conservación de los bienes individuales y colectivos.

Contrariamente a la imagen que los críticos del liberalismo presentan de las sociedades liberales, Baurmann sostiene que en éstas no se ha "liberado" únicamente la esfera económica, sino también la asociación humana autodeterminada y la unión voluntaria en comunidades y grupos. Y la libertad de asociación permitiría, a su vez, la formación de un "mercado de la virtud" en el que surja "una demanda estable de personas 'virtuosas' y, por lo tanto, adecuadas para la cooperación y la colaboración" (p. 55). Puesto que este "mercado de la virtud" asegura, por medio de su "mano invisible", que no se seque la fuente de moralidad y de sentido comunitario que una sociedad requiere para su existencia, Baurmann resuelve el problema de la "productividad moral" de las sociedades liberales que hemos indicado como una dificultad central. Con este planteo parecen desdibujarse también las críticas comunitaristas, pues una sociedad liberal no estaría condenada a provocar el completo aislamiento de sus ciudadanos ni a disolver los compromisos éticos de todo tipo. Por el contrario, tal como demuestra el presente libro, una sociedad liberal es un presupuesto indispensable para el surgimiento de una moral en la que los intereses de todas las personas sean tenidos igualmente en cuenta, de modo que se conjuguen la libertad política y el bienestar económico.

El debate entre liberalismo y comunitarismo ha sido una constante en la filosofía política desde hace al menos dos décadas. En ese marco, Baurmann construye un argumento que renueva ciertas ideas fundamentales del liberalismo. Pero, al mismo tiempo, intenta trascender el punto de vista del liberalismo clásico según el cual una sociedad basada en el autointerés requiere sólo "un mínimo de moralidad", concediendo a los detractores de la tradición liberal que la virtud y el sentido comunitario son indispensables para su funcionamiento. Hasta donde alcanzo a ver, la novedad de su planteo radica en sostener que puede existir un "mercado de la virtud" análogo al mercado económico que asegure un comportamiento moral desde los cimientos mismos de una sociedad liberal. Y, aunque la exposición podría haberse completado con un análisis

más detallado del concepto de virtud y de la manera en que el autor lo comprende, esta es una obra que indaga de un modo original antiguos y actuales problemas de la tradición liberal.

Julio Montero
Universidad de Buenos Aires

SUMARIO DEL VOLUMEN

ARTÍCULOS

- D. Brauer, La "memoria productiva" - Acerca de la concepción de Hegel del recuerdo y la memoria..... 319
- A. Cassini, Bunge, la lógica inductiva y el grado de confirmación de las teorías 303
- J. Comesaña, En defensa de un externalismo epistémico 173
- E. Cresto, Creer, inferir y aceptar: una defensa de la inferencia a la mejor explicación apta para incrédulos 201
- O. Lombardi, Caos, ergodicidad e internalismo 7
- C. A. Martins, A teoria kantiana da *afecção*..... 81
- M. J. Müller, Itinerário da problemática da expressão na filosofia Merleau-Pontyana: continuidades e *descontinuidades* 55
- M. Nalli, Foucault fenomenólogo? Fenomenologia da *significação* na proto-arqueologia foucauldiana 35
- S. Pinto, Dos argumentos en contra del lenguaje de pensamiento..... 275
- L. A. Rossi, El tedio y el terror: la crisis de 1929 en el curso de Heidegger *Los conceptos fundamentales de la metafísica* 247
- C. Sartorio, El comportamiento colectivo y la relación entre responsabilidad y causación 231

SIMPOSIO SOBRE RAZÓN PÚBLICA

- O. Guariglia, Razón pública y democracia deliberativa 107
- M. V. Costa, Las virtudes ciudadanas en el liberalismo político rawlsiano..... 119
- M. Garreta Leclercq, Neutralidad estatal, igual respeto y falibilismo moral 131

DISCUSIÓN

- M. Bunge, La lógica inductiva es inútil 339

COMENTARIOS BIBLIOGRAFICOS

M. Baumann, <i>El mercado de la virtud: moral y responsabilidad social en la sociedad liberal</i> (J. Montero)	353
A. Brancacci, <i>La filosofía in età imperiale</i> (G. Livov).....	158
S. Cabanchik, <i>Introducciones a la filosofía</i> (M. L. Levinas).	151
N. L. Cordero (comp.), <i>Ontologie et Dialogue. Hommage à Pierre Aubenque</i> (O. Kulesz).	347
A. Ferraro, <i>La ética del positivismo jurídico</i> (H. Arrese Igor).	161
Z. Loparic, <i>A semântica trascendental de Kant</i> (H. Pringe).	344
A. López Fernández, <i>Conciencia y juicio en Kant</i> (M. Caimi).	163
S. Maidana, <i>El papel terapéutico de la filosofía</i> (M. B. Contratti)	166
V. Palacios, <i>Lenguaje y pacto en Thomas Hobbes</i> (A. Pac).	343
W. V. O. Quine, <i>Acerca del conocimiento científico y otros dogmas</i> (A. Cassini).	147
J. de Salas y F. Martín (comp.), <i>David Hume</i> (M. Rodríguez Bustamante).	155
A. Salles y M. J. Bertomeu (edit.), <i>Bioethics: Latin American Perspectives</i> (S. Woods).	349
J. L. Tasset, <i>La ética y las pasiones. Un estudio de la filosofía moral y política de Hume</i> (M. Costa).	145
G. Vidiella, <i>El derecho de la salud</i> (C. de la Sierra).	153