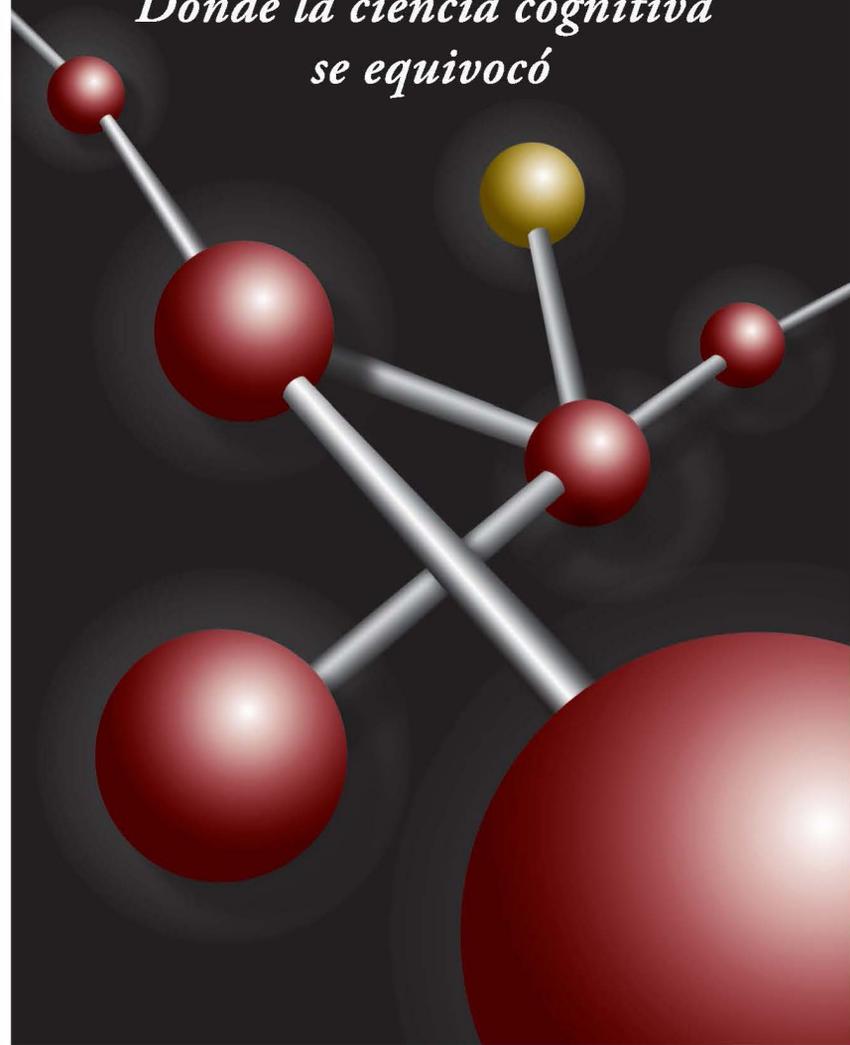


Ciencias cognitivas

Jerry A. Fodor

# Conceptos

*Donde la ciencia cognitiva  
se equivocó*



gedisa  
editorial

**Jerry A. Fodor**

---

**CONCEPTOS**

**Serie: CLA•DE•MA**  
**CIENCIAS COGNITIVAS**

Editorial Gedisa ofrece  
los siguientes títulos sobre

# CIENCIAS COGNITIVAS

- |                                      |  |
|--------------------------------------|--|
| AQUILES ESTÉ                         | <i>Cultura replicante</i>  |
| PASCUAL F. MARTÍNEZ-FREIRE           | <i>La nueva filosofía de la mente</i>  |
| NICHOLAS HUMPHREY                    | <i>Una historia de la mente</i>  |
| DANIEL C. DENNETT                    | <i>Contenido y conciencia</i>  |
| PAUL WATZLAWICK Y PETER KRIEG        | <i>El ojo del observador</i>   |
| STEPHEN R. GRAUBARD<br>(COMP.)       | <i>El nuevo debate sobre<br/>la inteligencia artificial</i>                                    |
| PAUL M. CHURCHLAND                   | <i>Materia y conciencia.<br/>Introducción<br/>contemporánea a la<br/>filosofía de la mente</i> |
| P. ENGEL (COMP.)                     | <i>Psicología ordinaria<br/>y ciencias cognitivas</i>  |
| F. VARELA, E. THOMPSON<br>Y E. ROSCH | <i>De cuerpo presente</i>  |
| DANIEL C. DENNETT                    | <i>La libertad de acción</i>   |
| HILARY PUTNAM                        | <i>Representación<br/>y realidad</i>   |
| DANIEL C. DENNETT                    | <i>La actitud internacional</i>  |
| FRANCISCO J. VARELA                  | <i>Conocer<br/>Realidad mental<br/>y mundos posibles</i>                                       |
| JEAN-PAUL DESPINS                    | <i>La música y el cerebro</i>  |
| EDGAR MORIN                          | <i>Introducción al<br/>pensamiento complejo</i>  |
| MARGARET A. BODEN                    | <i>La mente creativa</i>   |

# CONCEPTOS

*Donde la ciencia cognitiva  
se equivocó*

por

Jerry A. Fodor

gedisa  
editorial

Título original en inglés: *Concepts. Where cognitive science went wrong*  
Publicado por Clarendon Press, Oxford

*This translation of Concepts originally published in English in 1998 published by arrangement with Oxford University Press.*

La presente traducción de *Conceptos*, obra publicada originalmente en inglés, en 1998, se publica por acuerdo con Oxford University Press.

Traducción: Liza Skidelsky  
Revisión técnica: Eduardo Rabossi

Primera edición, septiembre de 1999, Barcelona  
Reedición, 2014

Derechos reservados para todas las ediciones en castellano

© 1998 Jerry A. Fodor

© Editorial Gedisa, S.A.  
<http://www.gedisa.com>

eISBN:978-84-1891-414-0

Queda prohibida la reproducción total o parcial por cualquier medio de impresión, en forma idéntica, extractada o modificada, en castellano o en cualquier otro idioma.

*Para Janet, KP y Anthony; familia nuclear*

**Obras de Jerry A. Fodor  
publicadas en español**

*La explicación psicológica*  
Madrid, Cátedra, 1980

*El lenguaje del pensamiento*  
Madrid, Alianza, 1985

*La modularidad de la mente*  
Madrid, Morata, 1986

*Psicosemántica*  
Madrid, Tecnos, 1994

*El olmo y el experto: el reino de la mente  
y su semántica*  
Barcelona, Paidós Ibérica, 1996

CHORUS: Zurück!

TAMINO: ...Zurück?

Da seh ich noch ein Tur,

Vielleicht find ich den Eingang hier.

*Die Zauberflöte*

CORO: ¡Atrás!

TAMINO: ...¿Atrás?

Ahí veo otra puerta,

Tal vez aquí encuentre la entrada.

*La flauta mágica*



# Índice

ABREVIATURAS Y CONVENCIONES TIPOGRÁFICAS .....	12
PREFACIO .....	13
1. Introducción filosófica: La teoría de fondo .....	17
2. Introducción no filosófica: Lo que los conceptos tienen que ser .....	45
3. La defunción de las definiciones. Parte I: El relato lingüístico .....	67
4. La defunción de las definiciones. Parte II: El relato filosófico .....	105
5. Prototipos y composicionalidad .....	129
Apéndice 5A: Postulados de significado .....	152
Apéndice 5B: La “teoría teoría” de los conceptos .....	157
6. Innatismo y ontología, Parte I: El argumento estándar .....	171
Apéndice 6A: Similitud .....	198
7. Innatismo y ontología, Parte II: Conceptos de clase natural .....	203
Apéndice 7A: Cuadrados redondos .....	224
BIBLIOGRAFÍA .....	229
ÍNDICE DE AUTORES .....	235

## **Abreviaturas y convenciones tipográficas**

Se adoptan las siguientes convenciones:

Los conceptos están contruidos como particulares mentales. Los nombres de los conceptos están en mayúsculas. Así, 'ROJO' nombra el concepto que expresa la *rojez* o *la propiedad de ser rojo*. Las fórmulas con mayúsculas no son, en general, descripciones estructurales de los conceptos que denotan. Véase el capítulo 3, nota 1.

Los nombres de las expresiones en español aparecen entre comillas simples. Así, 'rojo' es el nombre de la palabra homófona en español.

Los nombres de los valores semánticos de las palabras y los conceptos están en itálicas. Así, 'ROJO expresa la propiedad de *ser rojo*' y "Rojo' expresa la propiedad de *ser rojo*' son ambas verdaderas.

Las siguientes abreviaturas se utilizan frecuentemente (especialmente en los capítulos 6 y 7):

TRM: Teoría representacional de la mente

SRI: Semántica del rol informacional

MDP: Modo de presentación

RM: Representación mental

AI: Atomismo informacional

AE: Argumento estándar (en favor del innatismo radical de los conceptos)

AIS: Atomismo informacional con un suplemento (= AI más una teoría enlace de la posesión de conceptos)

Problema p/P: Problema picaporte/PICAPORTE

# Prefacio

En verdad, estoy un poco preocupado por el subtítulo. Hay una gran cantidad de estudios que han revisado lo que está mal en la ciencia cognitiva, dedicados a tirar, junto con el bebé no sólo la bañera, sino además la ducha, la toalla de baño, el cuarto de baño, muchos mirones inocentes y grandes sectores del bajo Manhattan\*. Los diagnósticos que ofrecen estos libros difieren bastante entre sí, y existe una preocupación real de que el paciente pueda morir por sobremedicación. Lo que está mal en la ciencia cognitiva es que, hablando estrictamente, no hay en absoluto estados mentales. O, estrictamente hablando, no hay estados mentales excepto los conscientes. O, estrictamente hablando, la intencionalidad está en el ojo del observador. O del intérprete. O del traductor. O es sólo una actitud. O es una malla gruesa sobre una red neuronal. O lo que sea.

Considero que estos puntos de vista simplemente no son creíbles y no tengo deseos de agregar nada a sus listas. En la mayoría de los temas, este libro está completamente comprometido con el programa tradicional de la ciencia cognitiva: los organismos superiores actúan a partir del contenido de sus estados mentales. Estos estados mentales son representacionales; en efecto, son relaciones con representaciones mentales. Por ende, el objetivo científico de la psicología es comprender qué son las representaciones mentales y explicitar las leyes y los procesos causales que las subsumen. En verdad, nada de esto cambió mucho desde Descartes.

Por eso esta es una crítica interna; es lo que a la tía Ana le gusta llamar crítica "constructiva". Por otro lado, dado el amplio consenso tradicional acerca de los objetivos y la arquitectura de las teorías de la cognición, pienso que algo salió bastante mal con respecto a cómo se llevó a cabo el programa. Por razones que intentaré aclarar, el núcleo de una ciencia cognitiva es su teoría de

los conceptos. Y pienso que la teoría de los conceptos que la ciencia cognitiva clásicamente adoptó, en cierta manera está seriamente equivocada. A diferencia de prácticamente todos los demás que trabajan o trabajaron en esta tradición, considero que la teoría de los conceptos debe ser atomista. Me siento un poco solitario estando aquí fuera sin compañía, pero eso me da espacio para manio-brar. En este libro me ocupo de por qué la teoría de los conceptos debe ser atomista. Y por qué le trajo problemas a toda la ciencia cognitiva el no haber sido atomista. Y cómo podrían ser la psicología, la ontología y la semántica de una teoría atomista de los conceptos.

La discusión se agrupa en tres secciones. Los capítulos 1 y 2 son en gran parte expositivos; están dedicados a esbozar lo que considero como la estructura general de las teorías cognitivas clásicas, y a ubicar en ese marco los temas relacionados con los conceptos. En particular, quiero establecer algunas restricciones sobre una teoría aceptable de los conceptos que, argumentaré, deben ser admitidas por quienquiera que desee desarrollar una teoría representacional de la mente. Los capítulos 3-5 analizan a la luz de estas restricciones las principales teorías de los conceptos que circulan actualmente en lingüística, filosofía y psicología cognitiva. Sostengo que todas son variantes de la explicación del “rol inferencial” del contenido conceptual. Argumentaré que esta concepción del rol inferencial del *contenido* de los conceptos y la concepción antiatomista de la *estructura* de los conceptos sobrevivieron por mucho tiempo una a costa de la otra y que ambas tendrán que marcharse. Con ellas desaparecerán todas las teorías estándar actuales acerca de qué son los conceptos: definiciones, estereotipos, prototipos, abstracciones a partir de sistemas de creencias, etc. Espero que este material crítico será de interés para los labradores empíricos en las viñas de la ciencia cognitiva. Espero que los ataques a las teorías estándar de los conceptos los mantengan alertas, aun si no aprueban mis propuestas para una alternativa atomista. Pienso que la mayoría de lo que la ciencia cognitiva contemporánea cree acerca de los conceptos es, radical y *demos-trablemente*, falso; y que se necesita hacer algo bastante drástico al respecto.

Los capítulos 6 y 7 exploran la alternativa atomista. Resultará, no tan sorprendentemente, que el atomismo con respecto a la estructura de los conceptos tiene implicaciones profundas para cuestiones psicológicas acerca de cómo se adquieren los conceptos,

para cuestiones metafísicas acerca de cómo se individualan los conceptos, y para cuestiones ontológicas acerca de qué clases o propiedades expresan esos conceptos. Antes de terminar tendremos mucho que decir, desde una perspectiva revisionista, acerca del innatismo, la información y los picaportes. Aunque las motivaciones para todo esto surgen desde adentro de la ciencia cognitiva, pasarse al atomismo conceptual requiere algo parecido a un cambio de punto de vista acerca del mundo. Si es así, que así sea.

Tuve muchos problemas en relación al tono de voz. Algunos de los argumentos que tengo en oferta son manifiestamente filosóficos; algunos dependen de datos experimentales y lingüísticos; muchos son metodológicos; y otros son sólo apelaciones al sentido común. Parte de lo que hace que trabajar en ciencia cognitiva sea tan difícil es que no hay una manera de hablar que sea adecuada para todas estas clases de lenguajes. A la larga, desistí; escribí simplemente como los temas en estudio parecían justificarlo. Si esto no suena exactamente como filosofía, no importa, en tanto no suene exactamente como psicología, lingüística o inteligencia artificial.

Una versión condensada de este material fue presentada en las Conferencias John Locke realizadas en la Universidad de Oxford en 1996. Estoy agradecido, más de lo que puedo expresar, a amigos y colegas en Oxford por darme la ocasión de exponer este material, por haberlo soportado y por sus críticas, discusión e inagotable hospitalidad. Estoy especialmente obligado, en todos estos aspectos, con Martin Davies, Chris Peacock y Galen Strawson; y con el All Souls College por darme una oficina, hospedaje y una cuenta de correo electrónico.

Otras obligaciones intelectuales (la lista es ciertamente incompleta): con Kent Bach y Ken Taylor por comentarios detallados y útiles a una versión anterior. Con Paul Boghossian por almuerzos filosóficos. Con Ned Block, Paul Bloom, Noam Chomsky, Jim Higginbotham, Ray Jackendoff, Ernie Lepore, Joe Levine, Steven Pinker, Zenon Pylyshyn, Georges Rey, Stephen Schiffer, Gabe Segal, Barry Smith, Neil Smith. Y con muchos, muchos otros.

El Departamento de Filosofía de Rutgers University me concedió tiempo libre para hacer el viaje a Oxford. Eso fue amable y cortés, y estoy contento de tener esta oportunidad de hacerle extensivo mi agradecimiento.

Pues bien, aquí está el libro. Resultó divertido armarlo y espero que sea divertido leerlo. Espero que les guste. Espero que algo de él sea verdadero.

Nueva York, 1997

JERRY A. FODOR

\* El proverbio inglés dice: “No eches el bebé junto con el agua del baño” como advertencia contra la exageración. Mantenemos este dicho literalmente por su divertida ampliación. [N. del T.]

# 1

## Introducción filosófica: La teoría de fondo

No es necesario decirlo, ese sistema de creencias un tanto barroco dio lugar a explicaciones increíblemente complicadas por parte de los ancianos de la tribu...

WILL SELF

Mi tema es qué son los conceptos. Puesto que estoy interesado en esta cuestión, primariamente tal como surge en el contexto de las teorías “representacionales” de la mente (TRM), una manera natural de comenzar sería contarle algo acerca de las TRM y de cómo suscitan la pregunta de qué son los conceptos. Podría, entonces, dar mi respuesta y usted podría decirme lo que le parece que está mal. Sin duda, la discusión que seguiría sería abstracta y cargada de teoría; pero, con algo de suerte, filosóficamente inocente.

De hecho, este es casi el curso que propongo seguir. Pero, en el clima presente de opiniones filosóficas, para bien o para mal, tal vez no es posible simplemente aventurarse a hacerlo. Las TRM presentan toda clase de problemas, tanto de esencia como de forma. Muchos pueden suponer que el proyecto completo de tratar de construir una está irremediablemente mal encaminado; si lo está, entonces ¿a quién le preocupa lo que dicen las TRM acerca de los conceptos? De manera que creo que les debo algún tipo de argumento general para mostrar que el proyecto no está irremediablemente mal encaminado.

Pero parece que he envejecido escribiendo libros que defienden las TRM; se me ocurre que si dejara de escribir libros defendiendo las TRM, quizá dejaría de envejecer. Por eso, en su lugar, les contaré una broma. Es una *vieja* broma y vale la pena contarla.

*Vieja broma:* Había una vez un discípulo que fue a ver a su gurú y le dijo: “Gurú, ¿qué es la vida?” A lo que el gurú contestó, des-

pués de pensarlo mucho: “Hijo, la vida es como una fuente”. El discípulo estaba indignado. “¿Es esto lo mejor que puede hacer? ¿Es esto lo que llama sabiduría?”. “Está bien”, dijo el gurú; “No te alteres. Quizá no es como una fuente”.

Este es el final de la broma, pero no es el final de la historia. El gurú se dio cuenta de que seguir esta línea era perder clientes, y los gurúes también tienen que comer. Así, la siguiente vez que un discípulo le preguntó: “Gurú, ¿qué es la vida?” su respuesta fue: “Hijo, no te lo puedo decir”. “¿Por qué no puede?” quiso saber el discípulo. “Porque”, dijo el gurú, “la pregunta ‘¿qué es tener una vida?’ es lógicamente anterior”. “¡Ah!”, dijo el discípulo, “esto es muy interesante”; y se inscribió para el curso académico completo.

No voy a emprender una defensa completa de la TRM; pero sí quiero empezar con algunas cuestiones metodológicas acerca de si tener un concepto es lógicamente anterior a ser un concepto, y qué consecuencia, si la hay, esto tiene para la teorización acerca de la representación mental.

Es una verdad general que si se sabe *qué es un X*, entonces también se sabe *qué es tener un X*. Y asimismo lo contrario. Esto se aplica a los conceptos en particular: la pregunta de qué son y la pregunta de qué es tenerlos están lógicamente ligadas; si se está comprometido con una, se está *de este modo* comprometido con la otra, se quiera o no. Supongamos, por ejemplo, que su teoría es que los conceptos son calabazas. Entonces, muy bien, tendrá que ser parte de su teoría que tener un concepto es tener una calabaza. E, inversamente: si su teoría es que tener un concepto es tener una calabaza, entonces tendrá que ser parte de su teoría que las calabazas son lo que los conceptos son. Supongo que todo esto es una perogrullada.

Ahora bien, hasta hace poco (en cualquier caso, hasta el siglo xx) prácticamente todos dieron por sentado que la explicación de la *posesión* de conceptos tiene que ser parasitaria de la explicación de la *individuación* de conceptos. Primero se dice qué es para algo *ser* el concepto *X*—se dan las “condiciones de identidad” de los conceptos— y entonces *tener* el concepto *X* es sólo *tener lo que sea que el concepto X resulte ser*. Pero las modas filosóficas cambiaron. Casi sin excepción, las teorías actuales acerca de los conceptos invierten la dirección clásica de análisis. Sus aspectos esenciales radican en lo que dicen acerca de las condiciones para *tener* el concepto *X* y tratan como derivativa la historia acerca de *ser* el concepto *X*.

El concepto *X* es sólo: *lo que sea que constituya tener en el concepto X*. Más aún, el nuevo consenso es que se tienen que tomar realmente las cosas en este orden; se dice que si se comienza por el otro lado, las sanciones en las que se puede incurrir son terribles. (Similarmen-*te*, *mutatis mutandis* para ser el significado de una palabra vs. conocer el significado de una palabra. Aquí y en otros lugares, propongo movernos, de manera bastante libre, entre los conceptos y los significados de las palabras; como sea que esto pueda resultar a largo plazo, para los propósitos de la presente investigación los significados de las palabras simplemente son los conceptos.)

Se puede preguntar, de manera razonable, cómo podría posiblemente darse esta rígida asimetría metodológica. Hemos estado viendo que la relación entre “es un *X*” y “tiene un *X*” es conceptual; se fija una y con eso se fija la otra. Entonces, ¿cómo podría haber una cuestión de principio acerca de por cuál se tiene que comenzar? La respuesta es que cuando los filósofos toman una línea dura en un tema metodológico casi siempre hay un subtexto metafísico. El presente caso no es una excepción.

Por un lado, las personas que empiezan de una manera tradicional preguntando “¿qué son los conceptos?” generalmente sostienen una metafísica tradicional de acuerdo con la cual un concepto es un tipo de particular mental. Espero que esta idea se aclare cada vez más a medida que avancemos. Es suficiente por ahora que la tesis de que los conceptos son particulares mentales está pensada para implicar que *tener* un concepto está constituido por tener un particular mental, y por ende excluir la tesis de que tener un concepto está constituido, en algún sentido interesante, por tener rasgos o capacidades mentales.<sup>1</sup> Podría decirse que tener el concepto *X* es tener la habilidad de pensar acerca de los *X* (o mejor, tener el concepto *X* es, “cuando mucho”, ser capaz de pensar acerca de los *X*). Pero, a pesar de que esto es suficientemente verdadero, no altera la situación metafísica tal como se la concibió tradicionalmente. Porque pensar acerca de los *X* consiste en tener pensamientos acerca de los *X*, y se supone que los pensamientos también son particulares mentales.

Por otro lado, las personas que comienzan preguntando “¿Qué es la posesión de conceptos?” generalmente tienen en mente como respuesta un cierto tipo de pragmatismo. Tener un concepto es un asunto de lo que se es capaz de *hacer*, es una cierta clase de ‘saber-cómo’ epistémico. Quizá, tener el concepto *X* sea algo así como *ser capaz, de mane-*

*ra fiable, de reconocer los X y/o ser capaz, de manera fiable, de extraer inferencias correctas acerca de la Xdad.*<sup>2</sup> De cualquier manera, una explicación que interprete tener conceptos como tener capacidades está pensada para excluir una explicación que interprete a los conceptos como especies de particulares mentales: las capacidades no son clases de cosas; a fortiori, no son clases de cosas *mentales*.

Así, reitero, la doctrina metodológica de que la posesión de conceptos es lógicamente anterior a la individuación de conceptos, frecuentemente manifiesta una preferencia hacia una ontología de disposiciones mentales más que hacia una ontología de particulares mentales. Este tipo de situación es familiar para los filósofos; proponer un análisis disposicional en ayuda de reducciones ontológicas es el método de la filosofía crítica que el empirismo nos enseñó. Si se les tiene inquina a los gatos, redúzcalos a posibilidades permanentes de sensación. Si se les tiene inquina a los electrones y protones, redúzcalos a posibilidades permanentes de resultados experimentales. Y así en más. Sin embargo, hay una diferencia notable entre el reduccionismo acerca de gatos y el reduccionismo acerca de conceptos: quizás algunas personas creen que *tienen* que pensar que los gatos son constructos a partir de experiencias posibles, pero seguramente nadie piensa eso realmente; se tolera un poco de *mauvaise foi* en metafísica. Sin embargo, aparentemente muchas personas piensan que los conceptos son constructos a partir de capacidades mentales (específicamente epistémicas). En consecuencia, y esta es una consideración que tomo bastante en serio, mientras que nadie construye teorías biológicas bajo el supuesto de que los gatos son sensaciones, gran parte de la ciencia cognitiva actual, y prácticamente toda nuestra actual filosofía de la mente, está construida bajo el supuesto de que los conceptos son capacidades. Si este supuesto es erróneo, se necesitarán revisiones muy radicales. Al menos así argumentaré.

Resumiendo: es enteramente plausible que una teoría acerca de qué son los conceptos deba responder a la pregunta “¿qué es tener un concepto?” y, *mutatis mutandis*, que una teoría del significado deba responder a la pregunta “¿qué es comprender un lenguaje?”. Sin embargo, vimos que esta demanda metodológica no tendenciosa a menudo concuerda con un sustantivo programa metafísico: a saber, la reducción de los conceptos y los significados a las capacidades epistémicas.

Así, Michael Dummett (1993a:4), por dar un ejemplo ilustre, dice que “cualquier teoría del significado que no sea o que no lleve

inmediatamente a una teoría de la comprensión, no cumpliría el propósito para el cual, filosóficamente, necesitamos una teoría del significado”. Tal como se destacó antes, hay una lectura en la que esto es verdad pero es inofensivo, puesto que *cualquier* constructo ontológico del *significado de una expresión* que establezcamos, automáticamente proveerá un constructo correspondiente tanto para *comprender la expresión* como para *captar* su significado. Sin embargo, Dummett no está recomendando esta verdad trillada. Más bien, él tiene en mente que una semántica aceptable debe explicar el contenido lingüístico haciendo referencia a las capacidades “prácticas” que los usuarios del lenguaje tienen qua usuarios de ese lenguaje. (De manera correspondiente, una teoría que explicara la noción de contenido conceptual lo haría sólo haciendo referencia a las capacidades prácticas que confiere tener los conceptos: Más aún, Dummett intenta, si lo interpreto bien, imponer esta condición de una forma muy fuerte: las capacidades sobre las cuales el significado lingüístico superviene deben ser tales que puedan manifestarse en la conducta, de manera individual y determinada. “Un axioma gana su lugar en la teoría [del significado]...sólo en la medida en que se lo requiere para la derivación de teoremas, la adscripción de este conocimiento implícito al hablante *se explica en términos de habilidades específicas que manifiestan ese conocimiento* (1993b:38, la bastardilla es mía).

No sé con certeza por qué Dummett cree eso, pero sospecho oscuramente que es víctima de ansiedades escépticas atávicas con respecto a la comunicación. Párrafos como el siguiente son recurrentes en sus escritos:

¿Qué ... constituye la comprensión de las oraciones de un lenguaje por parte de un sujeto...? ¿[Es] el tener internalizada una cierta teoría del significado para ese lenguaje?... entonces, por cierto cuando el sujeto toma parte en intercambios lingüísticos, su conducta puede, a lo sumo, ser evidencia fuerte pero falible en favor de la teoría internalizada. Sin embargo, en ese caso la suposición del oyente de que comprendió al hablante nunca puede ser definitivamente refutada ni confirmada. (1993c:180; nótese cuánto peso tiene aquí la palabra “definitivamente”).

Así, aparentemente, la idea es que las teorías acerca del contenido lingüístico tienen que reducirse a teorías acerca del uso del lenguaje y las teorías acerca del uso del lenguaje tienen que reducirse a teorías acerca de las capacidades lingüísticas del hablante

y las teorías acerca de las capacidades lingüísticas del hablante están restringidas por el requisito de que cualquier capacidad que es constitutiva del conocimiento de un lenguaje se puede manifestar, de manera abierta y específica, en el uso del lenguaje por parte de un usuario. Todo esto debe ser para ayudar a diseñar un antiescepticismo a prueba de balas con respecto a la comunicación, puesto que parecería que para *otros* propósitos que no sea el de refutar a los escépticos, todo lo que requiere una teoría de la comunicación es que las emisiones del hablante causen de manera fiable ciertos “procesos internos” en el oyente; específicamente, procesos mentales que hacen que acontezca en el oyente tener el pensamiento que el hablante intentó que tuviese.

Sin embargo, si el escepticismo es realmente el secreto vergonzante de Dummett, me parece que la preocupación está doblemente fuera de lugar: primero porque las preguntas de las que se ocupan básicamente las teorías del significado son metafísicas más que epistémicas. Esto es como tiene que ser; comprender qué es una cosa es invariablemente anterior a comprender cómo sabemos lo que es. Y, en segundo lugar, porque no hay una razón obvia de por qué las atribuciones que están basadas en la conducta como las de conceptos, significados, procesos mentales, intenciones comunicativas y demás, tengan que estar más libres del riesgo inductivo normal que las que están basadas en la percepción, como las atribuciones de colas a los gatos. Lo mejor que obtenemos en cualquiera de ambos casos es “evidencia fuerte pero falible”. Por cierto, las verdades contingentes son como estas, según nos lo enseñó Hume hace algún tiempo. Sin duda, esta es la actitud que Dummett intenta rechazar como inadecuada para los propósitos por los cuales necesitamos “filosóficamente” una teoría del significado. Mucho peor, quizás, para la probabilidad de que los filósofos obtengan de una teoría del significado lo que Dummett dice que necesitan. Yo, por nombrar alguno, no esperaré que una buena explicación de qué son los conceptos refutara el escepticismo acerca de otras mentes mucho más de lo que esperaré que una buena explicación de lo que son los gatos refutara el escepticismo acerca de otros cuerpos. En ambos casos, estoy muy dispuesto a aceptar teorías que sean meramente *verdaderas*.

Entonces, una vez desembarazados de las inhibiciones metodológicas, he aquí cómo propongo organizar nuestro viaje. De manera muy general, los conceptos son los constituyentes de los estados

mentales. Así, por ejemplo, creer que *los gatos son animales* es un estado mental paradigmático, y el concepto ANIMAL es un constituyente de la creencia de que *los gatos son animales* (y de la creencia de que *los animales a veces muerden*; etc. Dejo pendiente si el concepto ANIMAL es, de la misma manera, un constituyente de la creencia de que *algunos gatos muerden*; enseguida plantaremos esta pregunta). Por eso, el hogar natural de una teoría de los conceptos es ser parte de la teoría de los estados mentales. A lo largo de este libro supondré que la TRM es la teoría correcta acerca de los estados mentales (cognitivos). Pues bien, comenzaré con una exposición de la TRM, es decir con una exposición de una teoría acerca de lo que son los estados y procesos mentales. Resultará que los estados y procesos mentales son típicamente especies de relaciones con representaciones mentales; los conceptos son típicamente las partes de estas últimas.

En efecto, seguir este curso es suponer que es correcto que a una teorización acerca de la naturaleza de los conceptos le preceda una teorización acerca de la posesión de conceptos. Tal como vimos, dejando de lado un subtexto metafísico, esta suposición tiene que ser inofensiva; las teorías de la individuación y las teorías de la posesión son trivialmente intertraducibles. Sin embargo, una vez que tengamos la TRM adecuada, argumentaré en favor de una versión muy fuerte del atomismo psicológico, según la cual qué conceptos se tienen es independiente, desde un punto de vista conceptual y metafísico, de qué capacidades epistémicas se tienen. Si esto es así, entonces, de manera manifiesta, los conceptos no podrían *ser* capacidades epistémicas.

Espero no cometer ninguna petición de principio al proceder de esta manera; o al menos que no me atrapen cometiendo alguna. Pero estoy de acuerdo en que si hay un argumento a priori demoledor de que los conceptos son constructos lógicos a partir de capacidades, entonces mi punto de vista acerca de la ontología de los conceptos no puede ser correcto y tendré que renunciar a la ciencia cognitiva, tal como la concibo. Está bien. Si hay un argumento a priori demoledor de que los gatos son constructos lógicos a partir de sensaciones, entonces, mis puntos de vista acerca de la ontología *de los gatos* tampoco pueden ser correctos y tendré que renunciar a la biología, tal como la concibo. Realmente, ninguna de las posibilidades me preocupa demasiado.

Pues bien, empecemos por fin.

## TRM

La TRM es realmente una vaga confederación de tesis; le falta, para decirlo suavemente, una formulación canónica. Para los propósitos presentes, dejemos que sea la conjunción de lo siguiente:

Primera Tesis: *La explicación psicológica es típicamente nómica y es completamente intencional.* Las leyes que invoca la explicación psicológica expresan, de manera típica, relaciones causales entre *estados mentales que se especifican bajo una descripción intencional*; es decir, entre estados mentales que se escogen por referencia a sus contenidos. Los paradigmas son las leyes acerca de las relaciones causales entre las creencias, deseos y acciones.

Reconozco que existen quienes piensan (la mayoría en el sur de California, por supuesto) que la explicación intencional a lo sumo es temporaria, y que las teorías de la mente eventualmente se formularán (o, de todas formas, tendrían que formularse) en el putativo idioma puramente extensional de la neurociencia. Pero no hay ninguna razón en el mundo para considerar seriamente esta idea y, en lo que sigue, no lo haré.

También existen quienes, aunque son entusiastas de la explicación intencional, niegan la posibilidad metafísica de leyes acerca de los estados intencionales. Tampoco propongo considerar seriamente esto en lo que sigue. Por un lado, considero que los argumentos que se supone que muestran que no puede haber leyes intencionales son muy difíciles de seguir. Por otro lado, si no hay leyes intencionales, entonces no se puede hacer ciencia a partir de explicaciones intencionales; en cuyo caso, no entiendo cómo la explicación intencional *podría* ser algo mejor que meramente temporaria. A través de los años, un número de filósofos se encargaron de explicarme amablemente para qué serían buenas las explicaciones intencionales no nómicas. Aparentemente, tiene que ver con que el dominio intencional (o tal vez, el dominio racional) sea autónomo. Pero me temo que también considero todo ese discurso acerca de los dominios muy difícil de seguir. Me pregunto cuál es el problema conmigo.<sup>3</sup>

Segunda Tesis: *Las "representaciones mentales" son las portadoras primitivas del contenido intencional.*

Tanto desde un punto de vista ontológico como en el orden explicativo, la intencionalidad de las actitudes proposicionales es anterior a la intencionalidad de los lenguajes naturales; y tanto desde un punto de vista ontológico como en el orden explicativo, la intencionalidad de las representaciones mentales es anterior a la intencionalidad de las actitudes proposicionales.

Sólo para el propósito de construir intuiciones, piénsese las representaciones mentales con el modelo de lo que los filósofos empiristas a veces llamaron 'Ideas'. Esto es, piénsese en ellas como particulares mentales dotados de poderes causales y susceptibles de evaluación semántica. Así, está la Idea PERRO. Es satisfecha por todos y sólo por los perros, y tiene relaciones asociativas-causales, por ejemplo, con la Idea GATO. Así, PERRO tiene condiciones de evaluación semántica y tiene poderes causales, como se requiere que tengan las ideas.

Puesto que mucho de lo que quiero decir acerca de las representaciones mentales incluye lo que los empiristas dijeron acerca de las Ideas, podría ser práctico y respetuoso hablar de Ideas más que de representaciones mentales. Pero no propongo hacer esto. La idea Idea está conectada históricamente con la idea de que las Ideas son imágenes, y no quiero tomar ese compromiso. Entonces, como una primera aproximación, la idea de que hay representaciones mentales es la idea de que hay Ideas *menos* la idea de que las Ideas son imágenes.

La TRM afirma que las representaciones mentales se relacionan con las actitudes proposicionales como sigue: para cada evento que consiste en que una creatura tenga una actitud proposicional con el contenido  $P$  (un evento tal como la creencia de Pérez en el tiempo  $t$  que  $P$ ) hay un evento correspondiente que consiste en que la creatura esté relacionada, de una manera característica, con una instancia de una representación mental que tiene el contenido  $P$ . Por favor, nótese la escrupulosidad con la que se mantiene la neutralidad metafísica. *No dije* (si bien estoy muy inclinado a creerlo) que tener una actitud proposicional *consista en* estar relacionado (en una u otra de las "maneras características" mencionadas más arriba) con una representación mental.

También soy neutral en relación a qué son las "maneras características" de estar relacionado con las representaciones mentales. Adoptaré una escabullida útil que inventó Stephen Schiffer: asumo que cualquiera que tenga creencias tiene una caja de creencias en su cabeza. Por consiguiente: