

Explicación psicológica y arquitectura cognitiva: Dos formas de descripción y explicación de la conducta inteligente en el debate Fodor-Churchland

Trabajo de grado para optar por el título de filósofo

Presentado por

Juan M. Gaitán Sánchez

Dirigido por

Dr. John Anderson Pinzón Duarte

Programa de Filosofía

Facultad de Filosofía y Ciencias Humanas

Universidad de La Sabana

Septiembre, 2021

“Talking about the mind, for many people, is rather like talking about sex: slightly embarrassing, undignified, maybe even disreputable. “Of course it exists,” some might say, “but do we have to talk about it?” Yes, we do.”

Daniel Dennett-

“For the story I have to tell is not one of solitary cognition but of an odyssey through many fields, and the solutions to the puzzles are inextricably woven into a fabric of dialogue and disagreement, where we often learn more from bold mistakes than from cautious equivocation.”

Daniel Dennett-

Agradecimientos

Hace un par de años que empecé que con el proceso de escritura para obtener el grado de filósofo. Algunos —me incluyo— hubieran preferido que el proceso hubiera sido más rápido. Sin embargo, he de admitir que no me hubiera gustado evitar o saltarme alguno de los momentos que me han traído hasta aquí. He tenido el privilegio de conocer personas increíbles que me han hecho mejor persona y me han acercado un poco más a una labor filosófica más rigurosa. A todos ellos quisiera agradecer en estas palabras, a todas aquellas personas que de alguna u otra forma aportaron a que pudiera encontrarme en el momento y lugar en el que ahora me encuentro.

Por supuesto, nombrar a cada una de estas personas sería una labor infinita, pues las personas son numerosas y las palabras limitadas. Por ello, quisiera hacer especial mención a aquellas que jugaron un rol activo y fundamental para ayudarme a cumplir con esta meta:

Quisiera empezar agradeciendo a mi director de tesis, Anderson Pinzón. Gracias por el incontable tiempo dedicado a este proyecto. Gracias por la rigurosidad y objetividad con las que leyó estas páginas, y la amabilidad de un amigo con la que me hizo los comentarios apropiados para hacer un mejor trabajo.

Al Padre Euclides Eslava y Dr. Jaime Martínez, por siempre estar pendientes de mí y hacer mucho más agradable y sencillo mi paso por la universidad.

A mi familia, por apoyarme siempre y brindarme todos los recursos necesarios para siempre seguir avanzando. Especialmente a mi madre, Sandra, mi tía Nahydú y mi abuela, Mercedes, por la paciencia incondicional que me mostraron y el cariño que solo una madre, una tía y una abuela pueden mostrar.

Finalmente, pero no menos importantes quiero agradecer a mis amigos más cercanos. A Margarita, por los momentos inolvidables durante la carrera y que aún persisten. A Cristian, Felipe, Camila y Daniela, por su más sincera amistad y las incontables sesiones en las que les presenté una y otra vez los avances de este documento y recibí a cambio una escucha atenta y las correcciones más apropiadas. Adicionalmente, un agradecimiento especial a Tatiana, por su sincera amistad y la atenta lectura del documento y sus correcciones precisas. Y Con especial cariño a Mafe, por los ánimos en los últimos y más críticos momentos para finalizar este proyecto.

Contenido

Introducción	1
Capítulo I	4
La psicología popular como una teoría científica de la conducta	4
Primera Parte	5
Psicología popular	5
La psicología popular, las conductas motoras como acciones y la estructura de las acciones	5
Las acciones estructuradas	10
El éxito de la psicología popular	14
Segunda parte	16
Psicología popular como teoría	16
Son semánticamente evaluables	17
Poderes causales y el funcionalismo en la psicología popular	19
La preservación de las generalizaciones	21
Clausulas ceteris paribus	24
Capítulo II	28
Modularidad y Simbolismo	28
Primera Parte	31
Modularidad de la mente y sistemas modulares	31
Los sistemas modulares son informativamente encapsulados	33
Los sistemas modulares son específicos de dominio	36
Los sistemas modulares son obligatorios y rápidos	37
Segunda Parte	40
Lenguaje de Pensamiento	40
Computación y Semántica	46
Capítulo III	53
Un ataque a esta descripción de la conducta inteligente: la discusión entre Fodor, Churchland y Clapp	53
Primera parte	54
Churchland en contra de la psicología popular	54
Cómo funciona la explicación	56
El poder representacional de los mapas estado-espacio	60
Segunda Parte	66
Fodor Responde: en defensa de la psicología popular	66
Respuesta de Fodor a Churchland: semántica por estipulación	66
Respuesta de Churchland: no hay conceptos, solo <i>prototipos</i>	69
Fodor y Lepore responden: redes neuronales y composicionalidad	72
Tercera Parte	74
Clapp entra al debate: ni composicionalidad, ni LOT	74
El argumento de la esencia	76
El argumento de la explicitud	78
Capítulo IV	84
Una defensa de la psicología popular: rehabilitando LOT en un mundo no tan composicional	84
¿Cuál era el argumento de Clapp?	86

Mi plan de defensa	88
Una respuesta a los dos argumentos de Clapp	89
En contra del primer argumento: PLOT	90
En contra del segundo argumento: la tesis de la indeterminación de la representación	95
La indeterminación de las representaciones: la fuerza de esta respuesta	100
Conclusiones	103
Referencias	109

Introducción

La *psicología popular* es el nombre con el que se suele referir a las intuiciones en las que implícitamente nos apoyamos para predecir y racionalizar las acciones de otras personas. No es extraño encontrarnos en situaciones donde hacemos planes con nuestros amigos: por ejemplo, cuando vamos a la universidad o simplemente nos sentamos a hablar con ellos. Si bien todos estos eventos los solemos desarrollar de manera fluida y, a veces, sin pensar mucho en ello, también requieren de una aguda comprensión y precisa predicción de lo que las demás personas hacen, esperan, buscan, creen, etc.

Sin embargo, nosotros mismos no estamos seguros de qué hacemos, cómo lo hacemos o la explicación de cómo podemos realizar tal hazaña. Las ciencias cognitivas y parte de la filosofía han encontrado en estas intuiciones una prometedora explicación del funcionamiento de la psicología humana.

Este mismo espíritu ha motivado mi investigación. A lo largo de estas páginas, me propongo explorar un poco más afondo la psicología popular como una forma de descripción y explicación de la conducta humana. El argumento que pretendo exponer a lo largo de este documento podría resumirse de la siguiente forma:

- I. Si la psicología popular es exitosa al explicar (y describir) la conducta de las personas, entonces tenemos buenas razones para creer que este éxito se da en virtud de que la psicología popular funciona como una teoría científica, es decir, apela a causas y generalizaciones basadas en estas causas.
- II. La Teoría representacional de la mente explica el carácter científico de la psicología popular.
- III. Hay buenas razones para creer que la psicología popular es exitosa en su explicación de la conducta. [Por lo tanto,]
- IV. Hay buenas razones para comprometernos y apelar a la teoría representacional de la mente.

Así las cosas, en el primer capítulo justificaré la premisa I.

En el primer capítulo presentaré lo que entiendo por psicología popular y cómo estas intuiciones nos presentan una descripción razonable de la conducta de las personas. Luego, apoyándome en el éxito que parece exhibir la psicología popular, argumentaré que ciertos elementos que son la base este éxito —tales como composicionalidad, entidades

semánticamente evaluables que preservan poderes causales y generalizaciones sobre estas entidades— permiten considerar a la psicología popular como una teoría científica.

En el segundo capítulo me encargaré de justificar la premisa II. Allí, argumentaré que una de las formas que puede tomar la teoría representacional son las hipótesis conjuntas presentadas por Jerry Fodor, a saber: *la modularidad de la mente y el lenguaje de pensamiento*, las cuales considero logran explicar y dar cuenta de los elementos esenciales del éxito de la psicología popular mencionados en el párrafo anterior.

En el tercer y cuarto capítulo justificaré la premisa III, de tal forma que la implicación de IV a partir de las premisas I, II y III sea la conclusión de este documento.

El tercer capítulo tiene la estructura de un debate alrededor de dos temas: la tesis de que la psicología popular describe adecuadamente la conducta de las personas y la tesis de que los recursos teóricos ofrecidos por Fodor están justificados. De acuerdo con las dos primeras premisas de mi argumento, creer que la psicología popular describe (y explica) correctamente la conducta humana nos da razones para creer que las hipótesis de Fodor son correctas. Pero Churchland no cree que la psicología popular describa adecuadamente la conducta humana y, por esta razón, él cree que no tenemos razones suficiente para creer en las hipótesis fodorianas. En este sentido, el debate del tercer capítulo es acerca de las razones que tenemos para creer que la psicología popular describe adecuadamente/correctamente la conducta humana.

En la primera parte de este capítulo, basado en la crítica desarrollada por Paul Churchland, evalúo críticamente la psicología popular como una adecuada descripción de la conducta y, por lo tanto, evalúo los mecanismos fodorianos que la explican. Basado en esta crítica, presento la explicación conductual ofrecida por Churchland que, según este autor, supera las falencias de la psicología popular.

En la segunda parte del capítulo presento las respuestas fodorianas a estas críticas y las contrarrespuestas del lado de Churchland. Todo esto con el fin de determinar si la psicología popular y, por lo tanto, la explicación fodoriana, aún pueden sostenerse como una teoría viable —y potencialmente la mejor disponible— para describir la conducta de las personas. La discusión en esta parte del capítulo gira sobre la naturaleza estructurada de la conducta. Fodor cree que la conducta ésta estructurada composicionalmente y argumenta que la teoría propuesta por Churchland es insuficiente para dar cuenta de ésta. Si los argumentos ofrecidos por Fodor en el debate con Churchland son suficientes —como considero lo son—, debemos mantener una explicación como la ofrecida por Fodor: la teoría representacional de la mente y el lenguaje de pensamiento.

En la tercera parte de este capítulo, pondré más presión a los argumentos *fodorianos* al apoyarme en los argumentos de Lenny Clapp (2012). De acuerdo con Clapp, Fodor ha aceptado que la conducta humana (en particular, la lingüística) no está compuesta de partes, no es composicional en este sentido. Sin embargo, aún cree que los recursos explicativos de la conducta inteligente que estaban disponibles aún son composicionales. Esto es, Fodor cree que, pese a que la conducta humana no sea composicional, el pensamiento humano aún es composicional y, por lo tanto, aún hay razones (más débiles) para creer en la hipótesis del lenguaje de pensamiento.

Los argumentos de Clapp apuntan a mostrar que no hay razones para creer que el pensamiento humano es composicional en cuanto que muestran que los recursos explicativos de la hipótesis del lenguaje de pensamiento no pueden dar cuenta de lo *sensible* que es el pensamiento (y también la conducta humana) al *contexto*. Si estos argumentos son correctos, logran mostrar que no hay razones para creer que la conducta y el pensamiento humanos sean composicionales, y esto sería suficiente —de acuerdo con mi argumento— para desmotivar la creencia en la hipótesis del lenguaje de pensamiento y, por ende, en la teoría representacional de la mente.

Considero que los argumentos de Clapp cumplen con su cometido y presentan buenas razones para rechazar las hipótesis *fodorianas*. Esto pone en tela de juicio la psicología popular y brinda una fuerza renovada a las pretensiones *churchlandianas* de abandonar la psicología popular como descripción y explicación conductual.

En el cuarto capítulo, finalmente, ofreceré un esquema de respuesta a los argumentos de Clapp. Mi intención es ofrecer una defensa a favor de una versión mejorada de la teoría representacional de la mente que nos permita evadir los problemas presentados por Clapp y, a su vez, ofrezca una renovada explicación del éxito de la psicología popular. Esta defensa, gira en torno a dos elementos principales: El primero, será sugerir un sistema representacional libre de los problemas computacionales mencionados por Clapp. El segundo, allanar el camino para una defensa más robusta de la composicionalidad basado en la idea de que hay ‘*indeterminación de las representaciones*’

El objetivo de esto será una defensa de los dos conceptos claves atacados por Churchland y Clapp en el tercer capítulo, a saber, el lenguaje de pensamiento y la composicionalidad de la conducta inteligente. Si los argumentos que ofrezco logran ser efectivos, entonces tenemos aún razones para creer en el éxito de la psicología popular en la descripción (y explicación) de la conducta humana, y en que dicho éxito se explica en términos representacionales.

Capítulo I

La psicología popular como una teoría científica de la conducta

La pregunta que se busca responder en este documento se encuentra esencialmente conectada con la conducta corporal de los sujetos. Como espero señalar en las secciones siguientes de este capítulo, la manera de describir la conducta corporal puede comprometernos con una teoría implícita acerca de su naturaleza y de los recursos teóricos requeridos para dar cuenta de ella.

En este sentido, son dos los objetivos que guían esta investigación. El primero concierne a la naturaleza de la conducta. Intentaré ofrecer una respuesta a la pregunta: ¿Cuál es la forma correcta o más apropiada de describir lo que los sujetos hacen? Mi intención es postular la psicología popular (concepto que se explicará más adelante en este capítulo) como respuesta a dicha pregunta. En otras palabras, diré que la manera adecuada de describir la conducta es usando las categorías que constituyen la psicología popular. El segundo objetivo consiste en mostrar la fuerza o éxito de describir la conducta de esta manera, y cómo esta descripción puede tornarse en una explicación científica de la conducta.

Para cumplir con los objetivos del capítulo, en primer lugar, argumentaré que es razonable pensar que la conducta motora humana es inteligente. Y, siguiendo las intuiciones de la psicología popular, la conducta inteligente se puede racionalizar mediante el uso de deseos, creencias y predicados mentales en general. En segundo lugar, mostraré que las acciones se encuentran estructuradas composicionalmente en virtud de que los deseos y las creencias lo están. En tercer lugar, señalaré, a través de algunos ejemplos, que esta forma de descripción es particularmente exitosa en la racionalización de la conducta debido a la gran cantidad de éxitos predictivos con los que cuenta. Finalmente, argumentaré que las generalizaciones que explican, y son la base del éxito de la psicología popular, son de carácter científico.

Primera Parte

Psicología popular

La psicología popular, las conductas motoras como acciones y la estructura de las acciones

Es común toparnos con la idea de que hay cierta clase de eventos corporales o motores que *hacemos* en contraposición a eventos que nos *pasan*. La intuición que nos lleva a marcar esta distinción —según(lo que se conoce como la “*folk psychology*”¹— yace en la idea de que podemos encontrar eventos o conductas corporales que se atribuyen a sujetos como un todo, en contraposición a aquellos eventos corporales que están atribuidos a partes de sujetos, como, por ejemplo, alguno de sus sistemas vitales. Así, ejemplos de lo que hacemos podrían encontrarse en las actividades de la vida diaria tales como pasear, saludar, trabajar, enamorarse. Mientras que ejemplos de lo que nos pasa son el funcionamiento del hígado, el fluir de la sangre, el latido del corazón, la digestión, etc.

Esta tesis no es acerca de los eventos motores que nos pasan, sino que el centro de mi interés filosófico es sobre lo que hacen los sujetos. Un avance en la delimitación parcial del objeto de estudio de esta tesis consiste en responder lo siguiente: ¿Cuál es la mejor forma de describir lo que hacen los sujetos? Responder esta pregunta requiere tener en cuenta otra intuición pre-teórica de la psicología popular, a saber, lo que hacen los sujetos (esto es, la conducta corporal humana) es esencialmente diferente a otros tipos de eventos a lo largo de la naturaleza. Esto es así porque, de acuerdo con estas intuiciones, la conducta humana es también *inteligente*.

¹Me referiré a esta como psicología popular; psicología de deseos y creencias, psicología de sentido común. Todos estos términos hacen referencia al mismo conjunto de ideas o intuiciones: “The final sense of “folk psychology” is closely associated with the work of David Lewis. On this view, folk psychology is a psychological theory constituted by the platitudes about the mind ordinary people are inclined to endorse” (Ravenscroft, 2019). O de manera más robusta: “Folk psychology’ denotes the prescientific, common-sense conceptual framework that all normally socialized humans deploy in order to comprehend, predict, explain, and manipulate the behaviour of humans and the higher animals. This framework includes concepts such as *belief, desire, pain, pleasure, love, hate, joy, fear, suspicion, memory, recognition, anger, sympathy, intention*, and so forth” (Churchland en Mills, 2001, p. 502).

Si bien ha habido un gran debate desde la filosofía y la psicología respecto a lo que es la inteligencia, es razonable pensar que la inteligencia es una propiedad que acompaña a las conductas corporales cotidianas que se hacen normalmente (por un humano adulto incrustado en un contexto cultural). No hace falta ser muy atento para notar que las personas suelen hacer diversas cosas hábilmente. Incluso si tornamos la vista hacia nosotros mismos no es extraño vernos trabajando, estudiando, leyendo un libro, o almorzando. Todas estas son conductas motoras que los humanos hacen inteligentemente. Veamos un ejemplo de esto.

Ejemplo (1): supongamos, por un momento, que sabemos que Felipe tiene una reunión importante temprano en la mañana en el trabajo a la cual definitivamente tiene que asistir. Si le preguntásemos a Felipe cuáles son sus planes para mañana, no sería una sorpresa que respondiera que debe asistir a una reunión, entre otras cosas. Como observadores e interlocutores de Felipe, podemos decir que él tiene la intención de asistir a la reunión; e incluso podríamos adelantarnos en el tiempo y prever que pasará algo así como que Felipe probablemente despierte al día siguiente muy temprano, se arregle, desayune, salga de su casa; y tome un autobús hasta el lugar donde será su reunión, todo esto con el fin de llegar a la hora acordada. Además, suponga que Felipe en su camino al trabajo se topa con un conocido. Felipe levanta su mano, la agita, o estrecha la mano de su conocido y emite un sonido: “Hola”. Luego, el conocido (digamos Andrés) también levanta su mano, la agita o la estrecha con la de Felipe y responde con un “Hola” de vuelta.

En este ejemplo la conducta (inteligente) de Felipe es ordinaria en el sentido en que probablemente en las mismas circunstancias nos comportaríamos de la misma forma. Pero ¿qué forma es esa? En términos generales, podría decirse que haríamos lo que *creemos* que es correcto o lo que deberíamos hacer para llegar a la reunión (o lo que debería hacerse cuando nos topamos con un conocido). Una parte importante de esto consiste en apelar al conocimiento de, por ejemplo, cómo funcionan los autobuses; qué bus se debe tomar para llegar al lugar designado para la cita; cuánto tiempo aproximado nos tardaremos en llegar y, por lo tanto, a qué hora deberíamos salir de casa, junto con un par de convenciones sociales; etc. Por otro lado, podemos aplicar esto mismo en el caso del saludo. Cuando saludamos a otra persona hacemos lo que hacemos (ya sea dar un apretón de manos, un beso, etc.) porque consideramos correcto hacerlo, porque deseamos hacerlo o solo porque queremos llamar su atención. En todo caso, el elemento que se mantiene es apelar a esos conocimientos,

creencias o deseos. En este sentido, las cosas que hacen los sujetos pueden ser descritas como *conducta motora inteligente* en cuanto que para describirla así necesitamos apelar a rasgos mentales (inteligibles para nosotros) que la hagan inteligible y razonable a nuestros ojos.

Existe otra motivación para apelar a este tipo de descripción. La relevancia de describir la conducta motora como conducta inteligente recae en el rasgo normativo del concepto de inteligencia. La inteligencia es una categoría normativa en cuanto que suministra un criterio de evaluación a partir del cual es posible distinguir cuáles conductas son satisfactoriamente inteligentes y cuáles no lo son. Este tipo de distinción no parece ser soportada por la descripción puramente física o anatómica del evento sucedido. Por ejemplo, no hay una forma de señalar cuándo se mueve un músculo cardíaco inteligentemente y cuándo no².

Con esta idea en mente, describir la conducta inteligente parece capturar esta normatividad en la medida en que podemos dar criterios más o menos claros para cuando una persona —por ejemplo, Felipe— logra llegar a tiempo a una reunión de trabajo; cuando saluda a alguien o cuando está desayunando —en general, con cualquier conducta que sea inteligible a la luz de los deseos o creencias del sujeto—; y también cuando lo hace hábil o torpemente.

El **Ejemplo 1** permite resaltar un punto de vital importancia para el desarrollo de este trabajo. La pregunta que me interesa en esta investigación tiene que ver, más precisamente, con la explicación de la conducta corporal inteligente de los sujetos o agentes. En particular me interesa la naturaleza de esta explicación. No solo me interesa la pregunta ¿qué hace Felipe y por qué hace lo que hace?, sino también cómo respondemos a esa pregunta. En otras palabras, también me llama la atención saber a qué tipo de entidades o mecanismos debemos apelar para decir por qué hace lo que hace.

Es por esta razón que la manera de describir la conducta inteligente de un sujeto resulta relevante. Inclinarsé por una u otra descripción puede llevar a una diferencia en las entidades o mecanismos; y obtener explicaciones de la conducta totalmente diferentes. Esto

² Hay, además, un rasgo particular con este tipo de eventos, a saber, hay un sentido en que el latir del corazón no es una acción producida por nadie, quizá en el mejor de los casos por un cerebro. Sin embargo, no es claro en qué sentido un cerebro puede ser inteligente sin apelar a la noción de un todo o un “sujeto”.

será lo que se discutirá en el capítulo III cuando se presente la discusión entre Fodor y Churchland, y qué consecuencias tiene esto para la psicología popular³.

Por ahora, volvamos sobre el evento descrito en el ejemplo 1. Existen otras maneras de describir *verdaderamente* el mismo evento. Podríamos decir que fue el evento que sucedía simultáneamente mientras un planeta en algún lugar se alineaba con el centro de la galaxia. O, sin tener que mencionar entidades extra planetarias, podríamos hacer una descripción física de lo que sucedió entre Andrés y Felipe: dos cuerpos se estaban moviendo a una cierta velocidad en una cierta trayectoria, luego se detienen y ciertas ondas son emitidas por ellos. Incluso, si queremos, podemos refinar aún más la descripción y hablar de conjuntos de átomos que se mueven e interaccionan entre ellos⁴.

Ciertamente podríamos tener todas estas descripciones. Algunas de ellas pueden no ser excluyentes con otras, o incluso podrían ser todas verdaderas simultáneamente. La decisión que tomemos sobre las descripciones que puede recibir un fenómeno recae en la pertinencia de esa descripción para los objetivos que tenemos entre manos. Si lo que nos interesa explicar es la conducta motora inteligente, para ello debemos centrarnos en la descripción que mejor se adapte a esta empresa⁵.

¿Cuál es esa descripción? La respuesta que quiero defender es la conducta inteligente descrita apelando a deseos, creencias, intenciones, etc. Veamos otro ejemplo que apoya la elección de este tipo de descripción.

Ejemplo (2): En *Sueño de una noche de verano* Shakespeare nos presenta a dos enamorados que planean escapar ante la imposibilidad de estar juntos según la ley ateniense:

Lisandro: Bien dicho, Entonces óyeme, Hermia. Tengo una tía viuda, dama de grandes méritos y muchos bienes, que no tiene hijos, y como hijo único me respeta a mí. De Atenas a su casa hay algo más de siete leguas.

³ Dar cuenta de los mecanismos y entidades necesarias para la explicación remite a lo que debería responder una teoría de la mente, a saber, los principios subyacentes a la conducta inteligente de Felipe que nos permiten explicarla y realizar predicciones correctas sobre esta.

⁴ Para un trato más detallado de las descripciones que puede tener un evento véase Davidson (1963). Allí, Donald Davidson defiende la identidad entre eventos mentales y físicos arguyendo que se puede hacer una descripción verdadera tanto física (puramente), como mentalista de un evento. La importancia de mencionar esto es el viejo problema mente-cuerpo, o la naturalización de la mente, es decir, cómo podemos hablar de entidades mentales sin salir del dominio de lo físico.

⁵ Esto puede llegar a sonar pretensioso, pero un claro ejemplo de esto son los números. Hay varias formas posibles de notación numérica como lo son la algebraica, binaria, e incluso podríamos usar números romanos. Sin embargo, cada forma de notación puede servir a propósitos diferentes para las cuales las demás no son aptas. Imagínese lo difícil que sería sumar con números romanos. Definitivamente sería posible, pero otras notaciones son más adecuadas para la tarea.

Allí, gentil Hermia, me he de casar contigo, donde la áspera ley ateniense nada podrá contra nosotros. Si de veras me amas, huye de la casa de tu padre mañana por la noche, y búscame en el bosque, a una legua de la ciudad, donde una vez te encontré con Helena para las celebraciones del florido mes de mayo.

Te estaré esperando.

Hermia: Buen Lisandro, te juro por el arco más templado de Cupido, por la flecha de punta más dorada, por la candidez de las palomas de Venus, por todo lo que une a las almas e impulsa a los amores, y por el fuego en que ardió la reina de Cartago cuando el falso hijo de Troya fue visto izando velas; por todos los juramentos rotos por los hombres -tanto más que los nunca hechos por mujeres- que en ese lugar al que me citas mañana verdaderamente estaré esperando.

Lisandro: Mantén tu promesa, amor. Mira aquí viene Helena.

Entra Helena

Muy romántico, y también muy plausible que suceda. Como sabemos, Hermia y Lisandro terminarán encontrándose en el bosque el día y la hora acordada. ¿Cómo pasó esto? Ciertamente el diálogo da una explicación de lo que sucedió, Lisandro y Hermia hicieron la promesa para encontrarse y actuaron en consecuencia. Sin embargo, aquí hay más oculto de lo que parece. Sería razonable pensar que quizá Hermia o Lisandro no hubieran ido al bosque si creyeran que se les estaba jugando una broma, o si creyeran que su amante no aparecería, o incluso si no deseaban hacerlo. Parece que el éxito de que el encuentro se llevara a cabo se debió no solo a creencias correctas sobre lo que haría su amante, sino también a las palabras que usaron para expresarlas.

Por supuesto, las emisiones de los dos amantes jugaron un papel fundamental en el evento. A partir de ellas ambos pudieron predecir que su compañero estaría en el bosque sin falta, sabían qué debían hacer, a donde ir y cuándo. No obstante, aún puede preguntarse ¿cómo un conjunto de sonidos pudo asegurar tal éxito? Nuevamente la psicología popular viene en nuestra ayuda: sería sensato decir que los sonidos emitidos por Hermia y Lisandro corresponden a la expresión de sus pensamientos, creencias y deseos; en general, de sus intenciones de huir y casarse. Podría decirse que la significatividad de sus expresiones está dada o se explica en virtud de las intenciones de Hermia y Lisandro.

Este aspecto de la conducta motora inteligente, que podríamos llamar *conducta lingüística*, parece ser una buena motivación que apoya la idea de que los sujetos actúan en virtud de sus intenciones. Esto puede ser así en la medida en que no parece haber una correlación entre algún tipo de propiedad física de los sonidos emitidos por los dos individuos

y la conducta luego realizada. Piense que ante el mismo dialogo, la respuesta de Hermia o Lisandro pudo haber sido diferente si sus intenciones fueran distintas: si Lisandro quisiera engañar a Hermia y Hermia lo supiera, es probable que no se encontraran en el bosque, y sus afirmaciones de amor significaran otra cosa. En este aspecto Davidson (2001) en su artículo “Mental Events” señala:

Suppose we try to say, not using any mental concepts, what it is for a man to believe there is life on Mars. One line we could take is this: when a certain sound is produced in the man's presence ('Is there life on Mars?') he produces another ('Yes'). But of course this shows he believes there is life on Mars only if he understands English, his production of the sound was intentional, and was a response to the sounds as meaning something in English; and so on. For each discovered deficiency, we add a new proviso. Yet no matter how we patch and fit the non-mental conditions, we always find the need for an additional condition (provided he notices, understands, etc.) that is mental in character. (Davidson, 2001, pp. 217)

Así, podemos volver a preguntar ¿cuál es la descripción más adecuada para describir la conducta inteligente de los sujetos? Luego de repasar los ejemplos anteriores parece razonable pensar que la descripción adecuada necesariamente involucra la apelación a las intenciones, deseos y creencias de los sujetos. Este tipo de descripción de su conducta motora parece ser de un tipo diferente al simple movimiento de músculos, funcionamiento de órganos, etc. También parece que la significatividad de las *emisiones* lingüísticas de Hermia y Lisandro y el éxito de su reunión se explican en virtud de sus intenciones y no a partir de propiedades físicas o biológicas de sus organismos.

La aproximación de la psicología popular ofrece una respuesta que suministra más información, lo que Felipe hace (y Hermia y Lisandro), principalmente, son *acciones*, es decir, conductas motoras que se explican o son inteligibles en virtud de sus creencias, deseos, temores, preferencias, etc. Pues, como se puede notar en el ejemplo, lo que da sentido a todo el evento es el hecho de que apelemos a las creencias de Felipe acerca de su cita, el conocimiento que tiene del sistema de trenes, su habilidad con, y conocimiento de, normas sociales para saludar a Andrés, etc.

Las acciones estructuradas

Las *acciones* parecen exhibir una organización o un cierto tipo de orden. Esto se desprende de la misma noción de inteligencia. Si tomamos en consideración el tipo de eventos presentados en los Ejemplos 1 y 2 (es decir, acciones como saludar, planear una cita,

etc.) es intuitivo pensar que lo que hace exitosamente Felipe (o Hermia), y la habilidad con que lo hace, depende de la manera misma como está estructurada la acción. Parte del éxito en la inteligibilidad de las acciones recae en la idea de que las acciones pueden ser complejas y estar compuestas de acciones más simples, donde estos simples mantienen su identidad.

Here's a way to put it. Practically everybody thinks that the objects of intentional states are in some way complex: for example, what you believe when you believe that John is late for dinner is something composite whose elements are –as it might be– the concept of John and the concept of being late for dinner (or –as it might be– John himself and the property of being late for dinner). And, similarly, what you believe when you believe that P&Q is also something composite, whose elements are –as it might be– the proposition that P and the proposition that Q. (Fodor, 1987, p. 136)

Piense en el sencillo caso de alguien bailando mientras canta una canción.

Intuitivamente podría pensarse que dicho evento consiste en una acción compuesta: por un lado, tenemos el baile que se realiza; y, por otro lado, tenemos el cantar la canción. Ambas acciones ocurren simultáneamente dando lugar a la acción compuesta de bailar y cantar. Sale a relucir la idea de que las acciones no solo pueden agruparse, sino que mantienen una estructura particular a partir de la cuál acciones simples pueden constituir acciones más complejas (es decir, compuestas). Se dice, entonces, que la conducta es composicional, en este sentido amplio, puesto que puede tener elementos que se agrupan y componen conductas más complejas.

El ejemplo paradigmático de la estructuración de la conducta puede verse claramente en la conducta lingüística. Por ejemplo: cuando alguien dice la oración (a) “John ama a Mary”, está realizando una serie de acciones que pretenden expresar algo (quizá una creencia). ¿Qué acciones se realizan? Bien, al igual que el caso de cantar y bailar, probablemente la acción de decir (a) consiste en una serie de acciones atómicas o simples separables: usar el nombre propio “Mary”, usar el nombre propio “John” y usar el verbo “amar”, para expresar una cierta relación entre Mary y John.

Considérese que alguien luego dice la oración (b) “John ama a Mary”, es razonable pensar que la persona que dijo ambas oraciones realizó las mismas acciones atómicas en ambos casos (suponiendo que los nombres sigan refiriendo a las mismas personas). Por su puesto el orden de las acciones se vio alterado, lo cual podría explicar que las oraciones sean distintas y signifiquen cosas distintas. Sin embargo, permanece la idea de que las acciones que componen la conducta compleja de decir la oración (b) son las mismas realizadas en la

oración (a), lo que nos ayuda comprender, por ejemplo, que usemos los mismos nombres y podamos referirnos a la misma Mary, a saber, se realiza la misma acción.⁶

Otro caso que puede señalarse a este respecto —y por mucho más sofisticado— es el caso de las inferencias. Si volvemos de nuevo al Ejemplo 2, podemos ver que en el fragmento Lisandro no le dice explícitamente a Hermia que la ama. Sin embargo, Hermia lo infiere a partir de la propuesta de Lisandro. Incluso aunque no se menciona explícitamente lo que Lisandro siente por Hermia (más allá de su intención de casarse con ella), es posible pensar que el fragmento no tendría sentido si, por ejemplo, Hermia no supiera lo que significa (o usara en la inferencia el concepto de) ‘amar a alguien’. Lo que nos lleva de vuelta a la intuición de que las acciones que realizamos (en este caso podría decirse que las inferencias que Hermia hace son acciones mentales) son las mismas acciones simples o atómicas que pueden ocurrir (y ocurren) a lo largo de diferentes actividades o acciones más complejas.

Thought is productive: in principle, you can entertain and produce an infinite number of distinct representations. How can you do this? Our brains have a limited storage capacity, so we can't possibly possess a mental phrase book in which the meaning of each sentence is encoded. The key is that the thoughts are built out of familiar constituents and combined according to rules. It is the combinatorial nature of thought that allows us to understand and produce these sentences on the basis of our antecedent knowledge of the grammar and atomic constituents (e.g., China, Italy). This allows for the construction of the potentially infinitely many thoughts given a finite stock of primitive expressions (Chomsky 1975; Fodor 1975, 31; Fodor and Pylyshyn 1988, 116; Fodor 1985, 1987). (Schneider, 2011, p. 10)

Esta característica de la conducta sugiere rasgos más profundos sobre la etiología y los mecanismos subyacentes y responsables de las acciones. Ya se ha estipulado que éstas —es decir, la conducta motora inteligente— son el resultado de las creencias, deseos y otros elementos mentales. Si este es el caso, como parece serlo a los ojos de la psicología popular, y las acciones están organizadas de manera sistemática, productiva y composicional, es

⁶ Lo mismo aplica para las habilidades que permiten este tipo de acciones. Podría decirse que la capacidad de entender/producir la oración “Mary ama a John” está íntimamente relacionada con la habilidad necesaria para entender/producir la oración “John ama a Mary” (Para una presentación más completa del ejemplo véase Fodor, 1987, p. 149). A esta propiedad de la conducta se le conoce como sistematicidad y puede describirse de la siguiente manera: Notwithstanding constraints on human hardware capacities, the productivity of thought refers to our capacity to entertain or grasp an indefinitely large number of thoughts. Thought processes, on the other hand, are systematic to the extent that our capacity to entertain or grasp a thought appears to be intrinsically connected with our capacity to entertain or grasp a number of other semantically related thoughts. Likewise, thought exhibits inferential coherence. Our capacity to follow a pattern of inference appears intrinsically connected to our capacity to draw certain other inferences. The productivity, systematicity, and inferential coherence of thought strongly suggest that mental representations possess a constituent structure without which it is difficult to come to terms with the interconnections among thought-related capacities. (Calvo y Simons, 2014, p. 9)

preciso suponer que los deseos, creencias y demás entidades mentales también lo están. Esto en la medida en que los deseos, creencias e intenciones en general, racionalizan o explican la conducta

La estructuración de la conducta y su composicionalidad señalan un aspecto fundamental, a saber, si la conducta puede llegar a tener constituyentes es razonable pensar que existan reglas sobre la forma en que estos constituyentes pueden agruparse para dar lugar a conductas más complejas —es decir—, compuestas. Lo que nos lleva al último elemento importante por mencionar en esta sección: debido a la estructuración composicional tanto de las intenciones y de la conducta, pueden elaborarse generalizaciones sobre ambas.

La composicionalidad de la conducta y las intenciones permite hacer generalizaciones en la medida en que nos permite distinguir entre diferentes conductas y estados mentales. Por supuesto, las generalizaciones son de carácter empírico, sin embargo, apelar a acciones simples y distinguibles soporta la idea de que las acciones se repiten, es decir, son las mismas en diversas ocasiones; lo cual nos permite relacionarlas entre ellas.

Piense en el caso de bailar y el deseo de bailar. Si nos topamos con una persona bailando, es razonable pensar que esa persona baila *porque quiere* bailar. Si esto aplica a todas las personas que bailan cuando sienten el deseo de hacerlo, entonces bailar es una acción que puede ser realizada por diversos sujetos en diversas ocasiones. Esta característica de las acciones, es decir, que se repitan y se haga la misma cosa en cada ocasión, sienta las bases para realizar generalizaciones del tipo:

(g) Si x quiere (o tiene el deseo, o desea, etc.) bailar, entonces x realizará la acción de bailar (si no hay nada que se lo impida o entre en conflicto con su deseo, etc.).

O incluso generalizaciones más sofisticadas como:

(h) Si x desea P , entonces x hará (o intentará conseguir, o buscará, etc.) P .

Veamos en las siguientes secciones cómo estas generalizaciones llevan a un éxito predictivo de la conducta y cómo funcionan más detalladamente.

El éxito de la psicología popular

Como se ha visto en las secciones anteriores y particularmente en el Ejemplo 2, la racionalización de una acción de forma inteligible para nosotros nos permite, a su vez, hacer predicciones sobre la conducta de los demás. El éxito de que Hermia y Lisandro se encontraran en el bosque no solo recae en racionalizar su conducta como una promesa o una cita; también recae en que gracias esa racionalización hicieron la predicción (verdadera) de que su amado estaría en dicho lugar esperándolos. Veamos otro ejemplo para ver de forma más detallada como las racionalizaciones se conectan con las predicciones (por lo general verdaderas) de la conducta.

Ejemplo (3): otro ejemplo de lo que me interesa analizar es presentado por Fodor (1987) en su texto “la persistencia de las actitudes” a partir de un fragmento de la obra *Sueño de una noche de verano* de William Shakespeare:

“*A Midsummer Night 's Dream*, act 3, scene 2.

Enter Demetrius and Hermia.

Dem. O, why rebuke you him that loves you so?

Lay breath so bitter on your bitter foe.

Herm. Now I but chide, but I should use thee worse;

For thou, I fear, hast given me cause to curse.

If thou hast slain Lysander in his sleep,

Being o'er shoes in blood, plunge in the deep,

And kill me too.

The sun was not so true unto the day

As he to me: would he have stol'n away

From sleeping Hermia? I'll believe as soon

This whole earth may be bor'd; and that the moon.

May through the centre creep, and so displease

Her brother 's noontide with the antipodes.

It cannot be but thou hast murder 'd him;

So should a murderer look; so dead, so grim.” (Fodor, 1987, Pág. 1)

La inferencia debió haber sido de esta manera: dejando de lado varias proposiciones subsidiarias, continúa diciendo Fodor:

Hermia has reason to believe herself beloved of Lysander. (Lysander has told her that he loves her- repeatedly and in elegant iambs- and inferences from how people say they feel to how they do feel are reliable, *ceteris paribus*.) But if Lysander does indeed love Hermia, then, a fortiori, Lysander wishes Hermia well. But if Lysander wishes Hermia well, then Lysander does not voluntarily desert Hermia at night in a darkling wood. (There may be lions. " There is not a more fearful wild -fowl than your lion

living.") But Hermia was, in fact, so deserted by Lysander. Therefore, not voluntarily. Therefore involuntarily. Therefore, it is plausible that Lysander has come to harm. At whose hands? Plausibly at Demetrius 's hands. For Demetrius is Lysander's rival for the love of Hermia, and the presumption is that rivals in love do not wish one another well. Specifically, Hermia believes that Demetrius believes that a live Lysander is an impediment to the success of his (Demetrius 's) wooing of her (Hermia). Moreover, Hermia believes (correctly) that if x wants that P, and x believes that not -P unless Q, and x believes that x can bring it about that Q, then (ceteris paribus) x tries to bring it about that Q. Moreover, Hermia believes (again correctly) that, by and large, people succeed in bringing about what they try to bring about. So: Knowing and believing all this, Hermia infers that perhaps Demetrius has killed Lysander. And we, the audience, who know what Hermia knows and believes and who share, more or less, her views about the psychology of lovers and rivals, understand how she has come to draw this inference. We sympathize. (Fodor, 1987, pp. 1-2)

Dado que la conducta esta esencialmente ligada a lo que los sujetos creen, desean/intentan e infieren, los lectores y Hermia apelamos a esto para explicar y predecir lo que harán los demás. Lo que hace Hermia —y nosotros como lectores— es una inferencia no demostrativa de las intenciones de Demetrio y Lisandro que explica su supuesta conducta conforme al estado actual de las cosas. Por supuesto, sabemos que Hermia de hecho se equivocó, que Demetrio no mató a Lisandro y que Lisandro no eran tan fiel como Hermia creía. Sin embargo, ese pudo ser el caso y la estrategia de Hermia de apelar a deseos y creencias. Es lo que le da significado a la conducta y termina explicando la conducta de Lisandro y es la razón por la que nosotros y Hermia podemos comprender a otros, y también coordinar nuestras acciones con otros.

No es difícil para nosotros comprender lo que piensa Hermia. De hecho, podría decirse que es perfectamente natural llegar a las conclusiones a las que ella llega, pues no es diferente a lo que solemos hacer en nuestra vida diaria en las interacciones con nuestra familia, amigos, compañeros, entre otros. Por supuesto, todos apelamos a la psicología popular para hacerlo, y quizá lo hacemos porque de hecho funciona. La psicología popular funciona tan bien como lo hace en parte porque no son simples predicciones sin sentido o al azar:

Commonsense psychology works so well it disappears. It's like those mythical Rolls Royce cars whose engines are sealed when they leave the factory; only it's better because it isn't mythical. Someone I don't know phones me at my office in New York from -as it might be- Arizona. 'Would you like to lecture here next Tuesday?' are the words that he utters. 'Yes, thank you. I'll be at your airport on the 3 p.m. flight ' are the words that I reply. That's *all* that happens, but it's more than enough; the rest of the burden of predicting behavior – of bridging the gap between utterances and actions- is

routinely taken up by theory. And the theory works so well that several days later (or weeks later, or months later, or years later; you can vary the example to taste) and several thousand miles away, there I am at the airport, and there he is to meet me. (Fodor, 1987, p. 3)

Los logros de la psicología popular están por todas partes y ocurren todo el tiempo. O al menos eso es lo que parece. Cada uno de los ejemplos presentados anteriormente muestran características o elementos clave en el fenómeno conductual. Por ejemplo, la comprensión de la conducta de otros, la habilidad de realizar predicciones verdaderas sobre lo que harán otros sujetos y actuar acorde a ello, etc. Cada uno de estos aspectos conductuales o habilidades debería ser explicado en virtud de la descripción o teoría de lo que es la conducta y los principios subyacentes a ella.

Es por esta razón que tanto he insistido por la pregunta sobre *lo que hacen los sujetos*. Es aquí donde la psicología popular se muestra como una primera aproximación sobre el fenómeno conductual. La psicología popular está constituida por intuiciones que usamos para describir y racionalizar el fenómeno, y, en esta medida, ofrece los cimientos de lo que puede llegar a ser una explicación de cómo somos capaces de coordinarnos, interpretar y predecir la conducta de otros.

A estas alturas, podríamos empezar a hablar de la psicología popular como una teoría explicativa de la conducta en la medida que ofrece una descripción del fenómeno conductual de tal manera que lo hace inteligible y racional. Por supuesto aún no es lo suficientemente robusta para ser una teoría científica, o incluso una teoría completa, pero puede serlo. Veamos las implicaciones de esta idea.

Segunda parte

Psicología popular como teoría

Describir la conducta de un modo u otro puede acarrear (en este caso implícitamente) una teoría. Esto es justo en lo que está pensando Fodor cuando se refiere al sentido común o a la *folk psychology* como un esquema de teoría que delimita y explica el fenómeno en la medida en que la teoría conecta las creencias y deseos con las acciones y es —según nuestro ejemplo— en lo que Hermia se apoya para inferir lo que infiere, lo que usamos como

audiencia para entender los pensamientos y las acciones de Hermia, y lo que Shakespeare emplea para manipular nuestras simpatías (Fodor, 1987).

Llamar al sentido común una teoría parece pretencioso, sin embargo, funciona como el esquema de una, pues a partir de nuestra atribución de intenciones a los agentes o hablantes pueden ofrecerse generalizaciones que abarcan intenciones, acciones y predicciones sobre la conducta de los sujetos:

Or rather, since there probably are no generalizations which connect preferences to actions irrespective of beliefs, what Hermia must be relaying on is an implicit theory of how beliefs, preferences, and behaviors interact; an implicit decision theory, no less. (Fodor, 1987, p.7)

Hasta ahora se ha presentado a la psicología popular como un conjunto de buenas intuiciones que nos ayudan a comprender la conducta de los demás. Para hacer de la psicología popular una teoría hace falta refinar algunos de los elementos antes ya mencionados. Empecemos hablando un poco más sobre las entidades mentales a las que apela la psicología popular, es decir, deseos, creencias, miedos, esperanzas, entre otras. Como ya se ha mencionado en párrafos anteriores, la psicología del sentido común apela a estas entidades; entidades que son las responsables de la conducta de las personas. Sin embargo, no hemos dicho nada respecto su naturaleza. ¿Qué clase de entidades son? Presumiblemente no hay una única respuesta. Aun así, parecen haber ciertas características que poseen estas entidades que son propicias para la explicación psicológica, tal como se ha presentado:

- I. Son semánticamente evaluables
- II. Tienen poderes causales
- III. Las generalizaciones de la psicología de creencias/deseos del sentido común son, en gran parte, verdaderas de tales estados.

Al aclarar estos elementos pretendo esbozar de manera breve y simple las características que considero definen el tipo de entidades a las que apela la psicología popular. Adicionalmente, estas características sirven como criterios que cualquier teoría de la mente debe cumplir si pretende mostrarse como defensora de la psicología popular.

Son semánticamente evaluables

Toda creencia puede ser evaluada en términos de verdadero/falso, los deseos pueden evaluarse en satisfechos/frustrados, incluso los presentimientos pueden ser correctos/incorrectos. La hipótesis que se asume es que la evaluación de tales entidades se

hace en relación con el ambiente: algo del mundo no psicológico es lo que hace verdaderas a las creencias, satisface los deseos o hace correctos los presentimientos. A esta evaluación se le llamará *evaluación semántica*.

Hay que hacer una distinción importante entre la *evaluación semántica* de la creencia y el objeto de la creencia. Un ejemplo en el que se ve claro es el siguiente:

Sometimes I'll talk of the *content* of a psychological state rather than its semantic evaluability. These two ideas are intimately interconnected. Consider -for a change of plays- Hamlet's belief that his uncle killed his father. That belief has a certain semantic value; in particular, it's a *true* belief. Why true? Well, because it corresponds to a certain fact. Which fact? Well, the fact that Hamlet's uncle killed Hamlet's father. But why is it *that* fact that determines the semantic evaluation of Hamlet's belief? Why not the fact that two is a prime number, or the fact that Demetrius didn't kill Lysander? Well, because the *content* of Hamlet's belief is *that* his uncle killed his father. (If you like, the belief 'expresses the proposition' that Hamlet's uncle killed his father.) *If you know what the content of a belief is, then you know what it is about the world that determines the semantic evaluation of the belief*; that, at a minimum, is how the notions of content and semantic evaluation connect. (Fodor, 1987, p. 11)

En términos generales, podría establecerse la siguiente estructura o forma:

(1) X cree que P

La oración (1) especifica la actitud proposicional designa un estado mental, en este caso, la creencia que P . Aquí podríamos fácilmente cambiar la expresión “cree” por “desea”, “presiente”, “teme”, etc., cambiando así la actitud proposicional a la que nos referimos o la actitud proposicional que es especificada. Por otro lado, según la cita anterior, lo que P denota es el contenido de la creencia. A menudo, dicho contenido resulta ser una proposición. Luego, esa proposición también nos da pistas sobre cuál es el aspecto del mundo que debemos tener en cuenta para la evaluación semántica de la actitud proposicional. Así, tenemos que P es el contenido de la creencia, y la relación entre P y el mundo es lo que nos da la evaluación semántica, es decir, si es verdadera o falsa. Tenemos entonces tres elementos relevantes para distinguir:

- i. La relación entre el sujeto x y la proposición p , refiere al tipo de actitud proposicional en cuestión (ya sea de creencia, deseo, etc.)
- ii. P es reemplazada por una proposición que especifica el contenido de la actitud proposicional, es decir, qué es lo que se cree, se desea, etc.

- iii. La evaluación semántica se establece entre p y el mundo, es decir, si la proposición expresada por p corresponde a un hecho y es, por lo tanto, verdadera (o falsa).

Hacer estas distinciones nos brinda herramientas para la identificación de actitudes proposicionales. Teniendo solo dos cosas en cuenta podemos lograrlo: (a) qué clase de actitud proposicional es, ya bien sea una creencia, deseo, temor, esperanza, etc.; y (b) cuál es el contenido de la actitud o qué proposición expresa⁷.

Podría decirse que cualquier otra posición que intente capturar la psicología de deseos y creencias o sentido común debe aceptar entidades que pueden individuarse de esta manera. Esto es así en la medida en que la distinción parece capturar, por un lado, el carácter normativo característico de la conducta inteligente en la medida en que la verdad y la falsedad de las creencias se corresponden con el éxito o el fracaso de la acción. Por otro lado, la relación entre lo que la persona piensa, siente, percibe, etc., respecto al ambiente. Por ejemplo, imagine un caso en el que Felipe cree que los leones son peligrosos. Se tiene, entonces, que el contenido de su creencia es “los leones son peligrosos”. Esto sumado a la actitud de Felipe frente a ello —en este caso de *creencia*— explica la manera en la que Felipe se comporta, a saber, probablemente intente escapar o protegerse. Esto es relevante porque frente al mismo contenido un mismo sujeto puede tener diferentes actitudes —creencia, deseo, duda, temor— y así mismo, su conducta puede cambiar dependiendo de la actitud en la que se encuentre.

Poderes causales y el funcionalismo en la psicología popular

Ya se ha dicho que las entidades mentales como los deseos y las creencias *explican* la conducta, o incluso que se puede racionalizar la conducta apelando a (o poniéndola en términos de) entidades mentales. No obstante, puede que esta forma de describir la relación entre acciones y entidades mentales sea demasiado ambigua o se preste para diferentes interpretaciones. No obstante, ya se ha insinuado la naturaleza de esta relación en los párrafos anteriores y ahora argumentaré a favor de ello.

A partir de la exposición de Davidson (1963), el argumento general que puede presentarse es que, si bien puede ofrecerse una descripción de la conducta en términos de

⁷ En esta distinción recae el poder explicativo de la psicología popular, pues las generalizaciones que se obtienen —como se verá más adelante— cuantifican sobre los contenidos de las actitudes proposicionales.

acciones e intenciones, las intenciones por sí solas no necesariamente explican o racionalizan la acción. En el ejemplo:

(a) “*x* hizo *y* porque tenía la intención *z*”,

podría pensarse que es *z* lo que racionaliza la acción y la explica. Sin embargo, como Davidson argumenta, la verdad de una intención —esto es, tener la intención de— no requiere que el evento sugerido ocurra. Así, alguien puede tener la intención de encender la luz, pero nunca encenderla. En este sentido, el “porque...” que aparece en (a) permanece sin análisis, es decir, no es totalmente claro que *x* hizo *y* porque *z*.

La solución a la que podemos apelar —a la que Davidson apela y a la que la psicología popular debería apelar— es que, así como para explicar⁸ los fenómenos solemos describirlos en términos de sus causas, lo mismo debemos hacer con la conducta. Así, cuando analizamos situaciones del tipo (a), debe asumirse que hay causalidad entre las intenciones y las acciones, porque así puede garantizarse que *x* hizo *y* porque tenía la intención *z* (Davidson, 1963, pp. 691-693).

Adicionalmente, esta explicación debe darse bajo los estándares de la ciencia empírica, como veremos más adelante con Fodor, con la esperanza de que se pueda hacer una ciencia de la conducta. Por esta razón, se considera que la postulación de generalizaciones o leyes debe darse en términos causales de, en este caso, las acciones y las intenciones.

What's important for now is this: It is characteristic of commonsense belief/desire psychology -and hence of any explicit theory that I'm prepared to view as vindicating commonsense belief/desire psychology- that it attributes contents and causal powers *to the very same mental things that it takes to be semantically evaluable*. It is Hamlet 's belief that Claudius killed his father - the very same belief which is true or false in virtue of the facts about his father's death- that causes him to behave in such a beastly way to Gertrude.

In fact, there's a deeper point to make. It's not just that, in a psychology of propositional attitudes, content and causal powers are attributed to the same things. It's also that causal relations among propositional attitudes somehow typically contrive to respect their relations of content, and belief/desire explanations often turn on this. (Fodor, 1987, p. 12).

La importancia de que las entidades semánticamente evaluables sean las mismas entidades que tienen un contenido y poderes causales se debe a que esto garantiza la conexión

⁸ Davidson sugiere que una forma de explicación consiste en ofrecer las causas de los efectos y podemos apelar a esta misma idea con la conducta. Sin embargo, si se quisiera ofrecer otro sentido de “explicación” para la conducta, hace falta mostrar en qué consiste y cómo operaría en el marco de la explicación conductual.

entre intenciones y acciones. Si se niegan los poderes causales a las entidades semánticas, se pierde la explicación —al menos en un sentido científico—. Es decir, no es posible que actuemos en virtud de nuestras creencias y deseos, porque nuestros deseos y creencias no causan nuestra conducta. Lo mismo ocurre en el caso de los contenidos: si nuestras y creencias y deseos no fueran sobre algún objeto o estado de cosas, las creencias y los deseos serían irrelevantes en la explicación conductual, dado que serían incapaces de establecer las relaciones pertinentes entre nuestra conducta y el mundo que explican por qué hacemos lo que hacemos, y a la larga, es lo que permite las generalizaciones conductuales.

La preservación de las generalizaciones

Parece sorprendente que las generalizaciones que valen para los contenidos de las actitudes proposicionales se preserven en las relaciones causales que dichos estados poseen. Es aquí donde parece recaer todo el poder explicativo de la psicología popular. Pues, como veíamos a lo largo de esta segunda parte, el hecho de poder cuantificar sobre los contenidos de los estados mentales es lo que garantiza la posibilidad de las generalidades conductuales. Por supuesto, estas generalidades no deben ser disparatadas, después de todo se espera que al menos gran parte de la psicología popular sea verdadera.

¿Qué tanta psicología de sentido común debe ser preservada? La pregunta es legítima en la medida en que no todo lo que podemos encontrar en el sentido común resulta ser verdadero. Un ejemplo podemos encontrarlo en el *éter* o el *élan vital*, conceptos que pretendían explicar otros fenómenos y terminaron por ser radicalmente falsos y descartados por sus respectivas ciencias (ver Churchland 1981)⁹. Sin embargo, parece deseable que al menos un poco de psicología de sentido común sea verdadera, y entre más, mejor.

No parece ser este el caso con la psicología popular. Después de todo, sus éxitos pueden evidenciarse en el día a día por montones, e incluso podría decirse que es la mejor opción a nuestro alcance para explicar y racionalizar la conducta de los demás. Sus generalizaciones nos ofrecen una visión sistemática e inteligible de lo que hacen otros

⁹ Estos conceptos surgieron del sentido común e intentaban ofrecer una explicación de fenómenos como la vida, y el material del que estaba hecho el universo. A medida que la física, la química, la biología y otras disciplinas se desarrollaban, estos conceptos fueron reemplazados y eliminados por teorías sofisticadas que ahora conocemos. El punto para resaltar es que los conceptos del sentido común resultaron ser falsos en su totalidad. No obstante, tenemos buenas razones para pensar que la psicología de creencias y deseos no se parece a estos conceptos eliminados por otras teorías en su campo, en parte porque la psicología popular parece funcionar y ser indispensable para la comprensión de los demás. Esto es lo que se intenta mostrar en este trabajo y se tratará principalmente en el capítulo 3.

sujetos; lo cual es un elemento fundamental para conformar una teoría explicativa. Es preciso que se diga algo más sobre las generalizaciones y qué papel juegan.

Como vimos en el Ejemplo 3, Hermia apelaba implícitamente a una generalización del sentido común para hacer su inferencia sobre Lisandro. Por supuesto, se espera que este tipo de generalizaciones respalde contrafacticos¹⁰ —siempre y cuando se busque una psicología científica— si la causalidad entre creencias/deseos del sujeto y sus acciones son los adecuados para satisfacer la generalización. Presumiblemente las generalizaciones hechas por la ciencia, especialmente las ciencias básicas, funcionan de esta manera; lo que quiere decir qué, si es posible para las generalizaciones de la psicología popular soportar contrafacticos, es posible establecer leyes “científicas” sobre la psicología popular, o, en otras palabras, hacer de la psicología popular, y la psicología por extensión, de una ciencia porque de esta forma se captura una clase natural¹¹ que correspondería a las intenciones, deseos, creencias, temores.

Cuando se acepta que es posible generar leyes de este tipo sobre las intenciones y las acciones, nos comprometemos con la existencia de este tipo de entidades (aceptando el realismo sugerido en el pie de página 11). Por esta razón, las generalizaciones de la psicología popular, de ser verdaderas y legaliformes, nos comprometen con la existencia del tipo de entidades sobre las cuales cuantifican; si es posible ofrecer generalizaciones legaliformes respecto a las intenciones y acciones de los sujetos, estamos capturando algo respecto del mundo.

¿Cómo se logran entonces esta clase de generalizaciones? Evidentemente, como cualquier otra clase de explicación, busca generalización mediante la cuantificación sobre los sujetos que actúan por deseos y creencias; adicionalmente, es importante llamar la atención

¹⁰ La discusión respecto a los enunciados legaliformes y los enunciados con forma de leyes ha sido ampliamente discutido en la historia de la filosofía. No pretendo entrar a ese debate o traer dicho debate al tema que ahora nos ocupa. Por ahora baste decir que un enunciado legaliforme es un enunciado condicional cuantificado universalmente, esto es, la relación que se establezca en la ley vale para todos los objetos que abarcan los predicados implicados. Adicionalmente, que estos enunciados soporten contrafacticos refiere a que los predicados pueden proyectarse a mundos posibles cercanos, o mundos nomológicamente cercanos (mundos donde un cierto conjunto de leyes se aplique, por ejemplo, las de la física). Con esto se intenta que las relaciones establecidas en las leyes capturen propiedades esenciales o ‘necesarias’, es decir capturen una parte del mundo y cómo funciona, y no solamente una cuestión accidental.

¹¹ No es mi intención aquí discutir sobre lo que es una clase natural, o si la ciencia logra identificarlas efectivamente o no. Para propósitos de texto, asumiré que las clases naturales existen. Una breve idea de lo que son: “To say that a kind is *natural* is to say that it corresponds to a grouping that reflects the structure of the natural world rather than the interests and actions of human beings. We tend to assume that science is often successful in revealing these kinds; it is a corollary of scientific realism that when all goes well the classifications and taxonomies employed by science correspond to the real kinds in nature.” (Para un buen resumen de lo que este debate implica véase la entrada *Natural Kinds* de la Stanford Encyclopedia (Bird y Tobin, 2018).

sobre el hecho de que las generalizaciones también se consiguen mediante la cuantificación sobre los contenidos.

Considérese el siguiente enunciado “ x desea P y P a menos que Q ”. Si se pretende que este enunciado sea una generalización —por ejemplo— de la psicología popular, la cuantificación pretende valer para todos los P y Q . Esto se debe a algo que hemos dicho anteriormente, a saber, las relaciones causales entre las creencias/deseos y la conducta usualmente suelen respetar las relaciones semánticas entre los contenidos.

Analicemos esto más cuidadosamente. Tómese como ejemplo el predicado ‘...cree que P ’, puede parecer, a primera vista, un predicado extraño o *sui generis*, pero no es diferente a otros predicados en disciplinas como la física que se forman a partir de sustituir “ P ”. Paul Churchland pone un ejemplo bastante claro en su artículo “Eliminative Materialism and the Propositional Attitudes” (1981). El paralelismo va de la siguiente manera:

Imagine expresiones “creadoras de predicados” en el marco conceptual de la física, por ejemplo, “... tiene una masa_{kg} de n ”, “... tiene una velocidad de n ”, “... tiene una temperatura_k de n ”, etc. Lo interesante de estas expresiones es que una vez se reemplaza el término singular n por un número se forma un predicado determinado. Adicionalmente, estos predicados permiten cuantificación, haciendo así posible la expresión de generalidades que intentan capturar las relaciones legaliformes de los predicados en la naturaleza. Por ejemplo:

$$(x) (f) (m) [((x \text{ tiene una masa de } m) \ \& \ (x \text{ sufre una fuerza neta de } f)) \supset (x \text{ acelera a } f/m)]$$

Pensemos en las expresiones formadoras de predicados propias del sentido común: “... cree que p ”, “... desea que p ”, “... teme que p ”, etc. En este caso, encontramos que el término singular p es sustituido por una proposición formando un predicado psicológico determinado. Resulta interesante que las actitudes proposicionales formadas al reemplazar p mantengan las relaciones que estas proposiciones guardan entre sí, relaciones como implicación, equivalencia, inconsistencia. Aun así, lo que resulta más interesante es que las proposiciones que reemplazan los términos particulares son susceptibles también de cuantificación, lo que permite expresar generalizaciones o “leyes” entre actitudes proposicionales. Algunos ejemplos de lo anterior son:

- $(x) (p) [(x \text{ teme que } p) \supset (x \text{ desea } \sim p)]$
- $(x) (p) [((x \text{ cree que } p) \ \& \ (x \text{ cree que (si } p \text{ entonces } q))) \supset (\text{si no es el caso de confusión, distracción, etc., } x \text{ cree que } q)]$

Como bien se ha mostrado, este paralelismo entre las características estructurales de la psicología de sentido común y la física revela al sentido como una teoría capaz de explicar y predecir la conducta humana. No hay una diferencia significativa en la forma en como la teoría de la psicología popular funciona respecto a como lo hace una teoría científica. La diferencia radica en el tipo de entidades sobre las cuales se cuantifica, es decir, la psicología popular lo hace sobre actitudes proposicionales, acciones y personas, mientras que la física sobre números u objetos teóricos como átomos, partículas, moléculas, campos de fuerza, entre otros.

Podría, sin embargo, decirse que las generalizaciones de la física son diferentes a la psicología popular. Después de todo, la física no tiene excepciones, y cabría esperar que sus leyes tampoco las tengan. Por otro lado, la psicología popular parece susceptible de —quizá demasiadas— excepciones, pues las personas piensan, creen, sienten y desean cosas diferentes; no sería de extrañar que las generalizaciones de la psicología popular no aplicaran en algunas ocasiones.

No obstante, este no es el caso, esto no resulta ser un problema para —ni un argumento contra— la psicología popular. Pues las generalizaciones apelan a lo que se denomina cláusulas *ceteris paribus*. Veamos lo que esto significa.

Clausulas ceteris paribus

Las predicciones hechas por esta forma de psicología podrían verse como espurias o vacuas debido a que usualmente hace falta apelar a una cláusula *ceteris paribus* (o *en igualdad de condiciones*) para que la predicción se cumpla. Tómese como ejemplo la preferencia de x de la oración “llegaré al aeropuerto a las 3:00”, entonces “ x tiene la intención de estar en el aeropuerto a dicha hora”. Para que la predicción de la acción funcione correctamente es probable que deban agregarse cláusulas del tipo “la predicción es verdadera siempre y cuando el hablante no mienta; o no sea un hablante del francés que accidentalmente profirió la oración, o esté usando la oración como ejemplo de oraciones falsas, etc.”

Fácilmente podría arreglarse la predicción añadiendo una cláusula “en igualdad de circunstancias (*ceteris paribus*)”. Así, la predicción sería algo como “en igualdad de circunstancias, si x dice que llegará al aeropuerto a cierta hora, entonces x tiene la intención de estar en el aeropuerto a esa hora”. A pesar de esto, podría contraargumentarse que, decir “si alguien dice que llegará al aeropuerto, entonces realmente tiene la intención de estar allí a

menos que no la tenga”. Ciertamente parece predictivamente adecuado, dado que nada de lo que suceda puede hacer falsa la predicción. Sin embargo, es vacua debido a que no es explicativa y no brinda información real respecto a las acciones de las personas. Aun así, como vimos en el caso de Shakespeare, usualmente las predicciones del sentido común suelen ser verdaderas, ¿cómo es posible entonces que las generalizaciones sobre las que basamos nuestras predicciones sean vacuas?

La respuesta es simple: no son vacuas. Esto podría entenderse fácilmente si lo hacemos desde el punto de vista de una comparación. Fodor propone un ejemplo bastante claro: la geología. Compárese las generalizaciones que pueden obtenerse de la psicología de sentido común y las generalizaciones de la geología. Es ampliamente aceptado que la geología es una ciencia natural seria y consolidada y que sus generalizaciones no son vacuas y mucho menos falsas. Tómese como ejemplo “un río con meandros erosiona su margen”. Esta predicción parece ser verdadera, sin embargo, es verdadera siempre y cuando el clima no cambie y el río se congele; o el clima cambie y el río se seque; etc. Por supuesto pueden hacerse este tipo de salvedades añadiendo una cláusula *ceteris paribus*: “en igualdad de circunstancias un río con meandros erosiona su margen”. No obstante, quizá esto no quiera decir nada más que “un río con meandros erosiona su margen a menos que no lo haga”.

Claramente, este no es el caso, “en igualdad de circunstancias un río con meandros erosiona su margen” no es una generalización vacua y mucho menos falsa. Del mismo modo, es lícito para la psicología del sentido común apoyarse en cláusulas *ceteris paribus*, como también lo hace la geología. Podría, de igual manera, argumentarse que podría especificarse las excepciones a las generalizaciones de la geología de tal manera que puedan eliminarse las cláusulas *ceteris paribus*, haciéndola así diferente a la psicología de sentido común dado que dicha psicología no puede especificar de manera satisfactoria para todos los eventos y circunstancias las excepciones que aplican.

Sin embargo, si es así, bajo dicho criterio la única ciencia real sería la física básica. La razón es simple. No es cierto que puedan especificarse todas las excepciones a las generalizaciones de la geología desde un vocabulario puramente geológico. Esto es en parte porque los fenómenos que suelen ser las causas de las excepciones no son en sí mismos eventos geológicos. Un ejemplo de ello es “un río con meandros erosiona su margen a menos que el clima cambie y el río se seque; o el mundo se termine; etc.”, pero “clima”, “el mundo se termine” no son conceptos de la geología, y esto aplica para gran cantidad de descriptores (*descriptors*). Todo lo que puede decirse que aporte utilidad es que, si la generalización dejó de valer, de alguna manera no fueron satisfechas las idealizaciones requeridas. Pero lo mismo

aplica para la psicología de sentido común: si *x* no llegó cuando tenía la intención de llegar, algo debió de salir mal (Fodor, 1987); quizá el vuelo se retrasó o hubo un problema con la aerolínea, pero es poco probable que dudemos de la teoría. Es decir, negar que la conducta se explique con las intenciones de los sujetos.

En resumen, las cláusulas *ceteris paribus* son necesarias para la psicología popular, pero también lo son para otras ciencias como la geología. Esto quiere decir, que no por ello debemos renunciar a las generalizaciones legaliformes sobre la conducta. En parte, porque también queremos rescatar la realidad de las acciones, las creencias, deseos, etc. En este sentido, la psicología popular intenta capturar el fenómeno con todo lo que implica, sus entidades y su realidad ontológica.

Todo lo anterior apunta a que la psicología popular no solo puede llegar a ser una teoría científica, sino que incluso se comporta como una, pues ofrece generalizaciones, predicciones e incluso predicciones correctas sobre la conducta. Por supuesto, todo lo que se ha mencionado sobre la psicología popular proviene de un conjunto de intuiciones pre-teóricas, por lo que quizá haga falta un poco más de robustez conceptual y teórica para que de hecho la psicología popular sea una teoría científica factible. Es posible que haya varias formas diferentes de lograr esto, sin embargo, lo que pretendo presentar en el siguiente capítulo es una forma de —y quizá una de las más famosas y robustas— hacer de la psicología popular una teoría científica en toda regla bajo las exigencias de la ciencia empírica y sus postulados filosóficos.

Como breve recapitulación de lo que se ha mencionado hasta el momento, me he centrado en exponer lo que entiendo por psicología popular y las intuiciones centrales en lo que respecta a la racionalización y explicación de la conducta humana. Esta racionalización se hace apelando a los deseos, creencias e intenciones por las cuales un determinado agente realiza una acción. Adicionalmente, mostré que las acciones se encuentran estructuradas y pueden formarse acciones más complejas a partir de acciones más simples. El caso paradigmático es la conducta lingüística, donde parece haber una clara estructuración sistemática y composicional. Luego, mostré cómo las intuiciones de la psicología popular y su aparente estructuración dan como resultado un éxito explicativo y predictivo de la conducta, llegando a obtener el esquema de una teoría. Finalmente, en las últimas secciones del capítulo me centré en mostrar por qué la psicología popular puede considerarse no solo una teoría, sino una teoría de carácter científico. Argumenté que para lograr este objetivo

hacía falta apelar a una explicación más robusta del éxito explicativo de la psicología popular. En el siguiente capítulo pretendo elaborar esta explicación.

Capítulo II

Modularidad y Simbolismo

La psicología popular, como se dijo en el capítulo anterior, es un conjunto de intuiciones acerca de cómo es la conducta humana. Para explicar su éxito y darle la forma de una teoría científica, hace falta apelar a algunas hipótesis subsidiarias. A lo largo de este documento asumiré como verdaderas: la hipótesis del *funcionalismo* por un lado; y por otro, la hipótesis de la *computacionalidad*. Aunque no es obvio que la tesis funcionalista y computacional de la mente sean correctas, en esta investigación y en el desarrollo de este documento asumiré la verdad del funcionalismo y la computacionalidad de la mente. En parte porque el debate que quiero presentar y evaluar dependen de estas hipótesis (los protagonistas del debate creen en ellas y no las ponen en cuestión) y en parte porque la tesis que quiero defender en esta investigación depende igualmente de ellas.

De manera general, podría decirse que el funcionalismo es la tesis que describe los estados mentales y la mente en general por su descripción funcional:

Functionalism is the doctrine that what makes something a thought, desire, pain (or any other type of mental state) depends not on its internal constitution, but solely on its function, or the role it plays, in the cognitive system of which it is a part. More precisely, functionalist theories take the identity of a mental state to be determined by its causal relations to sensory stimulations, other mental states, and behavior. (Levin, J., 2018)

A lo largo de este capítulo apelaré a una de las formas que puede tomar esta hipótesis, a saber, *la computacionalidad de la mente*.¹² Pretendo mostrar que las hipótesis empíricas que explican el éxito de la psicología popular de forma computacional son la *modularidad de la mente* y el *lenguaje de pensamiento*.

Una forma general de presentar la computacionalidad es de la siguiente manera:

The intuitive notions of computation and algorithm are central to mathematics. Roughly speaking, an algorithm is an explicit, step-by-step procedure for answering some question or solving some problem. An algorithm provides routine mechanical instructions dictating how to proceed at each step. Obeying the instructions requires no special ingenuity or creativity. (Rescorla, M, 2020)

¹² No es mi intención en esta investigación argumentar sobre la plausible relación entre el funcionalismo y la computacionalidad y el tipo de relación que sea esta. Para fines de la discusión asumiré el funcionalismo como cierto y la computacionalidad de la mente como una hipótesis más específica acerca de cómo describir las relaciones causales entre los elementos de la teoría.

En las siguientes secciones de este capítulo me centraré en presentar cómo las hipótesis de la modularidad de la mente y del lenguaje de pensamiento ofrecen una explicación computacional de la mente: es decir, ofrecen una caracterización de los algoritmos y el medio sobre el cual operan estos algoritmos. Y a partir de esto, espero señalar cómo estas dos hipótesis logran explicar el éxito de la psicología popular. La intuición central en este punto es que la hipótesis de la *modularidad de la mente* y el *lenguaje de pensamiento* ofrecerán una descripción de las relaciones causales adecuadas que plausiblemente son las responsables de la conducta inteligente tal como la describe la psicología popular.

Por esta razón, este capítulo contará con dos partes. La primera consiste en presentar esquemáticamente lo que considero son los argumentos principales y las hipótesis centrales de la teoría *fodoriana* y su posición respecto a la mente. En primera instancia, Fodor apela a la *modularidad de la mente* como una tesis con la que busca mostrar y explicar los mecanismos cognitivos que dan cuenta de la conducta inteligente como parece ser exhibida por los sujetos. En la segunda parte presentaré de manera breve algunos argumentos a favor de un sistema representacional y una caracterización del lenguaje en el que se llevan a cabo las computaciones o sobre el cual operan los algoritmos. De acuerdo con Fodor, el lenguaje en el que ocurren estas computaciones es muy similar al lenguaje natural y lo llama el lenguaje de pensamiento, **LOT** (por sus siglas en inglés). Este lenguaje, LOT, es esencialmente proposicional y será la base del sistema representacional en el cual se computa y se representa la información del sistema. Este sistema representacional se asemeja a los lenguajes naturales debido a que estos ofrecen la riqueza y versatilidad requerida para garantizar que la explicación conductual sea exitosa.

Finalmente, mostraré cómo estas hipótesis logran dar cuenta de la conducta lingüística, y, a la larga, de toda la conducta inteligente. Expondré la explicación que esta hipótesis ofrece a la propiedad que parece ser central a la conducta lingüística e inteligente, tal como se mencionó en la sección “Las acciones estructuradas” en el capítulo II, se trata de la composicionalidad.

Para el final de este capítulo pretendo haber explicado algo parecido a la siguiente imagen¹³:

¹³ En primera instancia, es importante aclarar que la imagen representa mi propia versión de cómo puede plasmarse la *modularidad de la mente*. Por otro lado, las flechas indican el flujo de la información que “fluye” de los módulos y es tomada luego por los sistemas centrales. También la falta de divisiones en el cuadro de los sistemas centrales indica que de una u otra forma la información es unificada en estas instancias.

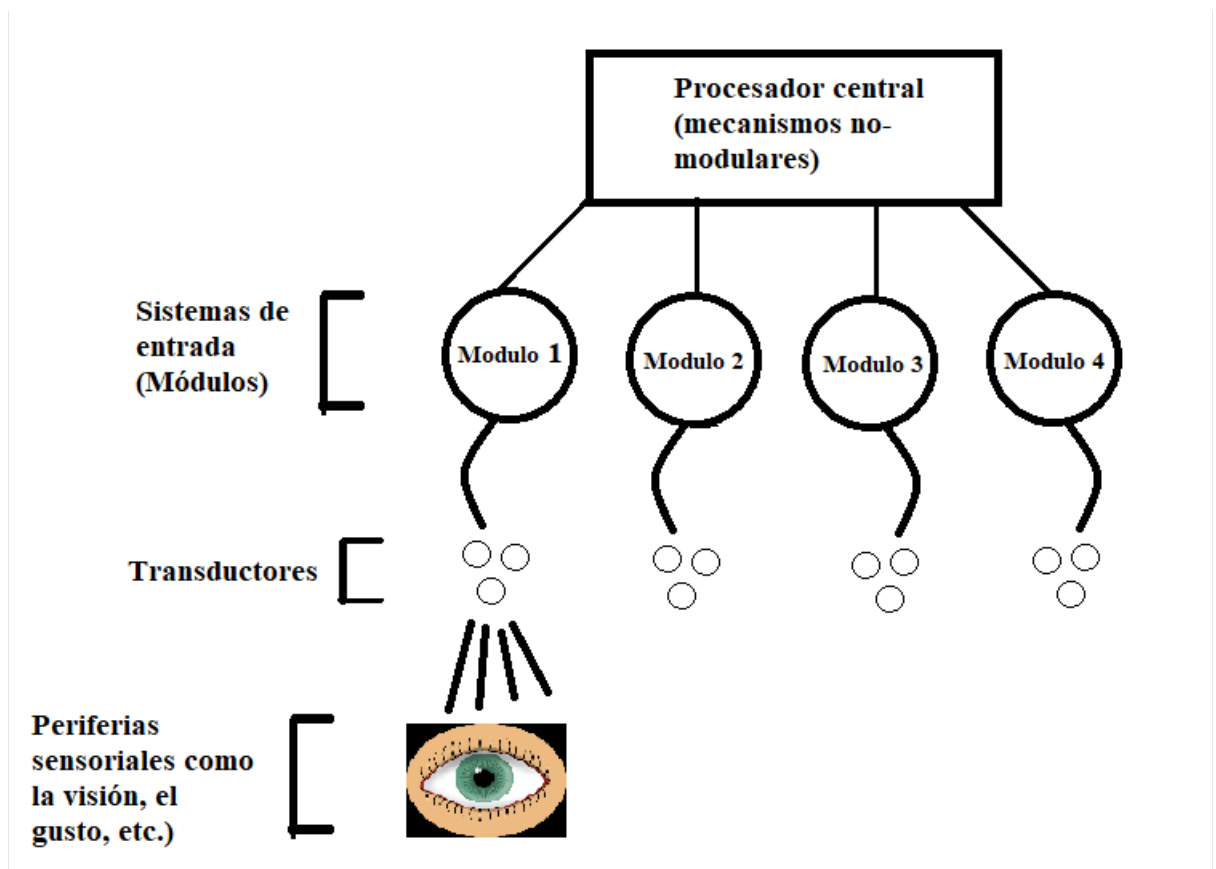


Fig. 1¹⁴. En esta imagen la ‘modularidad’ es representada por la jerarquía entre las periferias sensoriales, cada uno de los módulos y el procesador central. El ‘lenguaje de pensamiento’, por otro lado, es representado por las conexiones de los módulos al procesador central (representadas por las líneas rectas delgadas) y lo que plausiblemente ocurre dentro de este.

¹⁴ Fodor establece una taxonomía funcional de la mente: (1) Transductores, (2) sistemas de entrada y (3) Procesadores centrales. A grandes rasgos, los transductores y los sistemas de entrada se encargan de suministrar información (ambiental) a los sistemas centrales. Los transductores introducen la información al sistema (proveniente de las irritaciones nerviosas); el resultado final de los transductores u *output* es luego tomado por los analizadores perceptivos o sistemas de entrada. Los sistemas de entrada se encargan de generar hipótesis, con base en la información de la estimulación proximal, respecto de los objetos distales que pudieron causar dicha estimulación. En este sentido, deben entenderse a los sistemas de entradas como generadores de inferencias no demostrativas (hipótesis) a cerca de los objetos externos teniendo como premisas la información de la estimulación proximal. Además de esto, se encargan de transducir la información proveniente de los transductores a un lenguaje entendible para la máquina, o puesto en otras palabras, un nivel de representación que sea accesible a los procesos cognitivos centrales. La transducción antes mencionada consiste en la transformación de la información en diferentes niveles de representación hasta tal punto que sea accesible para los mecanismos cognitivos de orden superior no modulares, que como se explicará después, pueden ser el pensamiento, la memoria o la fijación de creencias.

Primera Parte

Modularidad de la mente y sistemas modulares¹⁵

La primera hipótesis que quiero esbozar es llamada por Fodor mismo como *modularidad de la mente*. Presentaré lo que considero son las nociones más importantes respecto a esta hipótesis con la intención de hacer una breve reconstrucción que muestre cómo el éxito de la psicología popular puede ser explicado apelando a esta hipótesis. La conclusión que quiero dibujar en esta sección es que la mente se encuentra dividida en sistemas modulares y no-modulares. Esta distinción entre los mecanismos cognitivos nos permitirá dar cuenta de la explicación de fenómenos —tanto cognitivos como conductuales— de diferente naturaleza, principalmente los siguientes fenómenos: (1) procesos que son llevados a cabo por el sistema de los sujetos, esto es, de manera automática, sub-personal, y relativamente indiferentes a las necesidades o deseos del sujeto; (2) procesos que son atribuidos al sujeto como un todo y son explicados en términos de mecanismos no-modulares que dan cuenta de los aspectos cognitivos relativamente superiores (es decir, pensamiento, memoria, solución de problemas, fijación de creencias, etc.). Finalmente, mencionaré que las computaciones en estos sistemas no-modulares ofrecen la explicación subyacente a la conducta intencional de los sujetos y la conducta lingüística que se encuentra en nuestro punto de interés en esta investigación.

La hipótesis de la *modularidad de la mente* consiste en una taxonomía funcional de la mente, es decir, divide la mente en sistemas que se individualizan por su función de cara a la producción de la conducta y su papel en otros procesos cognitivos. De esta forma, la *modularidad de la mente* nos brinda herramientas útiles para distinguir los procesos cognitivos que son hechos por el sujeto (como las acciones, pensamiento, memoria, solución de problemas, etc., que es lo que nos interesa) y aquellos procesos que son realizados por el sistema solo y no dependen en ningún sentido del sujeto (como la digestión, el latido del corazón, etc.).

¹⁵ La hipótesis de la modularidad podría generar toda una discusión aparte de la investigación que ahora nos concierne. Aun así, no es de mi interés entrar en dicha discusión, por esta razón, intentaré presentar los aspectos de la *modularidad de la mente* que considero son los más importantes y deben tenerse en cuenta para la presente discusión a riesgo de dejar por fuera elementos si bien interesantes, secundarios o menos importantes en la presente discusión.

La idea de postular mecanismos que son los responsables causales de la conducta se encuentra estrechamente ligada a la idea de que la conducta se encuentra organizada: La conducta se encuentra organizada, aunque su organización es un fenómeno derivado; la estructura de la conducta guarda la misma relación con la estructura de la mente que la de un efecto con su causa. En esto se resume la doctrina mentalista ortodoxa, que es la que vamos a asumir a lo largo de la exposición que ahora se inicia: las explicaciones psicológicas ortodoxas dan cuenta de la organización de la conducta apelando a unos principios que, según ellas, explican la estructura de la mente. (Fodor, 1986, p. 21)

Con esta cita en mente, incluso si aceptamos que la conducta se encuentra organizada, queda por explicar la forma en la que está organizada, o, para ser más precisos, los mecanismos que la organizan. Veamos un ejemplo. Si yo le pregunto, a usted lector, “¿cuánto es 3 más 5?”, y usted hace bien las cuentas, entonces responderá “8”. En ese caso, su conducta se encuentra estructurada de la forma correcta. Sin duda su conducta puede ser explicada a partir de lo que sabe y cree sobre los números. Pero, hace falta algo más en la explicación, después de todo, el conocimiento y las creencias relevantes no se hacen presentes solo por el hecho de ser relevantes. Hacen falta algunos mecanismos que pongan en marcha lo que conocemos o, en otras palabras, que organicen la conducta para que esté en concordancia con aquello que sabemos¹⁶.

Como se mencionó en el capítulo anterior, podríamos señalar al menos dos tipos de conductas diferentes, a saber, la conducta intencional e inteligente, y los eventos motores realizados por el sistema. Hay aquí una intuición más que explotar. Si la conducta intencional es diferente a otros tipos de eventos motores, es racional pensar que la diferencia recae en su causa. Es plausible que los mecanismos destinados a la producción de la conducta intencional sean diferentes que los mecanismos que dan lugar a otros eventos motores no-intencionales o sub-personales. ¿Qué razones tenemos para apoyar esta intuición? Incluso si se necesita de mecanismos cognitivos para organizar y estructurar la información, ¿por qué hace falta postular dos clases de mecanismos diferentes? ¿En qué consiste la diferencia?

Tomemos por ejemplo los procesos perceptuales más básicos —tales como visión o audición—. Estos procesos, podría decirse, ocurren de manera automática y de manera relativamente independiente al sujeto, esto es, no podemos elegir escuchar o ver algo a menos que interrumpamos el proceso externamente cubriendo nuestros ojos u oídos.

¹⁶ Este ejemplo ha sido tomado de Fodor (1986, pp. 26, 29). Sin embargo, Fodor allí lo utiliza para ilustrar un punto respecto al neo-cartesianismo, por lo que no resalta de manera explícita lo que ahora yo intento resaltar. Me permitiré manipular ligeramente el ejemplo para resaltar un punto de interés para esta investigación, a saber, la estructuración de las acciones.

Adicionalmente, es razonable pensar que procesos como la visión son relativamente independientes a las necesidades del sujeto en tanto que no suele ser el caso que veamos lo que *esperamos* o *queremos* ver; y es aquí donde podemos encontrar una diferencia importante con la conducta intencional.

La diferencia, como pretendo mostrar a lo largo de este capítulo, recae en los mecanismos que producen ambas conductas. Por un lado, los mecanismos responsables de la conducta intencional deben ser sensibles a la información y necesidades del sujeto. Por otro lado, los mecanismos responsables de procesos —principalmente perceptuales— como la visión o audición parecen requerir de mecanismos menos sofisticados y con acceso más limitado. Veamos que clase mecanismos pueden cumplir estas tareas.

La hipótesis de la *modularidad de la mente* propuesta por Fodor ofrece una respuesta a esta problemática. Fodor clasifica los mecanismos mentales en dos clases: modulares y no-modulares. Así, los sistemas modulares responden causalmente a los procesos cognitivos no-intencionales y los mecanismos no-modulares son los responsables de la conducta intencional. Veamos esto en más detalle.

La característica principal de los sistemas modulares o, más bien, la propiedad que los hace modulares es el *encapsulamiento informativo*. En este sentido, debemos explicar principalmente qué conlleva que un sistema sea encapsulado informativamente y cómo se relacionan las propiedades que se siguen de esta.

Los sistemas modulares son: (i) informativamente encapsulados; (ii) tienen un dominio específico; (iii) son automáticos y obligatorios. Veamos.

Los sistemas modulares son informativamente encapsulados

Un ejemplo claro de módulos informativamente encapsulados son los sistemas de entrada (ver, *supra*. Fig. 1). Cuando se dice que los sistemas de entrada¹⁷ se hallan informativamente encapsulados nos referimos a la *inaccesibilidad* de los sistemas de entrada (o el sistema modular) a la información general de la que dispone el sujeto.

Piense por ejemplo en la ilusión de Müller-Lyer donde, a pesar de que el sujeto sabe explícitamente que las flechas tienen la misma longitud por haberlas medido en presencia de él, el sujeto sigue percibiéndolas como una más larga que la otra, haciéndonos pensar en la relativa indiferencia del análisis perceptivo respecto de la información contextual del sujeto.

¹⁷ A lo largo de este texto tomaré “Sistemas de entrada” y “módulos” o “sistemas modulares” como la misma clase de mecanismos. Esta identidad, sin embargo, no es obvia y hace parte de la hipótesis defendida por Fodor en *La modularidad de la mente*, para un mejor desarrollo véase Fodor (1986).

El producto de los sistemas perceptivos es, en su mayor parte, insensible a lo que el receptor supone o desea. [...] una condición que la percepción debe satisfacer para ser fiable, al menos en el caso de los organismos que pueden equivocarse, es que en general vea lo que hay, y no lo que desea o espera que haya. El organismo que no funcione así no podrá sobrevivir. (Fodor, 1986, pp. 102-103)

Esto no implica que no puedan presentarse casos de información descendente¹⁸, puesto que también existen casos evidentes en los que se da la influencia en el resultado del análisis debido a la información contextual como lo es, por ejemplo, el “efecto de restauración de fonemas” donde se omiten fonemas en un enunciado y, sin embargo, el sujeto es capaz de percibirlos como no faltantes en el enunciado¹⁹.

Lo que esto sugiere es la capacidad de los sistemas de entrada de dar respuesta a un determinado grupo de problemas a partir de una información de naturaleza específica y limitada, e independiente de información general que tiene el sujeto conscientemente. Los sistemas de análisis perceptuales, es decir, aquellos sistemas que se encargan de analizar la información proveniente de las periferias sensoriales, al realizar las inferencias no demostrativas o hipotéticas acerca de los objetos distales es suficiente con la información provista por los transductores y no hace falta ningún otro tipo de información contextual o de orden superior. Desde otra perspectiva, también puede verse como la inaccesibilidad por parte de los sistemas de entrada a información accesible o disponible a otros procesos cognitivos²⁰.

¹⁸ Piense por ejemplo en aquel cuadro de su casa que ha visto tantas veces con antelación. Cuando pasa cerca del cuadro rápidamente, digamos por prisa; y apenas logra capturar una forma muy difuminada, la percepción, sin embargo, es capaz de reconstruir el percepto con mejor detalle debido a la información que ya se tiene respecto al objeto.

¹⁹ Este fenómeno se conoce como *Phoneme Restoration Effect*. Consiste en la recuperación de fonemas que son inexistentes o se han perdido en los inputs auditivos. Podría describirse de la siguiente manera: “Primero se graba una palabra (por ejemplo, la palabra “legislatura”) y a continuación se borra uno de los sonidos que integran dicha palabra (por ejemplo, el sonido “s”), sustituyéndolo por el ruido de una tos “de fondo”. La estructura acústica de la señal resultantes es, pues, /legi (tos) latura/. Sin embargo, lo que el sujeto oye al escuchar la cinta es una emisión de la palabra /legislatura/ con una tos “de fondo”. Según parece, lo que ocurre es que la constitución fonética percibida de esta emisión vocal viene determinada no sólo por la información transducida (esto es, la información definida en niveles *subfonéticos* de análisis), sino también por información de orden superior acerca de la posible representación léxica de dicha emisión vocal (es decir, por la suposición del sujeto de que la palabra emitida debía ser /legislatura/)” (Fodor, 1986, pp. 99). Para una muestra más gráfica de lo que sucede véase Knapp, 2012.

²⁰ Pylyshyn (1980) se refiere a esta propiedad como “impenetrabilidad cognitiva (Fodor, 1986, pp. 102). Vale la pena aclarar que ‘encapsulamiento de la información’ e ‘impenetrabilidad cognitiva’ no son estrictamente la misma propiedad. Podría decirse que son dos caras de la misma moneda, me explico. La impenetrabilidad cognitiva a la que Pylyshyn hace referencia se refiere a la poca o nula accesibilidad que otros sistemas -como los sistemas centrales- tienen al módulo y a lo que sucede en su interior. Por otro lado, el encapsulamiento de la información al que Fodor se refiere limita la información a la que el módulo apela para llevar su procesamiento. La información que un sistema informacionalmente encapsulado requiere para

Por otro lado, es de esperar que los mecanismos no-modulares (como el sistema central, ver Fig.1) o los procesos superiores también tengan un acceso restringido a lo que sucede en el interior de los módulos. Con “acceso” quiero decir —siguiendo a Fodor— la capacidad del procesador central de representar de manera explícita el contenido informativo que se encuentra en las representaciones mentales.

El aspecto que quiero destacar es que el sujeto no tiene idéntico acceso a todos estos niveles ascendentes de representación —al menos si tomamos como criterio de accesibilidad la posibilidad de disponer de un suministro explícito de la información que estas representaciones encierran. De hecho, como señalé anteriormente, los niveles inferiores (los que se hallan más próximos a los productos de los transductores) parecen ser completamente inaccesibles para cualquier propósito. (Fodor, 1986, p. 88)

Un ejemplo en el que se puede percibir más claramente esto es en el caso de la visión. El sujeto tiene un acceso a las representaciones relativamente superiores, por ejemplo, si estamos viendo una naranja, podemos dar cuenta explícita de dicho objeto. Sin embargo, si se nos preguntara cómo es que construimos la imagen tridimensional de la naranja a partir de las imágenes 2D que capta nuestra retina, sería un completo misterio para nosotros, aunque este proceso de hecho sucede en el análisis relativamente temprano de la visión.

El acceso del procesador central respecto de los diferentes niveles de computación de los sistemas de entrada puede describirse de la siguiente manera: el acceso desde el procesador central es directamente proporcional a la superioridad del nivel. A partir de esta idea podemos inferir que entre más bajo sea el nivel de computación menos accesible será para el procesador central. Esto no quiere decir que la información se pierda o no juegue ningún rol importante. No obstante, el sujeto no es capaz de notarlo o de tener acceso explícito a la información que hace que cambie su conducta respecto a ciertos estímulos perceptivos²¹.

funcionar es relativamente mínima y se encuentra dentro del mismo módulo (por decirlo de alguna manera), no se necesita apelar a la información de otros módulos o de los sistemas centrales para operar.

Por supuesto ambas ideas se complementan, a la larga lo que se expresa es que el módulo está relativamente aislado de los demás subsistemas (incluyendo los procesadores centrales) que componen la máquina. Los módulos no necesitan acceso a otros módulos o a información externa al mismo; y de igual manera, los demás módulos y sistemas centrales no tienen acceso al procesamiento de información que ocurre dentro del módulo.

²¹ Estos términos pueden sonar algo técnicos y conceptualmente oscuros, pero no pretenden designar más que fenómenos cercanamente conocidos como lo que es la generación de creencias o el llegar a creer algo. Lo mismo ocurre con el pensamiento que, podría entenderse —bajo la descripción de la psicología popular— como cadenas de contenidos semánticamente relacionados. En general, los sistemas centrales, según Fodor, deberían encargarse de procesos que pueden describirse como ‘globales’, es decir, que no se centran en información particular, sino que, por el contrario, tienen acceso o depende de toda la información de la que dispone el sujeto. En este sentido, procesos como el pensamiento, fijación de creencias, memoria, atención, etc.,

Los sistemas modulares son específicos de dominio

Una forma trivial y casi cómica de “explicar” la especificidad de dominio es la forma según la cual la función de un sistema de percepción de vacas -por ejemplo- es la de identificar tipos de vacas a partir de la información suministrada por los transductores y transducir dicha información a fórmulas en un determinado código central. Otra manera de ponerlo y por mucho menos trivial, es la capacidad de un sistema de entrada para responder a una cierta clase de problemas a partir de un conjunto relativamente determinado de información. La noción de especificidad de dominio es la idea Galliana [de Gall] de que existen mecanismos diferenciados que corresponden a dominios estímulares diferenciados. (Fodor, 1986).

Los sistemas modulares son susceptibles —es decir, se activan— solo a cierto tipo de información. Piense que los sistemas modulares que se encargan de procesar, por ejemplo, la información visual, solo se activarán frente a la información relevante, es decir, plausiblemente a la luz. En este sentido, por ejemplo, si usted, lector, estuviera observando una naranja, la información no activaría el módulo dedicado al procesamiento de información auditiva no-verbal, sino el módulo encargado de la información visual. Otro ejemplo intuitivo que puede ofrecerse —guardando las debidas distancias y diferencias— puede encontrarse en *De Anima* en la exposición del “sensible propio de cada sentido:

Llamo por lo demás, ‘propio’ a aquel objeto que no puede ser percibido por ninguna otra sensación y en torno al cual no es posible sufrir error, por ejemplo, la visión del color, la audición del sonido y la gustación del sabor. (*De Anima* II, 6, 418^a10-15)

Con esto quiero señalar que un conjunto reducido de estímulos de una naturaleza específica es la que se espera que ponga en funcionamiento un determinado mecanismo de análisis perceptual. Hay, por ejemplo, varios estudios empíricos que muestran con sorprendente precisión la especificidad de dominio de mecanismos de análisis verbal frente a mecanismos de análisis no-verbal, es decir, la manera en que le suena el sonido al oyente depende ampliamente del contexto. Cuando el contexto sugiere que lo que se está escuchando puede ser un sonido verbal, los analizadores perceptivos puestos en marcha son diferentes a los que se activan cuando el contexto indica que un sonido es no-verbal y esto conlleva a que el oyente los escuche de manera diferente:

no tienen acceso directo a los módulos y solo de manera parcial a los niveles más elevados de representación dentro de estos.

Expresado a grandes rasgos, la misma señal que se escucha como el comienzo de una consonante, cuando el contexto especifica que el estímulo percibido es lenguaje, se percibe como un ‘silbido’ una ‘semivocal’ cuando es aislado de la cadena hablada. De aquí se desprende la importante conclusión de que los sistemas computacionales que entran en juego durante el análisis perceptivo del habla son distintivos, en la medida en que operan *únicamente* sobre las señales acústicas que se identifican como enunciados verbales (Véase Liberman y cols., 1967; para un estudio más detallado, véase Fodor, Bever y Garrett, 1974). (Fodor, 1986, pp. 78-79, referencias y énfasis en el original)

Los sistemas modulares son obligatorios y rápidos

Una vez el módulo ha sido activado por la información pertinente que intenta ingresar al sistema, el sistema se vuelve un paso necesario y obligatorio en el proceso. Esto refiere, por supuesto, al ingreso de la información al módulo. Es imposible —por ejemplo— *evitar percibir/escuchar un enunciado verbal como tal*; es imposible dejar de percibir los estímulos visuales como un conjunto de objetos distribuidos en el espacio tridimensional (Recordemos la ilusión de Müller-Lyer). Con esto se quiere decir que la computación hecha por parte de los sistemas de entrada es indiferente a las necesidades o deseos del sujeto²². Los sistemas de entrada no se apagan o desconectan porque así el sujeto lo quiera; no dejan de recibir información del ambiente ni tampoco procesarla para integrarla al sistema.

En lo que respecta a la rapidez, esta hace referencia a la velocidad con que se analiza la información²³. No es coincidencia que el procesamiento de información en los módulos sea rápido, ya que al ser la entrada obligatoria se gana efectividad computacional. Esto justamente se debe a que no hay que invertir recursos en discriminar lo que es pertinente de lo que no, sino que automáticamente se lleva a cabo el análisis de la información; será luego (cuando la información llegue a los procesos centrales) cuando se decida si la información entrante es pertinente a las necesidades o deseos del sujeto.

²² Es cierto que hasta cierto punto y en un sentido suave podemos tener control sobre los sistemas de entrada. Por ejemplo, si nos tapamos los oídos y dejamos de atender a los estímulos auditivos centrando nuestra atención en otra cosa. Sin embargo, esto no es exactamente que la información no sea procesada o el módulo desactivado, pues son medios de control externos al sistema mismo.

²³ La velocidad de los módulos se contrasta con la relativa lentitud de los procesos centrales. Es fácil notar que podemos pasar horas o días pensando sobre un problema matemático o filosófico, pero el análisis de la información visual, en contraste, tarda solo unos milisegundos. Esto señala que el procesamiento de la información está de alguna manera desprovisto de toda inteligencia y pone en funcionamiento solo una pequeña parte de los recursos computacionales del sistema (Fodor, 1986, p. 97) y esto libera recursos para los procesos centrales.

* * *

Hasta aquí lo que respecta a la explicación de las propiedades más importantes de los sistemas cognitivos modulares. Estas no son todas las propiedades descritas por Fodor para considerar a un sistema modular, pero son necesarias: si una de estas —o en palabras más apropiadas, la mayoría de estas— propiedades llegase a faltar en un sistema, difícilmente podría verse como modular. Las propiedades que faltan mencionar son: (1) los sistemas de entrada se refieren a aspectos ‘superficiales’²⁴, (2) los sistemas de entrada se hallan asociados a una arquitectura neural fija, (3) los sistemas de entrada presentan unas pautas de deterioro características y específicas, y (4) la ontogénesis de los sistemas de entrada presenta un ritmo peculiar y una característica sucesión de estadios²⁵.

Si bien los módulos dan cuenta de problemas relativos a su especificidad de dominio, no bastan por sí solos para dar cuenta de toda la economía cognitiva. Postular únicamente mecanismos que se encuentran relativamente aislados los unos de los otros no es suficiente para dar una explicación satisfactoria de la mente en tanto que nos topamos con procesos como la imaginación, el pensamiento, la memoria o la fijación de creencias, etc., que no son específicos de dominio y tienen acceso a una variada cantidad de información proveniente de diferentes dominios y módulos. Esto a su vez nos lleva a rechazar la idea del encapsulamiento informativo por parte de los mecanismos responsables de los procesos antes mencionados.

Por esta razón, Fodor postula un *procesador central*, esto es, un ‘mecanismo’ donde se unifica la información proveniente de los diferentes módulos y se llevan a cabo las computaciones que sí son sensibles a las necesidades e intereses del sujeto y tienen acceso a información proveniente de otros módulos. Es lo que sucede en este procesador central: los procesos cognitivos que éste lleva a cabo son plausiblemente explícitos al sujeto.²⁶

²⁴ Respecto a (1) en la sección siguiente (cuando se exponga el *lenguaje de pensamiento*) se dirá un poco más sobre ella y como entra a jugar dentro de toda la explicación. Por ahora baste decir que esta propiedad es la que relaciona o conecta los módulos con los procesos centrales, después de todo los resultados u *outputs* de los módulos es lo que el procesador central toma y sobre lo cual opera. Esto es lo que sería el lenguaje de pensamiento.

²⁵ Fodor considera todo este conjunto de propiedades como descriptivas de los sistemas modulares, sin embargo, no son propiedades que considero centrales y necesarias para la modularidad, por lo que no considero que sean relevantes para la discusión.

²⁶ En este punto no quisiera comprometer a Fodor con la *transparencia de la mente* a sí misma. Fodor no es muy claro frente a la posición que toma en este respecto; parece sugerir que, en el procesador central, de una u otra manera, las representaciones son accesibles a la conciencia, por lo que me limitaré a señalar la plausibilidad de que los procesos llevados a cabo en el procesador central *sean explícitos* para el sujeto.

El pensamiento o la imaginación serían procesos llevados a cabo en el procesador central, manipulable por el sujeto, susceptible a sus intereses, deseos, necesidades, etc.; que además tiene acceso a la información proveniente de los diferentes módulos perceptuales y a la información almacenada en el sistema —es decir, la memoria—. Lo mismo podría decirse de la *fijación de creencias*, solución de problemas, etc.

* * *

Finalmente, y a modo de conclusión provisional para esta sección, puede decirse lo siguiente. La hipótesis de la modularidad de la mente consiste en la explicación de la mente a partir de la postulación de mecanismos modulares y no-modulares. Los modulares son mecanismos funcionalmente distinguibles que cumplen las propiedades de ser específicos de dominio, están informativamente encapsulados y, por último, son automáticos y rápidos. El flujo de la información en estos mecanismos debe ser obligatoriamente analizada y el acceso a ésta por parte de sistemas no-modulares de orden superior es relativamente restringido. Adicionalmente, dada la naturaleza de algunos procesos cognitivos, la mente debería contar con un procesador central no-modular que lleve a cabo procesos de orden superior que integran la información de los diferentes módulos, la almacenen y, a partir de esta información integrada y almacenada de lugar a la conducta intencional y lingüística del sujeto.

Habiendo dicho esto, es necesario señalar que la hipótesis de la modularidad de la mente no es suficiente por sí sola para ofrecer una explicación completa de la psicología popular, dado que la modularidad se centra en los mecanismos modulares y dice poco respecto del procesador central no-modular. Para una mejor comprensión del procesador central, hace falta ver más de cerca la clase de computaciones que se llevan a cabo dentro de este.

Para ello, presentaré, a continuación, la hipótesis del *lenguaje de pensamiento* desarrollada por Fodor, la cual busca ofrecer una caracterización de esos algoritmos y símbolos sobre los que se llevan a cabo las computaciones; y ofrece una respuesta a la manera en que los módulos están conectados y se comunican con el procesador central.

Segunda Parte

Lenguaje de Pensamiento

Recapitulando las hipótesis ofrecidas por Fodor hasta el momento, tenemos que: la mente puede describirse de forma funcional, específicamente, de forma computacional. La hipótesis de la modularidad nos brinda dos tipos de mecanismos para hacer esta descripción: los *modulares* que, por un lado, se encargan de analizar y transducir la información ambiental y “ponerla” en un lenguaje entendible para la máquina; y el segundo tipo de mecanismo, un procesador central presumiblemente *no-modular* que se encarga de unificar y almacenar la información proveniente de los módulos y las periféricas sensoriales.

La propuesta de Fodor en este punto consiste en describir el tipo de computaciones que ocurren en el procesador central. Para ello, necesita ofrecer una descripción funcional de la mente más robusta que cumpla con las exigencias impuestas por la psicología popular para garantizar el éxito explicativo; exigencias presentadas en la parte final del primer capítulo de este documento.

En todo caso, para el final de este capítulo, Fodor dirá que las computaciones que ocurren en el procesador central pueden verse codificadas como si se tratara de un lenguaje natural de carácter proposicional, composicional y sistemático; definido, principalmente, por sus propiedades sintácticas que, aun así, logran capturar los contenidos sobre los que cuantifican las generalizaciones de la psicología popular. En lo que resta del capítulo ofreceré algunos argumentos a favor de la necesidad de dicho sistema representacional para, finalmente, ofrecer una breve presentación del sistema que logra cumplir con estos requerimientos explicativos y su lugar en la explicación de la conducta.

* * *

La hipótesis de la computacionalidad nos dice, a grandes rasgos, que la información proveniente de los módulos son cadenas de representaciones que son transformadas en otras representaciones a partir de algoritmos. Para esto, es necesario un sistema representacional que cumpla con los requisitos impuestos por la psicología popular (ver segunda parte del capítulo I) —si es que se intenta ofrecer una explicación de esta—.

Fodor ofrece una caracterización que, desde ahora, la llamaremos como la *hipótesis del lenguaje de pensamiento*: el sistema representacional compatible con la modularidad y la psicología popular debe ser de carácter funcional, computacional y proposicional para dar

cuenta de propiedades como la composicionalidad de la conducta lingüística e inteligente. Veamos en más detalle cómo la caracterización de Fodor cumple con esta descripción y cómo se relaciona con la hipótesis de la modularidad y la psicología popular.

Empecemos con la caracterización funcional y computacional del sistema representacional. Fodor la presenta de la siguiente manera:

1. Los únicos modelos psicológicos de procesos cognitivos que parecen ser remotamente plausibles representan esos procesos como computacionales.
2. Dicha computación presupone un medio de computación, es decir, un sistema representacional.
3. Tener una teoría que sea remotamente plausible es mejor que no tener ninguna teoría en absoluto.
4. Estamos, entonces, comprometidos provisionalmente con atribuir dicho sistema representacional a los organismos. “Provisionalmente comprometidos” quiere decir que estamos comprometidos en el mismo sentido en que nos comprometemos con atribuir procesos cognitivos a organismos y tan lejos como tomamos en serio dichas teorías y son actualmente aceptables.
5. Es una meta razonable de investigación tratar de caracterizar el sistema representacional con el que estamos provisionalmente comprometidos.
6. Es una estrategia razonable de investigación tratar de inferir dicha caracterización a partir de teorías psicológicas que parecen demostrar ser verdaderas. (Fodor, 1975, p. 27)²⁷

Los numerales 1-6 parecen ser un argumento a la mejor explicación donde las premisas también son de carácter empírico. (1.) puede verse como la tesis de la modularidad presentada en la sección anterior cuya conclusión o implicación es (2.). La premisa (3.) revela el carácter empírico del argumento; tal como parece sugerir la evidencia —es decir, la conducta inteligente—. Lo que nos compromete con (4.), es decir, debemos aceptar que la teoría propuesta en (1.) no solo es plausible, sino que probablemente es la forma correcta de describir y explicar los procesos cognitivos de los organismos. Lo que nos obliga a ofrecer una mejor caracterización del sistema representacional que probablemente figure en las computaciones de los organismos —es decir, (5.)—. Donde esta caracterización puede ser inferida a partir de las teorías psicológicas que parecen ser verdaderas, a saber, el sistema representacional debe ser compatible con la computacionalidad y la *modularidad de la mente*.

²⁷ Tanto estos numerales como los siguientes (7-15) funcionan como el esquema de una teoría. Por supuesto debe darse un valor a las variables para obtener la predicción y explicación de un episodio conductual en una ocasión determinada.

El argumento señala la necesidad de un sistema representacional y su caracterización a partir de otras teorías plausiblemente verdaderas. Se sigue entonces que las computaciones que operan sobre el sistema representacional son las causas de la conducta de los sujetos.

For according to the model, deciding is a computational process; the act the agent performs is the consequences of computations defined over representations of possible actions. No representations, no computations. No computations, no model. (Fodor, 1975, p. 31)²⁸

La clave radica, según Fodor, en que la explicación de la conducta suele partir de las mismas creencias que los organismos tienen respecto a lo que hacen o al cómo se comportan. Así, la explicación conductual toma su forma advirtiendo las creencias que los organismos tienen sobre lo que hacen —y sobre lo que los demás hacen— (Fodor, 1975)²⁹.

Si aceptamos la intuición de Fodor, el siguiente modelo puede ser aproximadamente plausible sobre la forma en que una conducta es decidida por los sujetos que la producen:

7. Un agente se encuentra a sí mismo en una situación S.
8. El agente cree que un cierto grupo de conductas ($B_1, B_2, B_3 \dots B_n$) están disponibles para él en S. De esta manera, dado S, B_1 hasta B_n son las cosas que el agente cree que puede hacer.
9. La probable consecuencia de llevar a cabo B_1 hasta B_n es predicha. Así, el agente computa un grupo de situaciones hipotéticas de la forma “si B_1 es llevada a cabo en S, entonces, probablemente, C_1 . Las situaciones hipotéticas son computadas y las probabilidades son asignadas de acuerdo con, y dependiendo de lo que el organismo sabe o cree sobre situaciones como S.
10. Un orden de preferencia es asignado a las consecuencias.
11. La elección de conducta del agente es determinada como una función de sus preferencias y las probabilidades asignadas. (Fodor, 1975, pp. 28-29)

Los numerales 7-11 representan un esquema *idealizado* de lo que un sujeto plausiblemente tiene en cuenta para actuar. En términos de la psicología popular, de forma idealizada, cuando un sujeto cree que se encuentra en una determinada situación cree que tiene cierto número de conductas disponibles y actúa según la que ha considerado o valorado como más deseable según sus consecuencias. Tal como la premisa (9.) sugiere, la propuesta

²⁸ Es importante mencionar que, si una parte central de varios intentos de explicación psicológica parten de la idea de que los sujetos u organismos son capaces de considerar su conducta como perteneciente a cierta clase o tipo, entonces, ya sea implícita o explícitamente, dichos intentos de explicación deben reconocer un sistema de representaciones que posee el sujeto dueño de la conducta que le permita organizarla de la manera en que lo hace.

²⁹ De ahí que se tome como punto de partida la psicología popular. Si se toma la psicología popular como el conjunto de creencias que tienen los sujetos sobre la forma en que se caracteriza la misma conducta que producen, no es de extrañar que la explicación psicológica deba también explicar cómo es que los sujetos son capaces de representarla como lo hacen.

de Fodor es describir —y explicar— el proceso como computacional. De esta manera, la conducta final realizada por el sujeto puede verse como el producto de algoritmos aplicados sobre variables que toman el lugar de las situaciones, conducta y valoraciones que el sujeto proyecta.

El esquema nos permite saber aproximadamente de qué manera y en qué sentido las conductas son funciones de las creencias y preferencias del agente, es decir, cómo están conectadas y cómo el sujeto se las representa. En parte, argumenta Fodor, aceptar que este modelo se aplica a un sinnúmero de casos nos compromete con el tipo de explicaciones que ofrece. Por ejemplo, dado el hecho de que un organismo *a* produjo una conducta *B*, la explicación podría ser algo como:

12. El organismo *a* creía que se encontraba en una situación *S*.
13. *a* creía que producir una conducta del tipo B_i en *S* probablemente conduciría a una consecuencia C_i .
14. C_i era una (o la) consecuencia mejor valorada por *a*.
15. *a* creyó e intentó *B* para ser una conducta del tipo B_i . (Fodor, 1975, p. 30)

La explicación que va de 12-15 y que se deriva de los numerales anteriores es ciertamente plausible y de hecho parece una explicación precisa y acertada de por qué *a* hizo *B*. Sin embargo, esto descansa en el punto central que ya se planteó anteriormente, a saber, este esquema de teoría y sus explicaciones derivadas solo son efectivas si el sujeto se representa su conducta como siendo de cierta clase o tipo:

The point to notice is that it is built into this pattern of explanation that agents sometimes take their behavior to be of a certain kind; in the present case, it is part of the explanation of *a*'s behavior that he believed it to be of the B_i kind, since it is behavior of that kind for which highly valued consequences are predicted. To put it briefly, the explanation fails to be a (full) explanation of *a*'s behavior unless that behavior was B_i and *a* believed it to be so. (Fodor, 1975, p. 30)

Por otro lado, comprometerse con que los sujetos se representan su propia conducta implica aceptar que el organismo posee un medio representacional o los medios para hacerlo:

The moral I want to draw, then, is that certain kinds of very central patterns of psychological explanation presuppose the availability, to the behaving organism, of some sort of representational system. I have emphasized, for purposes of exposition, the significance of the organism's representation of its own behavior in the explanation of its considered actions. But, once made, the point is seen to be ubiquitous. It was, for example, implicit in the model that organism has available means for representing not only its behavioral options but also: the probable consequence of acting on those options, a preference ordering defined over those consequences and, of course, the original situation in which it finds itself. To use this

sort of model is then to presuppose that the agent has access to a representational system of very considerable richness. (Fodor, 1975, p. 31)

Otro caso en el que podría verse la utilidad (y necesidad) de un sistema representacional es el caso de la atribución de actitudes intencionales. El sujeto no solo representa y clasifica su propia conducta al racionalizarla —y explicarla—, sino que también representa la conducta e intenciones de otros sujetos al racionalizarlas y explicarlas. Esto implica que el sujeto tiene acceso a un sistema representacional lo suficientemente rico como para albergar y significar cuidadosamente cada una de las situaciones, posibles conductas y consecuencias³⁰.

La intuición que entra en juego ahora es que los sujetos al representarse su conducta lo hacen como siendo de cierta manera, a saber, usan una representación con una estructura proposicional. Como se mencionó en la sección de la modularidad, la estructura de la conducta puede verse como derivada de sus causas, por lo que es razonable suponer que el sistema representacional mantiene una estructura similar. Un ejemplo que puede refrescar la memoria y entablar una relación más clara y directa con acciones es el caso de bailar mientras se canta una canción (presentado en el capítulo I, sección “Las acciones estructuradas” de este documento). A partir de las intuiciones de la psicología popular, describimos este tipo de eventos como acciones que ocurren en simultáneo. Esto, desde luego, no es obvio, bien podría tratarse de una sola acción completamente diferente a otras, podríamos llamarla “bailo-cantar”, y ser esta acción completamente diferente a “bailar” y “cantar”. La intuición, entonces, de que las acciones pueden agruparse para dar lugar a otras más complejas recae en la idea de que los simples mantienen su identidad y esta forma de estructurarse es lo que hace posible conductas como la lingüística, que, por excelencia, parece ser la conducta compuesta más compleja y sofisticada.

Finalmente, hay otro elemento que funciona como requisito y motivación para un sistema representacional. Se ha hablado de los estados mentales como las creencias y los deseos, y las intenciones como las causantes de la conducta. No obstante, no se ha dado una explicación clara respecto a cómo estas entidades mentales pueden causar o afectar objetos físicos.

³⁰ Si bien aún no se ha dicho en qué sentido o cómo las representaciones son significativas el argumento se mantiene si cambiamos la palabra ‘significados’ por ‘valores’; valores que son relevantes para el organismo.

El sistema representacional puede —y probablemente debe— cumplir con esta tarea. La noción de representación que se presentará a continuación provee una explicación de cómo lo anterior es posible. Es decir, cómo las creencias, deseos e intenciones (entidades psicológicas en general) logran obtener poderes causales sobre el mundo físico. La respuesta, como se mostrará, consiste en que las representaciones son esencialmente físicas y plausiblemente sensibles a propiedades sintácticas. En esta medida, las representaciones cumplen con mantener los poderes causales que los contenidos de las intenciones mantienen semánticamente. En resumidas cuentas, las representaciones ofrecen una respuesta a cómo las entidades psicológicas pueden tener efectos y ser parte del mundo físico³¹.

El sistema representacional, por tanto, debe cumplir con estas exigencias. De ser así, el sistema cognitivo posee la suficiente versatilidad para cumplir con la integración constante de nueva información ambiental y ser lo suficientemente robusto para mantener las propiedades semánticas a partir de su sintaxis³². Por supuesto, esto exige que las representaciones de la conducta sean también de carácter proposicional, dado que la conducta humana está organizada de esta misma manera siguiendo a Fodor y las intuiciones de la psicología popular³³.

Podríamos decir, entonces, que hasta ahora Fodor se compromete con: 1. Si aceptamos la teoría computacional y modular de la mente, debemos aceptar que los principales procesos cognitivos deben ser transformaciones de representaciones, dichas representaciones deben ser parte de un sistema de una naturaleza particular. 2. El sistema representacional es de carácter proposicional porque, 3. la conducta está organizada o estructurada proposicionalmente en la medida en que el sujeto se la representa de esta manera.

Ya se ha argumentado brevemente a favor de los tres numerales, más bien, se ha argumentado a favor de su plausibilidad. Sin embargo, en lo que sigue, presentaré con más

³¹ Este tema es de vital importancia en la pretensión de hacer de la psicología popular una teoría científica. Después de todo, las ciencias toman como verdadera la afirmación de que la única substancia o naturaleza de la que está compuesta el mundo es de carácter material. Así las cosas, si no es posible ofrecer una respuesta a cómo es posible que entidades psicológicas puedan pertenecer al mundo material, es difícil divisar un escenario en el que las entidades psicológicas hagan parte de una ciencia de la conducta o de la mente.

³² Sin olvidar aún que dicho sistema representacional debe dar cuenta de las propiedades que exhibe la conducta como lo son la sistematicidad y la composicionalidad.

³³ Esta idea puede ser confusa como esta presentada. Es cierto que, si aceptamos un sistema representacional proposicional, debemos, por extensión, aceptar que los outputs del sistema también son proposicionales. Sin embargo, la justificación es en sentido contrario: dado que la conducta inteligente es de carácter composicional (Fodor está pensando principalmente en nuestra conducta lingüística) es que nos vemos en la necesidad de apelar a una explicación (y por tanto a un sistema representacional) que dé cuenta de la composicionalidad y proposicionalidad de la conducta.

detalles cómo este conjunto de hipótesis *fodorianas* logran dar cuenta de la psicología popular y, *a fortiori*, de la conducta inteligente (principalmente, aunque no exclusivamente la conducta lingüística).

Computación y Semántica

Hasta el momento, todo el peso de la explicación fodoriana ha recaído en la hipótesis de la modularidad y, sobre todo, en el sistema representacional. A continuación, ofreceré más detalles sobre este sistema representacional, cómo funciona su explicación y cómo se conecta con las exigencias de la psicología popular.

Fodor, tras aceptar que el sistema representacional es proposicional y cercano a un lenguaje lo llama ‘lenguaje de pensamiento’ (*Language of Thought - LOT*). A partir de lo que Susan Schneider (2011) expone en su libro *The Language of Thought – A New Philosophical Direction*, se pueden establecer tres afirmaciones claves y básicas en LOT:

1. Los procesos cognitivos consisten en secuencias causales de *tokens*³⁴ de representaciones en el cerebro.
2. Estas representaciones internas tienen una sintaxis y semántica combinatoria, y, además, la manipulación de símbolos preserva las propiedades semánticas de los pensamientos:
 - a. Sintaxis combinatoria: Las representaciones complejas de LOT están constituidas de símbolos atómicos, junto con la gramática de LOT.
 - b. Semántica combinatoria: El significado o contenido de las oraciones de LOT es una función de los significados de los símbolos atómicos, junto con su gramática.

³⁴ Es importante mencionar que la teoría apela a una distinción *type/token*, donde los *tokens* son las instancias particulares que pertenecen a una misma clase *type*. Las representaciones de LOT sobre las que aquí se discute, son representaciones-tipo (*type*), sin embargo, la teoría cuantifica de esta manera sobre las instancias particulares en la cabeza de cada sujeto individual. Un ejemplo intuitivo puede ser el caso del ajedrez:

“Ultimately, the rules are what determine the types of the *tokens*. The rules specify what moves would be legal in what positions. If interchanging two particular *tokens* wouldn’t make any difference to which moves were legal in any position, then it couldn’t make any difference to the game at all; and those *tokens* would be of the same type. To see the significance of this, consider chess again. In some fancy chess sets, every piece is unique; each white pawn, for instance, is a little figurine, slightly different from the others. Why then are they all the same type? Because, in any position whatsoever, if you interchanged any two of them, exactly the same moves would be legal. That is, each of them contributes to an overall position in exactly the same way, namely, in the way that pawns do. And that’s what makes them all pawns. (Haugeland 1989, 52)” (Schneider, 2011, p. 114).

- c. Pensar, entendido como una forma de manipulación de símbolos, preserva las propiedades semánticas de los pensamientos involucrados.
3. Las operaciones mentales en las representaciones internas son causalmente sensitivas a la estructura sintácticas de los símbolos³⁵.

En este punto podemos establecer que cuando un sujeto se encuentra en un estado de creencia, por ejemplo, equivale a decir que se encuentra en un estado informacional tal que en su '*caja de creencias*'³⁶ están activas representaciones que son instancias de cierta clase de representaciones *tipo*, estas representaciones están dispuestas en relaciones causales determinadas tanto con la información entrante al sistema y la salida conductual —junto con otras representaciones—. Ha de añadirse que es justamente este tipo de conexiones causales lo que determina la naturaleza de un estado mental a nivel computacional, es decir, lo que diferencia un deseo de una creencia (y así con cualquier otro ejemplo de actitud proposicional); no es el contenido de la representación, sino lo que causa dicha representación y el estado funcional del sistema. Si el estado ha sido debidamente causado, es de esperar que preserve su adecuadamente su semántica, esto es, el estado mental individuado por su contenido³⁷.

Ahora bien, estas representaciones son objetos caracterizados mediante sus propiedades sintácticas y, aun así, pueden ser evaluados semánticamente. En este sentido, el mejor sistema que conocemos que satisface estas exigencias es el lenguaje natural. Por tal razón, Fodor apela a un lenguaje del pensamiento, un sistema representacional lo suficientemente complejo para expresar una gran cantidad de contenidos (potencialmente infinitos). Y, sin embargo, es recursivo, es decir, una vez clarificados los elementos más básicos del sistema y las reglas de combinación entre ellos, puede darse cuenta de representaciones cada vez más complejas. ¿Cómo funciona y cómo se dan las relaciones entre los diferentes componentes de este lenguaje de la mente?

³⁵ Los tres ítems y sus sub-ítems son tomados de (Schneider, 2011, p. 12).

³⁶ 'Caja de creencias' es una expresión tomada de Fodor (1987) para referirse al estado funcional que corresponde a una creencia, deseo, o cualquier otra actitud proposicional. Dado que uno puede tener diferentes actitudes hacia el contenido que expresan, hablar de cajas de creencias o cajas de lo que sea, solo se refiere a la actitud que se mantiene hacia dicho contenido, en dicho momento.

³⁷ Esto puede ser confuso, lo que se quiere decir es que la descripción relevante para las generalizaciones de la psicología popular se da a nivel semántico, es decir, los estados mentales se individuán por sus contenidos. Aun así, estos estados deben tener un correlato computacional en cada sujeto, dicho correlato computacional de los estados mentales, en el sistema representacional, se individua por sus propiedades sintácticas y causales.

Lo primero que hay que advertir es que, si bien dicho sistema representacional funciona a través de sus propiedades sintácticas³⁸, las generalizaciones de la psicología popular, las que capturan la conducta y las intenciones de los sujetos, son de orden semántico. Esto establece que las generalizaciones que buscamos se dan en virtud de los contenidos de las representaciones. En este punto, lo que se espera es que el sistema sintáctico mapee correctamente las relaciones que se dan entre los contenidos de tal forma que los estados mentales mantienen tanto las propiedades sintácticas de sus representaciones, como el contenido que los distingue y los individúa: “There is no tokening of a (cognitive) mental state or process (by a creature, at a time) unless there is a corresponding tokening of a mental representation (by that creature, at that time)” (Fodor, 2008, pp. 6).

En este sentido, las relaciones semánticas entre los diferentes estados psicológicos logran ser capturadas a través de cadenas causales de operaciones sobre representaciones que se mapean con los estados mentales del sujeto. Sin embargo, no hay identidad entre las propiedades informacionales y las propiedades computacionales y ninguna clase de propiedades es condición necesaria de la otra³⁹. Esto se debe a que las propiedades computacionales son esencialmente internas, es decir, dependen enteramente del sistema al que pertenecen y no pueden darse fuera del sistema. Por otro lado, las propiedades informacionales o las propiedades del contenido de las representaciones son esencialmente externas al sistema, es decir, no dependen del sujeto y dependen enteramente de cómo es el mundo.

¿Cómo es posible que las propiedades informacionales o semánticas y las relaciones entre ellas puedan ser mapeadas por un sistema representacional cuyas propiedades son esencialmente internas? O puesto en los términos relevantes para la investigación ¿cómo es posible que las leyes psicológicas que rigen sobre el contenido semántico puedan ser capturadas por propiedades sintácticas de un sistema representacional interno de los sujetos⁴⁰?

³⁸ En este sentido, es necesario que los estados mentales como las creencias y los deseos, el procesamiento de información que se da a través de representaciones tenga una implementación física que soporte las relaciones causales pertinentes para la explicación. En virtud de esta exigencia, se revela una importante característica de LOT, a saber, su implementación es física y esto asegura los poderes causales necesarios para que ocurran cambios en el mundo físico que garantizan la explicación.

³⁹ A grandes rasgos, esto significa que no hay identificación de ningún tipo entre los contenidos y las representaciones computacionales. *A fortiori* no hay reducciones de ningún tipo.

⁴⁰ Recordemos que la idea central de la psicología popular —y que Fodor comparte— es que lo que hace un sujeto y su explicación implican a apelar a lo que el sujeto cree, desea, conoce, etc., en general, apelar a

Fodor resuelve esta aparente tensión haciendo una distinción entre lo que suele llamarse ‘contenido amplio (*broad content*)’ y ‘contenido estrecho (*narrow content*)’. Por un lado, el contenido estrecho es todo el nivel computacional que se ha descrito en las últimas secciones. Cuando se dice que un elemento *x* pertenece o hace parte del contenido estrecho no es diferente a decir que juega un papel dentro de las computaciones del sistema. Decir, por ejemplo, que una creencia tiene contenido estrecho, es referirse al estado funcional y computacional en el que el sistema se encuentra cuando cree en algo.

Por otro lado, el contenido amplio comprende la parte semántica de la explicación y es lo que es capturado por las generalizaciones de la psicología popular⁴¹. El contenido amplio, como mencionamos antes, está determinado por el mundo y son la clase de elementos de los cuales se predica verdad, significado, referencia, etc., es decir, es el contenido de los estados mentales. Por ejemplo, la creencia de que “El gato está sobre la mesa” es la relación que tiene un sujeto con un estado de cosas o una proposición; es una relación con el mundo. Y, en este sentido las propiedades del contenido son ajenas al sujeto y dependen, más bien, de que la relación que el sujeto mantiene con dicho estado de cosas sea la adecuada, a saber, esté debidamente causada por el mundo.

Así las cosas, LOT juega un doble papel en la explicación. El contenido de las representaciones captura un nivel puramente semántico (y quizá el más mentalista según la psicología popular). En dicho nivel se capturan las intenciones del sujeto y con ello mismo el contenido de las creencias, deseos, miedos, etc., que explican no solo causalmente, sino semánticamente la conducta del sujeto. Por otro lado, LOT captura dicho contenido semántico a través de un sistema representacional de naturaleza proposicional, y sobre todo computacional. La transformación de símbolos se lleva a cabo mediante las propiedades sintácticas de las representaciones, y dichas representaciones preservan los poderes causales que dan lugar a otras representaciones en el proceso computacional, o a la conducta inteligente.

contenidos mentales. Es por ello, que debe contarse también una historia que conecte el aspecto semántico antes mencionado con los aspectos sintácticos, internos de la explicación.

⁴¹ En caso de que no se halla mencionado con suficiente claridad anteriormente, los contenidos de los estados mentales son las entidades relevantes que explican la conducta intencional y lingüística. En este sentido, es de interés para la psicología que las generalizaciones relevantes cuantifiquen sobre estas entidades. En este punto podemos encontrar un viejo problema conocido como “casos de tierras gemelas” o “casos de gemelos”. Las primeras versiones de este problema son presentadas por Putnam (1975). En resumidas cuentas, el ejemplo muestra que no es suficiente la parte semántica para capturar todos los matices de las generalizaciones psicológicas que pueden establecerse. No es mi intención ofrecer una respuesta o posible defensa de la teoría fodoriana frente a este problema, para ello puede consultarse Fodor (1987). Sin embargo, en el capítulo IV volveré sobre este problema y una posible respuesta para defender la psicología popular.

Como mencioné anteriormente, el contenido de las representaciones es de carácter externo, depende enteramente de la relación del sujeto con el mundo. Así que el contenido de las representaciones se corresponda con el estado de cosas depende de que las representaciones del sujeto sean correctamente causadas por el mundo. Es aquí donde los dos contenidos se unen. Dado que el mundo causa ciertos contenidos, el sujeto se los representa de una determinada manera. La representación que el sujeto se forma depende de la relación causal que hay entre el mundo y la representación generada. Dado que la representación captura el contenido semántico en el mundo, el sistema genera una representación particular que preserva los poderes causales relevantes para el proceso computacional⁴². Dichos poderes causales dependen de las propiedades de la representación y esas propiedades están sujetas a la relación causal del mundo con el sujeto. En este orden, a un contenido le corresponde una cierta representación y no es posible que haya una misma representación para dos contenidos diferentes.

En resumen, la relación con el mundo hace que el sistema genere representaciones y dichas representaciones mantienen su doble papel representando el contenido semántico y cumpliendo un papel computacional dentro del sistema. Adicionalmente, cuando las relaciones causales con el mundo son adecuadas, las propiedades computacionales se mapean adecuadamente con las propiedades informacionales, manteniendo así ambos aspectos del contenido (tanto el amplio como el estrecho) sin tensiones entre ellos. Un ejemplo puede ilustrar mejor el punto:

The basic idea of the STM (Syntactic theory of the mind) is that the cognitive states whose interaction is (in part) responsible for behavior can be systematically mapped to abstract syntactic objects in such a way that causal interactions among cognitive states, as well as causal links with stimuli and behavioral events, can be described in terms of the syntactic properties and relations of the abstract objects to which the cognitive states are mapped. More briefly, the idea is that causal relations among cognitive states mirror formal relations among syntactic objects. If this is right, then it will be natural to view cognitive state tokens as tokens of abstract syntactic objects. (Stich, 1983, p. 151)

Finalmente, todo lo mencionado hasta este punto nos brinda los insumos para responder a lo que podría ser el reto central de una teoría psicológica —desde el punto de vista de la psicología popular, al menos—. En las últimas secciones del primer capítulo se

⁴² Este es plausiblemente el proceso que se lleva a cabo dentro de los módulos. Transducir y procesar la información no es otra cosa que crear las representaciones respectivas, en el lenguaje de la máquina, para la información que ingresa al sistema; y así pueda ser manipulada por el procesador central y tomar parte en los procesos cognitivos.

mencionó la composicionalidad de la conducta inteligente. ¿Cómo el lenguaje de pensamiento da explicación a esto?

La respuesta: LOT también es composicional. El hecho de que el sistema representacional sea de carácter proposicional nos ayuda a dar cuenta de la estructuración necesaria para dar cuenta de la composicionalidad de la conducta lingüística e inteligente. Nótese que las representaciones de LOT permiten la combinación de las representaciones simples en complejas a través de reglas y su posición sintáctica en la oración. Veamos el siguiente ejemplo: Si una persona enuncia la oración “El gato está sobre la mesa”, probablemente su conducta pueda ser explicada por la intención de expresar su creencia de “El gato está sobre la mesa”. Ahora sabemos, por la hipótesis del lenguaje de pensamiento, que las creencias de los sujetos deben tener un correlato computacional que preserva los poderes causales necesarios en la producción de la conducta. Sabemos también que las representaciones sirven al sistema —y al sujeto— para representarse su ambiente, por lo que sería razonable suponer que las representaciones involucradas en este caso son el fruto de la interacción del sujeto con el mundo. En este sentido, el sujeto tiene representaciones como ‘GATO’, ‘PERRO’, ‘ESTAR DEBAJO’, ‘ESTAR SOBRE’, ‘MESA’, etc. En el proceso de la producción de la oración ‘El gato está sobre la mesa’ plausiblemente fueron seleccionadas las representaciones ‘GATO’ – ‘ESTAR SOBRE’ – ‘MESA’. Luego de ser tomadas las representaciones pertinentes y ser ubicadas sintácticamente en el lugar correspondiente, el sujeto plausiblemente formó la representación compleja ‘EL GATO ESTÁ SOBRE LA MESA’ constituida por los elementos simples antes mencionados⁴³.

Volviendo al caso presentado en la sección “las acciones estructuradas” en el capítulo I. Casos de acciones como bailar y cantar deberían recibir la misma explicación, es decir, las representaciones involucradas “CANTAR” y BAILAR” deberían figurar en el proceso de deseo o creencia que explican el realizar la acción respectiva. Nótese que las representaciones funcionan como elementos básicos que mantienen su identidad del contexto. Es posible que la misma representación de “CANTAR” se vea involucrada en la producción de acciones, cuando, por ejemplo, se canta y no se baila o se realiza otra acción diferente.

⁴³ Esta es, por supuesto, una descripción muy general del proceso que probablemente se lleve a cabo. Después de todo, la descripción de este proceso también puede incluir diversos estados y procesos cerebrales que especifiquen cómo las representaciones son recuperadas de la memoria —o su lugar de almacenamiento—; para posteriormente generar las respuestas motoras apropiadas. Sin embargo, el ejemplo permite mostrar cómo las representaciones pueden agruparse para dar lugar a representaciones más complejas; y cómo las propiedades sintácticas de las representaciones pueden salvaguardar las debidas relaciones semánticas.

Finalmente, las diferentes representaciones y el estado funcional de creencia (o deseo, etc.) del sujeto, generan la conducta de enunciar la oración "El gato está sobre la mesa" (o la acción de bailar, la acción de cantar, etc.). Esto nos garantiza que son los mismos contenidos o conceptos que el sujeto sabe o conoce y que aparecen en diferentes episodios conductuales que involucren la producción de otras oraciones acerca de gatos o mesas, etc.⁴⁴

* * *

Aquí concluye la presentación de la teoría fodoriana y su explicación de la psicología popular y la conducta intencional mostrando además la clase de arquitectura mental que hace falta postular para que la explicación sea efectiva. Ya desde el primer capítulo se ha mencionado que la psicología popular —y por ende la explicación que Fodor hace de esta— no es obvia ni contundente. ¿Qué pasaría si pudiera ofrecerse una descripción de la conducta radicalmente diferente y que ello amerite una arquitectura mental diferente? ¿Sería la psicología popular y el aparato teórico de Fodor lo suficientemente robusto para mantenerse aún como la mejor explicación de la que disponemos?

Las respuestas a estas preguntas pueden ser variadas y hechas desde un amplio abanico de posturas. A continuación, en el capítulo siguiente se tomará la teoría postulada por Paul Churchland respecto a la conducta y la arquitectura cognitiva. Se revisará someramente en qué consiste su propuesta y luego algunos argumentos en contra de la psicología popular y la hipótesis de Fodor. Así, luego de presentar algunas respuestas y contra-respuestas en el debate Fodor-Churchland, presentaré algunos argumentos de Lenny Clapp en contra de la teoría fodoriana, en particular contra la composicionalidad de la conducta lingüística e inteligente, y sus consecuencias para nuestra versión de la composicionalidad de la mente. Finalmente, presentaré una pequeña conclusión de la discusión que será la forma en cómo estos argumentos afectan a la teoría fodoriana y cómo afectan la explicación, esto con el fin de analizar nuevamente la plausibilidad de la psicología popular como teoría y la clase de arquitectura mental a la que debería apelarse para responder a los argumentos de Clapp.

⁴⁴ Esto nos ofrece una explicación recursiva de cómo el sujeto es capaz de construir una cantidad potencialmente infinita de oraciones, a partir de conceptos o elementos simples finitos: "Once one knows how to generate a particular sentence out of primitive expressions, one can also generate many others that have the same primitives" (Fodor 1987; Fodor and Pylyshyn 1988; Fodor and McLaughlin 1990)" (Schneider, 2011, pp. 10).

Capítulo III

Un ataque a esta descripción de la conducta inteligente: la discusión entre Fodor, Churchland y Clapp

En los capítulos anteriores se presentó a la psicología popular como una teoría explicativa de la conducta inteligente. Adicionalmente, al presentar las tesis de la *modularidad* y el *lenguaje de pensamiento* se ofreció una explicación más robusta de por qué la psicología popular parece funcionar como lo hace; cómo sus generalizaciones pueden ser de carácter científico. Y, no menos importante, estas hipótesis nos comprometen con una arquitectura mental particular de carácter funcional, computacional y proposicional.

En este capítulo pretendo poner en tela de juicio ambas explicaciones ofreciendo razones para pensar que la psicología popular, y por tanto la arquitectura cognitiva tipo Fodor no es suficiente para explicar todos los rasgos de la conducta. Este capítulo tendrá, por tanto, tres partes: La primera consta de dos pasos: uno crítico y el otro constructivo. El primer paso consiste en la presentación —de manos de Paul Churchland— de algunas intuiciones que nos pueden hacer dudar de la efectividad de la psicología popular para explicar la conducta inteligente. Estas intuiciones consisten, principalmente, en señalar aspectos relativamente básicos e importantes de la conducta inteligente que permanecen sin una explicación adecuada por parte de la psicología popular o las tesis fodorianas. Luego, el paso constructivo consiste en la presentación de una descripción diferente de la conducta humana y, por lo tanto, en la postulación de una arquitectura cognitiva que se ajusta mejor a esta forma de descripción de la conducta, estos mecanismos recibirán el nombre de *mapas de estado-espacio*.

La segunda parte del capítulo consistirá en un breve debate entre Fodor y Churchland. Fodor señala que la explicación churchlandiana no es clara respecto a cómo los mapas de estado-espacio están dotados de contenido semántico y respecto a cómo estos mapas pueden garantizar que dicho contenido es el mismo a lo largo de los diferentes individuos. Lo que implica una falla explicativa y predictiva de la conducta inteligente. Finalmente, presentaré el argumento final de Fodor en contra de Churchland, según el cual su teoría deja por fuera uno de los aspectos centrales de la conducta inteligente mencionado en el capítulo I de esta tesis: la composicionalidad. De forma que la teoría propuesta por Paul Churchland es insuficiente para dar cuenta de la conducta inteligente que es lo que nos interesa en esta investigación.

Por último, basado en la discusión de Lenny Clapp (2012), en la tercera parte de este capítulo pretendo reconstruir un argumento en contra de Fodor: el argumento apuntará a mostrar que la discusión no es acerca de cuál es la mejor arquitectura cognitiva que permite dar cuenta de la composicionalidad de la conducta inteligente, sino que esta descripción de la conducta inteligente no es adecuada: hay razones para pensar que esta clase de conducta no es composicional. Por lo tanto, no tenemos buenas razones para sostener una arquitectura mental como la sostenida por Fodor (con un sistema central no-modular y LOT). Al final, señalaré las implicaciones y alcances de este argumento para pensar el proyecto de la psicología popular como el marco para la explicación de la conducta inteligente.

Primera parte

Churchland en contra de la psicología popular

La psicología popular ofrece una primera explicación de la conducta intencional e inteligente de los sujetos. El propósito de centrarse especialmente en el aspecto intencional de la conducta —teniendo como caso paradigmático la conducta lingüística— es ofrecer una caracterización y racionalización de la conducta en general (Ver primera parte del capítulo I). Sin embargo, no todos aceptan que la conducta inteligente se reduce a la conducta intencional, pues argumentan que ésta abarca algo mucho más básico que aquella y, a partir de esto, consideran que la psicología popular no es adecuada para dar cuenta de la conducta inteligente en general.

Uno de los autores que sigue este camino es Paul Churchland (1981). De acuerdo con Churchland, hay varias razones para afirmar que la psicología popular es una explicación insuficiente de la conducta inteligente en general.⁴⁵ La razón más pertinente para esta discusión tiene que ver con la capacidad explicativa de la psicología popular de fenómenos psicológicos fundamentales como la *coordinación sensoriomotora*:

Dentro de la temática perceptual se encuentra un viejo problema que no es tan conocido en el marco de la filosofía, pero es sumamente relevante en la psicología y la

⁴⁵ En parte porque considera que la psicología popular no es del todo coherente con disciplinas cercanas como la neurociencia o la biología: “An account of cognition that locates us on a continuum with all of our evolutionary brothers and sisters is thus a prime desideratum of any responsible epistemological theory. And the price we have to pay to meet it is to give up the linguaformal “judgment” or “proposition” as the presumed unit of knowledge or representation.” (Churchland, 2002, p. 28)

neurobiología, a saber, el problema de la coordinación sensoriomotora. La psicología popular solamente explica cómo la conducta intencional y las acciones se realizan a partir de intenciones, pero no ofrece una explicación plausible de cómo la percepción participa de la producción de la conducta o cómo el cerebro realiza satisfactoriamente estos movimientos. Piense, por ejemplo, en el movimiento que se realiza al esquivar un objeto. Bien puede ser cierto que lo que explica la acción general de moverse de la trayectoria del objeto es la intención de esquivarlo, pero es difícil ver cómo, a partir de deseos y creencias, se puede explicar la coordinación sensoriomotora entre la captación de la trayectoria del objeto y el movimiento de músculos y el cuerpo en general para tener éxito en la evasión.⁴⁶

Según Churchland, esta falla revela que desde la psicología popular es difícil formular una explicación plausible, apelando a deseos y creencias, que dé cuenta de fenómenos fundamentales vinculados con la percepción. En parte, porque según las intuiciones de la psicología popular parece que parte de estos fenómenos ocurren de manera sub-personal o son contrarios a los deseos y necesidades de los sujetos. Por ejemplo, es difícil pensar en una posible explicación de la construcción de imágenes 3D a partir de la irritación sensorial de la retina en términos de lo que el sujeto cree o desea, pues parecen ser fenómenos que no se ven afectados por las intenciones del sujeto.

Sin embargo, la coordinación sensoriomotora es un fenómeno que ocurre a un nivel personal de descripción de los sujetos y organismos en general. Que esté en este nivel es parte de lo que explica la supervivencia de los sujetos y organismos al navegar el ambiente. La percepción y la coordinación sensoriomotora juegan un rol fundamental, pues permiten a los organismos obtener información relevante de su entorno y posteriormente *actuar intencionalmente* acorde a ello. Todo esto, por supuesto, pasa en tiempo real, por lo que el

⁴⁶ La lista de elementos que requieren de mayor profundidad explicativa puede seguir. Pero no es mi intención tomar en consideración todos estos aspectos en el presente texto. No obstante, para tener presente la profundidad de este reclamo léase la siguiente cita de Churchland:

When one centers one's attention not on what FP can explain, but on what it cannot explain or fails even to address, one discovers; that there is a very great deal. As examples of central and important mental phenomena that remain largely or wholly mysterious within the framework of FP, consider the nature and dynamics of mental illness, the faculty of creative imagination, or the ground of intelligence differences between individuals. Consider our utter ignorance of the nature and psychological functions of sleep, that curious state in which a third of one's life is spent. Reflect on the common ability to catch an outfield fly ball on the run, or hit a moving car with a snowball. Consider the internal construction of a 3D visual image from subtle differences in the 2D array of stimulations in our respective retinas. Consider the rich variety of perceptual illusions, visual and otherwise. Or consider the miracle of memory, with its lightning capacity for relevant retrieval. On these and many other mental phenomena, FP sheds negligible light. (Churchland, 1981, p. 73).

sujeto se está enfrentando continuamente a situaciones que requieren solucionar problemas o tomar decisiones, etc.:

One of them is a problem relatively unfamiliar to philosophers: the problem of sensorimotor coordination. However unfamiliar it may be, this problem is of fundamental importance to cognitive theory *since the administration of appropriate behavior in the light of current experience is where intelligence has its raw beginnings*. Since sensorimotor coordination is the most fundamental problem that any animal must solve, a means of solution coupled to a general account of representation must surely arouse our curiosity. (Churchland, 1986, p. 279, énfasis añadido)

Churchland no tiene en mente una reducción o alguna clase de teoría mixta que dé lugar a la psicología popular. Después de todo, aceptarla conlleva a aceptar entidades que la física, la química, la biología e incluso la neurociencia no pueden dar cuenta, como lo son las intenciones, los deseos y las creencias. La agrupación de los vacíos explicativos de la psicología popular como la coordinación sensoriomotora, añadidos al problema de articular las entidades de la psicología popular con las entidades de las ciencias naturales, nos dan razones para dudar de la efectividad y la plausibilidad de la psicología popular como la más adecuada explicación de la conducta y la psicología de los organismos.

Por esta razón, Churchland propone eliminar completamente las entidades propuestas por la psicología popular (esto es, las actitudes proposicionales). Así, con miras a dar cuenta de las habilidades cognitivas que permiten la supervivencia y adaptabilidad en el ambiente, Churchland sugiere apelar a toda una *explicación diferente* de la conducta inteligente y, por lo tanto, una *arquitectura cognitiva diferente*. Veamos de manera breve en qué consiste esta explicación y por qué resulta prometedora.

Cómo funciona la explicación

Los ataques a la psicología popular no serían igual de efectivos sin una teoría que la remplace. Por ello, hace falta ofrecer una teoría diferente que dé cuenta de los mismos fenómenos que explica la psicología popular y que llene sus vacíos explicativos. En opinión de Churchland, hay una teoría disponible que cumple con los requisitos: a saber: la construida a partir de la noción de *red neuronal*. Veamos brevemente en qué consiste esta explicación.

Imagine a un pequeño cangrejo que posee un par de ojos rotatorios y un brazo extensible. El cangrejo, al toparse con algo que puede ser su alimento, naturalmente movería

su brazo para atraparlo y comerlo. Si asumimos que la respuesta motora (es decir, la forma en la que el brazo se desplaza para atrapar el objeto) está de alguna forma guiada por la percepción del objeto, es razonable suponer que hay una función entre los pares de ángulos que forman los ojos (cuando un objeto es localizado) y el movimiento del cuerpo del cangrejo (cuando el codo y el hombro del cangrejo coordinan para alcanzar el objeto).

Los mecanismos representacionales a los que apela Churchland permiten representar el input (los pares de ángulos de los ojos) como un punto de un plano de dos dimensiones de un sistema sensorial o sistema *sensorial* de estado-espacio, y también el output (el par de ángulos formados por el codo y el hombro del brazo extensible) como un punto apropiado en un sistema *motor* separado de estado-espacio. Necesitamos una función que nos lleve desde cualquier punto del sistema sensorial de estado-espacio al punto adecuado del sistema motor de estado-espacio⁴⁷. Dicha función permitirá la coordinación de la posición del brazo con la posición de los ojos de tal manera que, a partir de la posición de los ojos, el brazo pueda adquirir una posición adecuada para alcanzar el objeto identificado por el sistema sensorial en las coordenadas obtenidas en la triangulación de ambos mapas.

Una vez ubicados en el plano de dos dimensiones los puntos que representan tanto los ojos del cangrejo, como el codo y el hombro, es posible encontrar una función trigonométrica que dé como resultado la posición final del cuerpo del cangrejo. En este sentido, al describir la posición final del brazo (output) como una función específica de la posición de los ojos (input) se tiene como resultado un sistema efectivo y funcional de un sistema sensoriomotor (ver figura 2).

⁴⁷ Nótese que la noción de ‘representación’ que ahora está en juego es de diferente carácter a la sostenida por Fodor. La noción de ‘representación’ que ahora se presenta no es simbólica, es decir, no consiste en la construcción de un símbolo que captura algún contenido. Más bien, la noción de representación que ahora está en juego suele recibir el nombre de representación ‘extendida’ o ‘distribuida’. Se habla de representación ya que hay un sentido en que el objeto es ‘representado’, es decir, adquiere un valor dentro del sistema, es ubicado en un mapa de estado-espacio y, por esta razón, adquiere valor computacional. Sin embargo, las neuronas o la red neuronal involucrada en esta forma representacional no es un símbolo asociado al objeto o input. Por extensión, la noción de ‘computación’ también difiere de la Fodor, pues no es serial, ni simbólica ni proposicional (o altamente estructurada).

Figura 2

Imagen tomada de

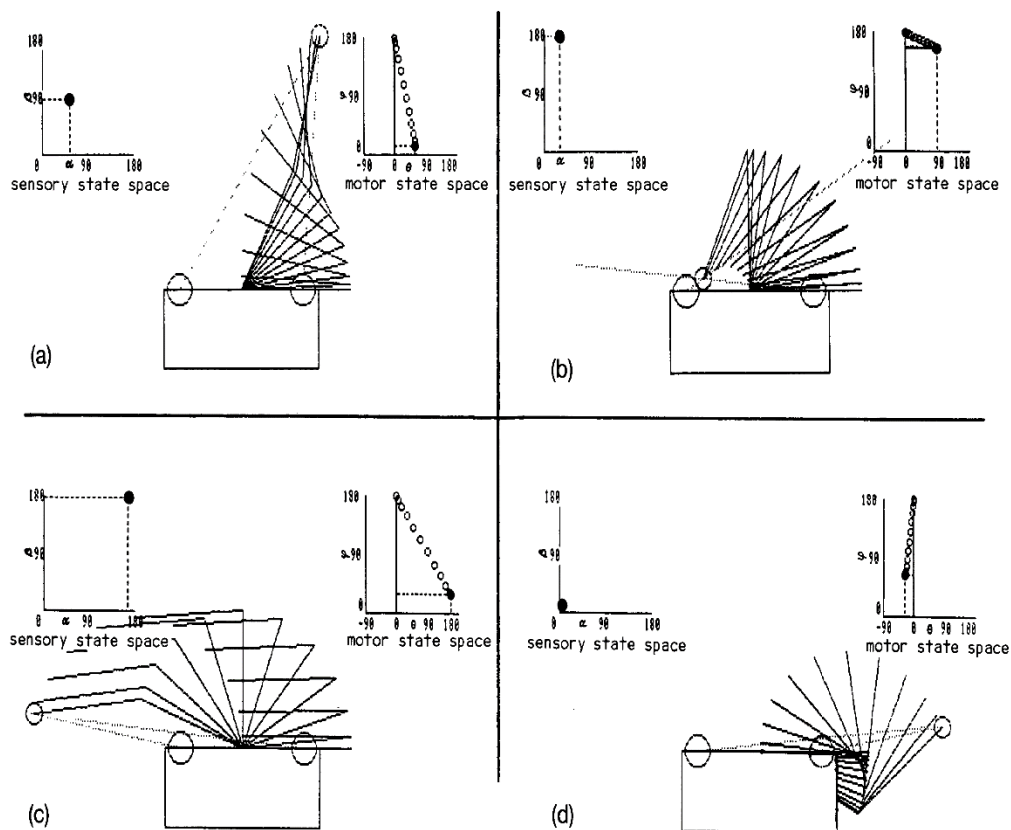


Fig. 4.

Imagen tomada de Churchland, P. (1993). State-Space Semantics and Meaning Holism. *Philosophy and Phenomenological Research*, 53(3), 667-672. doi:10.2307/2108090

Ahora bien, ¿qué clase de sistemas o mecanismos pueden ofrecer el tipo de representación y computación antes descrita? ¿Cómo se instancian o implementan los mapas dimensionales sobre los que se ubican los inputs sensoriales? La respuesta ha sido dada por la misma naturaleza. La solución se encuentra en el cerebro, o mejor aún, en las redes neuronales que conforman el cerebro. Veamos una breve descripción de cómo funcionaría neuronalmente.

Llamemos a los mapas dimensionales 'mapas de estado-espacio'. Permítase entonces que el mapa de estado-espacio sensorial esté representado por una cuadrícula de fibras que

permiten el paso de señales; y dejemos que el mapa de estado-espacio motor esté representado por otra cuadrícula de fibras ortogonal al primer mapa. Imagine ahora que el primer mapa de estado-espacio se encuentra ubicado sobre el segundo mapa y se conectan a través de pequeñas fibras verticales de tal modo que una coordenada en el mapa sensorial (es decir, el primer mapa) se conecte con la coordenada más próxima del mapa motor que yace debajo del mapa sensorial (es decir, el segundo mapa).

Si el ojo izquierdo, digamos del cangrejo, activa ciertas fibras desde la parte derecha del mapa, y el ojo izquierdo activa ciertas fibras desde la parte izquierda del mapa, la activación de ambos ojos dará como resultado una coordenada específica donde se intersecan las señales. Así, la estimulación visual dará como resultado una coordenada en el plano motor debido a las fibras ortogonales que comunican ambos planos. En este sentido, una vez las fibras motoras son estimuladas, podemos decir que inducen al brazo a tomar una posición acorde con la coordenada del estímulo.⁴⁸

Las conexiones verticales funcionan como una especie de puente o *switch* que se activan únicamente si un punto de intersección es estimulado por ambas señales simultáneamente. Churchland denomina a esta clase de sistema “sándwich de estado-espacio” (*state-space sandwich*). (Churchland, 1986)⁴⁹.

Es necesario subrayar que los mapas de estado-espacio no necesitan ser —y probablemente no sean— mapas de regiones anatómicamente obvias, como la superficie de la retina, o de la piel. Debería esperarse que los mapas sean de cosas quizá más abstractas, y por esto mismo más oscuros para el observador, pero de vital importancia funcional para el sistema. No debería resultar extraño que un mapa de estado-espacio tenga por inputs lo que constituyen outputs de otro mapa de estado-espacio, lo que crea “cadenas” interrelacionadas

⁴⁸ Para Churchland, esta solución al problema sensoriomotor parece ser bastante realista. No es difícil imaginar que este tipo de estructuras neuronales crezcan en el cerebro a través de procesos de morfogénesis y se den a lo largo de la naturaleza. Si bien diferentes creaturas pueden tener distintas formas de localizar objetos y diversos sistemas motores para efectuar contacto con ellos, todos deben enfrentar el problema de ubicar y coordinar los diferentes puntos ubicados en el plano sensorial y motor. Así, los “sándwiches estado-espacio” constituyen un método biológico realista para ejercer cualquier computación de coordenadas de 2D a 2D sin importar su complejidad matemática.

⁴⁹ Presumiblemente no es un problema ver cómo el sistema de sándwiches de estado-espacio podría resolver problemas que impliquen más de dos capas laminares de neuronas. Sabemos que el cerebro humano tiene por lo menos seis capas distintas. Una mayor cantidad de capas permite la solución de problemas más sofisticados al permitir una mayor relación de información dentro del sistema. Estos sándwiches de estado-espacio de tres o cuatro capas constituyen lo que Churchland denomina *transformador de coordenadas multimodal*. Adicionalmente, no hay razones para pensar que esa clase de sistemas se limite únicamente al sistema sensorial y motor; podría ser el caso que esta clase de sistema sea útil o incluso esencial para tareas cognitivas más elevadas. Adicionalmente, los sándwiches de estado-espacio no se limitan únicamente a dos dimensiones, como se presentará más adelante, las capas de neuronas podrían ser tantas como dimensiones requiera la tarea cognitiva.

de mapas. Por esta razón, si quisiéramos decodificar la totalidad de los mapas ubicados en la corteza, debemos entender una gran parte de mapas de estado-espacio, pues debido a que entender un mapa de estado-espacio de la cadena de mapas requiere que entendamos los mapas previos a este.

Ahora se presenta una pregunta fundamental en esta investigación. Incluso si los mapas de estado-espacio presentan una forma de representación más cercana al funcionamiento del cerebro ¿qué garantiza que los mapas de estado-espacio son lo suficientemente poderosos como para representar la riqueza de la conducta? A continuación, se ofrecerán algunas razones para pensar que, de hecho, los mapas de estado-espacio pueden ser una alternativa genuina al simbolismo fodoriano.

El poder representacional de los mapas estado-espacio

Los mapas de estado-espacio ofrecen una comprensión prometedora de los procesos cognitivos sub-personales, tal como el procesamiento de información perceptual y las conductas motoras no intencionales. Es aquí donde Churchland se muestra como una teoría alternativa. Churchland pretende que esta forma de explicación pueda expandirse a todos los procesos que, de hecho, se llevan a cabo en el cerebro y eliminar los conceptos de la psicología popular que oscurecen los conceptos de la neuropsicología.

Tal como se mencionó en el último párrafo de la sección anterior, aún hace falta mostrar las capacidades representacionales de esta nueva teoría presentada por Churchland. Como ya se vio en el caso del cangrejo, los mapas de estado-espacio ofrecen una explicación de las conductas motoras de organismos como el cangrejo que, plausiblemente, no poseen medios representacionales demasiado sofisticados, pero ¿qué lugar ocupan los conceptos o las categorías con las que los humanos categorizamos el ambiente, y, a partir de la cual, solemos tomar decisiones, resolvemos problemas y actuamos inteligentemente?

Empecemos con un caso que podría considerarse como “más mentalista” o que podría involucrar una capacidad representacional más sofisticada, por ejemplo, la experiencia de los colores.⁵⁰ En la representación de los colores los mapas de estado-espacio juegan un rol

⁵⁰ En este punto Churchland toma una posición más fuerte, no solo se está presentando una teoría explicativa y representacional de los procesos cognitivos, sino que con esta se busca la eliminación de los conceptos mentalistas. Churchland (1986 y 2007) argumenta cómo los mapas de estado-espacio puede explicar y reemplazar conceptos fundamentalmente mentalistas como los *qualia* (los colores han sido un caso paradigmático de los *qualia*. Véase por ejemplo Jackson (1982 y 1986).). Sin embargo, no es de mi interés presentar esta discusión y no pretendo defender el éxito o fallo de Churchland en la eliminación o reducción.

fundamental. Imaginemos, como lo propone Edwin Land (como se cita en Churchland, 1986, p. 300), un cubo de tres dimensiones donde cada color discriminable ocupa una posición o volumen único. Cada uno de los ejes es responsable de reconstruir la reflectancia objetiva del objeto visto por una de las tres longitudes de onda de las que nuestros conos son respectiva y selectivamente responsables. Así, podemos decir que un color es cercanamente similar a otro cuando sus posiciones en el espacio-estado estén cercanas la una de la otra, y, por otro lado, dos colores no son similares cuando sus posiciones son distantes. En este sentido, podría hablarse incluso de grados de similitud. No es difícil ver cómo podría explicarse que un color se encuentre entre otros dos, a saber, cuando la posición de un color se encuentra entre las posiciones de otros dos relativamente cercanos (ver figura 3).

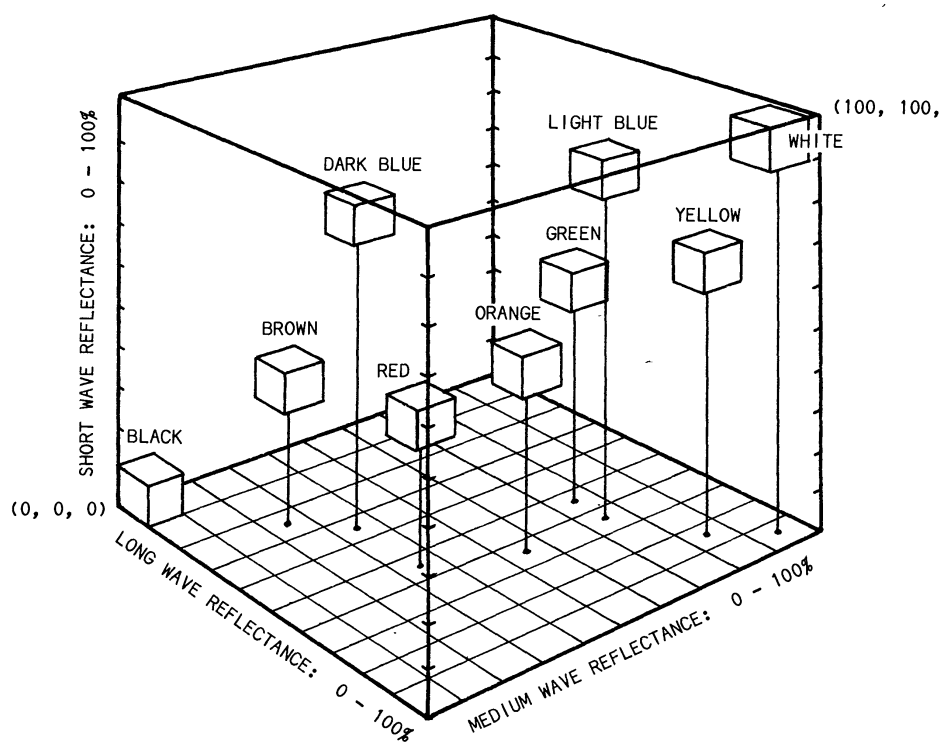


Fig. 11. Colour state space.

Fig.3: La imagen muestra una posible representación espacial de los colores ofrecida por los mapas de estado-espacio. La representación contaría con tres lados que adquieren los valores de cada uno de los tipos de células retinales y la posición de cada color corresponde al valor de activación de cada tipo de células frente a cada color. Tomado de Churchland, P. (1993). State-Space Semantics and Meaning Holism. *Philosophy and Phenomenological Research*, 53(3), 667-672. doi:10.2307/2108090

Imagine que cada uno de los ejes es dividido en diez posiciones distintas, por consiguiente, un sistema de tres ejes debe ser capaz de representar 10^3 colores distintos. Quizá este número sea subestimar nuestras capacidades, pero cualquier número mayor no tendrá ningún problema en su implementación. Cada uno de los ejes corresponde a la

cantidad de clases de conos ubicados en la retina que permiten la captación del color. En este sentido, toda percepción de color surge de exactamente estas tres clases de conos retinales. Lo anterior sugiere la hipótesis de que cualquier sensación de color es literalmente idéntica a un triplete de frecuencias específicas en un sistema del cerebro. Así, como se ha mencionado antes, la diferencia entre dos colores y su similitud está dada por la posición que adquieren en el sistema de espacio-estado.

De esta manera, los conceptos de los diferentes colores que podemos captar y discriminar, plausiblemente se corresponderían con una coordenada dentro del mapa de estado-espacio destinado⁵¹. En este caso la propuesta de Churchland —respecto a los colores— podría interpretarse de la siguiente manera: los conceptos tales como “rojo”, “amarillo”, etc., corresponden o no son otra cosa más que conjuntos de activaciones retinianas y patrones de activación neuronales ubicadas en coordenadas de mapas de estado-espacio destinada a la representación de los estímulos retinales en el cerebro activadas por la luz y los objetos físicos⁵².

Ciertamente la reducción tanto teórica como empírica esta justamente motivada. Sin embargo, hasta ahora están empezando a hacerse significativos avances en este tema. Y, como bien señala Churchland (1986), el mapeo del cerebro aún no ha sido satisfactoriamente completado. De hecho, el éxito práctico que han tenido estos sistemas de representación y computación podría extenderse desde las funciones perceptuales hasta procesos cognitivos más elevados. Puede pensarse, por ejemplo, en el proceso de reconocimiento de rostros, donde cada una de las cualidades del rostro pueda necesitar de un sistema de hasta veinte dimensiones; que al tener en cuenta la posibilidad de distinguir entre un vasto número de diferencias de cada característica, da como resultado la posibilidad de distinguir entre cientos de millones de rostros humanos. De igual forma podría aplicarse para distinguir y explicar la adquisición de otros conceptos como ‘gato’, ‘perro’, ‘vaca’, etc.

⁵¹ Los casos anómalos, por supuesto, aún obtienen explicación apelando a la misma teoría. En el caso del daltonismo, por ejemplo, si por un infortunio una de las tres clases de células es dañada y detiene su funcionamiento, uno de los tres caminos de información se vería afectado impidiendo el flujo de información. Si esto llegara a pasar, podría esperarse que el mapa de estado-espacio también se viera afectado y representara la información de solo dos de los tres caminos neuronales. Esto podría hacer que el sistema colapse en un sistema de 2D causando que algunas posiciones se solapen entre sí, de tal manera que los colores que ocupaban posiciones distintas en un mapa de estado-espacio estándar ocupen la misma posición, dando como resultado la imposibilidad de discriminación entre ellos.

⁵² Churchland (2007) presenta un argumento en la que se puede dar cuenta de los colores percibidos por el ser humano a partir del funcionamiento cerebral y de los órganos visuales perceptivos. Si el argumento es exitoso podría capturarse los colores como cotidianamente los entendemos mientras que corrige los “errores” o ingenuidades de la psicología popular sobre los colores.

* * *

Tracemos algunas conclusiones provisionales de lo que Churchland intenta defender y que serán importantes en lo que sigue del capítulo y la discusión:

- i. La psicología popular no es la única forma de descripción plausible de la conducta. Churchland nos presenta una descripción basada en la coordinación sensoriomotora que no apela en ningún sentido a entidades como deseos, creencias, intenciones, etc. Comprometerse con este tipo de descripción de la conducta, también nos conduce a apelar a otro tipo de mecanismos cognitivos adecuados para este tipo de descripción y explicación.
Churchland intenta desligarse de la explicación fodoriana en términos atomistas y composicionales y de las representaciones simbólicas definidas sintácticamente. En parte, porque considera que no es necesario ofrecer un criterio tan robusto como la *mismidad de contenido* en las representaciones de los sujetos, sino que basta con una noción de *similaridad*. Por ello, apela a mecanismos basados en la noción de *red neuronal* y mapas de coordenadas de activación neuronal que procesan la información sensorial y envían dichas coordenadas a los mapas motores correspondientes para generar respuestas motoras (conductas) apropiadas o acorde a lo que se percibe.⁵³
- ii. Un caso en el que mejor se puede apreciar el poder representacional de los mapas de estado-espacio es el caso de la percepción de los colores. Los mapas de estado-espacio parecen presentar una forma de representación bastante plausible (puramente física, esto es, sin apelar a entidades como creencias, experiencias, *qualia*, etc.) de un fenómeno (los colores) que tradicionalmente se ha tratado como mentalista o psicológico. Si esta explicación de la percepción del color resulta adecuada (como Churchland pretende defender), puede sentar las bases para considerar la explicación de procesos cognitivamente más “elevados” que se han visto tradicionalmente como explicados únicamente por la psicología popular tales como: el pensamiento, la imaginación, etc.
- iii. Uno de los procesos cognitivos más relevantes (y de mayor interés para esta investigación) al que se puede avanzar una explicación en términos de mapas de estado-espacio es el lenguaje. Podría decirse que la producción del lenguaje podría

⁵³ Por si el párrafo no es demasiado claro, las redes neuronales no son replicas idénticas (uno a uno) de otras redes, sino que los patrones de activación pueden llegar a ser parecidos o cercanos, pero jamás iguales.

llegar a ser mapeada de tal manera que a cada palabra u oración le corresponda una posición en un mapa de estado-espacio; y que la producción de las posiblemente infinitas oraciones se explique a través de la computación de diferentes posiciones de los mapas de estado-espacio que intervienen en el proceso y el output como una oración.

Briefly, Churchland invites us to view concepts as points in a partial state space of a dynamical system. These points correspond to the tips of the vectors determined by the levels of activation of the different units in hidden layers. The semantic characteristics of a concept can then be seen as a function of the place that that concept—i.e., point—occupies in a geometrically characterized hyperspace. In this way, Churchland proposes, we may talk of semantic similarity between concepts in terms of the proximity of their respective absolute positions in state space, as identified in relation to a number of semantically relevant dimensions. (Calvo Garzón, 2000, pp. 77-78)

Churchland (1998) presenta el caso de cómo una red neuronal puede ser entrenada para representar a partir de estructuras jerárquicas y relaciones de similaridad las vocales y consonantes del alfabeto. Algo parecido puede ser el caso de la estructuración de conceptos o incluso oraciones que se acomodan como puntos de activación en mapas de estado-espacio que se conectan a capas de neuronas motoras responsables de, por ejemplo, la producción lingüística.

Los mapas de estado-espacio podrían ofrecer una alternativa para la conducta lingüística o uso del lenguaje frente a la teoría fodoriana —que nos interesa particularmente en este documento—. Apelar a una explicación de mapas de estado-espacio implica —al menos para Churchland— abandonar la psicología popular, pues no haría falta apelar a entidades como deseos, creencias o intenciones en la explicación. Después de todo, hablar o la producción lingüística *sería* el resultado computacional —u output motor— de los mapas de estado-espacio y los patrones de activación neuronal.

Si la explicación del lenguaje es plausiblemente exitosa en estos términos, tenemos razones para pensar que una teoría ofrecida por la neurociencia —como la teoría ofrecida por Churchland— puede también ofrecer una explicación de la conducta independiente a la psicología popular y que se muestre no solo como plausible, sino también como favorita en tanto que muestra una mejor coherencia con teorías cercanas como la biología, química e incluso la física.

En este sentido, a grandes rasgos, Churchland no solo no está de acuerdo con el módulo central propuesto por Fodor en *la modularidad de la mente*, sino que rechaza el tipo

de computaciones que allí se llevan a cabo, es decir, LOT. La razón: aceptar un procesador central y las computaciones que Fodor presenta es aceptar la descripción —y explicación— conductual en términos de entidades mentalistas; entidades que se deben abandonar con el propósito de aceptar una teoría proveniente de la neurociencia y que apela a redes neuronales y lo que conocemos sobre el cerebro⁵⁴. La propuesta de Churchland representa un ataque fuerte a la teoría de Fodor en el sentido en que Churchland propone toda una reinterpretación y descripción de la conducta inteligente y por tanto una arquitectura cognitiva acorde para su explicación.

A continuación, presentaré las objeciones de Fodor a esta propuesta y la manera en la que Churchland las afronta con el propósito de señalar lo que está en juego en esta discusión: a saber, cuál es la arquitectura mental correcta para dar cuenta de la conducta inteligente de los seres humanos. En esta discusión, Fodor resultará victorioso. Sin embargo, no creo que este resultado sea suficiente para decidirse a favor de la teoría de Fodor.

El éxito de Fodor descansa en el hecho de que quienes entran al debate aceptan sin cuestionar que la conducta inteligente (en particular, la conducta lingüística) está estructurada (esto es, es composicional en el sentido que está explicado en el capítulo I de esta tesis) y, dado este hecho, la arquitectura mental más adecuada para dar cuenta de la conducta humana es probablemente la fodoriana.

Sin embargo, no creo que sea obvio que la conducta inteligente (y en particular la conducta lingüística) tenga este tipo de estructura. Por esta razón, al final de este capítulo, presentaré un argumento desarrollado por Lenny Clapp (2012) que ofrece buenas razones para rechazar esta suposición sobre la conducta y que puede enriquecer el debate respecto de cuál es la arquitectura cognitiva más adecuada para dar cuenta de la conducta humana. Y finalmente trazaré una conclusión sobre el resultado del debate a la luz de dichos argumentos.

Al final de la tesis en el cuarto capítulo, pretendo defender la propuesta de Fodor en cuanto que quiero defender la tesis de que la arquitectura mental adecuada para dar cuenta de la conducta inteligente tiene la forma de las hipótesis de Fodor, con algunas modificaciones. Por esta razón, deseo presentar algunos argumentos en contra de los argumentos de Clapp (2012) y sugerir algunas modificaciones a la hipótesis de Fodor basadas en el trabajo de dos filósofas, Claudia Picazo (2019) y Susan Schneider (2011).

⁵⁴ En parte porque no hay que dar cuenta de la proposicionalidad de las representaciones y porque Churchland pretende ofrecer una explicación de la conducta lingüística que no se caracteriza como proposicional, sistemática, composicional, etc.

Empecemos, por ahora, con la respuesta de Fodor al proyecto de Churchland y su debida defensa

Segunda Parte

Fodor Responde: en defensa de la psicología popular

Respuesta de Fodor a Churchland: semántica por estipulación

Los mapas de estado-espacio presentan una arquitectura basada en lo que presumiblemente es el funcionamiento del cerebro. Sin embargo, las redes neuronales son una idealización, o mejor aún, una descripción funcional de los cerebros particulares. El problema de esto es que cada cerebro individual es altamente idiosincrático, es decir, las conexiones neuronales no son una copia idéntica o uno-a-uno de las demás. En este sentido, apelar al funcionamiento del cerebro no es suficiente para mostrar que los mapas de estado-espacio expresan los *mismos* contenidos o representan lo mismo para todos los sujetos. Siendo, en el mejor de los casos, posible hablar únicamente de *similitud* de contenido o similitud de representación entre diferentes sujetos.

Churchland, si quiere rechazar la teoría planteada por Fodor y abandonar un sistema representacional simbólico y composicional, debe ofrecer un criterio claro y robusto de *similaridad de contenido* para explicar cuándo dos sujetos diferentes poseen el mismo concepto —lo que permite la comunicación, atribución de estados mentales, etc. Esto es, permita una explicación generalizada de la conducta de los individuos tal como la buscada en el capítulo I de esta tesis (Fodor y Lepore, 1992).

La idea de Churchland es sencilla: dos conceptos son similares en tanto que ocupan posiciones cercanas en el *mismo estado-espacio*. La pregunta que debe ser respondida, por lo tanto, es ¿Cuándo S_1 y S_2 son el mismo mapa de estado-espacio? Es decir, ¿Cuándo, por ejemplo, el espacio semántico de usted lector es una instancia del mismo tipo de mapa de estado-espacio semántico que el mío?

La condición suficiente de identidad entre mapas de estado-espacio podría entenderse como la identidad entre las dimensiones que componen el mapa. Ahora bien, los diferentes conceptos deberían representarse en un punto de coordenadas específico, según la relación que mantienen con las dimensiones en cuestión. La pregunta relevante es la siguiente, ¿cuándo m y n son las mismas dimensiones para diferentes mapas?

Una respuesta empirista identificaría las dimensiones semánticas con propiedades observables por el sujeto. Si esto es así, los criterios de identidad de las dimensiones son claros, permitiendo, a partir de allí, rastrear la identidad entre los conceptos no observables del lenguaje del sujeto. Sin embargo, en el caso de la tesis avanzada por Churchland, las dimensiones no siempre se corresponden con propiedades observables, sino con cualesquiera propiedades que el cerebro se represente. Esto pone un problema sobre la teoría de Churchland. Según Fodor y Lepore, dado que el criterio empirista de identidad entre dimensiones semánticas ha sido abolido en la teoría de Churchland (ver nota al pie 55)⁵⁵ se sigue que no hay criterio alguno que pueda asumirse para la identidad entre dimensiones semánticas. El argumento/crítica puede presentarse de la siguiente manera:

To repeat, we have a robust notion of semantic similarity only if we have a criterion for the identity of state spaces. We have a criterion for the identity of state spaces only if we have a criterion for identity of dimensions of state spaces. And we have a (no empiricist) criterion for the identity of dimensions of state spaces only if we have a criterion of "property expressed by a dimension of state space" that works for arbitrary properties, not just for observable properties. But a criterion for "property expressed" that works for arbitrary properties just is a criterion for identity of meaning. So Churchland's proposal for a robust theory of content similarity fails to avoid the problem of robust content identity – and, of course, fails to solve it. (Fodor y Lepore, 1992, pp. 198-199)

Parece, a ojos de Fodor y Lepore, que Churchland es incapaz de ofrecer un criterio claro de similaridad (o mismidad) de contenido. En Churchland podemos ver esto en la medida en que necesitamos una noción robusta de identidad de dimensiones para obtener así un criterio plausible de similaridad de contenido. Así, el problema de Churchland parece ser que da por garantizadas las etiquetas de las dimensiones de los mapas de estado-espacio. Dichas etiquetas solo nos dicen cómo las posiciones a lo largo de dicha dimensión deben ser interpretadas expresando grados de F-idad (*F-ness*). Sin embargo, cabe la pregunta, ¿en virtud de qué la dimensión expresa grados de F-idad (*F-ness*) y no de G-idad (*G-ness*)? Y más importante aún, ¿qué hace que una dimensión que expresa la propiedad F en su mapa de

⁵⁵ Churchland rechaza el criterio empirista debido a que no quiere comprometerse con la idea de que el cerebro se representa propiedades observables obvias:

Experience with artificial networks has taught us, however, that the many interlocking sub-features detected at lower levels-whose collective impact at higher levels ultimately produces the activation vector for "kitten"-are rarely the features that one would list in a definition of "kitten" (e.g., small, furry, four-legged, young, feline). More typically they will be diffuse, opaque, inarticulable features whose only obvious significance is that, in conjunction with hundreds of other similarly diffuse coding features passing through the labyrinth of the massively parallel network, they occasionally participate in the selective activation of the higher-level "kitten" vector. (Churchland, 1993, p. 670)

estado-espacio exprese la misma propiedad F en una dimensión en mi mapa de estado-espacio?

Tal como se expresó en el capítulo I, los estados mentales se clasifican esencialmente por su contenido o sobre lo que tratan. Es decir, representamos y clasificamos cosas rojas, por ejemplo, porque son rojas; o nos representamos a *tías* o *tíos* como tías o tíos y no como objetos que poseen ciertas propiedades psicofísicas empleadas para la detección de tías o tíos (si es que tales propiedades existen). Incluso si los mapas de estado-espacio de Churchland lograran representar estas propiedades psicofísicas, la teoría de Churchland no ofrece explicaciones de qué cosas corresponden a regiones en los mapas de estado-espacio, y qué cosas representan sus ejes. Así, la semántica propuesta no ofrece una explicación robusta de lo que se supone debería, esto es, por qué y cómo los mapas de estado-espacio significan lo que significan. Dado que no hay una explicación convincente sobre ello, parece que la semántica ofrecida por Churchland “*it’s just semantics by stipulation*” (Fodor y Lepore, 1992, p. 200).

Este problema, derivado de la falta de un criterio robusto de similaridad de contenido (o mismidad de contenido), junto con la falta de explicación de lo que los mapas de estado-espacio de hecho representan, trae graves consecuencias a la teoría de la mente de Churchland (es decir, a su explicación) y, por ende, al tipo de descripción conductual que propone Churchland. Un ejemplo de esto es que Churchland no puede ofrecer un criterio claro de lo que significa estar en un estado mental u otro, lo que significa compartir un concepto con otros hablantes e incluso cómo una conducta particular puede ser de cierto tipo u otro.

La imposibilidad de generalización conduce a la teoría a un fracaso explicativo, pues no permite la elaboración de una semántica de los estados mentales aplicable a tipos (*types*), sino elaborada sobre *tokens* de mapas de estado-espacio, es decir, cada caso particular recibe una explicación diferente porque ni siquiera es posible saber cuándo dos casos pueden llegar a ser el mismo tipo de caso. Debido a esto, la posible divergencia de posiciones que puede tomar un mismo concepto debido al conocimiento idiosincrático da lugar a una imposibilidad de generalización en la teoría y por tanto una insuficiencia explicativa en el lenguaje y por extensión en los procesos cognitivos.

En este sentido, la propuesta de Churchland no puede leerse como una teoría efectiva de la mente en tanto que no ofrece un criterio robusto de lo que consiste estar en un determinado estado mental para uno y varios individuos. Esto es particularmente problemático, en la medida en que no se ofrece una explicación en lo absoluto dada la

imposibilidad de generalización sobre individuos y, principalmente, sobre las *mismas* circunstancias, las *mismas* causas y *conductas*.

Respuesta de Churchland: no hay conceptos, solo *prototipos*

Churchland, por supuesto, no puede negarse a ofrecer una explicación de la conducta lingüística. Incluso si la conducta inteligente está ligada a la adaptación al ambiente y la supervivencia (como pretende mostrar en la primera parte de este capítulo), en los seres humanos la conducta lingüística es fundamental para este propósito; y es, como se señaló en el capítulo I, el paradigma de conducta inteligente debido a su versatilidad y sofisticación.

El argumento que Churchland intenta elaborar para defenderse de las críticas de Fodor y Lepore consiste en señalar que los mapas de estado-espacio no funcionan a partir de la semántica que Fodor tiene en mente, es decir composicionalmente. Churchland intentará mostrar que la semántica de mapas de estado-espacio no ubica conceptos en sus mapas sino ‘*prototipos*’ de activación neuronal, y estos *prototipos* son los vehículos semánticos.

Churchland propone un ejemplo: imagine que usted y yo, lector, estamos en una habitación. En esa habitación hay un sofá y sobre el sofá se encuentra un gato. Probablemente más de 200 millones de células retinianas se encuentran activas en nuestras retinas, y probablemente una cantidad más grande de neuronas de encuentren activas mientras observamos al gato sobre el sofá ronroneando y moviendo su cola.

Simultáneamente, dichas células están captando millones de otras características ambientales independientes a las del gato. De hecho, podría decirse que las características del gato parecieran estar “enterradas” o mezcladas entre la inmensa pluralidad de características del ambiente que estamos percibiendo. Adicionalmente, la información ambiental que captamos está cambiando continuamente en la medida en que caminamos a lo largo del cuarto y cambiamos de posición; otros sonidos y objetos entran y salen de nuestra escena sensorial. Usted, lector, que se encuentra en la misma habitación, digamos detrás del sofá, posee su propia cacofonía de estimulaciones neuronales que variaría con respecto a la mía dependiendo también de la posición que tome y por sus propias idiosincrasias perceptuales. ¿Cómo es posible que ambos logremos percibir e identificar sin esfuerzo alguno el gato ubicado en el sofá?

Lo que hace posible tal hazaña son los *prototipos* de activación neuronal formados a partir de la computación de vectores, es decir, de los patrones en los que la red neuronal se

activa al entrar en contacto con los objetos y el ambiente. El entrenamiento de la red —es decir, el contacto continuo y prolongado de la red con los objetos— ha transformado tendencias activacionales de poblaciones de neuronas en familias finitas de activación, por ejemplo, de gatos y sofás; de voces y maullidos; y en general el resto de las categorías que amueblan nuestra ontología. De esta manera, la red neuronal estandariza la activación de tal manera que el patrón de activación se vuelve prototípico (llegando incluso a que las neuronas de más bajo nivel o más cercanas al input se vuelvan como buscadoras de los elementos que encienden dicho patrón.). Es a esto que se le conoce como *prototipo* de activación neuronal.

Esto deja entrever un aspecto importante respecto a los prototipos de activación neuronal. Si lo pudiéramos en términos fodorianos, los prototipos no son composicionales, es decir, no hay prototipos simples que se combinen para dar lugar a prototipos más complejos. La razón de ello recae en lo antes mencionado: las representaciones ofrecidas por las redes neuronales no están estructuradas composicional o proposicionalmente debido a que la percepción y computación neuronal no parece estar estructurada de dicha forma. En este sentido, la representación ‘El gato está sobre la mesa’ puede ser *simple* en sí misma y no estar compuesta del prototipo de activación neuronal para ‘gato’, y otro para ‘mesa’, etc.

Por otro lado, cuando un sujeto percibe repetidamente gatos —al entrenar la red neuronal fortaleciendo los ‘pesos sinápticos’ que se dan entre las neuronas que se activan frente a la presencia de gatos— forma un ‘*prototipo*’ de activación de gatos. Hace falta mencionar que los *prototipos* de activación no son activados o formados por propiedades obvias, sino por micro-rasgos del contexto perceptual:

Experience with artificial networks has taught us, however, that the many interlocking *sub-features* detected at lower levels-whose collective impact at higher levels ultimately produces the activation vector for "kitten"-are rarely the features that one would list in a definition of "kitten" (e.g., small, furry, four-legged, young, feline). *More typically they will be diffuse, opaque, inarticulable features whose only obvious significance is that, in conjunction with hundreds of other similarly diffuse coding features passing through the labyrinth of the massively parallel network, they occasionally participate in the selective activation of the higher-level "kitten" vector.* (Churchland, 1993, pp. 670, énfasis añadido)

El proceso de entrenamiento convierte a las neuronas de los niveles bajos en detectores de sutiles patrones o sub-patrones de activación en niveles más bajos de neuronas hasta llegar al nivel sensorial (Churchland, 1993). En otras palabras, mediante el

entrenamiento en la etapa de configuración de conexiones sinápticas se han configurado prototipos de patrones de activación que pueden rastrearse hasta los niveles sensoriales y son los que explican que podamos identificar ciertos objetos entre la inmensa cantidad de información ambiental debido al reconocimiento de dichos patrones de activación generados por los objetos.

Bajo esta concepción del proceso de computación de funciones puede ofrecerse ahora una respuesta a la identidad de vectores entre diferentes individuos. La respuesta ofrecida por Churchland a este respecto es que el vector que representa gatos, es decir, que tiene como contenido “gatos” no significa la identidad de los vectores de los distintos individuos, ni es la igualdad de dimensiones en los niveles neuronales de procesamiento entre los niveles sensoriales y motores. Lo que determina el contenido del vector “gato” es el rol en conjunto que dicho vector juega dentro del extenso sistema económico cognitivo y motor del cual hace parte.

What gives this vector the content "kitten" is the overall role that this vector plays in the larger cognitive and motor economy of which it is an interlocking part. Thanks to an almost magical process of sensory filtering and vector completion, that prototype vector is typically activated in me by -any one of a wide range of possible sensory impacts on me by a kitten. But much more importantly, that prototype vector has rich causal and computational sequelae of its own. Because of its own transformational significance to the extensive and well-trained network downstream from that prototype vector, it prompts a family of kitten-specific perceptual expectations and prepares one to engage in a family of kitten-specific behaviors (Churchland, 1993, pp. 670-671)

Es este aspecto del rol computacional de los prototipos que es de vital importancia para reconocer mismidad de contenido:

It is this downstream aspect of the vector's computational role that is so vitally important for reckoning sameness of cognitive content across individuals, or across cultures. A person or culture that discriminated kittens reliably enough from the environment, but treated them in absolutely every respect as a variant form of wharf-rat, must be ascribed some conception of 'kitten' importantly different from our own. On the other hand, an alien, person or species whose expectations of and behavior towards kittens precisely mirror our own must be ascribed the same concept "kitten,"

even though they might discriminate kittens principally by means of alien olfaction and high-frequency sonar beamed from their foreheads. (Churchland, 1993, pp. 671)

En resumen, si lo que nos interesa entender es la significación o el contenido de dichas categorías prototípicas, entonces se debe mirar el rol computacional que desempeñan dentro de *todo* el conjunto de la economía cognitiva y la conducta motora que dicho prototipo puede llegar a generar. Es posible para Churchland explicar en qué consiste la identidad de contenido entre individuos teniendo en cuenta las posibles variaciones idiosincráticas de los vectores en casos particulares de individuos. Por otro lado, también muestra cómo es posible evadir el principio empirista en la medida que no necesita apelar a propiedades observables del ambiente distal o proximal sino a la misma conducta y economía cognitiva que el individuo desempeña desde cada concepto o categoría que posee. Esto le permite a Churchland responder las preocupaciones de Fodor y Lepore sobre la generalización entre individuos y el rol del conocimiento idiosincrático de cada individuo.

Fodor y Lepore responden: redes neuronales y composicionalidad

Mediante la noción de ‘Prototipo’ Churchland pretende explicar cómo los mapas de estado-espacio pueden recibir una interpretación semántica que logre dar cuenta de los contenidos expresados mediante la conducta lingüística y, por lo tanto, inteligente. Sin embargo, Fodor argumentará, como se presentará a continuación, que la noción de ‘*prototipo*’ no logra dar cuenta de una de las propiedades fundamentales de la conducta lingüística e inteligente, a saber, la composicionalidad. La razón de esto es que los *prototipos* no son composicionales.

Hay una diferencia radical entre la clase de representaciones que acepta Fodor (o las teorías clásicas o simbólicas de la mente) versus las representaciones postuladas por Churchland (o las que apelan a las redes neuronales). Las representaciones tipo Fodor, como se mostró en el capítulo II, son símbolos sintácticos que capturan o expresan un contenido semántico determinado. Por ejemplo, la representación de ‘GATO’ es un símbolo sintáctico que captura o expresa el concepto/contenido ‘Gato’ y preserva ciertos poderes causales relacionados a la conducta motora o lingüística del sujeto que posee dicha representación.

Por consiguiente, en tanto su naturaleza estructurada, puede combinarse con otros símbolos para dar lugar a representaciones más complejas y por lo tanto a conductas más complejas como producir/pensar la oración ‘el gato esta sobre la mesa’. De esta manera, su

carácter sintáctico les da a las representaciones simbólicas una naturaleza composicional, donde el significado o contenido de las representaciones complejas está dado por el contenido de sus componentes y las relaciones sintácticas adecuadas entre ellos. Esto le da a la teoría simbolista el poder explicativo suficiente para dar cuenta de dos factores que son de vital importancia en la explicación de la conducta lingüística e inteligente en general, a saber:

- a) La composicionalidad (En un sentido general, donde la conducta se encuentra estructurada de tal forma que hay conductas más simples que pueden agruparse y componer conductas más complejas, véase la sección “Las acciones estructuradas” en la primera parte del Capítulo I).
- b) Las relaciones explicativas que los sujetos establecen entre diferentes conductas (por ejemplo, la interpretación que Hermia hace de la conducta de Lisandro, véase el Ejemplo 2 en la primera parte del capítulo I).

La teoría de Churchland puede ofrecer una respuesta plausible a la manera en que los procesos cognitivos se instancian en el cerebro; e incluso como se desarrollan procesos modulares y sub-personales. Sin embargo, la teoría de los conceptos como prototipos parece no ofrecer una explicación satisfactoria de los factores antes mencionados. Dando como resultado una arquitectura mental insuficiente para explicar la conducta lingüística e inteligente.

En resumen, las representaciones mentales que hacen parte de la explicación de la conducta inteligente no pueden ser prototipos, porque los prototipos no son composicionales y la conducta inteligente y lingüística sí. Lo que la representación “GATO” contribuye al contenido de, por ejemplo, la producción de la oración “el gato esta sobre la mesa” es el contenido ‘gato’. Sin embargo, en el caso de Churchland, la conducta de producir la oración “el gato esta sobre la mesa” puede ser simple en sí misma. En este sentido, el prototipo de “GATO” no añade nada. Por esta razón, apelar a los prototipos no puede explicar rasgos fundamentales de la conducta inteligente y lingüística como los mencionados en (a) y (b), dejando así sin explicar una gran parte de la conducta humana, y según Fodor, la más importante.

Sin embargo, no todo está dicho en este debate. En este punto aún se le puede responder a Fodor. La propuesta fodoriana y sus contrargumentos están fuertemente anclados en la composicionalidad de la conducta lingüística e inteligente y, por tanto, la composicionalidad del pensamiento. Pero ¿qué pasaría si se demostrara que la conducta lingüística e inteligente no es composicional, esto es, que no está compuesta de partes

inidentificables? ¿Sería esto una razón suficiente para abandonar la teoría de Fodor y la psicología popular?

En la siguiente parte de este capítulo argumentaré a favor de esta posibilidad, apelando a los argumentos de Lenny Clapp (2012). Y en el siguiente capítulo, capítulo IV, evaluaré si esta posibilidad mina nuestras motivaciones para explicar la conducta lingüística e inteligente en términos de la psicología popular y lo que ésta implica.

Tercera Parte

Clapp entra al debate: ni composicionalidad, ni LOT

Los argumentos que se presentarán a continuación buscan hacer un punto a favor de la idea de que los mecanismos cognitivos responsables de la conducta no deben ser de carácter proposicional. Esto se da en el marco de una discusión que, por desgracia, no presentaré en este documento. Los pragmatistas radicales tienen buenas razones —y buenos argumentos, por supuesto— para mostrar que la conducta lingüística no es composicional. La razón, a grandes rasgos, es que el lenguaje es lo suficientemente sensible al contexto hasta el punto de que no permite hablar de elementos simples y su posición sintáctica para dar cuenta del contenido o el significado de una oración.

Esto, por supuesto, trae implicaciones para la teoría fodoriana. Si volvemos al argumento ofrecido en el capítulo II a favor de un sistema representacional, si la conducta lingüística e inteligente carece de este rasgo composicional, perdemos las motivaciones iniciales para suponer que sus causas o los mecanismos cognitivos responsables son también composicionales. Sin embargo, de acuerdo con Clapp (2012, pág. 301) en este punto Fodor acepta una posición mixta, es decir, acepta la no-composicionalidad de la conducta, pero argumenta que el pensamiento (y por lo tanto LOT) debe seguir siendo composicional. Considérese el siguiente ejemplo:

Ejemplo (Travis-Clapp): Recall Travis' (1997, p. 89) example concerning two utterances of the sentence 'The leaves are green.' One of these utterances expresses truth conditions which are satisfied if the leaves of the plant in question have been painted green, but are dead-brown underneath the paint. The other utterance expresses truth conditions which would not be satisfied if the leaves were dead and brown, but

had been painted green. Travis here presents a counterexample (or at least an apparent counterexample) against the traditional view that English is truth-conditionally compositional: He describes a case in which there are two utterances of the same indexical-free sentence (type) that express different truth conditions. If such a case is possible, then the truth conditions of utterances of 'The leaves are green' cannot be a function of (a) the meanings of the constituents in 'The leaves are green', and (b) the syntactic structure of 'The leaves are green'. (Clapp, 2012, pp. 302)

Como el ejemplo sugiere, parece ser que la totalidad del significado o las condiciones de verdad expresadas por la oración no son una función del contenido semántico de sus componentes y su sintaxis. Fodor está de acuerdo en que este tipo de situaciones no solo son posibles, sino que gran parte de nuestras interacciones lingüísticas son sensibles al contexto, y por ello mismo, minan la caracterización que se ha ofrecido de la composicionalidad de la conducta lingüística e inteligente.

Así las cosas, la posición que toma Fodor luego de la discusión con los pragmatistas radicales respecto al lenguaje y la conducta inteligente es que: si bien el lenguaje no es composicional, entonces el pensamiento debe serlo.⁵⁶ Los argumentos de Clapp presentados a continuación pretenden mostrar que Fodor no tiene buenas motivaciones para sostener esta posición mixta y que los argumentos en contra de la composicionalidad de la conducta lingüística e inteligente también aplican en contra de la composicionalidad del pensamiento y LOT. Finalmente, al final de esta sección, y como conclusión del capítulo, pretendo ofrecer una valoración y conclusión sobre el estado del debate a la luz de los argumentos expuestos y dibujar más claramente mi posición frente a las opciones resultantes.

Antes de iniciar, considero relevante ofrecer una caracterización más precisa de lo que es la composicionalidad de un pensamiento en términos de una oración de LOT. Siguiendo a Clapp (2012, pág. 302) no hay una única y clara definición de lo que es la composicionalidad para un sistema representacional y Fodor tampoco ofrece una caracterización muy rigurosa de esta propiedad. En vista de esta situación, Clapp ofrece una caracterización de 'composicionalidad' que, a ojos de él, la mayoría podría aceptar y que es la que tiene en mente al desarrollar sus argumentos:

⁵⁶ Esta parte de la discusión no se centrará en las motivaciones que tiene Fodor para sostener esta posición mixta; o si sus motivaciones son lo suficientemente fuertes para sostener aún la composicionalidad del pensamiento, pero no del lenguaje. Me centraré en presentar los argumentos de Clapp contra estas motivaciones.

Truth-Conditional Compositionality: A language L is truth-conditionally compositional iff all declarative sentences S of L, the truth conditions expressed by S (relative to a context) are a function of (a) the semantic content of the words in S (where) the semantic content of indexicals and other context-sensitive words is determined relative to the context) and (b) the relevant syntactic structure (the LF) of S. (Clapp, 2012, pp. 301)

Nótese que este criterio de composicionalidad solo puede aplicarse a lenguajes naturales o formales, o sistemas representacionales, que acepten una distinción *type/token*. Esto en la medida en que son las instancias u ocurrencias las que expresan ciertas condiciones de verdad.

Por esta razón, este criterio es perfecto para aplicarse al lenguaje de pensamiento, pues las representaciones aceptadas por Fodor son de naturaleza proposicional y aceptan la distinción *type/token*. Por lo que puede hablarse de ‘composicionalidad del pensamiento’: donde ‘composicionalidad’ se entiende en el sentido adscrito por Clapp y el pensamiento como representaciones que son, en último término, oraciones de Mentalés.

Veamos a continuación dos argumentos presentados por Clapp para sostener que el pensamiento (tal como es entendido por Fodor) no puede ser composicional en el sentido anteriormente definido por Clapp.

El argumento de la esencia

Lo primero que podría decirse respecto al ejemplo Travis-Clapp es que el lenguaje no tiene su contenido semántico esencialmente y por ello es por lo que pueden darse casos en que una misma oración del lenguaje exprese dos condiciones de verdad diferentes en emisiones distintas. Sin embargo, y es la idea general de este argumento, Fodor responde que los pensamientos tienen su contenido semántico esencialmente, quizá porque solo son proposiciones. Por ello, sería imposible que se presenten casos tipo Travis-Clapp, donde un mismo pensamiento (*token*) exprese condiciones de verdad diferentes o proposiciones diferentes. De ser este el caso, se sigue que es imposible presentar un contraejemplo en contra de la composicionalidad del pensamiento.

Sin embargo, a ojos de Clapp, de esto no se sigue que el pensamiento sea composicional. Así como es imposible presentar un contra ejemplo a que los números miden menos de 6 metros, de ahí no se sigue que todos los números entonces midan más o igual a 6 metros. La razón de ello es que no tiene sentido hablar de que los números tienen altura. En

este mismo sentido, según Clapp, no tiene sentido decir que los pensamientos son composicionales. Si recordamos la caracterización ofrecida de ‘composicionalidad’ unos párrafos atrás, solo se puede hablar de composicionalidad en medios que acepten una distinción *type/token*. En este sentido, los contenidos veritativo-funcionales —es decir, las condiciones de verdad— no aceptan tal distinción; y por ello, no tiene sentido decir que dos instancias de un mismo pensamiento expresan dos proposiciones diferentes, porque es imposible para un pensamiento tener instancias.

¿Cómo debemos interpretar a Fodor frente a este argumento? No hace falta ir muy lejos para dar una respuesta. Fodor acepta una teoría representacional de la mente, LOT. Como ya se vio antes, este sistema representacional acepta la distinción *type/token*. Teniendo esto en mente, cuando Fodor se refiere al pensamiento quizá se refiere a las oraciones en Mentales. Esta es la sugerencia de Clapp interpretando el siguiente pasaje de Fodor:

There can't be more-or less- to a thought than there is to its content because a thought just is its content. If you put this in the language of a representational theory of mind, it comes out something like: A mental representation is ipso facto compositional with respect to the content that a correct semantics would assign to it. (2001, p. 14) (Fodor en Clapp, 2012, pp. 305)

De esta forma, se puede presentar la conclusión de Fodor de esta nueva manera: Si tener un pensamiento (o pensar) consiste en la instanciación de oraciones de mentales, es correcto decir que las representaciones de mentales tienen su contenido esencialmente; y por ello es imposible que dos instancias de una misma representación expresen proposiciones diferentes (o contenidos diferentes). Se sigue entonces que es imposible presentar contraejemplos a la composicionalidad de las representaciones (Clapp, 2012, pp. 305).

Sin embargo, algo ha ido mal. Los pensamientos no pueden ser dos cosas al mismo tiempo, es decir, no pueden ser las representaciones y (al mismo tiempo) el contenido de las representaciones. Fodor tiene que elegir una de las dos opciones: o bien los pensamientos son los contenidos de las representaciones o bien son las representaciones mismas (esto es, las oraciones del Mentalés). Si Fodor elige lo primero, entonces no tiene sentido hablar de ‘composicionalidad’ de los pensamientos, debido a que los contenidos no tienen instancias y la composicionalidad solo se aplica a instancias de tipos. Pero, si Fodor elige lo segundo, entonces ya no es obvio por qué el pensamiento tiene su contenido esencialmente, pues es posible presentar casos Travis-Clapp a las representaciones. (Clapp, 2012, pp. 306).

Por supuesto, esto no quiere decir que Fodor no tenga una respuesta que ofrecer, quizá pueda ofrecerse otro argumento a favor de la composicionalidad del pensamiento que no apele a la esencialidad del contenido. Este es el caso del argumento de la explicitud. Este argumento descansa en la idea de que las oraciones del lenguaje natural no necesariamente son explícitas respecto a su contenido, mientras que las representaciones de LOT sí lo son. En este sentido, las oraciones que son semánticamente indeterminadas, como las del ejemplo Travis-Clapp, les corresponde una oración en Mentales completamente explícita respecto al contenido (o condiciones de verdad) que dichas oraciones expresan. Clapp, tampoco cree que este argumento esté disponible, pues también puede presentarse un caso, diferente al caso Travis-Clapp, que nos lleve a ver más claramente por qué el pensamiento no puede ser composicional.

El argumento de la explicitud

Hagamos una breve recapitulación de las entidades relevantes que se han discutido a lo largo de los argumentos: Tenemos entonces oraciones de Mentalés, que es lo que Fodor plausiblemente entiende que son los pensamientos. Los pensamientos, que son representaciones, al igual que las oraciones del lenguaje aceptan una distinción *type/token* y ambas (oraciones del lenguaje natural y oraciones de Mentalés) expresan determinados contenidos.

Como se ha insinuado en las secciones anteriores, el argumento de la esencia, presentado de manera simple, defiende la idea de que el lenguaje no es composicional debido a que sus oraciones pueden ser implícitas respecto al contenido que expresan. Sin embargo, cree Fodor, el pensamiento no es susceptible de estos argumentos, porque un pensamiento no puede ser implícito respecto al contenido que expresa. En este sentido, Fodor argumenta:

... language is strikingly elliptical and implicit about the thought it expresses ... and [this] couldn't be true ... if language were compositional ... For if it were (and assuming that the content of a sentence is, or is the same as, the content of the corresponding thought) the structure of a sentence would indeed have to be explicit about the structure of the thought it expresses; in particular, the constituents of the sentence would have to correspond in a straightforward way to the thought's constituents. For, if there are constituents of your thought that don't correspond to constituents of the sentences you utter, then since compositionality requires that the content of a thought contain all of the content of its constituents, it must be that there

was something in the thought that the sentence left out. So you've said less than you intended. (2001, p. 12) (Fodor en Clapp, 2012, pp. 317)

Si volvemos al caso Travis-Clapp, expuesto al inicio de este capítulo, podría leerse de la siguiente manera: la oración tipo 'las hojas son verdes' puede expresar dos condiciones de verdad diferentes en diferentes instancias porque las oraciones del lenguaje no necesariamente son explícitas respecto al contenido que expresan. Es decir, la oración del lenguaje natural 'las hojas son verdes' puede tener menos constituyentes que el contenido que expresa dicha oración. En este sentido, aceptar que una oración del lenguaje es implícita respecto al contenido que expresa, implica aceptar que las condiciones de verdad de la oración no son una función (exclusivamente al menos) de sus constituyentes y su estructura sintáctica. Se sigue que el lenguaje natural no es composicional en ese sentido.

Así las cosas, cuando Fodor arguye que el pensamiento sí es composicional, rechaza lo anteriormente dicho en la medida en que defiende que los pensamientos no pueden ser implícitos respecto a su contenido. Sin embargo, como se analizó en el argumento de la sección anterior, apelar a la esencialidad del contenido de los pensamientos no parece ser una salida plausible.

Por suerte, Fodor puede presentar un argumento independiente al argumento de la esencia para defender aún la composicionalidad del pensamiento. Clapp lo presenta de la siguiente manera:

Fodor could challenge us as follows: "Look, once you grant (E1), that the sentence is implicit with regard to the thought, what motivation could you possibly have for supposing that in addition to S's implicitness with regard to T (*where T stands for thought*) (and hence P, *where P stands for proposition*), T is also implicit with regard to P? If you grant that one's mental representation is richer than the sentence uttered, then there could be no reason for supposing that the content of your mental representation is also richer than your mental representation. To put it crudely, since the implicitness of S with respect to P can be adequately explained by the implicitness of S with respect to T, there is simply no reason to suppose that T is also implicit with respect to P (Fodor en Clapp, 2012, pp. 319, *itálicas añadidas*)

Si bien la nomenclatura de la cita anterior puede parecer confusa, el argumento sugiere, a grandes rasgos que, si se acepta que las oraciones del lenguaje natural pueden ser implícitas respecto al contenido que expresan, este hecho puede explicarse en la medida en

que la oración puede ser implícita respecto al pensamiento que expresa. Ahora bien, una vez garantizado que la oración es implícita respecto al pensamiento que expresa, no hay razones para pensar que las representaciones también son implícitas respecto al contenido semántico que expresan.

Si Fodor tiene razón en el argumento presentado, los argumentos ofrecidos por los pragmatistas radicales en contra de la composicionalidad del lenguaje natural no dicen nada respecto a la composicionalidad del pensamiento. Los argumentos revisados hasta ahora aplicados a la composicionalidad del lenguaje natural podrían descartarse alegando que solo muestran que el lenguaje natural es implícito respecto al pensamiento, pero no hay razones para suponer que el pensamiento también sea implícito respecto a su contenido y por ello mismo no hay razones para pensar que sea no-composicional.

Hasta este punto, Fodor parece tener un buen argumento para defender la composicionalidad del pensamiento. Casos como el Travis-Clapp, mostrarían únicamente la no-composicionalidad del lenguaje natural, pero serían neutros respecto al pensamiento. Sin embargo, Clapp (de manos de Perry) intenta presentar un caso que mostraría que el pensamiento sí puede ser implícito respecto al contenido que expresa. Es decir, que sí tenemos buenas razones para pensar que las oraciones de Mentalés pueden ser implícitas respecto al contenido que expresan.

En este sentido, la intuición que guía el caso presentado podría ponerse así: si una oración del lenguaje puede ser implícita respecto al contenido que expresa, ¿por qué una oración de Mentalés no puede serlo también? La situación planteada por Perry y Clapp es la siguiente:

The Z-landers, recall, neither talk nor think about anywhere other than Z-land. They have no words for locations in their language. Perhaps it is best to imagine that the Z-landers cannot move, nor even think about moving. So there is no reason to suppose that the Z-landers have mental representations whose contents are Z-land, or any other location. Now, suppose a Z-weatherman sincerely report the weather by uttering 'It is raining'. Since, as Frege would remind us, the content of thoughts is objective, when we contemplate the truth of the content expressed and believed by the Z-weatherman, we must be contemplating the same content. And what do we judge these objective truth conditions to be? Specifically, where must it be raining for the Z-weatherman's assertion to be true? The intuitive response is that it must be raining in Z-land. Here we have, to put the point in Fodor's terms, a good reason for thinking that the implicitness of the Z-lander's utterance of 'It is raining' with respect to its content,

which contains Z-land is a “constituent,” is not adequately explained by supposing that the sentence ‘It is raining’ is inexplicit with respect to the Z-weatherman’s mental representation -there is no reason to suppose that the sentence is inexplicit with regard to the thought. But if the content we grasp is the content the Z-weatherman expresses, then there is no reason to think that the Z-weatherman’s thought is inexplicit with regard to this content. (Clapp, 2012, pp. 320)

El ejemplo muestra que no hay buenas razones para pensar que los pensamientos en términos de representaciones mentales como el lenguaje del pensamiento carezcan de componentes inarticulados, tal como las oraciones del lenguaje natural (Perry, 1986, p. 145. en Clapp, 2012, pp. 320). Esto abre la puerta a que las representaciones mentales, es decir, los pensamientos sean inexplicitos respecto al contenido que expresan.

Volviendo al ejemplo Travis-Clapp, cuando se nos pregunta de qué color son las hojas de la planta con la que adornamos nuestro escritorio (suponiendo que tenemos una), no hay motivos para pensar que nuestra creencia es algo como: ‘LAS HOJAS SON VERDES DEBAJO DE LA PINTURA QUE PUEDAN TENER’ o algo como ‘LAS HOJAS SON VERDES INCLUSO SI SOLO ESTAN PINTADAS CON PINTURA VERDE’; es probable que nuestra representación, como Fodor pretende, sea solamente ‘LAS HOJAS SON VERDES’. Si nos preguntáramos por las condiciones de verdad de nuestra respuesta, ¿sería *suficiente* con que las hojas estuvieran pintadas de verde, o se requeriría que las hojas fueran real y naturalmente verdes? A este respecto, Clapp argumenta que, si bien la pregunta tiene una respuesta, nuestra representación al momento de responder la pregunta es inexplicita respecto al contenido que intentaba expresar en la medida en que nuestro pensamiento se refiera a que simplemente son verdes.

If your assertion has truth-conditional content, if it expressed a proposition, there must be an answer to this question. But now you are in the same position with regard to the Z-weather man: If it is possible to this question to be answered correctly, then your mental representation at the time of utterance was inexplicit with regard to its content. But if it cannot be answered correctly because you cannot grasp the content of your past utterance, then we are in effect rejecting the notion of content altogether, for content is essentially shareable. (Clapp, 2012, pp. 321)

Así las cosas, por los argumentos antes analizados, parece que hay buenas razones para abandonar la composicionalidad del pensamiento (habiendo aceptado ya que el lenguaje

natural y que, por lo tanto, la conducta lingüística e inteligente no es composicional). Pues los argumentos de Fodor para sostener una posición mixta parecen no ser satisfactorios en la defensa de la composicionalidad del pensamiento.

El problema que surge a partir de ahora para la teoría fodoriana es que rechazar la composicionalidad del pensamiento implica perder un fuerte incentivo para el lenguaje del pensamiento. Es importante recordar que Fodor postula LOT, en parte, como respuesta a la pregunta de cómo representamos nuestra propia conducta (ver la segunda parte del capítulo II). La tensión que surge entonces es que, si bien la estructura de LOT es insuficiente para dar cuenta de nuestra conducta lingüística e inteligente, entonces el pesimismo respecto a esta teoría expresado por Churchland al inicio de este capítulo parece tomar mayor vigencia y quizá haga falta esperar una mejor teoría con mayor éxito explicativo. ¿Dónde nos deja esto?

* * *

A lo largo de la tercera parte de este capítulo se argumentó en contra de la posición mixta de Fodor, es decir, en contra de la posición que rechaza la composicionalidad de la conducta lingüística e inteligente en general, pero sostiene aún la composicionalidad del pensamiento. Los argumentos presentados se apoyaban en casos que mostraban que tanto los lenguajes naturales, como el lenguaje de pensamiento (LOT) dependían del contexto para dar cuenta de los contenidos expresados o involucrados, mostrando así que no hay composicionalidad ni en el lenguaje natural ni en el pensamiento.

Si los argumentos en contra de la composicionalidad del lenguaje natural aplican a la composicionalidad del pensamiento, como parece ser el caso, Fodor (y cualquiera que pretenda aceptar una teoría como LOT basada en la composicionalidad) pierde un fuerte incentivo para la necesidad de un sistema representacional tan altamente estructurado. En otras palabras, perder dicho incentivo equivale a no tener buenas razones para sostener una teoría del tipo antes mencionada en la medida en que la teoría no explica de manera adecuada (o en absoluto) el fenómeno entre manos: en este caso, la conducta inteligente.

Sin embargo, como pretendo argumentar en el siguiente capítulo, el éxito de las generalizaciones de la psicología popular no depende exclusivamente del éxito de la teoría fodoriana y puede ofrecerse una teoría alternativa (relativamente cercana a LOT) que defienda y explique el éxito de la psicología popular como teoría de la conducta, sin caer en los problemas de la teoría fodoriana presentados en este capítulo.

De esta forma, en el siguiente capítulo pretendo argumentar, de manos de las filósofas Susan Schneider y Claudia Picazo, que los argumentos de Clapp se basan principalmente en falencias del sistema representacional de Fodor, es decir, LOT. Sin embargo, puede ofrecerse una caracterización distinta del sistema representacional que resuelva o evada los problemas y situaciones mencionados por Clapp.

Capítulo IV

Una defensa de la psicología popular: rehabilitando LOT en un mundo no tan composicional

Hemos llegado al último capítulo de esta investigación. Como se ha mencionado al final del capítulo anterior, quisiera explorar una respuesta a los argumentos planteados por Clapp con el propósito de defender la necesidad de LOT (o un sistema representacional cercano a LOT). Sin embargo, antes de ofrecer esta respuesta, es preciso hacer una reconstrucción de la discusión que se ha presentado en los capítulos anteriores y contextualizar el por qué es deseable ofrecer una respuesta a Clapp.

En el capítulo I y II se presentó a la psicología popular como una forma (probablemente la mejor disponible) de descripción y explicación de la conductual lingüística e inteligente. Para hacer de la psicología popular una teoría científica y explicar su éxito en las predicciones y generalizaciones que ofrecía, debía apelarse a una teoría representacional. El motivo de esto descansa en que el éxito de la explicación descansaba principalmente en dos elementos: la composicionalidad de la conducta lingüística e inteligente y que los sujetos mismos parecían representarse su conducta como siendo de esta misma forma.

La hipótesis de la modularidad de la mente nos ofrecía la relación (y explicación) causal entre los procesos cognitivos y la conducta, y la hipótesis del lenguaje de pensamiento nos ofrecía la explicación de cómo los sujetos se representaban su propia conducta y explicaba la composicionalidad, propiedad esencial, según Fodor, de la conducta lingüística e inteligente.

Sin embargo, en el capítulo III, se desafió la idea de que la conducta inteligente era como la psicología popular la describía, y, *a fortiori*, la explicación ofrecida por la modularidad y el lenguaje de pensamiento eran inadecuadas (o incompletas como mínimo), y debían ser reemplazadas por una teoría que apele a la noción de red neuronal, donde las representaciones no eran simbólicas ni mantenían una estructura composicional.

Luego, se presentó la discusión Fodor-Churchland, donde se concluía que debía mantenerse una explicación como la sugerida por Fodor, porque solo apelando a una arquitectura cognitiva de naturaleza composicional podía ofrecerse una explicación de la composicionalidad de la conducta lingüística e inteligente.

Finalmente, en la última parte del capítulo III nos adelantamos un poco en el debate hasta un punto donde Fodor aceptaba los argumentos de los pragmatistas radicales en contra de la composicionalidad de la conducta, pero mantenía una postura mixta sosteniendo la composicionalidad, aún necesaria, del pensamiento (esto es, para las representaciones Mentalés). Los argumentos de Clapp mostraban que las razones de Fodor para sostener su posición mixta eran insuficientes y que las conclusiones en contra de la composicionalidad de la conducta lingüística e inteligente aplicaban en contra de la composicionalidad del pensamiento (esto es, de las representaciones Mentalés).

¿Qué nos queda de la discusión? Lo primero que hay que decir, es que no todo está perdido. Hay tesis o intuiciones acerca de la arquitectura mental que parecen no estar en cuestión (al menos directamente) por los protagonistas del debate, mientras que otras tesis sí lo están y son el centro de esta discusión. Para dibujar mi posición en este debate, quiero, en primer lugar, señalar cuáles tesis no están en cuestión y cuáles sí lo están, y luego mencionar cómo se conectan con este último capítulo.

Las tesis que parecen no estar en disputa por parte de los autores protagonistas de la discusión hasta el momento son las siguientes:

- A. *La tesis funcionalista acerca de lo mental*, junto con la afirmación de que la conducta inteligente es el resultado causal de los mecanismos responsables de la conducta.
- B. *La tesis de la modularidad de la mente*. La idea de que hay módulos o mecanismos “internos” especializados en tareas o en procesos los cuales producen y explican la conducta.⁵⁷

Las tesis que sí parecen estar en cuestión son las siguientes:

1. *La modularidad al estilo Fodor*: Churchland pone en tela de juicio la necesidad de un **procesador central**, como el defendido por Fodor, que integre la información proveniente de los módulos y sea el principal mecanismo que explica la conducta inteligente.
2. *La hipótesis del lenguaje del pensamiento*. Al dudarse de la plausibilidad del procesador central, se duda también de LOT que ofrecía una descripción de las computaciones que se llevaban a cabo en el procesador central.

⁵⁷ Tanto Fodor como Churchland (y quizá Clapp de manera menos explícita) están de acuerdo con estas ideas generales sobre la forma de la explicación. Lo que nos ha mostrado la discusión es en cómo entender estas ideas y a qué clase de mecanismos deben apelarse para explicar la conducta, cuya descripción, también está en cuestión.

3. *La afirmación o intuición de la composicionalidad de la conducta lingüística e inteligente.* Si bien es posible argumentarse a favor de 1 y 2 apelando a que estas tesis ofrecían explicación de la composicionalidad de la conducta lingüística e inteligente, pero con los argumentos de los pragmatistas radicales y Clapp, se duda de que la conducta inteligente y el pensamiento (como su causa) sean en algún sentido composicionales.
4. *La afirmación de que la psicología popular (tal como está explicada aquí) es una explicación correcta de la conducta inteligente.* Dudar de la tesis 3 da como resultado la ausencia de una teoría plausible que explique el éxito de la psicología popular y, por consiguiente, su plausibilidad como una adecuada descripción de la conducta inteligente.

Es en este punto en el que pretendo presentar una respuesta a los argumentos de Clapp. Si las hipótesis que sugiero a continuación son correctas, entonces puede ofrecerse una teoría cercana al LOT (tal como originalmente lo presentó Fodor) que conserva sus virtudes, es decir, la explicación apelando a las categorías y conceptos de la psicología popular, pero que logra evadir los problemas computacionales señalados por Clapp en sus argumentos, y, adicionalmente, se apoya en una noción de composicionalidad más austera que no parece ser rechazada por los argumentos de Clapp, lo que rescata la idea de la necesidad de un sistema representacional composicional.

¿Cuál era el argumento de Clapp?

Vale la pena hacer una reconstrucción de la crítica de Clapp a la composicionalidad del pensamiento. Lenny Clapp (2012) ofrece un argumento relativamente sencillo en contra de la composicionalidad del lenguaje de pensamiento, el cual está basado en la no-composicionalidad de la conducta lingüística e inteligente. Los pragmatistas radicales, se dice en el argumento, sostienen que la conducta lingüística no es composicional (o en otras palabras que el lenguaje mismo no es composicional) puesto que lo que dicen las personas al usar el lenguaje verbal está determinado profundamente por el contexto de emisión y no por elementos que son idénticos a través de diferentes contextos y pueden recombinarse (lo mismo podría decirse de la conducta inteligente en general). De acuerdo con Clapp, Fodor acepta estos argumentos y concede que la conducta lingüística no es composicional en este sentido. Sin embargo, Fodor no renuncia a la idea de que los pensamientos, aquellas

entidades que causan y explican la conducta lingüística e inteligente (esto es, las oraciones del lenguaje de pensamiento instanciadas en el cerebro), sean composicionales; para él, el pensamiento aún es composicional.

El argumento de Clapp apunta a mostrar que Fodor no tiene razones para evitar que las conclusiones del pragmatismo radical apliquen de igual forma al pensamiento. Si los argumentos del pragmatismo radical en contra de la composicionalidad del lenguaje natural aplican, de igual modo, al pensamiento, entonces el pensamiento tampoco puede ser composicional.

Según lo dicho en el capítulo anterior, esta conclusión es relevante para esta investigación por la siguiente razón. Si los pensamientos, entendidos como oraciones de LOT, no son composicionales, entonces no hay una razón para sostener que la mente use un tipo de lenguaje representacional como LOT (el cual es esencialmente composicional) y, por lo tanto, tampoco tenemos razones para creer en la existencia de un procesador central (al estilo Fodor) que anide estados internos (como creencias, deseos, percepciones conscientes) y que *compute* este tipo de representaciones para causar la conducta inteligente (sea lingüística o no). Renunciar a una arquitectura mental de este estilo, desmotiva el uso de la psicología popular para describir y dar cuenta de la conducta inteligente en general.

En este sentido, las implicaciones del argumento de Clapp son devastadoras para una teoría representacional de la mente que use LOT como el lenguaje en el que se computan las representaciones mentales. Podría argumentarse que la no-composicionalidad de la conducta lingüística no solo elimina las motivaciones para que LOT sea composicional, sino que también elimina las motivaciones para tener una explicación y descripción de la conducta lingüística e inteligente en las líneas mantenidas por la psicología popular.

Por esta razón, no puedo estar de acuerdo con las conclusiones de Clapp respecto a la composicionalidad de LOT. Pienso que aún es posible defender alguna versión de la composicionalidad del pensamiento a pesar de aceptar (junto con Fodor) que la conducta lingüística e inteligente *no es* (o no parece ser) *composicional*.

Para sostener mi posición, a lo largo de este capítulo, ofreceré una plausible respuesta a los argumentos de Clapp al mostrar cómo cada uno de ellos es insuficiente para probar su punto. Y, a partir de allí, espero trazar conclusiones diferentes respecto a la *plausible* composicionalidad de LOT y la necesidad de un sistema representacional que cause la conducta lingüística e inteligente, pese a aceptar que ésta no es composicional en un sentido tan robusto como Fodor pretende señalar.

Mi plan de defensa

El esquema de argumentación de este capítulo será el siguiente. Debido a que Clapp usa dos argumentos, principalmente, para desafiar la composicionalidad de LOT, mi estrategia será tratar y desarmar cada uno de los argumentos. El primer argumento de Clapp depende de la idea de que un mismo *tipo de oración* del lenguaje del pensamiento puede expresar dos contenidos distintos en contextos distintos. Y, si esto es correcto, este tipo de oración (y su contenido) no puede ser composicional en estricto sentido (al menos en el sentido razonable acuñado por Clapp). Mi estrategia para tratar y superar este primer argumento consiste en ofrecer un criterio de individuación de pensamientos en cuanto que ofrezco un criterio para individuar diferentes *tipos* de oraciones de LOT y diferentes *casos particulares* de oraciones de LOT. Es decir, un criterio diferente para individuar representaciones mentales. Así, de la mano de Susan Schneider (2011), ofreceré una concepción diferente de la noción fodoriana de LOT, la cual logra esquivar estos defectos computacionales que Clapp aprovecha para argumentar en contra de la composicionalidad de LoT.

Esta nueva noción se alimenta del pragmatismo presente en la discusiones de la composicionalidad del lenguaje natural y la conducta inteligente y me permite introducir el concepto de atomismo pragmático: un criterio de individuación de oraciones de LOT con un tono pragmatista. Por esta razón, la nueva versión de LOT será nombrada como PLOT siguiendo a Schneider (2011).

El segundo argumento de Clapp descansa en la idea de que las oraciones de LOT pueden ser *inexplícitas* respecto del contenido que expresan, esto es, que no tienen todos los elementos sintácticos para determinar cuál es el contenido específico o cuál es el pensamiento expresado en un caso particular. Si esto es correcto, entonces LOT no tiene los elementos suficientes para componer el contenido o pensamiento expresado a partir de partes sus partes: hay componentes que le hacen falta. Y, por lo tanto, LOT no puede ser composicional.

Mi estrategia parte del reconocimiento que de que PLOT no es suficiente para responder esta objeción, pero es aún útil en la explicación. Por tal razón, apoyándome en el trabajo de Claudia Picazo (2019), diré que las oraciones particulares de PLOT lucen inexplicitas porque el contenido expresado está indeterminado y, en este sentido, el contenido que se expresa no es completo. Para esto, uso la tesis de Picazo acerca de la '*Token-underdeterminacy*' de las representaciones mentales. De acuerdo con esta tesis, las oraciones particulares (esto es, las oraciones-caso u oraciones-token) del lenguaje de pensamiento

expresan contenidos indeterminados en cuanto que sus condiciones de verdad lo están y, su aparente determinación, esto es, la aparente estabilidad de las condiciones de verdad está ubicada solamente en el nivel del *tipo* de oraciones (esto es, en las oraciones-tipo o type-oraciones) del lenguaje de pensamiento. Si esto es correcto, pese a que la conducta lingüística e inteligente no sea composicional en un sentido fuerte (esto es, que tenga partes identificables en todos los contextos), aún podría sostenerse alguna noción de composicionalidad en el pensamiento (esto es, en las representaciones mentales).

Finalmente, pretendo describir cómo la conjunción de la tesis del atomismo pragmático, PLOT y la ‘token-underdeterminacy’ aún abren un espacio razonable para dar cuenta por qué, pese a que la conducta lingüística e inteligente no sea composicional, las intuiciones de la psicología popular que motivan nuestra explicación científica de la conducta inteligente (esto es, la explicación basada en relaciones causales y generalizaciones nomológicas) aún tienen lugar.

La tesis terminará con unas breves conclusiones tanto del cuarto capítulo, como del documento en general. Allí, pretendo decir algo respecto a la composicionalidad y diré algunas palabras respecto al estado final (a la luz de la discusión planteada en este documento) de PLOT y la teoría representacional como teorías explicativas de la conducta.

Una respuesta a los dos argumentos de Clapp

Para empezar, es de mi interés enfatizar, tal como lo sugerí en los párrafos anteriores, que los argumentos de Clapp tienen una función central: dejar al descubierto algunos problemas de índole fundamentalmente *computacional* atribuibles a LOT y que son explotados por Clapp para mostrar que LOT no es composicional en absoluto.

La intuición general de mi rechazo a los argumentos de Clapp es que estos se apoyan en una caracterización más bien pobre respecto a los *criterios de individuación* de los símbolos de LOT. Y esto es relevante en la discusión. Como lo dice Schneider:

Clearly, for the symbolic conception to be of value to cognitive research, symbols need to be well defined so that their relationship to other computational accounts can be clarified and so that predictions involving symbols can be made. (Schneider, 2009, pp. 239)

En este sentido, considero que ofrecer un criterio más robusto para los símbolos puede ser un punto de partida para ofrecer una respuesta a los argumentos de Clapp y una defensa de LOT.

En contra del primer argumento: PLOT

Veamos el primer argumento de Clapp. Recuérdese que el éxito de la explicación intencional propuesta por Fodor recae en la idea en que las generalizaciones sobre la conducta de los sujetos (que están determinadas por el contenido amplio, es decir de la relación del sujeto con el mundo) puedan ser exitosamente capturadas por un sistema representacional que expresa contenidos semánticos y está armonizado con el nivel computacional (ver la segunda parte del capítulo II). La razón es doble: por un lado, Fodor insiste en que las generalizaciones de la psicología popular son esencialmente intencionales: “*By definition, an intentional generalization is one that subsumes psychological states by reference to their intentional contents*” (Fodor, 1994, pp. 4). Y, por otro lado, también insiste en que las generalizaciones o leyes se encuentran instanciadas computacionalmente dentro de la cabeza de cada individuo, pues los procesos cognitivos son esencialmente computacionales. En este sentido, aunque un pensamiento se individúe por criterios intencionales -es decir, por el contenido del pensamiento- debe tener un correlato computacional, esto es, una oración en Mentalés que expresa o captura el contenido intencional.

El ejemplo Clapp-Travis, que es la base del argumento de Clapp, presenta la posibilidad de que falle la “armonía” entre el nivel semántico —*broad content*— y el nivel computacional —*narrow content*—.

El problema inicia con la idea de que es posible que dos instancias de la expresión-tipo de LOT: ‘LAS HOJAS SON VERDES’ expresen condiciones de verdad o contenidos diferentes. Por ejemplo,

Condiciones de verdad (a) ‘Las hojas *son naturalmente* verdes’; y

Condiciones de verdad (b) ‘Las hojas *han sido pintadas de* verde’.

Esto es un problema al intentar responder la pregunta: ¿cómo pueden dos tokens de oraciones de Mentales que pertenecen a un mismo *tipo* expresar dos contenidos diferentes? Como se señaló en la sección de ‘Lenguaje de Pensamiento’ en el capítulo II, se esperaría que

las representaciones-tipo determinen a los tokens que subsumen, pues parte de la idea de que un token sea un token de cierto tipo se explica porque preserva las propiedades de dicho tipo.

Este ejemplo parece ser un fallo en la teoría de Fodor. La razón es que las condiciones de verdad **(a)** y **(b)** son diferentes, lo que quiere decir que son pensamientos diferentes. Esto implica un fallo en la teoría en la medida en que dos instancias de oraciones de Mentalés –de un mismo tipo- parecen expresar dos contenidos –esto es, dos pensamientos- diferentes; sin embargo, esto es absurdo: no debería ser posible que dos instancias que pertenecen al mismo tipo de representación y, por tanto, al mismo tipo de pensamiento, expresen contenidos diferentes y aún sigan siendo dos instancias de un mismo tipo.

Tal como lo entiendo, lo primero que debemos resaltar es que el fallo parece darse en el nivel *computacional*. Mi intuición para decir esto es que la manera en la que individuamos las representaciones da lugar al problema.⁵⁸ Y, por esta razón, si mi intuición es correcta, la manera de sortear este problema es sugerido en su misma formulación. En otras palabras, hace falta revisar los criterios de individuación que Fodor ofrece para los símbolos de LOT con el propósito de señalar exactamente por qué LOT falla en circunstancias como las del ejemplo Travis-Clapp.

Clapp concluye que el pensamiento no puede ser composicional.⁵⁹ Sin embargo, dado que creo que el problema es computacional, sugiero que el argumento puede verse redescrito como un ‘caso de gemelos’ (*twin cases*), es decir, como un tipo de caso de los sugeridos por Putnam (1975). Y esto trae una ventaja: Fodor (1994) tiene una respuesta a los casos de gemelos que, según creo, puede extenderse al ejemplo Travis-Clapp.

Para no entrar en detalle y hacer más confusa la presentación, los casos tipo gemelos sugieren que dos individuos físicamente idénticos y, por lo tanto, computacionalmente idénticos, pueden tener o expresar contenidos diferentes. En otras palabras, dos instancias computacionales pertenecientes al mismo tipo (*type*) de representación, expresan o capturan contenidos amplios (*broad content*) diferentes.

⁵⁸ Creo que el caso presentado parece neutral respecto a la individuación del contenido amplio —aun cuando Clapp señala que el contenido amplio no puede ser composicional según los criterios establecidos. Imagine que la discusión sobre la composicionalidad del pensamiento y/o la conducta deja de ser el centro de la discusión. En ese caso, aún es posible presentar el ejemplo Travis-Clapp como un contraargumento a LOT en la medida en que señala que las representaciones no son lo suficientemente “cuidadosas” en su individuación para distinguir entre contenidos distintos, lo que da lugar a predicciones erróneas o a contraejemplos de dichas predicciones.

⁵⁹ Una conclusión más fuerte podría ser que las computaciones llevadas a cabo en el procesador central no son como LOT sugiere; o que quizá ni siquiera sean computacionales.

Si reinterpretáramos el ejemplo Travis-Clapp bajo esta perspectiva, podríamos decir que dos individuos (o el mismo individuo en dos momentos de tiempo diferentes) física y computacionalmente idénticos, cuando tienen el pensamiento ‘las hojas son verdes’, expresan contenidos diferentes. Sean S1 y S2 sujetos distintos:

S1: tiene una representación del tipo *LAS HOJAS SON VERDES* que expresa el contenido/pensamiento ‘las hojas son naturalmente verdes’.

S2: Tiene una representación del tipo *LAS HOJAS SON VERDES* que expresa el contenido/pensamiento ‘las hojas son artificialmente verdes’.

El caso presentado solo representa un ataque a LOT si se logra mostrar que en ambos casos las instancias de ‘LAS HOJAS SON VERDES’ constituyen dos tokens del mismo tipo. Hasta ahora no se ha mencionada mucho (si es que algo) respecto a la forma en que deben individuarse las representaciones de Mentalés más allá de que están causalmente relacionadas con el mundo y parecen expresar un cierto contenido (ver segunda parte del capítulo II).

Es en este punto donde el *atomismo pragmático* (o por sus siglas en inglés *pragmatic language of thought*, PLOT) ofrece una solución al problema. PLOT nos ofrece un criterio de individuación: el atomismo pragmático sostiene que las condiciones de individuación de los símbolos simples con los que la mente computa representaciones están determinadas por el rol causal que dicho símbolo exhibe en toda la red cognitiva. Dentro de este rol causal se encuentra, por ejemplo, las relaciones causales por objetos en el ambiente y las relaciones causales de otras representaciones que causan el símbolo simple; y las conductas u otras representaciones que son causadas por este simple. Así lo quiere mantener Schneider: “First, I’ve observed that LOT’s naturalistic program contends that intentionality is a matter of a nomic or informational connection between symbolic representations in the brain and entities in the world.” (Schneider, 2011, pp. 55)

Esto nos deja con un sistema representacional atómico porque apela a símbolos simples, sin partes, pero combinables en símbolos más complejos (en este sentido, no es como la teoría propuesta por Churchland) y es pragmático porque la individuación de los símbolos está determinada *principalmente* por el rol o uso que tenga dentro de toda la economía cognitiva. El criterio de individuación ofrecido por PLOT es el siguiente:

(CAUSAL): *A symbol is defined by the role it plays in the algorithm that describes the central system.*” (Schneider, 2009b, pp. 540-541)

Con esto en mente, volvamos a la distinción *type/token* presentada el Capítulo II. Los lenguajes representacionales como LOT y PLOT soportan dicha distinción, la cual brinda los primeros rasgos de individuación de los símbolos. Se mencionó (véase el pie de página 36 en la subsección 3 de la Parte II del Capítulo II) que, dos tokens pertenecen al mismo tipo cuando son intercambiables sin alterar de ninguna manera el proceso computacional:⁶⁰

1. “T1 and T2 are tokens of the same type if, for any machine process (i.e., for any operation compatible with the machine's table), the numerically same output state that was (or would have been) produced by the machine when it is given T1 as input would have been produced if the input had been T2; and vice versa.
2. If T_i and T_j belong to the same symbol type by criterion 1, then T_m and T_n belong to the same symbol type if any machine process that yielded (or would have yielded) T_i as output given T_m as input would have issued in T_j as output if the input had been T_n.
3. The type / token relation for primitive expressions is closed under 1 and 2.” (Fodor, 1995, pp. 109)

Estos mismos criterios pueden valer igualmente para PLOT, sin embargo, aún no son suficientes para dar cuenta del problema que deja al descubierto el caso de Travis-Clapp o el caso gemelos. Si bien los criterios son claros respecto a la identidad de los tokens, no son explícitos respecto a la identidad de los tipos. Es aquí donde PLOT va más allá y propone un criterio más robusto para individuar símbolos basada en la condición (CAUSAL) antes presentada. Schneider ofrece el siguiente argumento:

(T1): Any token can be substituted by another token of the same type in an operation without changing the computation.

(T2): Any token of the same type will generate the same (proximal) output and internal states of the machine, given that the same internal states of the machine are activated.

(T3): (x) (y) (if x and y are type-identical LOT tokens, the x and y will have the same total computational role).

(T4): (x) (y) (if x and y are type-distinct LOT tokens, then x and y have distinct total computational roles).

⁶⁰ Volviendo a la metáfora con el ajedrez, don piezas pertenecen al mismo tipo –por ejemplo, peón– cuando, al ser intercambiadas, siguen teniendo las mismas posibilidades de moviente según su posición en el tablero de juego.

(T5): (x) (y) (if x and y have the same total computational role, then x and y are type-identical) (Schneider, 2009b, pp. 527-528)

Ahora vemos que (CAUSAL) no es solo un criterio arbitrario, sino que los mismos criterios de identidad de los tokens parecen sugerirlo. Con este nuevo criterio de identidad para los símbolos ¿se puede ofrecer una respuesta al aso Travis-Clapp?

Bajo los términos introducidos en los párrafos anteriores, considero que el primer argumento de Clapp puede ser esquematizado de la siguiente manera:

1. x y y son dos tokens de LOT.
2. x y y son tipo-idénticos, es decir, son dos instancias del mismo tipo. [Sin embargo,]
3. x expresa el contenido ‘las hojas son naturalmente verdes’ y y expresa el contenido ‘las hojas son artificialmente verdes’. [Por lo tanto,]
4. Es posible ofrecer un contraargumento contra la explicación ofrecida por LOT de la mente. [Así, como corolario, se sigue]
5. LOT es insuficiente para explicar la conducta inteligente en la medida en que puede presentársele contraejemplos o no puede capturar todas las generalizaciones relevantes de la conducta inteligente.

Podemos ahora ver más claramente que es la segunda premisa la que parece ser falsa (al menos tiene que serlo para seguir manteniendo un sistema representacional como LOT o cercano como PLOT). Para que x y y sean dos tokens del mismo tipo, hace falta que compartan la totalidad de su rol computacional y la totalidad de los algoritmos en los que aparecen. Sabemos que este no es el caso, pues expresan contenidos diferentes y, por lo tanto, su rol en la conducta del sujeto es diferente:

The upshot: it seems that we have mental processing that is not determined by LOT syntax together with algorithms, contra CTM. Hence, to preserve CTM, we must regard a and b as being type-distinct symbols. The reduction leaves us with the following constraint on LOT property individuation:

(Principle P) (x) (y) (It is not possible for x and y to be tokens in the same system, of the same type, and differ in their computational roles.) (Schneider, 2009b, pp. 530)

Los nuevos criterios ofrecidos para la individuación de los símbolos de PLOT parecen excluir exhaustivamente los casos tipo gemelos, donde un mismo estado computacional puede expresar dos contenidos diferentes y aun así pertenecer al mismo tipo de símbolo. En

este sentido, la respuesta al problema presentado por el caso Travis-Clapp parece, en principio, ser imposible desde los nuevos criterios de individuación de los símbolos del sistema, al menos no es obvio cómo este tipo de casos puede presentarse (sobre todo sistemáticamente) bajo los nuevos criterios.

En este sentido, la respuesta a Clapp debería ser que la oración de Mentalés ‘Las hojas son verdes’ solo expresa un contenido o pensamiento ya bien sea ‘Las hojas son naturalmente verdes’ o ‘Las hojas son artificialmente verdes’; esto no quiere decir que solo podamos expresar o tener alguno de los dos pensamientos, más bien, la respuesta señala es que debería haber representaciones diferentes para expresar los contenidos antes mencionados.

Esto, sin embargo, nos lleva al argumento de la explicitud, es decir, el otro caso presentado por. Puesto que la respuesta antes presentada podría implicar que el pensamiento —o las representaciones de LOT— siempre son explícitas respecto al contenido que estas expresan. Por ejemplo, cualquier persona que tiene el pensamiento ‘Las hojas son verdes’ siempre sabría cuál es el contenido de dicho pensamiento y sabría si se está refiriendo a las hojas naturalmente verdes o artificialmente verdes. Sin embargo, Clapp nos presentó un caso en el que no siempre es explícito el contenido de las representaciones; y, por tanto, distinguir entre representaciones no termina de ser una solución definitiva al problema.

En contra del segundo argumento: la tesis de la indeterminación de la representación

Veamos ahora el caso de los Z-landers o el argumento de la explicitud. ¿Bastan los nuevos criterios de individuación de los símbolos para dar respuesta al problema que plantea este caso? Seguramente la respuesta de Clapp es **No**. Veamos cómo sería el argumento.

El caso de los Z-landers es un recurso para mostrar la posibilidad de que las representaciones sean implícitas respecto al contenido que expresan. ¿Por qué esto representa un problema para la explicación ofrecida por LOT y PLOT? Si el caso de los Z-landers, de hecho, resulta verdadero para PLOT, entonces es posible que la teoría no sea capaz de capturar todas las generalizaciones relevantes sobre el contenido. Después de todo, si las representaciones son implícitas, entonces PLOT no es suficiente para implementar las generalizaciones de la psicología popular y hace falta algo más de carácter no-computacional; y, por lo tanto, la PLOT es incompleta.

Esta explicación no puede permitirse tal etiqueta, pues sus postulados mismos señalan que los procesos cognitivos *son solo* computacionales. Por lo que se debe ofrecer una respuesta al caso mencionado. El Caso podría expresarse de la siguiente manera:

1. PLOT no puede ser inexplicito respecto al contenido que expresa
2. Los Z-landers —que no tienen representaciones para lugares— pueden expresar el contenido ‘está lloviendo en Z-landia’. [Se sigue que]
3. Las representaciones son implícitas respecto al contenido que expresan.
4. Hay un Z-lander en cada uno desde que es posible tener la representación ‘LAS HOJAS SON VERDES’ y que ésta pueda expresar varios contenidos diferentes. [Se sigue que]
5. Las representaciones de LOT y PLOT pueden ser implícitas respecto a su contenido. [Por lo tanto,]
6. La explicación ofrecida por PLOT es insuficiente (o incorrecta) [Por 1 y 5]

Con la defensa presentada en la sección anterior podemos decir que la premisa *cuatro* es la que plausiblemente sea falsa en el argumento. El ataque que Clapp intenta presentar no toma (y no debería tomar) su verdadera fuerza en el caso mismo de los Z-landers, pues a la psicología de nuestro mundo difícilmente le puede interesar la psicología y conducta de individuos tan nomológicamente improbables (como Fodor (1994) mismo dice respecto al XYZ en los casos tipo gemelos de Putnam):

Empirical theories are responsible only to generalizations that hold in nomologically possible worlds. [...]. So much, then, for XYZ; there isn't any, and there couldn't be any, and so we don't have to worry about it. Unless we've done our science all wrong, it's a moral certainty that the nearby worlds in which our computational psychology holds are proper subset of the nearby worlds in which our chemistry does. (Fodor, 1995, Pp. 29)

El ataque real parece estar en la sugerencia de que es posible —y aparentemente común— encontrarnos ordinariamente en la posición de los Z-landers, tal como el siguiente pasaje parece indicar:

Fodor might of course deny that the content we grasp is the content the Z-weatherman expresses. Z-landers are, we must admit, outlandish. The problem with this response is that, as Perry would remind us, *there is a little Z-lander in all of us*: When in your ordinary day you sincerely assert that the leaves are some ordinary looking plant are

green there is no reason to suppose you have either something like THE LEAVES ARE GREEN UNDERNEATH WHATEVER PAINT THERE MIGHT BE ON THEM or something like THE LEAVES ARE GREEN EVEN IF ONLY BECAUSE OF GREEN PAINT in your belief-box. (Suppose your utterance is in response to a query about the color of an ordinary plant in your office.) The possibility of the leaves looking green as a result of being painted green is as foreign to you at that moment as the possibility of its raining in Palo Alto is to the Z-weatherman. Now suppose that later I ask you what the truth conditions of your assertion are: Does your assertion require for its truth that the leaves be really green, underneath whatever paint there might be, or would it suffice for its truth that the leaves be merely painted green? If your assertion has truth-conditional content, if it expressed a proposition, there must be an answer to this question. But now you are in the same position with regard to your question to be answered correctly, then your mental representation at the time of utterance was inexplicit with regard to its content. (Clapp, 2012, pp. 320-321, énfasis añadido)

Si volvemos al argumento expuesto en la sección anterior, es posible decir que la pregunta planteada sí tiene una respuesta y no implica estrictamente que las representaciones sean inexplicitas, pues como ya se señaló deberían ser dos representaciones diferentes para los dos contenidos posibles a expresar. ¿En qué recae entonces la confusión que surge y nuestro aparente parecido con los Z-landers?

La respuesta puede ser mucho más amplia de lo que esta investigación permite. Pero para ilustrar mi respuesta, quisiera usar una de las características y roles más importantes que las representaciones de LOT y PLOT; a saber, las representaciones deben cumplir el papel de ser *Modos de Presentación* (o MOPs de ahora en adelante):

The most crucial philosophical role that symbolic representations are supposed to play concerns their role as MOPs. Consider: there are different ways of representing the same entity. Despite differences in ways of representing things, diverse ways can pick out, or refer to, the same entity. (Schneider, 2011, pp. 94)

Tal como la cita lo sugiere, las representaciones ofrecen la posibilidad de explicar que una misma cosa u objeto pueda ser representada de formas diferentes. Con esto en mente, quizá pueda ser cierto que tanto los Z-landers como los seres humanos puedan expresar el mismo contenido (broad content o condiciones de verdad), pero no hay razón para suponer

que compartimos el mismo símbolo (después de todo, un Z-lander se comportaría de manera bastante diferente a un ser humano).

Por ello, incluso si “hay un Z-lander dentro de cada uno de nosotros” tenemos las herramientas para explicar cómo el sistema representacional permite esto; ya sea porque el rol computacional de los símbolos varía —y aunque parezcan el mismo símbolo, son dos símbolos diferentes—, o porque la función de MOP que tenemos por representación varía de sujeto a sujeto.⁶¹

Apelar a la función de MOP de las representaciones nos puede ayudar, probablemente, solo con algunos de los casos en que los pensamientos parecen ambiguos respecto al contenido que expresan; incluso si el contenido es el mismo y la conducta difiere.⁶² Sin embargo, nosotros podríamos hacer un poco más de presión con el caso Travis-Clapp.

Cuando alguien tiene el pensamiento —es decir, cuando tiene la oración de Mentalés— ‘LAS HOJAS SON VERDES’, quizá podría decirse que *no hay condiciones específicas de verdad para dicho pensamiento*, es decir, quizá lo que el sujeto piensa con dicha oración de Mentalés es distinto tanto a S1 (Las hojas son naturalmente verdes), como a S2 (Las hojas son artificialmente verdes); quizá piensa *solamente* S3:

S3: ‘Las hojas son verdes’.

¿Qué podría decir LOT o PLOT frente a este tipo de casos? Apelar a diferentes representaciones ya no parece ser una solución viable. La razón es la siguiente: si, por ejemplo, poseemos distintas representaciones para cada posible condición en la que podemos aplicar el predicado ‘verde’, entonces no tendría sentido tener un concepto general de verde, por lo que se rechaza la composicionalidad que tanto se intenta defender en LOT y PLOT.

En todo caso, pese a que LOT y PLOT no parezcan tener una respuesta explícita a esta pregunta (o al menos tanto Fodor como Schneider no contemplan este problema), es

⁶¹ Cuando este tipo de casos suceden, probablemente lo mejor que podemos hacer es preguntar a la otra persona por aclaración. Sin embargo, esto no representa un reto a PLOT, por lo que no hay razón para suponer que son un contraejemplo y por ende no hay que rechazar la explicación.

⁶² Piense por ejemplo en los casos tipos Frege, donde la referencia es la misma, pero el individuo no sabe o no se percata de ello y, por lo tanto, actúa de formas distintas frente al mismo objeto. Una presentación y tratamiento interesante de estos casos por parte del simbolismo aparece en Fodor (1995).

posible ofrecer una respuesta que no se deriva de LOT ni PLOT, pero que es consistente con (no excluye a) estas tesis.⁶³

La respuesta depende de sostener que haya una especie de indeterminación de los tokens de las representaciones. Pero antes de ello, veamos en qué consiste que las representaciones-tipo sean indeterminadas.

Tal como está planteado el caso propuesto por Clapp, se parece sugerir que son las representaciones-tipo de LOT las que son implícitas respecto al contenido que expresan, donde esto se puede entender de la siguiente manera:

Type-Underdeterminacy: A non-indexical structured representational item S is type-undetermined if and only if there are tokens of S that have distinct truth-values.

(Picazo, 2019, Pp. 3)

En este sentido, el pensamiento u oración de Mentalés ‘LAS HOJAS SON VERDES’ sería indeterminada-tipo en la medida en que dos —o más— de sus tokens pueden llegar a expresar valores de verdad diferentes. De esta forma, podría seguirse la tesis de que cualquier sistema representacional —en este caso las oraciones de LOT y, por lo tanto, el pensamiento— que cumple el principio de ‘*Type-underdeterminacy*’ no puede ser composicional. En primera instancia, porque las representaciones-tipo no determinarían un único valor de verdad o contenido, por lo que el significado o contenido de sus representaciones complejas no termina de ser explicado por las representaciones simples constituyentes y su estructura sintáctica.

Recanati (2007, p. 13–14) asks whether it is possible to find thought-vehicles with different truth-conditions at different contexts. His affirmative answer is supported by the following example. Imagine a subject who entertains the thought corresponding to ‘This man is happy’ while looking at a certain man —Bob. His thought is true if and only if Bob is happy. Now, had the context been different, the truth-conditions of his thought-vehicle could have been different. Suppose that it is Bill, not Bob, the man that the subject is looking at. In this case, the truth-conditions of his thought-vehicle would depend on Bill’s properties, not Bob’s. It seems that the same thought-vehicle can have different truth-conditional content. (Picazo, 2019, pp.12)

⁶³ Esto quiere decir que tanto LOT y PLOT no implican los postulados que voy a presentar a continuación.

Podría, aún, haber una salida disponible. Esta es entender la indeterminación de las oraciones de Mentalés no al nivel de los tipos, sino al nivel de los tokens. Es decir, la tesis de la ‘*indeterminación-token*’ sostiene que son los tokens de las representaciones las que pueden llegar a ser indeterminadas respecto al contenido que expresan, explicando, de este modo, la posible inexplicitud de las representaciones:

A token of a structured representational item *S* is token-underdetermined if and only if for some possible states of affairs its truth-value is indeterminate (I.e., if and only if it determines a *partial* function from possible worlds to truth-values) (Picazo, 2017, pp. 17-18)

Si volvemos al ejemplo Travis-Clapp, se puede ofrecer ahora una doble respuesta que le brinda mayor poder explicativo a PLOT. El pensamiento “Las hojas son verdes” expresa un contenido que o bien está determinado y por tanto las representaciones involucradas están distinguidas mediante la totalidad del rol funcional; o bien las condiciones de verdad no están determinadas; y al no estar determinadas, los tokens de la representación ‘LAS HOJAS SON VERDES’ expresan solo una parte de las condiciones de verdad expresadas por las representaciones-tipo.

Veamos ahora donde nos deja esta respuesta.

La indeterminación de las representaciones: la fuerza de esta respuesta

Veamos ahora cómo estas hipótesis podrían ofrecer una explicación de la psicología popular y complementar PLOT. Volvamos nuevamente al ejemplo (1) del Capítulo I (Ver pág. 11). Cuando Felipe tiene el pensamiento —o creencia o deseo, etc.— “Debo llegar al trabajo” sabemos ahora que la explicación implica que esa creencia esté codificada en representaciones dentro de su cabeza que serán las responsables causales de su conducta. Incluso, Felipe puede decir algo como “Iré en el servicio público”, también producto de las representaciones en su cabeza.

Lo que plausiblemente Clapp y los pragmatistas radicales mencionen en este punto es que las representaciones en cuestión, es decir, las representaciones que Felipe y cualquier persona tiene en su cabeza en el momento de comportarse como lo hace, son inexplicitas respecto a su contenido. Por ejemplo: si le preguntara, a usted lector, “¿de qué color es el pasto?” Probablemente su respuesta sea algo como “verde, por supuesto”. Si recurrimos al lenguaje de la teoría del *lenguaje de pensamiento*, una cadena de computaciones sobre las

representaciones adecuadas es el proceso responsable por la conducta realizada, es decir, su emisión de la oración.

Sin embargo, ¿cuáles son esas representaciones adecuadas? ¿cuál es el criterio para determinar el contenido de ‘verde’, y, por tanto, la representación responsable de la oración emitida? Esto, podría decir Clapp, muestra nuevamente que PLOT o LOT no son suficientes para explicar la conducta inteligente.

Sin embargo, bajo la perspectiva que nos brinda la tesis del *token-underterminacy* podría responderse que las representaciones-token que juegan el papel causal relevante.⁶⁴ son las únicas (en contraposición a las representaciones-tipo) que expresan un contenido parcial o incompleto. Esto se puede entender de la siguiente manera: las representaciones-tipo son una *idealización* y generalización que expresa un contenido completo y totalmente determinado. Por ejemplo, la representación-tipo ‘GATO’ expresa el contenido ‘gato’, contenido que está determinado de manera causal por el mundo: “Proponents of LOT are generally *externalists* about mental content, holding that thought content is not solely determined by one’s internal states; instead, content depends upon entities in the world as well” (Schneider, 2011, pág. 15). Sin embargo, siguiendo y extendiendo el espíritu de la indeterminación de los tokens, no hay razones para pensar que una sola y misma instancia o token de la representación ‘GATO’ exprese la totalidad del contenido que expresa la representación-tipo bajo la cual es subsumida.

Queda, entonces, la pregunta ¿cómo es posible que las representaciones-token codifiquen o expresen solamente una parte limitada del contenido —o condiciones de verdad— que expresa una representación-tipo? Es aquí donde la explicación puede tornarse confusa. La hipótesis que quisiera considerar es la siguiente: las relaciones causales *actuales* que la representación-token exhibe son las que determinan el contenido parcial que la representación-token expresa.

Recuérdese que el criterio de individuación de un símbolo-tipo de PLOT es determinado por la totalidad de su rol causal. Sin embargo, no hay razones para pensar que cuando una persona piensa en su gato, por ejemplo, se activen todas las posibles relaciones

⁶⁴ En este punto el lector podría preguntar ¿Cómo es posible que una representación indeterminada cause una conducta? En este punto entra la parte de la explicación aportada por PLOT. Lo que tendría que decirse es que la parte relevante de las representaciones para explicar sus poderes causales es la parte sintáctica. PLOT define la identidad de las representaciones por su rol en la economía cognitiva y su relación con el mundo. En este sentido, incluso si el contenido de la representación ‘VERDE’ —por ejemplo— es indeterminado a nivel de los tokens, aún la representación cuenta con su rol causal especificado (o lo suficientemente especificado) para generar conductas asociadas al color verde.

causales que la representación ‘GATO’ puede tener. Imagine que a una persona se le pregunta “¿Su gato es tierno? Dicha persona, en tal momento particular podría simplemente pensar/responder ‘Mi gato es tierno’, por lo que la representación-token ‘GATO’ solo podría exhibir solo un par de relaciones causales, por ejemplo, con la representación-token ‘TERNURA’. Podría incluso preguntársele a este sujeto “Sí, pero ¿todos los gatos del mundo son tiernos?” y no tener una respuesta inmediata.⁶⁵

En este sentido, lo que creo que explica que haya una respuesta clara a la primera pregunta, pero no a la segunda es que los símbolos-token empleados en el proceso computacional de la persona en cuestión exhiben, a su vez, solo una parte de la totalidad posible de relaciones causales que el símbolo ‘GATO’ o ‘TERNURA’ pueden llegar a tener. Por lo que el contenido que expresan dichas representaciones es también parcial.

⁶⁵ A lo que se refiere este párrafo es que, si bien las representaciones están individuadas por la totalidad del rol funcional en la economía cognitiva, es razonable pensar que dicho rol no es desplegado actualmente y en su totalidad en cada ocasión que la representación ‘GATO’ juega un papel activo en la producción de la conducta. De esta manera, el contenido puede estar parcialmente determinado en la medida en que el sujeto deba aclararlo más dependiendo de los detalles que se pidan, por ejemplo, propiedades sobre todos los gatos del mundo.

Conclusiones

A lo largo de este documento, me propuse defender la psicología popular como una buena teoría de la conducta. Para ello, en el capítulo I presenté y argumenté a favor de la psicología popular como una descripción conductual que goza no solo de popularidad entre los sujetos, sino de un gran éxito tanto predictivo como explicativo de la conducta. Aceptar la psicología popular implica, entre otras cosas, describir la conducta inteligente como intencional, es decir, como el resultado o efecto de las intenciones de los sujetos, y sostener que es posible hacer generalizaciones que conectan dichas intenciones con la conducta.

Argumenté que sostener las intuiciones de la psicología popular nos compromete con una arquitectura mental; o, en otras palabras, con ciertos mecanismos que posibilitan la implementación y por tanto el exitoso funcionamiento de la psicología popular. De esta forma, para tener una teoría de la psicología popular —como fue la intención en esta investigación— estos mecanismos subyacentes debían dar cuenta de los elementos centrales con los que parece comprometerse la psicología popular (*Entidades semánticamente evaluables; Poderes causales exhibidos por estas mismas entidades, Generalizaciones sobre el contenido; y composicionalidad de dichas entidades*).

Con esto en mente, en el capítulo II ofrecí un argumento a la mejor explicación presentando las hipótesis de Jerry Fodor que explicaban cuáles eran estos mecanismos subyacentes a la psicología popular y cómo es que permitían y explicaban el éxito de esta misma. En este punto, acudí a las hipótesis de la *modularidad de la mente* y el *lenguaje de pensamiento*. Con estas hipótesis presenté la taxonomía funcional de los mecanismos mentales y el sistema representacional en el cual se codifica y sobre el cual operan los mecanismos mentales para dar lugar a la conducta tal cual es descrita por la psicología popular.

Habiendo presentado una teoría explicativa del éxito de la psicología popular, en el capítulo III presenté algunas intuiciones para dudar de la psicología popular —y por tanto de las hipótesis que explicaban su éxito. Las intuiciones en contra de la psicología popular nos conducían a una descripción alternativa de la conducta inteligente de la mano de Paul Churchland y una teoría plausible de los mecanismos mentales subyacentes a esta nueva descripción y su funcionamiento.

Al final del debate Fodor-Churchland, la intuición que permanece es que la conducta inteligente es paradigmáticamente la conducta lingüística y esta es fundamentalmente composicional, esto es, está compuesta de partes más o menos inidentificables. Con esta

discusión entre manos, presenté algunos argumentos a favor de la posición sostenida por Churchland y la posición sostenida por Fodor. Donde Fodor parecía ganar la discusión apelando a la composicionalidad de la conducta lingüística e inteligente. Finalmente, introduje los argumentos de Lenny Clapp como respuesta final a la discusión con Fodor.

Sus argumentos presentaban un fuerte ataque a la composicionalidad de la conducta lingüística e inteligente y, por este camino, también a la composicionalidad del pensamiento, y podían ser usados para eliminar las motivaciones para sostener una teoría —como la de Fodor— que ofrece una explicación centrada en esta propiedad. Así, sus argumentos parecían minar por completo las motivaciones para sostener el *lenguaje de pensamiento*; socavando la explicación del éxito de la psicología popular, y ofreciendo mayor fuerza a los argumentos de Churchland en contra de esta forma de descripción y explicación.

Finalmente, en el capítulo IV presenté lo que considero puede ser una respuesta viable a los desafíos que imponen los argumentos de Clapp a favor de LOT, PLOT o un sistema representacional lo suficientemente cercano a LOT que nos permita sostener y ofrecer una explicación plausible de la psicología popular. Mi punto es que los problemas aducidos a la hipótesis de LOT son principalmente problemas computacionales, relativos a las condiciones de identidad de los símbolos de LOT y de la capacidad expresiva de dichas representaciones respecto de lo que piensan o dicen los sujetos —es decir, los contenidos que expresan—. En este sentido, la nueva respuesta/explicación clarificaba de mejor manera las condiciones de identidad de las representaciones de LOT de tal forma que podía explicar el rasgo inexplicito de dichas representaciones respecto del pensamiento expresado. De esta forma, se presentaba una alternativa teórica y lo que puede ser el esbozo de una posible respuesta y defensa tanto de LOT como de la psicología popular.

Este es, pues, el esquema general del argumento que presenté en esta investigación. A continuación, presentaré algunas de las conclusiones y límites que considero permiten esbozarse de los argumentos y de las discusiones antes presentadas:

1. La primera conclusión que quisiera esbozar es respecto a la composicionalidad de la conducta inteligente/pensamiento. Como he intentado argumentar a lo largo de este último capítulo, tanto PLOT como la hipótesis de la *token-underdeterminacy* pueden brindarle recursos a la LOT para dar cuenta de casos que se mostraban especialmente problemáticos y tenían la forma de contraejemplos a la composicionalidad. Desde estas teorías, quizá sea posible defender un sentido más austero de composicionalidad

que aún pueda predicarse del pensamiento. Sin embargo, incluso si tuviera una respuesta de cómo PLOT, junto con las representaciones indeterminadas pueden ser composicionales, aún queda una pregunta relevante que responder: ¿por qué insistir en la necesidad de la composicionalidad del pensamiento (o los mecanismos cognitivos) si ya se ha mostrado (y aceptado) que la conducta lingüística e inteligente no lo es?

Me gustaría explorar brevemente una respuesta a esta pregunta. Estoy de acuerdo con Clapp, los pragmatistas radicales y Fodor que toda la evidencia apunta a mostrar que la conducta inteligente es no-composicional (al menos en el sentido fuerte de que tiene partes que mantienen su identidad a lo largo de todos los contextos). Sin embargo, es difícil deshacerse de todas las intuiciones que parecen recoger la composicionalidad de la conducta inteligente.

Quisiera explorar brevemente una intuición en particular. La forma en que *ordinariamente* describimos y generalizamos sobre la conducta inteligente permite hablar de composicionalidad en el siguiente sentido: los seres humanos, podría argumentarse, son capaces de una cantidad gigantesca de conductas diferentes. Es razonable pensar que estas conductas pueden ser, en principio, potencialmente infinitas. Las conductas potencialmente infinitas que cada sujeto puede realizar deben ser explicadas por la teoría de alguna manera (en particular, con recursos finitos). Una explicación razonable sería aceptar que la conducta puede ser, en algún sentido, recursiva. Y una forma en la que esto podría ser el caso es aceptar que la conducta está estructurada en términos de partes que pueden ser recombinables.

La explicación precisa de cuáles son estas partes y de qué manera son recombinables escapa a los límites de este trabajo; requiere mayor investigación determinar cómo es este tipo de composicionalidad y cuáles son las restricciones que impone sobre las arquitecturas cognitivas que intentan explicarla. En todo caso, si esta intuición es correcta, considero que PLOT puede ofrecer una aproximación a esta forma de composicionalidad de la conducta lingüística e inteligente.

2. La segunda conclusión que quiero trazar de esta investigación es que aún podemos sostener la psicología popular como una forma correcta de descripción/explicación de la conducta inteligente apelando a lo presentado en el capítulo IV. Si la defensa que presento sobre la necesidad de un sistema representacional cercano al *lenguaje de pensamiento* se encuentra bien encaminada, entonces aún tenemos buenas razones

para sostener una explicación simbólica y composicional de la conducta inteligente. Si este es el caso, siguiendo el espíritu de los capítulos I y II, tenemos aún razones para sostener la psicología popular y el tipo de descripción y explicación que ofrece de la conducta.

Eso es una consecuencia deseable, tal como se presentó en el capítulo I. La psicología popular captura intuiciones y explicaciones que solemos usar en la vida cotidiana para racionalizar la conducta de los demás y no solo entenderla, sino coordinar nuestra conducta con el resto del ambiente donde ésta se despliega (por ejemplo, al tomar un bus para ir a una reunión con otras personas). Es razonable pensar que el aparente éxito de estas intuiciones y generalizaciones descansa en la forma en que nos representamos las acciones y la conducta nuestra y de los demás. De allí la importancia de un lenguaje de pensamiento. PLOT nos ofrece una versión mejorada del LOT fodoriano que logra superar los problemas computacionales presentados por Clapp, dando lugar a una explicación plausible que nos permite apelar a representaciones para dar cuenta de la psicología popular y, *a fortiori*, de la conducta inteligente. Veamos brevemente un ejemplo de la forma que tomaría la explicación:

(1): '(x) si x desea q & piensa que no q a menos que p , entonces x desea p '

La oración (1) podría ser una de las generalizaciones de la psicología popular. La hipótesis de PLOT, nos ofrece una explicación de cómo dicha generalización se instancia en cada uno de los sujetos, donde los diferentes contenidos (q y p) se codifican a través de representaciones-token que mantienen los poderes causales relevantes, en este caso, entre q y p , y la posible conducta del sujeto. Adicionalmente, las representaciones-token que reemplazan q y p , son individuadas por el rol computacional que juegan en la conducta; y el contenido que expresan está parcialmente indeterminado.

De esta forma, mantenemos la explicación de la psicología popular en términos de deseos y creencias. Este tipo de entidades aún pueden entenderse como estados computacionales en los que se procesan, de determinada manera (según la actitud proposicional en cuestión), representaciones que dan lugar a la conducta y permiten instanciar las generalizaciones que cuantifican sobre contenidos y actitudes proposicionales.

3. Si las tesis de PLOT y la *token-undeterminedness* están bien encaminadas, y ofrecen una explicación del éxito de la psicología popular, entonces algo debe decirse de la tesis de la modularidad tal como es defendida por Fodor. Como se mencionó en la reconstrucción del argumento general al inicio de esta sección, la idea de apelar a mecanismos internos que producían la conducta no se encuentra en cuestión. Sin embargo, el cómo definir estos sistemas y a cuáles apelar si lo está.

El procesador central propuesto por Fodor es uno de los mecanismos que parecía perderse si se eliminaba la hipótesis del lenguaje de pensamiento. Después de todo, no hace falta en la explicación si se eliminan el tipo de computaciones que se supone ocurren dentro de este. Sin embargo, con la posible defensa de PLOT, no se ha ofrecido ningún argumento contundente en contra del procesador central. Pues no se ha rechazado de forma explícita: (i) la necesidad de apelar a un mecanismo en el cual se integre la información proveniente de los módulos, (ii) que la información es relativamente explícita al sujeto y (iii) que las representaciones relevantes (los deseos, creencias e intenciones del sujeto) causan las diferentes conductas inteligentes.

En otras palabras, esto nos deja en la siguiente posición a mi parecer. La hipótesis de la modularidad de la mente a la Fodor (es decir, la modularidad con un procesador central) sugiere fuertemente la idea de un sistema representacional altamente estructurado como LOT o PLOT. A su vez, comprometerse con LOT o PLOT, nos deja con la necesidad de mecanismos que permitan y favorezcan ese tipo de computaciones. Por lo tanto, la discusión sobre rechazar LOT o aceptarlo, nos compromete, también, con aceptar o rechazar un sistema central. Tenemos razones para sostener un sistema representacional como PLOT. Por consiguiente, tenemos buenas razones para sostener la modularidad de la mente a la Fodor, es decir, con procesador central de carácter no-modular.

4. Teniendo las conclusiones anteriores en cuenta, aún parece viable un proyecto representacional y computacional que explique el éxito de la psicología popular. Sin embargo, dicho proyecto se ha visto afectado a lo largo de la discusión y nos hemos visto en la necesidad de ciertas modificaciones para superar problemas y argumentos en su contra. Uno de los cabos sueltos que deben ser investigados en un futuro proyecto es la posibilidad de que los mecanismos modificados subyacentes a la

psicología popular (es decir, PLOT y la indeterminación de los tokens) aún ofrezcan una explicación científica de la psicología popular.

Una de las pretensiones (y puntos favorables) de la teoría presentada por Fodor, era que las intuiciones de la psicología popular implicaban una descripción de la conducta inteligente y una explicación (cuasi)científica de la misma. Las hipótesis empíricas de Fodor, LOT y la Modularidad de la mente, podían dar cuenta de estos rasgos de la psicología popular en cuanto que LOT y la modularidad permitían generalizaciones legaliformes que cuantificaban sobre los contenidos y sobre las representaciones de LOT. Surge ahora la pregunta ¿tanto PLOT como la indeterminación de las representaciones-token ofrecen aún la posibilidad de generalizaciones de carácter legaliforme? ¿Es posible hacer una ciencia de la psicología popular?

La respuesta no es obvia; y es oscuro cómo condiciones de verdad indeterminadas pueden dar lugar a generalizaciones legaliformes. Por esto mismo hace falta una investigación más extensa sobre estos elementos, dejando esto como un cabo suelto de esta investigación. Sin embargo, creo que pueden presentarse intuiciones para responder estas preguntas. PLOT es un sistema representacional más sensible al contexto, por lo que quizá sea relevante prestar atención a esto en las generalizaciones ofrecidas por la psicología popular.

Si apelamos a variaciones en las cláusulas *ceteris paribus* (ver la sección “cláusulas *ceteris paribus*, capítulo I), podrían hacerse más locales las generalizaciones de la psicología popular a tal punto de ser relativizadas a comunidades o grupos más o menos definidos en las sociedades. Esto permitiría apelar no solo a mundos nomológicamente cercanos, sino los “culturalmente cercanos”; definidos estos como los mundos donde las culturas sean cercanas y los valores culturales sean también cercanos. Estableciendo la posibilidad de una ciencia de la psicología popular más local, que reconozca la diversidad cultural, regional y epistémica de los sujetos y de la manera en la que ellos mismos entienden y representan su propia conducta inteligente (sea lingüística o no).

Referencias

- Aristóteles. (1978). *De Anima* (Traductor Calvo Martínez, T.). Madrid: Gredos.
- Bird, A. y Tobin, E. (2018). Natural Kinds. En N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2018 ed).
<https://plato.stanford.edu/archives/spr2018/entries/natural-kinds/>.
- Calvo Garzón, F. (2000). State Space Semantics and Conceptual Similarity: Reply to Churchland. *Philosophical Psychology*, 13(1), 77–95.
<https://doi.org/10.1080/09515080050002735>.
- Calvo, P. y Symons, J. (Eds.). (2014). *The Architecture of Cognition: Rethinking Fodor and Pylyshyn's Systematicity Challenge*. Cambridge: The MIT Press.
- Churchland, P. (1981). Eliminative Materialism and the Propositional Attitudes. *The Journal of Philosophy*, 78(2), 67-90. <https://doi.org/10.2307/2025900>.
- Churchland, P. (1986). Some Reductive Strategies in Cognitive Neurobiology. *Mind*, 95(379), new series, 279-309.
- Churchland, P. (1993). State-Space Semantics and Meaning Holism. *Philosophy and Phenomenological Research*, 53(3), 667-672. <https://doi.org/10.2307/2108090>.
- Churchland, P. (1998). Conceptual Similarity across Sensory and Neural Diversity: The Fodor/Lepore Challenge Answered. *The Journal of Philosophy*, 95(1), 5-32.
<https://doi.org/10.2307/2564566>.
- Churchland, P. (2002). Outer Space and Inner Space: The New Epistemology. *Proceedings and Addresses of the American Philosophical Association*, 76(2), 25-48. <https://doi.org/10.2307/3218627>.

Churchland, P. (2007). On the Reality (and Diversity) of Objective Colors: How Color-Qualia Space Is a Map of Reflectance-Profile Space*. *Philosophy of Science*, 74(2), 119-149. <https://doi.org/10.1086/519027>.

Clapp, L. (2012). Is even thought compositional? *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 157(2), 299-322. <https://doi.org/10.1007/s11098-010-9649-2>.

Davidson, D. (1963). Actions, Reasons, and Causes. *The Journal of Philosophy*, 60(23), 685-700. doi:10.2307/2023177

Davidson, D. (2001). *Essays on Actions and Events*. New York: Oxford University Press.

Fodor, J. A. (1975). *The Language of Thought*. New York: Thomas Y. Crowell.

Fodor, J. A. (1986). *La modularidad de la mente*. Madrid: Morata (V.O. 1983)

Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge: The MIT Press.

Fodor, J. A. (1994). *The Elm and the Expert: Mentalese and Its Semantics*. Cambridge: The MIT Press.

Fodor, J. A. (2008). *LOT 2: The Language of Thought Revisited*. New York: Oxford University Press.

Fodor, J. A. y Lepore, E. (1992). *Holism: A Shopper's Guide*. Cambridge: Blackwell.

Jackson, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly (1950-)*, 32(127), 127-136. <https://doi.org/10.2307/2960077>.

Jackson, F. (1986). What Mary Didn't Know. *The Journal of Philosophy*, 83(5), 291-295.
doi:10.2307/2026143

Knapp, W. [William Knapp]. (26 de marzo, 2012) Phonemic Restoration Effect [Video]. YouTube. <https://www.youtube.com/watch?v=ZyvyGMkzNQc>.

Levin, J. (2018). Functionalism. En N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2018 ed). <https://plato.stanford.edu/archives/fall2018/entries/functionalism/>.

Mills, S. (2001). The Idea of Different Folk Psychologies. *International Journal of Philosophical Studies*, 9, 501-519. <https://doi.org/10.1080/09672550110081285>.

Picazo Jaque, C. (2019). Are mental representations underdeterminacy-free? *Synthese* 196, 633–654. <https://doi.org/10.1007/s11229-017-1494-9>.

Putnam, H. (1975). The meaning of “meaning”. *Minnesota Studies in the Philosophy of Science*, 7, 131-193.

Ravenscroft, I. (2019). Folk Psychology as a Theory. En N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2019 ed). <https://plato.stanford.edu/archives/sum2019/entries/folkpsych-theory/>.

Rescorla, M. (2020). The Computational Theory of Mind. En N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2020 ed). <https://plato.stanford.edu/archives/fall2020/entries/computational-mind/>.

Schneider, S. (2009a). LOT, CTM, and the Elephant in the Room. *Synthese*, 170(2), 235-250.

Schneider, S. (2009b). The nature of symbols in the language of thought. *Mind and Language*, 24(5), 523-553.

Schneider, S. (2011). *The Language of Thought: A New Philosophical Direction*.
Cambridge: The MIT Press.

Stich, S. P. (1983). *From Folk Psychology to Cognitive Science: The Case Against
Belief*. Cambridge: The MIT Press.