

UNIVERSIDAD DE CHILE

FACULTAD DE FILOSOFÍA Y HUMANIDADES

ESCUELA DE POSTGRADO

DEPARTAMENTO DE FILOSOFÍA

SER O NO SER... INDIVIDUALISTA: ¿UN DILEMA?

Argumentos para el (anti)individualismo en Psicología.

TESIS PARA OPTAR AL GRADO DE

MAGÍSTER EN FILOSOFÍA CON MENCIÓN EN EPISTEMOLOGÍA

PROFESOR GUÍA

GUIDO VALLEJOS OPORTOT

ESTUDIANTE

PATRICIO VERGARA NELSON

SANTIAGO DE CHILE, DICIEMBRE 2004

AGRADECIMIENTOS

De distintas maneras, muchas personas han posibilitado la concreción de este trabajo; de un modo especial, quienes menciono a continuación:

Al profesor filósofo Don Guido Vallejos Oportot, agradezco la generosidad intelectual y personal con las que guió mi trabajo.

A Alejandra, compañera de viaje, agradezco tu apoyo constante, tu paciencia la mayor parte del tiempo, y tu amor incondicional.

A Inti, Martín y Javiera, luces en mi camino.

A Margarita, por ser mi madre.

A Luisa, por hacer que no sólo Fodor tenga una *güeli*.

Al Departamento de Postgrado y Postítulo de la Universidad de Chile, por concederme la **Beca de Financiamiento para Tesis de Postgrado PG/028/98**, que apoyó sustantivamente el desarrollo de esta investigación.

Índice

	Pag.
Índice.....	1
Resumen.....	3
Introducción.....	4
Capítulo 1. Individualismo y psicología.....	8
I. ¿Qué es el individualismo en psicología?.....	8
II. Individuación e individualismo.....	8
III. La visión individualista de la psicología.....	9
IV. Principio de Autonomía.....	13
IV.1. Duplicados moleculares o doppelgängers.....	13
IV.2. Superveniencia.....	17
V. Individualismo metodológico.....	19
VI. Síntesis.....	21
Capítulo 2. El argumento de los poderes causales.....	22
I. Individualismo metodológico: El argumento de Fodor.....	22
II. Argumentación de Wilson.....	27
II.1. Individualismo global.....	29
II.2. Dos sentidos para ‘poderes causales’.....	33
II.3. El caso del argumento modal.....	42
III. Conclusiones.....	46
Capítulo 3. El argumento computacional.....	48
I. Teoría computacional de la mente.....	48
I.1. Introducción.....	48
I.2. Concepción computacional de los estados mentales en la TCM.....	51
I.3. Procesos cognitivos.....	54
I.4. Formalización y computación.....	57
I.5. Concepción computacional de los procesos cognitivos.....	65
II. Wilson y el argumento computacional para el individualismo.....	67
II.1. Computacionalismo en psicología.....	67

II.2. El argumento computacional.....	69
II.3. Computacionalismo amplio.....	71
II.3.1. Coherencia del computacionalismo amplio.....	71
II.3.2. La noción de formalidad.....	78
II.3.3. Objeciones al computacionalismo amplio.....	82
II.4. Plausibilidad del computacionalismo amplio en psicología.....	86
II.5. Computacionalismo reconsiderado.....	92
III. Conclusiones.....	97
Capítulo 4 Argumentos metodológicos.....	99
I. Introducción.....	99
II. Argumentos metodológicos para el individualismo.....	101
II.1. Argumento contra el naturalismo en psicología.....	101
II.2. Crítica al argumento contra la psicología naturalista.....	104
II.3. Argumento de la modularidad al individualismo.....	106
III. Conclusiones.....	109
Capítulo 5 Conclusiones.....	110
Referencias.....	115

Resumen

La discusión central del presente trabajo gira en torno a la disputa entre dos modalidades de individuación de estados mentales en psicología (cognitiva), cada una de las cuales da lugar a taxonomías psicológicas diferentes, a saber: las modalidades individualistas, que consideran que la individuación debe hacerse sin consideración por los aspectos ambientales o contextuales —en otros términos, sin considerar propiedades históricas o relacionales— del organismo que instancia tales estados mentales, que originan taxonomías ‘estrechas’; y las anti-individualistas, que sostienen que la individuación debe realizarse atendiendo a las variables ambientales y que ven al organismo ‘anidado’ en un contexto que no puede obviarse sin pérdida sustantiva, que dan lugar a taxonomías ‘amplias’.

El debate acerca de cuál es la modalidad correcta de individuación en psicología se expone siguiendo la estructura propuesta por Wilson (1995), quien revisa tres clases de argumentos organizados en torno a ámbitos de discusión entre individualistas y anti-individualistas: la taxonomía en ciencia, particularmente en psicología; el computacionalismo en las ciencias de la cognición; y un par de argumentos metodológicos. La conclusión que Wilson pretende invitarnos a compartir es que hay razones para rechazar el individualismo.

La tesis se propone como objetivo evaluar los argumentos anti-individualistas planteados por Wilson y ponderar su capacidad para debilitar o afectar la posición individualista en psicología. La conclusión será que el carácter de la argumentación anti-individualista es lejos menos apremiante de lo que se sugiere y que, en definitiva, el individualismo aún se sostiene.

Introducción

Puede afirmarse que existe considerable acuerdo sobre el enfoque que debe darse al estudio empírico de las capacidades y habilidades humanas, estudio que corresponde realizar a la psicología. El acuerdo apunta, fundamentalmente, al empleo de un enfoque mentalista, que involucra la atribución a un organismo de estados, procesos y eventos intencionales (representacionales) del tipo de los deseos, las creencias, los pensamientos, los recuerdos y las percepciones; en síntesis, estados mentales. Se trata de los estados mediante los cuales el organismo conoce, representa y utiliza la información relativa a su entorno. (Burge 1986)

Si asumimos que la psicología (cognitiva) debe explicar la conducta de los organismos —al menos la conducta inteligente, esto es, aquella guiada por metas y propósitos—, entonces la posibilidad de una psicología (intencional) que pueda ofrecer explicaciones respecto a por qué un organismo dado hace lo que hace del modo en que lo hace, según los estados mentales en los que se halla, requerirá un procedimiento para individuar estos últimos. Un procedimiento para fijar sus naturalezas, establecer sus tipos, y generar una taxonomía psicológica de estados mentales tipo. Las explicaciones psicológicas que tal psicología ofrezca identificarán la conducta de un organismo dado como consecuencia de la instanciación de alguno(s) de esos estados mentales tipo.

La discusión para establecer la modalidad de individuación correcta para los estados mentales se ha polarizado entre quienes piensan que éstos son de naturaleza tal que pueden darse sus condiciones de identidad prescindiendo absolutamente de la relación entre el organismo que los tiene y su entorno, los *individualistas*, y quienes sostienen que las condiciones ambientales pueden producir, por sí mismas, variaciones (contrafácticamente sustentadas) en los estados mentales en los que se encuentra el organismo y, en consecuencia, que la individuación ha de ser relacional y considerar las variables del entorno, los *anti-individualistas* (o no-individualistas). Mientras los individualistas sostienen que la psicología de un organismo depende *sólo* de rasgos internos o intrínsecos a él, los anti-individualistas sostienen que depende también de factores externos. Estos puntos de vista contrapuestos originan

taxonomías diferentes, ‘estrechas’ en el caso de los individualistas y ‘amplias’ en el de los no-individualistas.

Otro punto de vista individualista emerge al aceptar el supuesto computacional, que sostiene que el cerebro es el equivalente funcional de un ordenador serial y que los estados y procesos mentales son computaciones sobre representaciones. Los procesos computacionales son simbólicos y formales. Simbólicos porque se definen sobre representaciones y formales porque se aplican a las representaciones en virtud de la sintaxis de las representaciones. La formalidad de los estados y procesos mentales impondría la condición de que las operaciones computacionales se especifiquen sin referencia a las propiedades semánticas de tales representaciones —como verdad, referencia y significado. La doctrina que sostiene que la individuación de estados mentales se realiza con independencia de su evaluación semántica, es denominada por Fodor (1980), siguiendo a Putnam (1975), *solipsismo metodológico*.

Por otro lado, si los estados mentales informan sobre el mundo, si su contenido es acerca del mundo, deben hallarse en algún tipo de relación con él. Esto nos lleva a dos temas importantes en la discusión sobre la individuación y (más o menos) relacionados entre sí, que se refieren a la causación y la superveniencia. Los individualistas plantean que los estímulos distales (el mundo) afectan causalmente en forma inmediata a los sistemas perceptivos de un organismo, pero que no hacen lo mismo con sus estados mentales. El origen causal de los estímulos y, por lo tanto, lo que los estímulos representan no tiene importancia causal inmediata para el funcionamiento de la mente. Esta concepción de la causación se vincula a una afirmación sobre la superveniencia, que afirma que las propiedades psicológicas deben reducirse a, o al menos supervenir en, las propiedades que están causalmente involucradas en forma inmediata en las reacciones a la estimulación perceptiva y en la producción de conducta. Hacerlo de otro modo supondría plantear procesos causales fantasmas en la explicación psicológica. Las propiedades base son internas al organismo, ya que son propiedades intrínsecas de estados cerebrales. En consecuencia, no habría diferencias psicológicas *sin* diferencias en las propiedades intrínsecas de los estados cerebrales. No obstante, autores como Tyler Burge señalan que estos argumentos confunden la causación con la individuación y presuponen que

el efecto causal del entorno estaría excluyendo que éste tenga, además, efecto sobre el cuerpo del sujeto. Burge afirmará que las variaciones en el entorno que no varían los impactos que ‘afectan’ causalmente al cuerpo del sujeto, pueden ‘afectar’ la individuación de la información que el sujeto está recibiendo, o de los procesos intencionales que está protagonizando, o la manera en que está actuando.

Finalmente, se plantea el problema respecto al tipo y alcance de las generalizaciones psicológicas que pueden obtenerse bien de una concepción individualista bien de una no-individualista, generalizaciones que serán invocadas en las explicaciones que dicha psicología ofrezca. En tal sentido, por ejemplo, Fodor afirma que una taxonomía individualista de estados mentales, una que los individúe por sus *poderes causales*, tendrá la capacidad de proporcionar generalizaciones válidas en todo mundo posible, permitiendo una psicología que no pierda generalizaciones, como sucede en el caso de una individuación relacional del contenido, que únicamente admite generalizaciones de alcance limitado, ya que valen sólo para, y en, un contexto (mundo) determinado. De este modo habría, para ilustrar el punto, generalizaciones psicológicas válidas sólo aquí en la Tierra y otras válidas sólo en la Tierra Gemela. Es evidente que esto limita enormemente el poder explicativo de la psicología y hace complejo el estatus de sus generalizaciones, pero salva el mundo que aparentemente el individualista habría dejado de lado u olvidado *sin* justificación suficiente.

Como lo señala Robert Wilson (1995), el tema del individualismo en psicología continúa ocupando una posición central en el debate dentro de la filosofía de la mente. Además del interés que reviste en sí mismo, tiene relevancia también para una serie de otros temas en la psicología filosófica contemporánea: ‘la naturalización de la intencionalidad, la autonomía de la psicología, la superveniencia de lo mental en lo físico, la naturaleza de la causación mental, la viabilidad de la “psicología popular” o de sentido común, y la forma que las ciencias cognitivas pueden y deben tomar’. (Op. cit., p xi). Estando el tema próximo a cumplir treinta años en la palestra filosófica, me ha parecido oportuno realizar una revisión crítica del estado actual de la discusión respecto a los fundamentos generales para una modalidad correcta de individuación de las clases naturales psicológicas y de su relación e incidencia en la forma de

concebir la explicación psicológica, intentando recuperar parte de lo que este tiempo ha dejado como resultado de los esfuerzos de filósofos, psicólogos, científicos cognitivos y otros por ofrecer una respuesta satisfactoria a este problema.

Aunque, como se ha dicho, la problemática de la individuación de estados mentales tiene ya una historia, que puede considerarse iniciada por Putnam en 1975,¹ no es un tema que parezca haber recibido hasta ahora mucha atención en nuestro medio —un rasgo que comparte con otros asuntos—, y sobre el que existe relativamente poco material actual, disponible, y en español, lo que actúa también como restricción para algunos de aquellos que podrían interesarse en conocer algo al respecto. Pienso que esto último puede hacer que este trabajo, además de lo que su contenido pueda ofrecer, sea una contribución al estudio de un tema apenas difundido entre aquellos que podemos tener alguna vinculación con la filosofía (de la mente), la psicología (cognitiva) o las ciencias cognitivas en general.

¹ La discusión sobre el significado, así como el haber acuñado el término solipsismo metodológico, en el sentido que tiene actualmente, parecen situar a Putnam en los inicios de la discusión de este tema (cf. Putnam 1975).

Capítulo 1 - Individualismo y psicología.

I. ¿Qué es el individualismo en psicología?

En palabras de Wilson, “el individualismo en psicología es una idea acerca de qué son los estados mentales, una idea acerca de cómo van a individuarse, clasificarse, taxonomizarse o tipificarse los estados mentales” (1995:1). Kim Sterelny nos ofrece otra manera de caracterizar al individualismo diciendo que se trata de “una doctrina sobre las clases naturales en psicología” (1990:81). También podemos responder la pregunta que abre esta sección, planteando que el individualismo es una visión respecto a la modalidad de individuación de las clases naturales para los propósitos de la explicación causal en psicología. Al prescribir la forma de individuar las entidades que aparecerán como *explanantia* en las explicaciones psicológicas, el individualismo resulta ser también un constreñimiento para la explicación psicológica. Con estos elementos básicos podemos formarnos una primera impresión sobre el individualismo. Antes de exponer con más detalle qué *dice* el individualismo sobre la individuación y la explicación en psicología, voy a detenerme un momento en unas breves consideraciones sobre los conceptos de individuación y explicación relevantes para la discusión que sigue.

II. Individuación e individualismo.

Hemos dicho que puede caracterizarse al individualismo como una visión sobre la individuación en psicología pero, dado que ‘individuar’ no tiene un uso unívoco entre los filósofos, importa hacer alguna puntualización al respecto antes de continuar con esta caracterización del individualismo psicológico.

Las discusiones filosóficas tradicionales sobre criterios para la individuación se refieren a particulares, usualmente objetos, y su foco principal se orienta al problema de determinar qué hace a una cosa dada la cosa que es. En este sentido, individuar un objeto supone especificar sus condiciones de identidad, es decir, las condiciones bajo las cuales puede decirse que una cosa particular existe o continúa existiendo. En tales discusiones, el foco está en identificar o precisar qué hace a una cosa en un tiempo particular la *misma cosa* que alguna cosa en algún

otro tiempo; el interés se concentra en la naturaleza de los vínculos entre etapas temporales sucesivas de uno y el mismo objeto, y en el rol que la continuidad tiene para el concepto de identidad. En este contexto, la identidad de los indiscernibles puede verse como un principio abstracto que proporciona una condición suficiente para la identidad e individuación de objetos en este sentido de ‘individuación’.

El individualismo en psicología *no* es una visión sobre la individuación de particulares al modo planteado más arriba; antes bien, se refiere a la individuación de clases, tipos o categorías de particulares; es una modalidad de clasificación de entidades particulares en clases o categorías. El individualismo psicológico propone criterios de organización para estas clases o categorías y, por esta vía, establece un constreñimiento para la taxonomía de estados mentales tipo. En general, el individualismo psicológico no se refiere directamente a la identidad de particulares mentales, sean estos objetos, estados o procesos. Como un constreñimiento para la explicación psicológica, el individualismo es una idea sobre la individuación con fines taxonómicos, ya que haciendo una afirmación sobre la individuación de clases de entidades plantea un constreñimiento relativo a qué rasgos de los estados mentales, aquellos que autorizan su tipificación en clases, van a figurar en las explicaciones psicológicas. Volvamos ahora a revisar con más de detalle la propuesta individualista en psicología.

III. La visión individualista de la psicología.

Dado que puede señalarse a Hilary Putnam como precursor de la discusión contemporánea sobre el individualismo en psicología, parece apropiado exponer su posición en primer término. Hay que decir, no obstante, que la caracterización del individualismo que hace Putnam no emerge en el contexto de una discusión sobre la naturaleza de los estados mentales, sino sobre la naturaleza del significado. En “Meaning of Meaning”, Putnam indica que “cuando los filósofos tradicionales² hablan sobre estados psicológicos (o “estados mentales”), suponen lo que podemos llamar *solipsismo metodológico*. Esta suposición es la suposición de

² Aquí, “filósofos tradicionales” incluye a aquellos que sostienen (a) que saber el significado de un término es estar en cierto estado psicológico, y (b) que el significado de un término (la *intensión*) determina la *extensión*.

que ningún estado psicológico, propiamente así llamado, presupone la existencia de ningún otro individuo aparte del sujeto a quien se adscribe dicho estado”. (1975:220).

En un contexto diferente, aunque relacionado, Jerry Fodor (1980) emplea el término *solipsismo* metodológico para describir un supuesto de la psicología (cognitiva), derivado del carácter computacional de los estados y procesos mentales. Sintéticamente, los procesos computacionales son *simbólicos* y *formales*: “simbólicos porque se definen sobre representaciones y formales porque se aplican a las representaciones en virtud de su sintaxis” (Fodor 1980:486). Lo anterior implica que, si los procesos y estados mentales son computacionales, la individuación de dichos estados y procesos no considera su semántica; dicho de otra forma, la individuación de estados y procesos mentales no considera propiedades tales como verdad, referencia y significado; propiedades que, en principio, suponen una relación con el mundo más allá del sujeto. En palabras de Sterelny, “la cognición es procesamiento de información, pero las operaciones cognitivas deben concordar con los rasgos físicos intrínsecos de los estados que *codifican* esa información, pues éstas no tienen acceso directo a las causas distales de esos rasgos intrínsecos” (1989:75).

En *Psicosemántica*, Fodor dirá, precisamente, que el solipsismo metodológico es la doctrina que sostiene que “la individuación de estados mentales se realiza con independencia de su evaluación semántica” (1987:73). Siguiendo esta línea de razonamiento de Fodor sobre la *inaccesibilidad* de las propiedades semánticas de los estados mentales, Colin McGinn dice:

Se plantea la siguiente pregunta [para Fodor]: ¿En virtud de qué las creencias cumplen un rol en la psicología del agente? Y, parece, Fodor supone que puede haber sólo una respuesta a esta pregunta: Las creencias cumplen un rol en la psicología del agente en virtud de propiedades intrínsecas de las representaciones internas implicadas —las *relaciones* semánticas entre representaciones y cosas en el mundo deben ser irrelevantes para el rol psicológico de las creencias.. Más precisamente, el rol causal de una creencia debe depender de, y sólo de, esas propiedades de las representaciones que pueden caracterizarse sin referirse a asuntos que yacen fuera de la cabeza del agente. (1982:208).

Una forma de entender la afirmación de Putnam de que los estados psicológicos no presuponen la existencia de otros individuos es pensar que lo que esos estados son y el modo en que son individuados y taxonomizados, para los propósitos de la explicación psicológica,

no nos *obliga* a presuponer nada sobre el mundo externo del individuo que los tiene. ¿Por qué? Porque las personas actúan de acuerdo con el modo en que conciben el mundo y porque “hay un importante sentido en el cual cómo es el mundo no hace diferencia en los estados mentales de uno” (Fodor 1980:486). La última idea conecta el argumento individualista de Fodor con los argumentos que Descartes ofrece al discutir la posibilidad de engaño permanente:

En este momento, verdaderamente me parece que es con los ojos abiertos que estoy mirando este papel; que esta cabeza que muevo no está dormida, que es deliberadamente y a propósito que extiendo mi mano y la percibo... Pero, al pensar sobre esto, me digo a mí mismo que en muchas ocasiones he sido engañado por similares ilusiones, y al extender esta reflexión veo tan manifiestamente que no hay indicios ciertos mediante los cuales pueda distinguir claramente la vigilia del sueño, que me encuentro perdido en el asombro. Y mi asombro es tal que casi es capaz de persuadirme de que ahora sueño. (Descartes 1931)

En una línea que, en cierta forma, resume lo dicho con respecto al solipsismo metodológico, Kent Bach señala:

Lo que está fuera de la mente de una persona es irrelevante para la psicología. Sin consideración por cómo sea el mundo en comparación a cómo es representado siendo y sin consideración por cómo pueda cambiar mientras los estados psicológicos de la persona permanecen los mismos, todo es igual en lo que a la psicología concierne. De acuerdo con el SM [solipsismo metodológico], la forma psicológicamente apropiada de individuar tipos de creencias, deseos y otros estados intencionales es por sus contenidos, no por relaciones que el sujeto tiene con las cosas en el mundo, que podrían haber sido diferentes sin afectar al contenido... Lo que importa es cómo el sujeto se representa el mundo, no cómo es realmente el mundo. (1982:123)

Para Tyler Burge —importante filósofo anti-individualista—, el individualismo considera que “las naturalezas mentales de los estados (y eventos) mentales de toda persona o animal son tales que no hay una relación de individuación profunda entre que el individuo esté en estados de esas clases y la naturaleza del ambiente físico o social del mismo” (1986:3-4). Según Burge, el individualismo sostiene que la individuación de los estados mentales no es afectada por factores externos al individuo que instancia esos estados.

Estas tres visiones sobre el individualismo —las de Putnam, Fodor y Burge— comparten un carácter negativo, señalan aquello que *no* es estrictamente relevante para la individuación de estados mentales. En tanto caracterizaciones negativas, debe indicarse que, como cuestión de

hecho, no constituyen *criterios de individuación*, pues no nos dicen cómo debemos individuar estados y procesos mentales, sino sólo qué *no* debemos considerar. No obstante, sí ponen de manifiesto una importante distinción, que la psicología debe tener en cuenta, entre los estados psicológicos de individuos y características o propiedades de los ambientes físicos y sociales donde esos individuos se hallan. Michael Devitt lo expone claramente:

En psicología, estamos interesados en explicar por qué, dado un estímulo en sus órganos sensoriales, una persona demuestra cierta conducta. Sólo algo que sea enteramente superveniente en lo que está dentro de su piel —en sus estados físicos internos intrínsecos, particularmente su cerebro— puede cumplir el rol explicativo requerido entre input periférico y output. Las causas ambientales de su estímulo y los efectos de su conducta están más allá del punto psicológico. La persona y todos sus duplicados físicos, incluso funcionales, deben ser psicológicamente iguales, cualesquiera sean sus ambientes. Los estados mentales deben individuarse de acuerdo a su rol dentro del individuo, sin considerar sus relaciones con un ambiente. (1990:377).

Para el individualista los estados psicológicos están, como a veces se dice, *en la cabeza* y, por eso, la distinción entre lo que está dentro de la cabeza y lo que está fuera de ella se hace relevante. Epistemológicamente, a los efectos de la taxonomía y la explicación en psicología, importa distinguir entre sujeto-que-conoce y mundo-conocido; esto podría considerarse una consecuencia del hecho de que los símbolos con los que operamos cognitivamente nos permiten tener pensamientos con relativa independencia de la experiencia. Por ejemplo, podemos pensar en objetos ausentes o incluso inexistentes.

Voy a exponer ahora otras dos formas de concebir al individualismo en psicología. Estas dos concepciones tienen un carácter positivo, precisamente en la medida en que, a diferencia de las tres visiones anteriores, sí proponen criterios de individuación para los estados y procesos mentales. Se trata del *principio de autonomía*, propuesto por Stephen Stich y del *individualismo metodológico*, planteado por Jerry Fodor. Me propongo tratar un poco más extensamente ambas ideas, para mostrar con algo de detalle las intuiciones y argumentos ofrecidos para apoyarlas. Veamos primero el principio de autonomía.

IV. Principio de Autonomía.

En “Psicología autónoma y la tesis de creencias y deseos”, Stephen Stich (1978)³ examina la tensión que existiría entre lo que llama *tesis de creencias y deseos* —la idea de que la acción humana va a explicarse, al menos en parte, aplicando en un sentido sustantivo términos como deseos y creencias— y un supuesto acerca de la naturaleza de las teorías psicológicas explicativas que, en su opinión, se ha sostenido ampliamente y que opera como principio regulativo fundamental en buena parte de la teorización psicológica contemporánea. Stich denomina a este supuesto *principio de autonomía psicológica* (abreviado, principio de autonomía) y la relevancia que él le otorga al mismo estriba en su sospecha de que “la mayoría de las explicaciones de la acción en términos de creencias y deseos son incompatibles con el principio de autonomía” (p. 590). Aquí solamente revisaremos la presentación que Stich hace del principio en “Psicología autónoma”, donde utiliza dos modalidades para explicar en qué consiste éste:

IV.1. Duplicados moleculares o doppelgängers.

Stich nos propone considerar el principio de autonomía, primero, bajo el prisma de un ejemplo de ciencia-ficción. Se trata de imaginar que contamos con la tecnología para duplicar personas, esto es, para “construir un ser humano que sea átomo por átomo y molécula por molécula réplica de algún ser humano dado” (p. 590): un *doppelgänger*. En la situación hipotética de hallarnos frente a un ser humano (o, para el caso, frente a cualquier animal) y su réplica exacta, el principio de autonomía afirmará que “estos dos humanos serán psicológicamente idénticos, que cualquier propiedad psicológica instanciada por uno de los sujetos también será instanciada por el otro” (p. 590). Cabe agregar que *psicológicamente idénticos* implica *conductualmente* idénticos.

Stich establece algunas condiciones que especifican lo que entiende por *identidad psicológica*. La identidad psicológica incluye a cualquier par de organismos, coexistan o no temporalmente, que sean réplicas átomo por átomo uno del otro. Además, es irrelevante que la

³ Los números de página señalados en las citas de “Psicología autónoma”, corresponden a la versión que aparece en Rosenthal (1991).

duplicación obedezca o no a un acto deliberado; la identidad psicológica es válida también en casos de duplicación accidental. Hay una aclaración especial ante la objeción de que un organismo y su réplica no comparten *todas* sus propiedades psicológicas. El punto se discute a partir del caso de un organismo que, por ejemplo, siguió por televisión las alternativas del proceso judicial efectuado en Londres para decidir la extradición o no a España del Senador Vitalicio y, por tanto, las recuerda, mientras su réplica *no* las recuerda. El duplicado podría *pensar* que las recuerda o tener una “huella mnémica” idéntica, pero en virtud de haber sido creado mucho tiempo después de terminado dicho proceso judicial —por ejemplo, el año 2005— no habría podido *verlas* ni, en consecuencia, *recordar haberlas visto*.

Como respuesta a esta observación crítica, Stich plantea que existen muchas clases de propiedades, que plausiblemente denominamos (o denominaríamos) “psicológicas”, que podrían ser instanciadas por un individuo y no serlo por su réplica. Ejemplos de propiedades de esta clase serían “recordar que *p*”, “saber que *p*”, “ver que *p*”. Según Stich, estas propiedades exhiben un carácter “híbrido”, pues pueden analizarse en términos de “una propiedad ‘puramente psicológica’ (como parecer recordar que *p* o creer que *p*), junto con una o más propiedades y relaciones no-psicológicas (como *p* es verdadera, o la huella mnémica es causada de cierta forma por el hecho que *p*)” (p. 591). Aunque reconoce que sería arbitrario insistir en el intento de estipular que las propiedades psicológicas “híbridas” no sean en lo absoluto propiedades psicológicas, Stich piensa que hay algo extraño respecto a éstas, una particularidad que se reflejaría en la distinción intuitiva que se hace entre estas propiedades “híbridas” y los componentes “puramente psicológicos” que les subyacen. Esta peculiaridad de las híbridas se expresaría, según él, en que “no esperamos que desempeñen ningún rol en una teoría psicológica explicativa. Más bien, esperamos que una teoría psicológica que apunta a explicar la conducta invoque sólo las propiedades “puramente psicológicas” *que son compartidas por un individuo y su réplica*” (p. 591, énfasis mío).

El punto que interesa, respecto al caso de recordar que *p* o parecer recordar que *p*, se centra en el carácter irrelevante que la diferencia entre ambas propiedades tiene *para la conducta* del individuo y la de su réplica. En consecuencia, el principio de autonomía, en tanto afirma la

identidad psicológica de las réplicas, debe entenderse limitándose a sí mismo sólo a las propiedades que pueden desempeñar un rol en una teoría psicológica explicativa. Así, el principio es una afirmación acerca de qué clases de propiedades y relaciones pueden desempeñar un rol en una teoría psicológica explicativa, esto es, un constreñimiento para las clases de propiedades y relaciones que una teoría psicológica explicativa debe considerar: aquellas compartidas por un sujeto y su réplica.

Advertido de no haber proporcionado una justificación suficientemente sólida del principio de autonomía, Stich retoma el punto en *From Folk Psychology to Cognitive Science* (Stich 1983). Me enfocaré solamente en las aclaraciones que Stich hace a propósito de la objeción de que no todas las conductas de un organismo y su réplica son idénticas, al menos en la forma en la que estas son descritas normalmente. Recordemos el caso del individuo que presencié por televisión las audiencias del proceso de extradición del General, mientras su réplica, por haber sido creada muchos años después, quizá incluso luego de la muerte del General, no pudo haberlas presenciado y, en consecuencia, no podría recordar haber sido testigo de éstas. Stich piensa que lo correcto es conceder que, efectivamente, es falso que una persona y su réplica se comporten siempre de la misma manera. Pero, agrega que “no debemos esperar que una teoría psicológica prediga o explique la conducta bajo *toda y cualquier descripción* expresada por el sentido común” (1983:166, énfasis mío).

Para captar el punto, Stich nos sugiere una analogía con la química: puede resultar verdadero que hervir el contenido de una botella de Chateau Lafitte cause una baja sustancial de su precio en el mercado. Pero sería inadecuado esperar que la sola química dé cuenta de esa situación. Lo que esperamos es que la química nos ofrezca una explicación de los efectos del hervor descritos en “un vocabulario propiamente químico, apropiadamente delimitado” (1983:166). Por otra parte, “no es probable que haya alguna especificación antecedente obvia del rango de descripciones apropiadas en los *explananda* de la química” (1983:166). Para explicar los efectos que el hervor del vino tiene en su precio, tendríamos que agregar a la explicación estrictamente química algunos hechos relativos a la relación entre el hervor del

mosto y algunos cambios en sus cualidades sensoriales, y entre estos cambios sensoriales y el precio del vino.

Según Stich, la situación es similar en el caso de la psicología. Sería inadecuado esperar que la psicología científica explique los eventos conductuales bajo todas las descripciones imaginables. Los psicólogos también deben seleccionar un lenguaje descriptivo apropiado para sus *explananda*; formular este vocabulario será parte fundamental en la construcción de la teoría psicológica. Dos autores, Fodor y Wilkes, son citados apoyando esta convicción:

Merece enfatizarse, que el sentido de “conducta” *es* propio, y que es mucho más de lo que usted esperaría. No toda descripción de un acto puede ser tal que una teoría de la causación mental de la conducta explicará el acto bajo esa descripción... Usted no puede tener explicaciones de todo bajo cualquier descripción, y es una cuestión de determinación empírica qué descripciones de la conducta revelan su sistematicidad *vis-a-vis* sus causas. (Fodor 1980:330-1, en Stich 1983:167).

Cada ciencia debe desarrollar una taxonomía de los eventos que caen bajo su dominio de discurso, y para esto tiene que desarrollar un vocabulario descriptivo de predicados teóricos y de observación. Dado que estos eventos pueden describirse variadamente, no toda descripción de una acción o una capacidad para la acción será una descripción en el dominio de la psicología. (Wilkes 1981:150, en Stich 1983:167).

La consecuencia de estas reflexiones es una distinción entre *descripciones conductuales autónomas* —cualquier descripción de la conducta que satisface la siguiente condición: si se aplica a un organismo en un entorno dado, entonces se aplicaría también a cualquier réplica del organismo en ese entorno— y descripciones conductuales no-autónomas. Esta distinción entre descripciones conductuales autónomas y no-autónomas apoya la que existe entre propiedades “puramente psicológicas” y propiedades y relaciones no-psicológicas, en el sentido de que podemos decir que las descripciones conductuales autónomas implican la instanciación de propiedades “puramente psicológicas”, esas que un organismo comparte con todas sus réplicas y que, en consecuencia, hacen idénticas sus respectivas conductas. La psicología (científica) debe ocuparse de explicar los eventos conductuales descritos autónomamente. Las descripciones no-autónomas serán consideradas, como se hizo con las propiedades y relaciones no-psicológicas, híbridos conceptuales o lógicos conteniendo una

descripción autónoma de la conducta y una mención, por ejemplo, a un hecho histórico o a una relación que no esperamos sea explicada por una teoría psicológica de la conducta.

Un último asunto, respecto de las descripciones de la conducta que interesan a la psicología, se refiere a dónde van a obtenerse las citadas descripciones conductuales autónomas. Sobre el particular, Stich dice que habrá ocasiones en las que el sentido común proporcionará tales descripciones, pero que este no tiene porqué ser siempre el caso. Como ya se indicó, la formulación de una terminología apropiada para la descripción de los *explananda*, es un paso esencial en el desarrollo de una nueva ciencia. Adicionalmente, hay que aclarar que no hay razón para esperar que la terminología adecuada para las descripciones conductuales autónomas sea una descripción puramente física de movimientos, tal como la que los conductistas quisieron ofrecer y buscaron sin éxito.

IV.2. Superveniencia.

La segunda modalidad que Stich emplea para explicar el principio de autonomía, se apoya en la idea de *superveniencia*, expresada de la siguiente forma:

“Supóngase que S y W son dos clases de propiedades, y que S* y W* son los conjuntos de todas las propiedades que pueden construirse a partir de las propiedades de S y W respectivamente. Entonces, siguiendo a Kim, diremos que la familia de propiedades S superviene en la familia de propiedades W (con respecto a un dominio de objetos D) sólo en caso de que, necesariamente, dos objetos cualesquiera en D, que comparten todas las propiedades en W*, compartan también todas las propiedades en S*” (1978:591).

Una manera menos formal de plantearlo es diciendo que “una clase de propiedades superviene en otra si la presencia o ausencia de propiedades en la primera está completamente determinada por la presencia o ausencia de propiedades en la última” (1978:591). Desde esta perspectiva, el principio de autonomía establece que las propiedades y relaciones que se invoquen en una teoría psicológica explicativa deben supervenir en “las propiedades y relaciones *físicas internas, actuales* de los organismos” (1978:591), que son precisamente aquellas que un organismo comparte con *todas* sus réplicas.

Stich también revisa el tema desde la perspectiva de lo que el principio excluye. En primer término, ya que el principio de autonomía exige que las propiedades psicológicas *explicativas* supervengan en las propiedades *físicas*, al menos algunas formas de dualismo resultan falsas. Se trataría de aquellas formas de dualismo que sostienen que no existe correlación *nomológica* entre las propiedades psicológicas o mentales y las propiedades físicas, pero que, no obstante, sostienen que las primeras deben invocarse al ofrecer una explicación de la conducta de un organismo; al hacerlo, estarían negando que haya superveniencia de los estados psicológicos explicativos en los estados físicos. Las formas de dualismo que acepten la existencia de una correlación nomológica entre propiedades psicológicas y propiedades físicas, no estarían en conflicto con el principio de autonomía.

Sin embargo, el principio afirma algo más fuerte que la sola superveniencia de los estados psicológicos en los estados físicos. Esto porque excluye, considerándolas irrelevantes, algunas propiedades y relaciones físicas, de modo que organismos que difieren *sólo* en estas propiedades son psicológicamente idénticos. Al especificar que sólo las propiedades físicas *actuales* son psicológicamente relevantes, el principio de autonomía excluye como irrelevantes todas las propiedades que tratan con la historia pasada o futura del organismo. Cómo, desde una perspectiva histórica, llegaron a estar un organismo y su réplica en un determinado estado psicológico, es irrelevante desde el punto de vista de la psicología explicativa, pues, como se dijo, esto no afectará su conducta. Recordar haber visto las peripecias del juicio del General o sólo parecer recordarlas no altera en nada la conducta. Por otro lado, ya que el principio de autonomía indica que sólo las propiedades y relaciones físicas *internas* son relevantes, las relaciones entre un organismo y su ambiente externo son irrelevantes para su estado psicológico (explicativo) actual. Los hechos históricos y ambientales adquieren relevancia psicológica sólo si influyen en un estado físico interno actual del organismo.

El principio de autonomía establece, entonces, una condición necesaria para clasificar dos instanciaciones de estados mentales como estados de diferentes clases psicológicas y, de este modo, ofrece una condición suficiente para la clasificación de dos instanciaciones de estados

mentales como estados de la *misma* clase psicológica. Si dos estados mentales se corresponden con diferencias físicas internas de los organismos en los que se instancian, deben clasificarse como pertenecientes a diferentes clases psicológicas; si dos estados mentales se corresponden con las mismas bases físicas intrínsecas en las respectivas criaturas en las que se instancian, entonces pertenecen a la misma clase psicológica.

V. Individualismo metodológico.

Consideremos ahora la propuesta de Fodor. Su argumentación en defensa del individualismo tiene un precedente en *Psicosemántica* (Fodor 1987). Allí, Fodor intentó mostrar que “los estados mentales que sólo difieren en las propiedades intencionales *amplias* no difieren *ipso facto* en los poderes causales y, por ello, que las meras diferencias en el contenido intencional amplio no determinan diferencias en las clases naturales a los efectos de la explicación psicológica” (Fodor 1990:332).

Fodor define al *individualismo metodológico* como “la doctrina según la cual los estados psicológicos se individualizan *con relación a sus poderes causales*” (1987:73) y esgrime una razón de principio para ofrecer el individualismo: él considera que “la individuación en ciencia es *siempre individualista*. De donde se sigue que los constructos científicos de la psicología deben ser individualistas” (1987:59).⁴ Así, lo que importa a los efectos de la explicación psicológica causal, desde una perspectiva individualista, es que los estados mentales que serán subsumidos por las generalizaciones de una psicología (cognitiva) científica sean individuados por sus poderes causales.

Esta sección ha estado dedicada a examinar las propuestas de individuación de estados mentales propuestas por Stephen Stich y Jerry Fodor. Como lo indicamos más arriba, estas dos formas de individuación tienen un carácter positivo, en tanto ambas especifican criterios para la identidad de estados mentales. En el caso de Stich, para establecer la identidad es condición

⁴ *Siempre*, pero no excluyente ni exclusivamente. El individualismo (metodológico) no prohíbe la individuación relacional de los estados mentales, solamente dice que ninguna propiedad de los estados mentales, relacional o de otra clase, cuenta taxonómicamente si no afecta a los poderes causales. (PS, p. 73-4)

suficiente la correspondencia de las bases físicas intrínsecas de superveniencia de los respectivos estados mentales en análisis, lo que implica la identidad psicológica de los doppelgängers. Para taxonomizar diferencialmente dos estados mentales, es necesario que éstos difieran en las propiedades físicas intrínsecas en las que supervienen respectivamente.

Un aspecto que me parece relevante señalar respecto de la concepción de Stich es que resulta compatible con una visión epifenomenalista de lo mental, dado que la superveniencia de lo mental en lo físico que propone no excluye la inercia causal de los estados mentales. De hecho, la superveniencia en general no parece excluir definitivamente o señalar una incompatibilidad palmaria con el epifenomenalismo de lo mental.⁵ Las clases de dualismo que excluye de plano sólo se refieren a las que proponen una independencia mental de lo físico, aún cuando lo mental se incluya en las explicaciones de la conducta que se formulen. Adicionalmente, al depender de la identidad de las bases físicas de superveniencia, las condiciones de identidad de estados mentales establecidas por el principio de autonomía requieren mecanismos para reconocer dicha identidad física, supeditando el avance de la psicología al de la ciencia que pueda hacer ese trabajo.

Por su parte, la posición de Fodor, al estipular condiciones de identidad de estados mentales basadas en la identidad de sus poderes causales, explícitamente se compromete con una afirmación sobre la eficacia causal de lo mental; no obstante, lo hace recurriendo también al expediente de la superveniencia, recurso que expone su concepción a las mismas dificultades metafísicas que afectan al principio de autonomía. Pese a esto, al vincular la determinación de los poderes causales de los estados mentales a sus propiedades de contenido estrecho, dado que propone que las diferencias en las propiedades intencionales amplias no constituyen diferencias de poder causal de dichos estados mentales, logra mantener un nivel de autonomía relativa respecto de las ciencias del cerebro, al tiempo que constriñe la influencia de propiedades históricas o relacionales no semánticas (aunque de un modo menos radical que Stich).

⁵ Al respecto véase Kim 1998.

VI. Síntesis.

A modo de síntesis, volvamos sobre algunas ideas centrales a las versiones de individualismo que hemos considerado hasta aquí: (a) los estados mentales son intrínsecos; (b) lo relevante para la psicología es la representación del mundo, no cómo sea el mundo; (c) los estados psicológicamente relevantes deben supervenir sobre los estados físicos internos actuales del organismo; y (d) lo relevante para la taxonomía y la explicación psicológica son los poderes causales de los estados mentales.

Lo que la psicología trata de identificar es la contribución cognitiva del individuo a su propia conducta; es decir, cómo los estados mentales de un individuo interactúan con otros estados mentales y cómo éstos, a su vez, causan su conducta. Esto no equivale a proponer que todos formamos nuestros estados mentales del mismo modo o que llegamos a ellos por las mismas vías, sino que estos aspectos no son relevantes para nuestra conducta subsiguiente; lo relevante para nuestra conducta es que estemos *en* esos estados antes que cómo llegamos a estar en ellos. Como Wilson dice, “dado que la psicología está interesada en gran parte en predecir y explicar la conducta cognitiva, debe ignorar cualquier diferencia entre individuos que no haga, ella misma, diferencia para el rol que algún estado mental cumple en la causación de conducta” (1995:7).

En lo que sigue, abordaré tres clases de argumentos que se habrían ofrecido para apoyar el individualismo, a saber: El argumento de los poderes causales; el argumento computacional; y dos argumentos metodológicos. Estas tres clases de argumentos serán expuestas y consideradas desde la perspectiva crítica que Wilson desarrolla en *Cartesian Psychology and Physical Minds: Individualism and the Sciences of the Mind*, libro publicado en 1995, cuya conclusión general es que el individualismo en psicología debe rechazarse. El objetivo de esta exposición de argumentos en pro y en contra del individualismo es someterlos a evaluación y, de acuerdo con los resultados de dicha evaluación, tomar posición respecto de la posibilidad fundada de sostener (o no) una visión individualista en psicología. En la primera estación de este recorrido (anti)individualista en psicología, someteremos a consideración el *argumento de los poderes causales*.

Capítulo 2 - El argumento de los poderes causales.

En el segundo capítulo de su libro, Wilson (1995) se ocupa de la afirmación de Fodor (1987:cap. 2) que propone que la individuación de estados mentales en psicología debe hacerse por sus poderes causales, dado que la naturaleza causal de la explicación científica exigiría individuación por poderes causales. Desde esta perspectiva, la psicología sólo sería un caso de un constreñimiento general en ciencia, un constreñimiento individualista. Inicio este capítulo exponiendo las ideas que, al respecto, Fodor presentó en *Psicosemántica*, para luego desplegar la crítica formulada por Wilson. Finalmente, presentaré una evaluación de lo expuesto.

I. Individualismo metodológico: el argumento de Fodor.

El contexto de la argumentación de Fodor en el Capítulo 2 de *Psicosemántica*, corresponde a la discusión sobre el tipo correcto de individuación de las actitudes proposicionales. Las alternativas, una individuación relacional o una no-relacional, se vinculan, además, con la violación o no de la superveniencia mente/cerebro. A favor de una individuación relacional de las actitudes proposicionales, que viola la superveniencia, se revisan argumentos de Putnam y Burge. Por su parte, Fodor apoya una individuación no-relacional o individualista —según un término usado por el propio Burge— que respeta la superveniencia. En el núcleo de esta discusión está la noción de contenido que necesita la psicología.

En *Psicosemántica*, Fodor se concentra el caso de la Tierra Gemela planteado por Putnam (1975), considerando que el argumento de Burge (1986) queda subsumido por el primero; en pro de la simplicidad y brevedad de la exposición, haré lo mismo. En pocas palabras, el asunto es aproximadamente así: existe un planeta idéntico en casi todo a la Tierra, la Tierra Gemela, salvo que allí hay XYZ donde nosotros tenemos H₂O. A simple vista ambas sustancias son indistinguibles, aunque una prueba química adecuada podría diferenciarlas sin problema. Además, existe allí un duplicado molecular suyo, su *gemelo*, que es exactamente idéntico a usted (incluyendo su microestructura neurofisiológica). Dadas estas condiciones, se supone que, en virtud de los hechos químicos, y a pesar de los neurofisiológicos, cuando ambos

emiten la forma de palabras “el agua moja”, ella significa cosas diferentes en vuestras bocas. De aquí se sigue que el *contenido* de su pensamiento y el de su gemelo cuando piensan que “el agua moja” es diferente. Finalmente, se llega a plantear que su gemelo no puede tener de ningún modo el pensamiento “el agua moja”, dado que lo que él piensa sería *algo así como* “el gagua [agua gemela] moja”.⁶

El caso de la Tierra Gemela mostraría que la diferencia en el contenido intencional de su pensamiento y el de su gemelo cuando tienen en la cabeza que “el agua moja” (*mutatis mutandis* que “el gagua moja”), no se acompaña de una correspondiente diferencia en la microestructura de sus respectivos cerebros, puesto que ambos son, siguen siendo, siempre y en todo momento, molecularmente idénticos. Esto es, sus pensamientos no supervienen en las respectivas microestructuras neurológicas. (Recordemos que, siendo *X* un estado superveniente e *Y* un estado base de superveniencia, hay superveniencia de los estados *X* en los estados *Y* si y sólo si *no hay* una diferencia en los estados *X* *sin* que haya una diferencia correspondiente en los estados *Y*. Para el caso que nos ocupa, los estados mentales de un organismo supervienen en sus estados cerebrales si y sólo si sus cerebros muestran diferencias siempre que lo hacen sus mentes). Pero el caso de los gemelos nos dice que las diferencias en el contenido intencional de sus pensamientos y los de su gemelo, diferencias determinadas por las respectivas diferencias de contexto (un planeta con XYZ y otro con H₂O), no afectan en modo alguno la identidad de vuestros estados cerebrales.

¿Por qué importa preservar la superveniencia mente/cerebro? En opinión de Fodor, se trata de la mejor idea para hacer inteligible la causación mental. La explicación psicológica causal científica descansaría, de acuerdo con él, en aceptar la superveniencia de los poderes causales de los estados mentales en los poderes causales neurológicos/cerebrales. Por otra parte, y dado que, en general, la empresa científica está interesada en ofrecer explicaciones *causales* de aquellos acontecimientos de la naturaleza que *admiten* ser explicados causalmente, una psicología (cognitiva) científica, en particular, también debería ofrecer explicaciones causales

⁶ Para todos los efectos, ambas **formas de palabras** se oyen y piensan igual; no pueden ser distinguidas por oyente o pensante alguno cuando son proferidas o pensadas por usted o su gemelo.

de los acontecimientos que la ocupan. Para dar estas explicaciones, la psicología (cognitiva) debe ser capaz de formular generalizaciones causales que subsuman a los estados mentales en virtud de sus propiedades causales para producir otros estados mentales y/o conductas. Y para hacer esto último, debe contar con un criterio que le permita individuar estados mentales en virtud de sus poderes causales relevantes para la producción de otros estados mentales y/o de conducta (Fodor 1987). Para completar el cuadro, Fodor afirmará que “la individuación en ciencia es siempre individualista” (1987:59), afirmación para la que considera necesario ofrecer “una especie de argumento metafísico” (p. 59), argumento cuyo desarrollo implicaría recortar con precisión la diferencia entre una individuación ‘no-relacional’ y una individuación ‘individualista’.

Para alcanzar el núcleo del argumento fodoreano, necesitamos familiarizarnos con el ejemplo que, se supone, ilustra el punto; el asunto va así⁷: Suponga que usted tiene una moneda genuina de diez pesos chilenos, esto es, una moneda fabricada por la Casa de Moneda de Chile, con dos configuraciones fijas, a saber: ‘Cara’ y ‘Sello’. Por bien de la teoría, se dejan de lado las monedas que se sostienen de canto; tal como indica Fodor, ‘¡ninguna teoría es perfecta!’ (1987:59). Pese a lo escaso de su valor, esta moneda le permite a usted tener el control sobre cada partícula física en el universo. ¿Cómo? Muy fácil. Definimos el predicado “es una partícula *C* en el tiempo *t*”, de modo que sea satisfecho por una partícula en el tiempo *t*, si y sólo si la moneda de diez pesos tiene la cara hacia arriba en *t*. Del mismo modo, definimos el predicado “es una partícula *S* en el tiempo *t*”, de modo que sea satisfecho por una partícula en el tiempo *t*, si y sólo si la moneda de diez pesos tiene el sello hacia arriba en *t*. Habiendo hecho esto, resulta que, a su entero arbitrio, dependiendo de la posición de su moneda, todas las partículas del universo son *C* o *S* según la cara de la moneda esté hacia arriba o lo esté el sello. Lo mismo es válido, por cierto, para objetos de un nivel superior, como podrían serlo los estados *C* o *S* de cada cerebro en el universo.

⁷ He cambiado algunos detalles para hacer más familiar el ejemplo. En cualquier caso, estos ajustes no tienen relevancia para lo que está en discusión.

Para Fodor no hay, en principio, ningún problema con todo lo anterior, en tanto no se pretenda generar una física de partículas que integre en su aparato explicativo partículas *C* o *S*. La razón para considerar inaceptable tal pretensión estriba en que la física está orientada a proporcionar explicaciones causales y ser una partícula *C* o *S* es, justamente, irrelevante para los poderes causales de tal partícula. La irrelevancia se evidencia mediante el siguiente contrafáctico: ‘Si un suceso *e* es causado por una partícula *C*, digamos por la partícula *p*, entonces el mismo suceso *e* es también causado por *p* en el mundo posible nomológicamente más próximo en el que *p* es *S* en vez de *C*.’ (Fodor 1987:61) Lo mismo vale, *mutatis mutandis*, para las propiedades de ser estados cerebrales *C* y *S*; *mutatis mutandis*, para las propiedades de ser estados mentales *C* y *S*; y, volviendo sobre lo discutido más arriba a propósito del contenido de los estados mentales, *mutatis mutandis* para la propiedad de ser un estado mental de una persona que vive en un mundo en cuyos charcos hay XYZ en lugar de H₂O. Desde el punto de vista de alguien interesado en la explicación causal de la conducta de estas personas, carece de sentido distinguir entre sus estados mentales, pues no habría diferencia en los poderes causales de uno y otro, como no lo hay entre estados mentales *C* o *S*.

De acuerdo con lo dicho más arriba, no habría diferencias en los efectos de las partículas *C* o *S*, tal como no lo habría entre sus comportamientos y los de su gemelo. No aceptar esto implicaría asumir que precisamente *los poderes causales*, no cualquier propiedad relacional, de la totalidad de las partículas físicas del universo, incluso de las que están muy alejadas de nosotros, dependen, *por estipulación*, de la orientación de su moneda de diez pesos. ¿Cómo podría esto ser posible? La respuesta de Fodor es que ‘o bien tendría que haber un mecanismo causal mediara esta dependencia, o bien tendría que estar mediada por una ley fundamental de la naturaleza. Pero no hay tales mecanismos, ni hay tales leyes. *Por supuesto* que no las hay.’ (Fodor 1994:68)

Para el caso de los estados mentales el razonamiento es similar. Si, volviendo al ejemplo de la Tierra Gemela esbozado más arriba, cuando usted dice “el agua moja” su comportamiento es diferente del comportamiento de su gemelo cuando dice “el gagua moja”, esto debería evidenciar una diferencia en los poderes causales de sus estados mentales correspondiente a la

diferencia en sus conductas. Para que esto sea así, debería haber algún mecanismo que conecte los poderes causales de sus estados mentales con la materia que hay en los charcos por donde usted vive, pero que los conecte sin afectar su fisiología (su cerebro) pues, si lo hiciera de otro modo, usted y su gemelo dejarían de ser molecularmente idénticos, algo que el caso de Putnam supone. ¿Pero cómo, si los estados mentales supervienen en los estados cerebrales, se podría afectar a los primeros sin afectar a los últimos? Según Fodor:

No se puede afectar a los poderes causales de los estados mentales de una persona sin afectar su fisiología. Por supuesto, eso no es hacer una afirmación conceptual, o una afirmación metafísica. Es un hecho contingente acerca de cómo Dios hizo el mundo. Dios hizo el mundo de tal manera que los mecanismos gracias a los cuales las variables ambientales afectan el comportamiento del organismo tienen lugar vía sus efectos en el sistema nervioso del organismo. O eso, al menos, es lo que me aseguran todos los fisiólogos que conozco. (Fodor 1994:69-70)

Si efectivamente las diferencias contextuales no pueden afectar a los poderes causales de los estados mentales sin afectar a los estados cerebrales, Fodor se pregunta: ‘¿Cómo podría nadie haber supuesto alguna vez que los criterios de individuación que son apropiados para los propósitos del psicólogo son otros que los individualistas?’ (1994:73) Una hipótesis que él propone alude a la posible confusión entre *solipsismo metodológico* e *individualismo*. El solipsismo metodológico, como lo indicamos en el Capítulo 1, es la doctrina que defiende que los estados psicológicos se individualizan sin relación alguna con su evaluación semántica. El individualismo es la doctrina según la cual los estados psicológicos se individualizan con relación a sus poderes causales. ¿Qué consecuencias tiene una eventual confusión? Si la evaluación semántica de un estado mental depende de propiedades relacionales (de cómo este se corresponde o no con el mundo), la individuación solipsista no es relacional. Pero, según Fodor, ‘no hay nada que ponga freno a que los principios de individuación sean, simultáneamente, relacionales e individualistas. *El individualismo no prohíbe la individuación relacional de los estados mentales*. Solamente dice que ninguna propiedad de los estados mentales, relacional o de otra clase, cuenta taxonómicamente a menos que afecte a los poderes causales. (Fodor 1994:73-4)

Lo anterior permite a Fodor, reconociendo la frecuencia con la que las categorías taxonómicas en ciencia son relacionales, afirmar que éstas respetan el constreñimiento individualista, pues

los criterios taxonómicos que emplean afectan a los poderes causales. El caso que Fodor presenta es el de la propiedad de “ser un planeta”. Al respecto señala:

Así, “ser un planeta” es una propiedad relacional por excelencia, pero es una propiedad que el individualismo permite utilizar en la taxonomía astronómica. El que algo sea un planeta afecta a su trayectoria y su trayectoria determina aquello con lo que ese algo se puede topar. De este modo, que algo sea un planeta afecta a sus poderes causales. Esto es todo lo que pide el individualismo. Equivalentemente, la propiedad de ser un planeta es una propiedad taxonómica, porque hay leyes causales que satisfacen las cosas en virtud de ser planetas.

... La observación operativa es, por supuesto, que donde no hay leyes causales con relación a una propiedad, el hecho de tener esa propiedad —o el de no tenerla— no tiene efecto en los poderes causales. (Fodor 1994:74)

Con estos antecedentes paso ahora a presentar los argumentos críticos de Wilson.

II. Argumentación de Wilson.⁸

Wilson denomina a la propuesta fodoriana que acabamos de esbozar respecto al individualismo de la empresa científica dedicada a la explicación causal *individualismo global* [en adelante, IG]. El núcleo de la crítica que Wilson dirige contra IG se ordena en torno mostrar que el individualista emplea la noción de ‘poder causal’ en dos sentidos diferentes e incompatibles, a saber: un sentido ‘extenso’, que debe admitir al menos algunas propiedades relacionales e históricas (concebidas como poderes causales extrínsecos); y un sentido ‘restringido’, que afirma la superveniencia de los poderes causales en las propiedades físicas intrínsecas de la entidad que los tiene. Estos dos sentidos de ‘poder causal’ no permitirían al individualista satisfacer simultáneamente las premisas (1) y (4) en lo que Wilson denomina *argumento de los poderes causales* [en lo que sigue, APC]:

(1) Las propiedades y entidades taxonómicas en las ciencias deben individuarse por sus poderes causales.

(2) Las ciencias cognitivas, particularmente la psicología, pretenden identificar causas *mentales* taxonómicas y formular generalizaciones acerca de esas causas.

(3) En las ciencias cognitivas, las causas mentales de la conducta de un individuo y la conducta misma deben individuarse en términos de los poderes causales de ese individuo.

⁸ Todos los números de página de las citas de Wilson corresponden a Wilson 1995.

Entonces, dado que

(4) Los poderes causales de cualquier cosa son determinados o fijados por las propiedades físicas intrínsecas de esa cosa,

(5) Cualesquiera causas de la conducta que vayan a ser taxonómicas en las ciencias cognitivas deben ser determinadas o fijadas por las propiedades físicas intrínsecas del individuo. En consecuencia,

(6) Las ciencias cognitivas, particularmente la psicología, se interesarán sólo en estados y procesos que en sí mismos sean determinados por las propiedades físicas intrínsecas del individuo. (p. 32; negrita mía).

Dado que Wilson asume que este argumento expone adecuadamente el núcleo del constreñimiento individualista *para* la psicología, si está en lo correcto APC no se seguiría, porque si (1) es verdadera de las ciencias, entonces los poderes causales (en su sentido ‘extenso’; es decir, cuando ‘poderes causales’ incluye propiedades relacionales o históricas) no supervienen localmente, esto es, no supervienen en las propiedades intrínsecas de las entidades en las que se instancian; y si (4) es verdadera, entonces las ciencias no individualizan sólo en términos de poderes causales (en sentido ‘restringido’; es decir, cuando ‘poderes causales’ sólo incluye propiedades que supervienen localmente). Esta ambigüedad en el significado y consecuente uso del término ‘poder causal’ señalaría una debilidad sustantiva en el argumento individualista.

Como fondo de la evaluación que Wilson hace de (1), está su crítica al carácter *a priori* de IG, crítica que él funda defendiendo una apelación metodológica a la práctica científica, a la ciencia *tal como se practica* que, en su opinión, debe operar como ‘un chequeo para tales afirmaciones sobre la naturaleza de la explicación científica’ (p. 41). Al apriorismo del criterio individualista, Wilson pretende oponer un criterio *a posteriori*, sustentado en lo que en una amplia gama de ciencias existe y se acepta como taxonomías relevantes. A continuación desarrollo los argumentos específicos de Wilson contra APC, donde estos puntos son tratados con mayor detenimiento.

II.1. Individualismo global [IG]

La premisa (1) en APC establece IG, la afirmación de que la naturaleza causal de la explicación científica supone individuación por poderes causales. La idea está tomada de un pasaje de Fodor en *Psicosemántica*:

Queremos que la ciencia dé explicaciones causales de las cosas (eventos, lo que sea) en la naturaleza, en tanto puedan ser causalmente explicados. Dar tales explicaciones involucra esencialmente proyectar y confirmar generalizaciones causales. Y las generalizaciones causales subsumen las cosas a las que se aplican en virtud de las propiedades causales de las cosas a las que se aplican. Por supuesto.

Brevemente, lo que necesitas en orden a hacer ciencia es un aparato taxonómico que distinga entre cosas en tanto tengan *diferentes* propiedades causales, y que agrupe las cosas en tanto tengan las *mismas* propiedades causales. (1987:34, omitida nota al pié).

En opinión de Wilson, la dificultad con esta afirmación nace porque no distingue adecuadamente entre *poderes causales* y *propiedades causales*. Una diferencia importante, según él, es que los poderes causales están orientados hacia adelante [*forward-looking*] de una forma en la que no lo están las propiedades causales. De modo que ‘las similitudes causales relevantes entre dos fenómenos en una disciplina dada pueden involucrar las *causas* de esos fenómenos o las relaciones causales en las que se hallan, antes que lo que esos fenómenos son capaces de causar’ (p. 33).

Esta puntualización de Wilson busca establecer una diferencia entre individualismo y funcionalismo; como constreñimiento, el individualismo sería más restrictivo que el funcionalismo, dado que, en la perspectiva funcionalista, dos estados mentales se individualizarían como pertenecientes a la misma clase según sus roles causales o funcionales totales, sin distinguir entre sus causas y sus efectos. Un funcionalista distinguiría dos estados mentales, S_1 y S_2 , si S_1 sólo pudiera ser causado por C_1 y C_2 , mientras S_2 sólo pudiese ser causado por C_3 , incluso si los efectos totales de S_1 y S_2 fuesen idénticos a través de todas las circunstancias posibles. Por el contrario, un individualista taxonomizaría bajo la misma categoría a S_1 y S_2 , dado que no hay diferencias en sus efectos; según el individualista, tienen poderes causales idénticos.

La distinción propuesta entre poderes causales y propiedades causales afectaría a APC, aun cuando la noción de propiedad causal se restringiese para incluir sólo poderes causales intrínsecos y extrínsecos, es decir, aun cuando no se permitiese que cualquier predicado relacional o histórico denominase a una propiedad causal, sino sólo aquellos predicados históricos o relacionales que fuesen, además, poderes causales (extrínsecos).⁹ En cualquier caso, el problema aquí es, según Wilson, que ‘si la individuación en ciencia es individuación por poderes causales *porque* la explicación científica **es** causal, entonces la naturaleza causal de otros tipos de explicación debe entrañar que las entidades a las que se refieren deban individuarse por sus poderes causales’ (p. 34; negrita mía). Si este no es el caso, la falsedad del individualismo como el modo característico de individuación en ciencia sería evidente.

Para ilustrar que, como cuestión de hecho, esto último es lo que efectivamente sucede, Wilson recurre primero a un ejemplo no-científico: la aplicación a un grupo de individuos del predicado ‘es una víctima de la bomba de Hiroshima’. Lo que determina la aplicación correcta de este predicado y, por ello, la pertenencia de un individuo a la clase que designa, son factores históricos de ese individuo o de sus padres, no los poderes causales que posee. Por haber estado en Hiroshima cierto día de agosto de 1945, o tener padres que estuvieron allí ese día, algunas personas padecen degeneración celular, tumores cancerosos y enfermedades genéticas, *causadas* por el ataque nuclear sobre esa ciudad.

Un aspecto que se destaca del ejemplo es que las generalizaciones verdaderas que pueden hacerse sobre estas personas tienen carácter sistemático y que resulta explicativo aludir a que se trata de una víctima de la bomba de Hiroshima al responder consultas respecto al estado de sus cuerpos. Todas estas personas comparten una propiedad causal, cuyo criterio de individuación no toma en cuenta sus poderes causales sino aspectos de su historia. Además, se

⁹ Wilson admite aquí la equivalencia entre ‘propiedades causales’ y ‘poderes causales’, a condición de que se incluyan poderes causales extrínsecos. Así, la distinción entre propiedades y poderes causales aparece menos sustantiva de lo que Wilson pareció afirmar inicialmente. El individualismo, como Fodor lo señala claramente, sólo exige que la individuación sea *por poderes causales*.

trata de un tipo de individuación relacional ubicua respecto de grupos de personas (por ejemplo, graduados universitarios, jubilados) y de sus integrantes.

La conclusión de lo anterior es que dar una explicación causal de un fenómeno no entraña la individuación por poderes causales de las entidades que constituyen el fenómeno o de aquellas a las que se hace referencia en la explicación. Del mismo modo, que la ciencia proporcione explicaciones causales y que individúe por propiedades causales no supone un compromiso con la explicación y la individuación por poderes causales. Antes de avanzar, Wilson evalúa dos posibles estrategias defensivas individualistas.

Una primera respuesta individualista a lo señalado, a partir del ejemplo de la ‘víctima de la bomba de Hiroshima’, podría apelar a que el hecho de que existan explicaciones de sentido común —que no individualizan por poderes causales—, no dice nada respecto a las explicaciones científicas. Apoyar una distinción entre explicaciones causales de sentido común y explicaciones científicas exige proporcionar criterios de demarcación. Algunos de esos criterios podrían indicar que las explicaciones científicas son proyectables, instancian leyes, y cuantifican sobre clases naturales en una forma particular. Wilson piensa que no es sencillo establecer el criterio de demarcación requerido, pues las explicaciones de sentido común satisfarían también estos criterios. Como ejemplo de estas dificultades examina la proyectabilidad, indicando que ser proyectable es, entre otras cosas, ser contrafácticamente rigurosa; pero las explicaciones de sentido común serían tan contrafácticamente rigurosas como las explicaciones científicas. Particularmente, afirma Wilson, ‘las generalizaciones acerca de entidades individuadas histórica o relacionalmente no son menos proyectables o científicas que cualesquiera otros tipos de generalizaciones causales’ (p. 37). En consecuencia, para el defensor del individualismo esta sería una vía poco prometedora de evitar la apelación a clases individuadas por propiedades históricas o relacionales.¹⁰

¹⁰ Quisiera mantener en la mente de quien esto lee, que el punto de vista individualista (al menos para Fodor) no prohíbe, en cuanto tales, propiedades históricas o relacionales en la explicación científica causal, siempre que éstas afecten a los poderes causales. La cuestión de si es o no posible conjugar ambos aspectos se discute más adelante.

Una segunda respuesta del individualista, en principio más prometedora, acepta que tanto en las explicaciones científicas como en las no-científicas figuran clases relacionales o históricas, pero indica que lo hacen siempre con un carácter preliminar respecto a taxonomías que sí satisfacen el constreñimiento que impone IG. El carácter preliminar de las taxonomías históricas y relacionales se entiende porque, en sí mismas, las propiedades históricas y relacionales no son causalmente responsables, en último término, de lo que hacen o pueden hacer las entidades que instancian tales propiedades. Muchas taxonomías que no individualizan entidades por poderes causales se corresponden con taxonomías que sí lo hacen. Que las taxonomías científicas agrupen entidades que comparten sus poderes causales sería una consecuencia de reconocer que son precisamente los poderes causales, en tanto propiedades intrínsecas de las entidades en las que se instancian, los causalmente responsables de los efectos que se les adscriben en las explicaciones causales.

Una revisión de este tipo es la ejemplificada por el estrechamiento del concepto de *peso* al de *masa* en la mecánica newtoniana (Stalnaker 1989:281). La propiedad extrínseca (peso) se descompone en una propiedad intrínseca (masa) más una relación (fuerza gravitacional). Aplicado al ejemplo puesto en consideración por Wilson, el asunto quedaría así:

Suponga que la siguiente generalización es verdadera: la mayoría de las víctimas de la bomba de Hiroshima sufrió los efectos de radiación de un tipo específico, digamos, alguna forma específica de cáncer. Aunque podríamos usar el predicado histórico ‘es una víctima de la bomba de Hiroshima’ para seleccionar un grupo particular de personas, lo que determina si la presente generalización es verdadera de *individuos particulares* es algo acerca de esos individuos en sí, no algo acerca de sus historias causales. La generalización se aplica a individuos particulares en virtud de algunos rasgos físicos intrínsecos, tales como la mutación de un gen particular o la destrucción de ciertas células. Lo que *realmente* explica por qué la gente con una propiedad histórica, relacional —estar en una cierta región espacio-temporal— tiene un tipo de cáncer específico es que estas personas ahora tienen la propiedad física intrínseca que *causa* el cáncer. (Wilson 1995:37-8)

De este modo, la taxonomía inicial, basada en una propiedad histórica, nos conduce a una taxonomía individualista, que identifica una propiedad intrínseca de los individuos afectados por el cáncer; una propiedad causalmente responsable en último término.

El cuestionamiento de Wilson a esta propuesta sobre el carácter preliminar de las taxonomías históricas y relacionales se apoyará en la revisión de algunos casos, especificados en la próxima sección, que él considera relevantes y lo bastante variados como para exponer su ubicuidad en la práctica científica. Su análisis pretende mostrar que la propuesta de revisión individualista es imposible en dichos casos, con la consecuencia de echar dudas sobre su plausibilidad general y dificultando su defensa como posibilidad de apoyar el caso de la psicología en particular. En síntesis, IG resultaría falso.

II.2. Dos sentidos para ‘poderes causales’.

Vamos a considerar ahora el núcleo del análisis que hace Wilson de APC. El inicio es la exposición de dos *puntos*, planteados por Fodor en *Psicosemántica*, que condensarían la discusión sobre IG:

Punto metodológico: la categorización en ciencia es característicamente taxonomía por poderes causales. Identidad de poderes causales es identidad de consecuencias causales a través de contextos nomológicamente posibles.

Punto metafísico: los poderes causales supervienen en la microestructura local. En el caso psicológico, supervienen en la estructura neuronal local. (1987:44)

Respecto de estos dos puntos, y sus correlatos en APC, las premisas (1) y (4), se plantea la imposibilidad para el individualista de satisfacerlos simultáneamente, generándose una ambigüedad de significado para ‘poder causal’, lo que permite a Wilson articular su propuesta de un sentido ‘extenso’ y uno ‘restringido’ del término. Pero, aun cuando pudiesen citarse casos de taxonomías científicas que no clasifican entidades por sus poderes causales, también podría decirse que las entidades así taxonomizadas figuran sólo como *explananda* en explicaciones causales y que es en tanto *explanantia* que reviste importancia la clasificación por poderes causales intrínsecos. En opinión de Wilson, decir que las clases individuadas en términos relacionales e históricos sólo aparecen como *explananda* es ‘la clase de afirmación reconstructiva acerca de la naturaleza de la explicación científica de la que uno debe estar precavido’ (p. 40).

Lo antes señalado le permite a Wilson, no obstante su carácter polémico, centrar la discusión en casos que considera *cruciales*, aquellos donde clases individuadas relacionales o históricamente sí figuran como *explanantia* en explicaciones causales y que no son revisables al modo que comentamos hacia el final de la sección anterior. Los casos en análisis son tres: *tabú*, en antropología; *criminal*, en sociología; y *especie*, en biología.¹¹ Veamos cada caso con algo más de detalle.

Según Wilson, la noción de un *tabú*, como actitud social o creencia, es central dentro de la disciplina antropológica. Pese a las diferencias que puede haber entre las prácticas que podrían caer bajo la denominación ‘tabú’, la afirmación que importa es que la individuación de los tabúes no se hace en función de sus poderes causales intrínsecos, no interesa la ‘microestructura local’ de esas acciones o prácticas, privilegiándose su localización en una red compleja de otras prácticas morales o sociales.

Para el caso de la categoría *criminal* la cosa no sería fundamentalmente distinta. Ser un criminal depende de las relaciones en las que un individuo entra en virtud de romper o no ciertas leyes, no es función de los poderes causales intrínsecos de ese individuo. A propósito de este caso en particular, Wilson discute dos posibles observaciones. Una, relativa a la existencia de teorías sobre la criminalidad que sí identifican factores intrínsecos para explicarla (por ejemplo, la teoría del hombre criminal de Lombroso) y que deberían ser falsas si él está en lo correcto; la otra, relacionada con la eventual vacuidad teórica de la clase criminal, resultante de que los criminales, en tanto criminales, no compartieran ninguna propiedad intrínseca. Sobre la primera observación señala que, aún si efectivamente hubiese teorías que identificaran propiedades que expliquen por qué algunas personas presentan conductas criminales, lo harían sólo para casos individuales, no proporcionando necesariamente criterios para individuar la clase criminal. Sobre la segunda observación dice que, incluso si ninguna teoría explicase adecuadamente la conducta criminal, esto no implicaría que no hubiese criminales ni generalizaciones interesantes sobre los criminales.

¹¹ No obstante revisar estos tres casos, Wilson menciona también *género y preferencia sexual* en Sociología; *clase*, en Economía e Historia.

Un par de mínimos comentarios sobre las últimas observaciones de Wilson. Respecto de la primera de estas: Siempre es posible que todo pueda salir mal y que, pese al esfuerzo, no se dé con la teoría que se busca. No me parece para nada claro la forma en que esto puede ir específicamente contra la afirmación individualista. Si una teoría no hace el trabajo requerido, en este caso identificar clases taxonómicamente relevantes por sus poderes causales, hay que seguir buscando. Respecto de la segunda observación, tampoco resulta evidente el modo en que la mera existencia de generalizaciones ‘interesantes’ o usos lingüísticos cualesquiera podrían *a priori* contar contra el individualismo. Como cuestión de hecho, se habla de brujas y podría haber generalizaciones interesantes sobre estas criaturas. ¿Y qué?

Finalmente, se analiza el concepto de especie en biología evolutiva. Una especie es concebida relacionamente como ‘una comunidad reproductiva de poblaciones (reproductivamente aislada de otras) que ocupa un nicho específico en la naturaleza’ (Mayr 1982:273). Para Wilson, entender a las especies como clases individuadas relacionamente implica dos cosas: Primero, que las propiedades intrínsecas de un individuo de una especie no son decisivas para determinar su calidad de miembro de la misma; segundo, que, en tanto clases, las especies no son individuadas a partir de otras por sus poderes causales, sino que lo son, en parte, por su historia filogenética.

La opinión final de Wilson es que las categorías examinadas, que corresponden a campos científicos diversos, son absolutamente centrales para las prácticas explicativas en las respectivas disciplinas y, además, no admiten un proceso de revisión que remita a una propiedad intrínseca responsable, en último término, de sus capacidades causales. Conceder esto último sería aceptar taxonomías que individúan por propiedades históricas o relacionales, propiedades que figuran como *explanantia* en las respectivas disciplinas y que, por eso, harían falso el punto metodológico de Fodor, (1) en APC, pues se trata de una individuación que no atiende a propiedades intrínsecas, los poderes causales. Y si (1) en APC debe ser verdadera, atendiendo a la práctica efectiva en ciencia, eso exigiría construir de modo ‘extenso’ la noción de poder causal, permitiendo que incluya algunas propiedades históricas o relacionales, en

cuyo caso (4) debería ser falsa, pues estas propiedades no supervienen en la microestructura física local.

Antes de avanzar, voy a plantear algunas consideraciones sobre lo presentado hasta aquí en esta sección. En primer término señalar que, pese a la insistencia de Wilson en afirmar que las clases analizadas por él figuran como *explanantia* en explicaciones ofrecidas por las disciplinas que las emplean, no ofrece ningún ejemplo de tal uso que pueda someterse a análisis. Uno podría imaginar, siguiendo el modelo de la propiedad de ‘ser una víctima de la bomba de Hiroshima’, que sería posible decir cosas como que el comportamiento de ciertos individuos es tal o cual *porque* en su comunidad existe tal o cual tabú (que, casualmente, prohíbe o prescribe *ese* comportamiento); o que una persona violó tal o cual ley *porque* es un criminal (que, casualmente, es en lo que *consiste* ser un criminal); o que un animal determinado realiza esta o esta otra conducta *porque* pertenece a tal o cual especie (que, casualmente, es una conducta *propia* de la especie). Si estos ejemplos de participación como *explanantia* son correctos, entonces, dado que tengo hijos que son sobrinos *porque* tengo hermanos, ser un hermano pudiese contar como la propiedad taxonómica *causalmente* responsable de mi tener hijos que sean sobrinos.¹²

En segundo término, es importante señalar que así como no se ofrecen casos de explicaciones donde figuren las clases analizadas por Wilson, tampoco ofrece argumentos que sustenten su afirmación de imposibilidad de revisión de tales clases, en términos de una descomposición que permita especificar alguna propiedad que satisfaga el criterio individualista. Si tal posibilidad no está bloqueada permanentemente, la afirmación de Wilson en este sentido se debilita.

En tercer lugar, la existencia de prácticas científicas que *de facto* pudiesen violar el constreñimiento individualista no hace *ipso facto* falso al individualismo, porque se requiere saber cómo van a interpretarse dichas prácticas, lo que motiva otra discusión.

¹² Fodor considera que este último caso (Cf. Fodor 1990), es similar al planteado con la moneda de diez pesos que controla las partículas del universo. En la sección siguiente revisaremos su posición al respecto.

En cuarto y último término, dado que la cuestión del individualismo se vincula estrechamente con la taxonomía apropiada a los fines de la explicación científica, se hace necesario plantear cuál es la idea que Fodor tiene en mente sobre la explicación científica. Al respecto, en *El olmo y el experto*, dice:

Doy por sentado que la explicación empírica es típicamente una cuestión de subsunción de acontecimientos [...] en el dominio de una ciencia bajo leyes que se articulan en su vocabulario teórico propio. (Fodor 1997:19)

Puede verse que la visión de Fodor sobre el modelo de explicación en ciencia guarda relación con la propuesta hempeliana de explicación por leyes cubrientes. Recordemos, además, que para Fodor estas leyes harán referencia precisamente a las propiedades relevantes (los poderes causales) y que, si no hay leyes que seleccionen determinada propiedad, tener o no tener esa propiedad no afecta a los poderes causales de las entidades que figuran en determinada explicación científica. En el mismo texto recién citado, Fodor amplía algo el tema señalando que su fidelidad a esta idea, aplicada al caso específico de la psicología, se basa en que

‘...es difícil dudar que por lo menos algunas regularidades psicológicas tienen aspecto legal (por ejemplo, que la luna parece más grande cuando está sobre el horizonte; que las figuras de Müller-Lyer se ven de diferente longitud; que todos los lenguajes naturales contienen nombres)’. (p. 19).

Hechas estas mínimas puntualizaciones, volvamos con Wilson que, tras dar por establecido que la taxonomía científica exhibe casos *cruciales* de individuación por propiedades relacionales o históricas no reductibles a propiedades intrínsecas ni supervenientes en la microestructura local, pasa a refutar la afirmación de Fodor que propone una compatibilidad entre individuación relacional e individuación por poderes causales. Recapitulando palabras de Fodor:

Tal como podría esperarse, las propiedades relacionales pueden contar taxonómicamente siempre que afecten a los poderes causales. Así, ‘ser un planeta’ es una propiedad relacional por excelencia, pero una que el individualismo permite que opere en la taxonomía astronómica. Pues el que usted sea un planeta afecta su trayectoria, y su trayectoria determina con qué puede toparse; entonces, el que usted sea un planeta *afecta a sus poderes causales*, que es todo lo que el individualismo pide. (1987:43; *itálica mía*).

Un primer paso contra esta afirmación se despliega a propósito de una crítica de Stalnaker (1989:307-8), en el sentido de un exceso de liberalismo en Fodor para construir su noción de

individualismo, de modo de permitir que las propiedades relacionales que ‘afectan’ a los poderes causales, cuenten como taxonómicamente válidas. Según Stalnaker, si el individualismo fuese la tesis de que la individuación es por *lo que afecta a los poderes causales*, esta modalidad de construcción del individualismo sería compatible con una individuación amplia en psicología, ya que los hechos ambientales claramente afectan en forma causal los poderes causales que tienen los objetos, incluyendo sus poderes causales intrínsecos. Siendo así, la versión fodoriana del individualismo sería insuficiente para posibilitar que IG sea una base sólida para apoyar el individualismo en psicología.

El segundo paso cuestiona puntualmente el significado de ‘afectar’ *los poderes causales*, según aparece en la última cita de Fodor. Habría dos formas en las que los poderes causales de algo podrían ser afectados: (a) Por las relaciones en las que entra ese algo en un momento dado; y (b) por la historia de ese algo, que determina que tenga los poderes causales que tiene. En opinión de Wilson, Fodor se orienta en el sentido de (a), y que denomina *afección contemporánea*. Básicamente, su argumento crítico se basa en proponer una modificación de (1) en APC para que incluya la afección contemporánea (incluso la histórica). Según Wilson, el resultado de esta modificación es que las propiedades que deberían aceptarse, modificando convenientemente (1) en APC para que incluya, además de poderes causales, propiedades que ‘afectan los poderes causales’, harían falsa (4), porque tales propiedades no supervienen sobre la microestructura local, e imposibilitarían inferir (5) y (6) en APC. Así expone Wilson su veredicto sobre este punto:

No hay un constreñimiento simple, el tener que ser ‘por poderes causales’, que se refleje en la práctica taxonómica real en ciencia y que especifique propiedades que supervengan en las propiedades físicas intrínsecas de las entidades en las extensiones de las clases resultantes. (Wilson 1995:49)

El ejemplo del propio Fodor, de un caso donde una propiedad relacional cuenta como propiedad taxonómica en ciencia, es el de la propiedad de ‘ser un planeta’. Algo es un planeta si mantiene una cierta relación con una estrella, más precisamente si la orbita. Esta relación distingue o puede distinguir dos trozos de materia idénticos. El punto es que un duplicado de la Tierra, incluso uno que pudiera moverse con la misma velocidad, no sería un planeta si no

orbita una estrella. Esta propiedad es claramente aceptada en astronomía y no parece revisable en virtud de algún carácter preliminar. Aún cuando haya casos en los que una propiedad relacional pueda revisarse y factorizarse en al menos un componente interno, planeta y especie (por citar dos casos) no serían, en principio, candidatos adecuados para dicha operación. De cualquier modo, el que no todas las clases científicas puedan revisarse en el sentido de identificar propiedades intrínsecas responsables de sus capacidades causales bastaría, según Wilson, para tener que reconocer el fracaso de APC sobre la base de IG.¹³

Este último párrafo amerita una pausa. En una nota a pie de página de *Psicosemántica*, Fodor se extiende sobre una crítica de Martín Davies (1986) que, me parece, va en una dirección semejante al análisis que Wilson presenta aquí sobre la propiedad de ser un planeta. De dicha nota, reproduzco lo más directamente vinculado con el asunto:

‘... Martín Davies señala que lo que digo acerca de ser un planeta “parece estar en tensión con la insistencia en que los poderes causales se deben comparar a lo largo de los contextos o ambientes. Porque no puede suceder *a la vez* que un planeta tenga poderes causales peculiares —y no simplemente los poderes causales de un trozo de materia que es físicamente similar, pero que no es un planeta—, y que los poderes causales tengan que ser comparados a lo largo de los contextos o ambientes de una forma completamente general”. Pero, para decirlo rudamente, esto es confundir la cuestión acerca de si *ser un planeta* es una propiedad taxonómica (que lo es; dos cosas que difieren en esa propiedad difieren *ipso facto*, en sus efectos en muchos contextos), con la cuestión de si *ser este trozo de roca* es una propiedad taxonómica (que no lo es; dos cosas que difieran en *esa* propiedad no difieren, a consecuencia de ello, en sus efectos en cualquier contexto).

Una vez más: una diferencia entre las propiedades P y P_I puede afectar a los poderes causales (puede que relevantes taxonómicamente) sólo cuando hay una situación S tal que la instanciación de P en S tiene, *ipso facto*, efectos diferentes de la instanciación de P_I en S . Según este criterio, la diferencia entre *ser un planeta* y *no ser un planeta* afecta a los poderes causales, porque hay situaciones en las que algo que es un planeta tiene, *ipso facto*, efectos diferentes de algo que no lo es... (Fodor 1987:n. 9)

¹³ Sobre lo anterior cabe una cierta restricción planteada por el propio Fodor. La taxonomía por poderes causales es exigible no de todas y cualquier ciencia, sino sólo de aquellas interesadas en la explicación causal. Expresamente, señala Fodor que ‘no hay necesidad de dogmatizar. Puede haber empresas científicas que no estén —o no lo estén primariamente— interesadas en la explicación causal como, por ejemplo, la historia natural.’ (Fodor 1994:Cap. 2, n. 3).

Retomando el argumento de Wilson, hay al menos dos aspectos que merecen comentario. Primero el más simple. De acuerdo con la última cita de Fodor, la existencia de un mundo posible donde un trozo de roca idéntico en todo a la Tierra, salvo en que no orbita una estrella, no constituye un problema para la afirmación individualista de que ‘ser un planeta’ es una propiedad taxonómica en astronomía, aun cuando sea una propiedad relacional. Ser un planeta y ser un trozo de roca idéntico a un planeta son propiedades diferentes. Ser un planeta afecta a los poderes causales, ser un trozo de roca, no.

El segundo aspecto, algo más dificultoso, se refiere a la crítica de Stalnaker sobre el exceso de libertad de Fodor para construir el individualismo, al punto de que admita como taxonómicas propiedades que ‘afectan’ a los poderes causales, en lugar de limitarse exclusiva y excluyentemente a poderes causales dando lugar de este modo, según Stalnaker, a una versión ‘débil’ de individualismo. Antes de proseguir, permítaseme un breve *excursus*, para situar cómo entiende Fodor la idea de algo que ‘afecte’ a los poderes causales. El precedente puede ubicarse en el debate entre Burge y Fodor sobre la individuación del contenido de los estados mentales. En este contexto, Fodor, citando a Burge, propone lo siguiente:

Así pues, para poner todas las piezas juntas, diré que se puede afectar de toda clase de formas a las propiedades relacionales de las cosas, incluso por estipulación. Pero para que una cosa afecte los poderes causales de otra debe haber una ley o un mecanismo que medie. Es un misterio qué podría ser esto en el caso de los Gemelos (o en el de los Óscar —el ejemplo de Burge). No es sorprendente que sea un misterio, dado que es ciertamente plausible pensar que los únicos mecanismos que *pueden* mediar los efectos ambientales sobre los poderes causales de los estados mentales, son neurológicos.

[...]

En un artículo reciente, Burge (1986) dice que razonar como yo lo he hecho «es confuso. La confusión está propiciada por la falta de cuidado en el uso del término “afectar”, lo que da lugar a mezclar la causación con la individuación. Las variaciones en el ambiente que no hacen variar los choques que “afectan” causalmente al cuerpo del sujeto, pueden “afectar” a la individuación de los ... procesos intencionales que está sufriendo él o ella... No se sigue que el ambiente afecte causalmente al sujeto de ninguna forma que burle el que el ambiente tenga efectos en el cuerpo del sujeto» (Burge 1986:16) (Fodor 1994: 71-2)

La oposición de Fodor a la visión de Burge supone rechazar la posibilidad de dividir las cuestiones relativas a la causación y las relativas a la individuación. De acuerdo con Fodor, si la causación local de los estados mentales implica la superveniencia local de los poderes causales, la primera exige individuación local: ‘Uno puede tener individuación contextual si insiste en ello. Pero no puede tenerla gratis. La etiología sufre.’ (Fodor 1994:73)

Bien, veamos si puedo hacer algo de todo esto. Lo primero, me parece, es conceder la diferencia entre individuar por poderes causales y hacerlo por aquello que afecta a los poderes causales, cuestión sobre la que, supongo, Fodor está suficientemente advertido. Si esto último es cierto, ¿cómo entender esta versión ‘débil’ del individualismo? Una primera respuesta sugiere que Fodor intentaría salvar el individualismo, revisando la formulación inicial ‘fuerte’ y *liberalizándola* para adecuarla a la práctica científica real, que no taxonomiza exclusivamente por poderes causales (asumo aquí que esto se aplica sólo a aquella práctica científica interesada en la explicación causal). El problema de la versión débil no parece centrarse en la liberalización *per se*, sino en sus consecuencias. Para Stalnaker, la versión débil del individualismo es compatible con la individuación amplia *en* psicología, porque los hechos ambientales afectan causalmente a los poderes causales que tienen los objetos, incluyendo sus poderes causales intrínsecos.

Este punto está precisamente en el centro de la discusión de Fodor con Burge sobre la individuación en psicología. Ambos coinciden en reconocer que los hechos ambientales afectan causalmente a los poderes causales intrínsecos, pero la posición de Fodor es, como vimos, que la individuación de los estados psicológicos no es independiente de los efectos ambientales sobre la neurología en la que dichos estados supervienen. Luego, la compatibilidad entre la individuación por lo que afecta a los poderes causales y la individuación amplia en psicología, requeriría la explicitación de tal mecanismo o ley, para satisfacer el requerimiento fodoriano del caso.

Para Wilson, el problema de esta forma débil de individualismo, desde la perspectiva de IG —esto es, aplicado no sólo al caso psicológico—, es que, al incluir propiedades relacionales o

históricas no reductibles, no satisfaría el requisito de superveniencia de las propiedades taxonómicas en la microestructura local. Al respecto, diré que Fodor no considera que las propiedades que afectan a los poderes causales sean ellas mismas poderes causales; luego, estas propiedades (relacionales o históricas) que afectan a los poderes causales no estarían sometidas a la exigencia de supervenir en la microestructura local. El problema planteado por Stalnaker, a diferencia del punto de Wilson, no es si las propiedades que afectan a los poderes causales supervienen o no sobre la microestructura local, que no lo hacen, sino que tales propiedades no constituyen poderes causales y, por eso, no satisfacen la exigencia de la versión fuerte del individualismo: individuación *por* —y sólo por— poderes causales.

No obstante aceptar esta última observación de Stalnaker, considero que la alternativa fodoriana busca mantenerse a la menor distancia posible de la propuesta fuerte; en tal sentido, la exigencia de que exista un mecanismo o una ley que medie la afección sobre los poderes causales se supone capaz de hacer el trabajo de conectar la propiedad relacional efectivamente taxonómica con un poder causal, aun cuando no se pueda especificar cuál. Así, para el caso de ser un planeta, se supone que existe un mecanismo o una ley que conecta dicha propiedad con alguna otra (su trayectoria o su velocidad) que sería un poder causal en sentido estricto. El problema se traslada entonces a la posibilidad de estar en condiciones de especificar tal mecanismo o ley en todos los casos relevantes. El propio Stalnaker (1989), citando a Fodor en un contexto diferente, indica que ‘Sin duda es correcto tener una estrategia de investigación que dice “espere un poco”, pero quién quiere esperar para siempre’ (Fodor 1981:248).

II. 3. El caso del argumento modal.

La discusión sobre poderes causales, previo aviso de posible saturación, se continúa en la consideración de las ideas que Fodor propone en ‘Un argumento modal a favor del contenido estrecho’ (1991). Se supone que aquí se hacen más patentes las tensiones de la apelación a los poderes causales y el apriorismo de Fodor en su compromiso con IG. El interés de Fodor en este artículo es proponer una forma de distinguir cuándo una diferencia en las causas *es* una diferencia en los poderes causales; esta distinción permitiría separar aquellas propiedades

relacionales (o amplias) taxonómicamente relevantes de las que no lo son, puesto que Fodor reconoce la importante presencia de propiedades relacionales en la taxonomía científica:

La taxonomía por propiedades relacionales es ubicua en las ciencias y no está en disputa que propiedades como *ser un meteorito* o *ser un planeta* —propiedades que pueden, nótelos, distinguir trozos de roca molecularmente idénticos— constituyen poderes causales. (p. 12).

Y la intuición acerca de rasgos de la historia causal es que algunos de estos *son* poderes causales (e.g., haber sido atropellado en el tránsito; haber sido vacunado contra la viruela) y algunos no lo son... (p. 18; énfasis agregado).

La propuesta de Fodor para establecer la diferencia de poderes causales entre propiedades causa se formula como una condición necesaria, la *condición C*:

Para que la diferencia entre tener PC1 y tener PC2 sea una diferencia de poderes causales, debe serlo al menos el que los efectos de tener PC1 difieran de los efectos de tener PC2. Pero, afirmo, se requiere, además, que esta diferencia entre los efectos no esté *conceptualmente* relacionada con la diferencia entre las causas. (p. 24).

Esta idea de Fodor de la vinculación conceptual tiene una base en el requerimiento humeano de que para que una propiedad sea un poder causal distintivo debe estar contingentemente (o no-conceptualmente) relacionada con su propiedad efecto distintiva.

Como era de esperarse, Wilson rechazará el argumento modal fodoriano. Primero, sobre la base de que, una vez más, Fodor hace un uso ambiguo de ‘poder causal’ al incluir propiedades relacionales, con los problemas ya indicados respecto a la validez de APC. En su opinión, Fodor presupone el sentido ‘extenso’ de poder causal en gran parte de ‘Un argumento modal’, lo que parecería conceder que duplicados moleculares puedan diferir en sus propiedades, propiedades que no podrían no supervenir en su microestructura local, bajo pena de incumplir el punto metafísico de Fodor. Nuevamente, la conclusión de Wilson es implacable:

Si los poderes causales reales deben supervenir en propiedades físicas intrínsecas, es falso que *sólo* los poderes causales reales taxonomizan clases científicas. Y si es verdadero que sólo los poderes causales reales se usan para taxonomizar clases científicas, entonces los poderes causales reales no supervienen en propiedades físicas intrínsecas y, así, no hay razón para que las propiedades de una *psicología* científica sean localmente supervenientes. (Wilson 1995:54)

En segundo lugar, Wilson cuestiona la invocación a la distinción analítico/sintética [A/S] que entrañaría la condición C, en el sentido de que A/S sería la herramienta que haría el trabajo de distinguir poderes causales reales de poderes causales *a lo mero Cambridge* (*sensu* Shoemaker 1980) mostrando cuándo la relación entre una propiedad causa y una propiedad efecto es conceptual. Esta crítica se dirige tanto a la utilización de A/S, de la que Fodor no sería precisamente un defensor, cuanto a su posibilidad de hacer el trabajo que éste quiere que haga, pues sería incapaz de demostrar que una propiedad amplia *no* es un poder causal real. La argumentación de Wilson apunta a establecer que no hay posibilidad de una distinción a priori entre propiedades relacionales taxonómicamente relevantes y aquellas que no lo son. Del mismo, que no hay posibilidad de separar a priori las propiedades en intrínsecas y relacionales, y afirmar que sólo las primeras pueden ser taxonómicas en ciencia.

En “Un argumento modal en favor del contenido estrecho” (Fodor 1990) Fodor trata de responder a la siguiente pregunta: ¿cuándo tener PC1 en vez de PC2 es —cuenta como— un poder causal? Su respuesta se formula en la Condición C. Esta condición, nos dice Fodor, ‘está motivada por nuestras intuiciones concernientes a ciertos ejemplos y por consideraciones respecto a que los poderes causales son poderes para entrar en relaciones causales contingentes. Entonces, por ejemplo, en virtud de tener hermanos soy capaz de tener hijos que sean sobrinos. Las personas que no tienen hermanos no son capaces de tener hijos que sean sobrinos. Pero intuitivamente la diferencia entre tener hermanos y no tener hermanos no es un poder causal en virtud de ser responsable por la diferencia entre las propiedades de los hijos’ (1990:347).

Fodor está interesado en las propiedades de las causas de las que puede decirse que satisfacen la Condición C, en tanto las diferencias entre ellas son responsables de diferencias en propiedades de sus efectos, y *ninguna de las cuales es realmente idéntica a un poder causal distintivo*, pero algunas de las cuales, intuitivamente, son poderes causales distintivos. Ser un planeta y ser un meteorito son poderes causales distintivos, pero ser un hermano y ser un cráter de meteorito no lo son. La condición C es aplicable a todas estas propiedades, y a las intencionales también, porque si son poderes causales, también lo son contingentemente. Sin

embargo, las propiedades intencionales fallan al incumplir la condición, junto con ser un hermano y ser un cráter de meteorito, y por la misma razón. Todas están conectadas conceptualmente a las propiedades de los efectos de los que son responsables. Tener un hermano está conceptualmente relacionado con ser un sobrino, ya que ser un sobrino es ser hijo de un padre que tiene hermanos.(Véase Macdonald y Macdonald 1995: cap. 9)

Del mismo modo, argumenta Fodor, ser un pensamiento-*agua* está conceptualmente relacionado con ser una conducta-*agua*. Es decir, aunque las propiedades intencionales son poderes causales contingentemente, no satisfacen la condición C porque las diferencias en las propiedades intencionales de las causas que son responsables por diferencias en las propiedades intencionales de sus efectos están relacionadas conceptualmente. Es decir, las propiedades intencionales no son poderes causales distintivos. Es importante notar, respecto de la afirmación que hace Wilson y que consignamos más arriba, que Fodor *no* está diciendo que las propiedades intencionales no son poderes causales, sino que no son poderes causales *distintivos*, ya que el tema del individualismo para él gira sobre si las *diferencias* en la conducta de los gemelos, descritas intencionalmente, son relevantes para la taxonomía de las clases intencionales por poderes causales.

Entonces, las diferencias en las propiedades intencionales amplias de sus pensamientos y los de su gemelo no satisfacen la Condición C, en consecuencia, no cuentan como una diferencia de poder causal en virtud de su responsabilidad por las diferencias intencionales de vuestras conductas. Y no cuentan como tal porque esas diferencias obedecen, primero, a la diferencia existente entre los contenidos de vuestros pensamientos y, segundo, a que usted está conectado con agua de un modo en que su gemelo no lo está. Sin embargo, las diferencias entre los efectos se relacionan conceptualmente con diferencias en las causas. Por estar conectado con agua en lugar de gagua es conceptualmente necesario que usted piense en agua antes que en gagua, y pensar en agua en lugar de en gagua conduce conceptualmente a conductas agua, a diferencia de lo que pasa al pensar en gagua. Según esto, y en virtud de la condición C, sus pensamientos-agua y los pensamientos-gagua de su gemelo son *el mismo pensamiento tipo* y, del mismo modo, su conducta-agua y la conducta-gagua de su gemelo son *la misma conducta*

tipo; en consecuencia, vuestros respectivos estados mentales_{agua o gagua o...etc.} SON taxonómicamente idénticos respecto de ser causalmente responsables de vuestras respectivas conductas_{agua o gagua o...etc.}, y la generalización que explique su conducta subsumirá también la de él.

Como último comentario sobre las críticas de Wilson, quisiera precisar algo con referencia a la acusación de que Fodor apela a la distinción A/S al establecer la condición C. Al respecto, remarco que la condición C es, según Fodor, una condición necesaria, en el marco de una argumentación a priori para cuándo una diferencia en las propiedades de las causas constituye una diferencia en sus poderes causales. Me parece que esto evita la objeción de que la condición C implica un recurso a la analiticidad. El carácter necesario de C, de acuerdo con Fodor (1991a), se sigue de que ‘después de todo, el compromiso con la explicación causal presumiblemente tiene *algunas* consecuencias metodológicas *qua* compromiso con la explicación causal; y debe ser posible deslindarlas reflexionando sobre qué clase de cosas son las explicaciones causales’ (p. 208).

III. Conclusiones.

La discusión de las secciones 2 y 3 me permite realizar ahora una evaluación de la crítica de Wilson a APC. Siguiendo su línea de la argumentación, me referiré primero a la cuestión de la verdad o falsedad de IG, para luego considerar la putativa ambigüedad de ‘poder causal’ en el argumento individualista.

Wilson argumenta la falsedad de IG con base en la existencia de clases taxonómicas en algunas ciencias, clases que figurarían como *explanantia* en explicaciones causales formuladas por esas ciencias, que no satisfacen el criterio individualista global de tener que ser por poderes causales. Adicionalmente, se supone bloqueada en esas ciencias la posibilidad de revisión de tales clases en términos de una propiedad intrínseca y otra relacional o histórica. Por último, se apela a la ciencia tal como se desarrolla, pues se supone que esto implica un criterio más válido que el apriorismo del constreñimiento individualista.

Como lo indiqué antes, los casos cruciales que Wilson ofrece para apoyar su intento de falsación de IG no se despliegan en términos de ejemplos que puedan ser objeto de un análisis pormenorizado. Lo mismo sucede con la presunta imposibilidad de revisión de las clases involucradas en esos casos cruciales, que no queda en caso alguno establecida como pretende Wilson. Es claro que el hecho de que, incluso si durante mucho tiempo no fuese posible realizar la revisión señalada, eso no constituiría una prueba de la imposibilidad de hacerlo. Finalmente, la apelación a la práctica científica no resulta por sí sola suficiente, pues debe(n) explicitarse el o los criterios que posibilitan su interpretación como prácticas válidas en uno u otro sentido, discusión que Wilson no aborda. En consecuencia, considero que los argumentos desplegados por Wilson no son lo bastante apremiantes como para aceptar la conclusión que él pretende sugerir.

Con respecto a la afirmación de que existe un uso ambiguo de ‘poder causal’ en el argumento individualista, me parece que la misma no es correcta. Por lo dicho precedentemente, creo que es razonable afirmar que: (i) ‘poder causal’ y ‘lo que afecta a los poderes causales’ no son la misma cosa, siendo este último término el que admite propiedades relacionales (o históricas) como taxonómicamente válidas, luego, no hay *dos* sentidos para el mismo término; (ii) la exigencia de supervenir en la microestructura local se aplica sólo en el primer caso —el de poder causal— y no en el segundo —lo que afecta a los poderes causales; (iii) si (ii) es correcta, (4) en APC no está en contradicción con (1), esto es, con la premisa que establece IG.

En consecuencia, todo marcha bien con APC y los individualistas pueden dormir tranquilos, aunque de cuando en cuando tengan pesadillas con el apriorismo.

Capítulo 3. El argumento computacional.

Varios importantes filósofos (Devitt 1990, 1991; Egan 1992; Fodor 1981, 1987; Segal 1989, 1991) han defendido la idea de que el individualismo en psicología encuentra apoyo en la Teoría Computacional de la Mente, una visión considerada por filósofos y científicos cognitivos como una suposición fundamental de la investigación contemporánea en ciencia cognitiva (Cummins 1989; Pylyshyn 1984). Básicamente, el apoyo se condensa en la afirmación de que la psicología teoriza sobre procesos mentales, en tanto procesos computacionales, y los procesos computacionales deben, por ello, ser individualistas. Wilson (1995) desarrolla una crítica a esta afirmación, que será objeto de análisis y discusión en este capítulo, con la finalidad de evaluar si constituye o no un argumento de peso contra el individualismo en psicología. Nuestro punto de partida estará constituido por una sección dedicada a presentar algunos aspectos centrales de la Teoría Computacional de la Mente.

I. Teoría computacional de la mente.¹⁴

I.1. Introducción.

Los pasados cuarenta años han sido testigos de la emergencia y ascendencia de un nuevo paradigma para comprender la mente: el paradigma de la máquina de computación, cuya influencia para el estudio de la mente ha sido profunda y de largo alcance, al punto de haberse constituido de hecho en una visión dominante en las disciplinas actualmente comprometidas en el estudio de la cognición.

Entre los filósofos de la mente y los investigadores empíricos existe acuerdo respecto a que el paradigma de máquina de computación puede tener implicaciones significativas. No obstante, lo mismo no puede decirse acerca de cuáles puedan ser esas implicaciones. Siendo relativamente claro que la modelación computacional puede ser una poderosa herramienta para los psicólogos, tanto como lo es para el físico y el meteorólogo, no todos los investigadores coinciden en que los procesos cognitivos que pueden modelarse en un computador sean, en sí mismos, computaciones; no más de lo que pueden ser computaciones

¹⁴ El contenido de esta sección se basa en Horst 1996:cap. 1.

las tormentas que modela el meteorólogo. Un punto de importante desacuerdo y discusión, particularmente entre los filósofos, se refiere al tipo de caracterización de la mente que posibilita el paradigma de máquina de computación. Pueden distinguirse al menos tres formas alternativas de evaluar la importancia del paradigma computacional:

a) La visión más modesta considera que la metáfora del computador funcionará como un catalizador que posibilite la generación de teorías en psicología, al modo en que otras metáforas han contribuido al desarrollo de otras ciencias. En este caso, la relevancia de la computación *per se* para las teorías resultantes será mínima o derechamente nula.

b) Una alternativa ligeramente más fuerte, supone que la maquinaria conceptual de la ciencia computacional proporcionará las clases de herramientas correctas para permitir que la psicología (o al menos parte de la psicología) devenga una ciencia rigurosa, al modo en que las herramientas conceptuales de la geometría cartesiana y el cálculo proporcionaron una base para la emergencia de la mecánica newtoniana, y la geometría diferencial hizo posible la física relativista que la reemplazó. Desde este punto de vista, el paradigma computacional dotaría a la psicología de las bases para su maduración, proveyéndola de medios adecuados para la matematización de sus explicaciones y posibilitando conexiones entre la explicación intencional y explicaciones desarrolladas al nivel de procesos de algún orden inferior (por ejemplo, neurológico), mediante los cuales se realizan los estados y procesos intencionales. El compromiso de esta concepción con la tesis de la mente como computador se limita a la afirmación débil de que las interrelaciones entre estados mentales tienen propiedades formales para las que el vocabulario asociado con la computación puede proporcionar una caracterización apta; o un isomorfismo entre estados/procesos mentales y un programa de computación; o los medios adecuados para la descripción de los estados y procesos mentales en términos computacionales. Desde este punto de vista, la mente sería “computacional” en el mismo sentido en que puede decirse que un universo relativista es “diferencial”.

c) La concepción más fuerte sobre la relevancia de la máquina de computación para la psicología plantea que nociones tales como “representación” y “computación” no sólo proveen al psicólogo de herramientas formales necesarias para hacer ciencia de un modo riguroso, sino que también proveen al filósofo de herramientas fundamentales para desarrollar un análisis de

la *naturaleza esencial de la cognición* y la solución de problemas filosóficos importantes y de larga data.

A los efectos de la discusión en este capítulo, en lo que sigue me referiré específicamente a una aplicación particular del paradigma de máquina computacional al estudio de la mente, a saber: la Teoría Computacional de la Mente (en adelante TCM) defendida en años recientes por Jerry Fodor (1975, 1980, 1981, 1987, 1990) y Zenon Pylyshyn (1980, 1984).

En las últimas dos décadas, la TCM ha emergido como la visión *dominante* sobre la significación de la computación en filosofía. Sus proponentes articulan y defienden una posición muy fuerte, a saber: que la cognición es *literalmente* computación y que la mente es *literalmente* un computador digital. La TCM está constituida por dos tesis. La primera es una tesis sobre la naturaleza de los estados intencionales, tales como las creencias y los deseos individuales, que se conciben como estados relacionales que involucran a un organismo (u otro ente conocedor) y representaciones mentales. Estas representaciones mentales se entienden según el modelo de representaciones almacenadas en un computador, esto es, como instanciaciones de símbolos que tienen propiedades sintácticas y semánticas. Estos símbolos incluyen primitivos semánticos y símbolos complejos cuyas propiedades semánticas son una función de su estructura sintáctica y los primitivos semánticos que contienen. La segunda tesis se refiere a la naturaleza de los procesos cognitivos —procesos como razonar hacia una conclusión, o formar y testear hipótesis, procesos que involucran cadenas de creencias, deseos y otros estados intencionales. Según la TCM, los procesos cognitivos son computaciones sobre representaciones mentales, es decir, son secuencias causales de instanciaciones de representaciones mentales en las que las regularidades causales relevantes están determinadas por las propiedades sintácticas de los símbolos y pueden describirse en términos de reglas formales (i.e., sintácticas).

Los defensores de la TCM también han hecho un caso muy persuasivo en el sentido de que ver la mente como un computador permite la solución de problemas filosóficos significativos. Particularmente, han argumentado que: (1) proporciona una concepción de la intencionalidad

de los estados mentales —los estados mentales tienen las propiedades semánticas que tienen en virtud de heredarlas de las que poseen las representaciones que son sus objetos—; y, (2) muestra que la psicología puede emplear explicaciones en el idiolecto intencional sin involucrarse a sí misma en dificultades metodológicas u ontológicas —al vincular, vía formalización, las propiedades semánticas (causalmente inertes) con las respectivas propiedades sintácticas (causalmente eficientes, en tanto poderes causales) de las representaciones que involucran los estados mentales que participan en los procesos cognitivos, las explicaciones psicológicas pueden ser, simultáneamente, intencionales y causales.¹⁵ Las afirmaciones hechas por parte de la TCM al respecto están comprendidas en la tercera y más “fuerte” categoría de actitudes hacia la promesa del paradigma computacional —alternativa *c*), más arriba.

I.2. Concepción computacional de los estados mentales en la TCM.

Se dijo que la primera tesis que constituye la TCM es una concepción representacional de la naturaleza de los estados intencionales. En la introducción de *Representations*, Fodor (1981) proporciona un esbozo claro de los supuestos básicos de esta concepción en las siguientes cinco afirmaciones:

- a) Los estados de actitud proposicional son relacionales.
- b) Entre los que se relacionan están representaciones mentales (con frecuencia llamadas “Ideas” en la literatura más antigua).
- c) Las representaciones mentales son símbolos: tienen propiedades formales y semánticas.
- d) Las representaciones mentales tienen sus roles causales en virtud de sus propiedades formales.
- e) Las actitudes proposicionales heredan sus propiedades semánticas de aquellas de las representaciones mentales que funcionan como sus objetos. (Fodor 1981:26)

Las afirmaciones (a) (b) y (c) muestran las concepciones de Fodor sobre la naturaleza de los estados intencionales, mientras (d) y (e) ofrecen los medios para conectar esta concepción representacional de los estados intencionales con una concepción computacional de los procesos cognitivos y una concepción de la intencionalidad de lo mental, respectivamente.

¹⁵ Estas afirmaciones serán examinadas con detalle a continuación, en I.2 y I.5, respectivamente.

Una concepción más formal de la naturaleza de los estados intencionales (actitudes proposicionales) de Fodor puede hallarse en *Psicosemántica*, donde plantea que:

Afirmación 1 (la naturaleza de las actitudes proposicionales):

Para cualquier organismo O , y cualquier actitud A hacia la proposición P , hay una relación ('computacional'-'funcional') R y una representación mental MP tal que

MP significa que P , y

O tiene A ssi O mantiene R a MP . (Fodor 1987:17)

Según la concepción de Fodor, la creencia de Pérez que dos es un número primo consiste en que Pérez está en una particular clase de relación funcional R con una representación mental MP . Esta representación mental MP es la instanciación de un símbolo, presumiblemente instanciado de alguna forma en el sistema nervioso de Pérez. MP tiene propiedades semánticas: en particular, MP significa que dos es un número primo. Y Pérez cree que dos es un número primo cuando y sólo cuando está en la relación R con MP . En el mismo texto Fodor brinda una formulación más precisa de esta concepción en lo relativo a las referencias a tipos e instanciaciones de actitudes:

Crear que tal y tal es tener un símbolo mental que significa que tal y tal instanciado en su cabeza de cierto modo; es tener tal instanciación 'en su caja de creencias', como algunas veces diré. Respectivamente, esperar que tal y tal es tener una instanciación del mismo símbolo mental instanciado en su cabeza, pero de una forma más bien diferente; es tenerlo instanciado "en su caja de esperanzas." . . . Y así para cada actitud que usted pueda mantener hacia una proposición; y así para cada proposición hacia la cual usted pueda mantener una actitud. (Fodor 1987:17)

Con base en esta glosa, Horst (1996) considera más razonable leer la formulación inicial de Fodor como sigue:

Naturaleza de las actitudes proposicionales (modificada)

Para cualquier organismo O , y cualquier instanciación-de-actitud a de tipo A hacia la proposición P , hay una relación ('computacional'-'funcional') R y una instanciación de representación mental t de tipo MP tal que

t significa que P en virtud de ser una instanciación- MP , y

O tiene una actitud de tipo A ssi O mantiene R a una instanciación de tipo MP . (Horst 1996:24)¹⁶

¹⁶ La glosa de Fodor parecería indicar también que creencias con contenidos diferentes son tipificadas por la misma relación funcional R , pero difieren respecto al tipo de representación con la que está conectado el

El punto principal de la concepción de Fodor es claro: lo que cuenta para el hecho de que los estados intencionales tengan las propiedades semánticas y la intencionalidad que tienen es la relación entre el organismo y sus representaciones mentales. En el pasaje de *Representations* ya citado, él escribe que los estados intencionales “heredan sus propiedades semánticas de aquellas de las representaciones mentales que funcionan como sus objetos” (Fodor 1981:26). Y en ese mismo ensayo escribe también que “los objetos de las actitudes proposicionales son símbolos (específicamente, representaciones mentales)” y que “este hecho cuenta para su intencionalidad y semánticidad” (Ibíd., 25).¹⁷

Un rasgo adicional importante de esta concepción descansa en el hecho de que los símbolos involucrados en la representación mental tienen propiedades semánticas y sintácticas, y pueden verse como instanciaciones en un “lenguaje del pensamiento”, también conocido como “mentalés”. Ver el sistema de representaciones mentales como un lenguaje con propiedades semánticas y sintácticas permite la posibilidad de *composicionalidad de los procesos mentales*. Los símbolos del mentalés no son todos primitivos léxicos y que, en lugar de eso, hay un conjunto finito de primitivos léxicos que pueden combinarse en varias formas —de acuerdo con las reglas sintácticas del *mentalés*— para formar una variedad potencialmente infinita de representaciones complejas, tal como en el caso de los lenguajes naturales es posible generar una infinita variedad de preferencias significativas a partir de un conjunto finito de morfemas y reglas composicionales. El *mentalés* se concibe entonces poseyendo los mismos aspectos generativos y creativos poseídos por los lenguajes naturales. Así, mientras

organismo. De este modo, si el que Pérez crea que dos es un número primo consiste en que Pérez está en la relación R con una representación MP que significa que dos es un número primo, entonces el que Pérez crea que está lloviendo debe consistir en que Pérez está en la misma relación R con una representación de diferente tipo, MP^* , tal que las instanciaciones de MP^* significan que está lloviendo. Respectivamente, cuando Pérez tiene diferentes estados cognitivos que comparten un contenido común, los diferentes estados son caracterizados por diferentes relaciones con instanciaciones del mismo símbolo tipo. Por ejemplo, si Pérez cree y espera que está(é) lloviendo, creer que está lloviendo consiste en estar en la relación R con una instanciación del tipo MP^* , y esperar que esté lloviendo consiste en estar en alguna otra relación R^* con una instanciación del tipo MP^* .

¹⁷ Fodor (1981:318, n. 2) reserva el término ‘intencionalidad,’ con c , para “contextos conectados con la intención”.

las propiedades semánticas de los estados mentales son “heredadas” de las representaciones que contienen, esas representaciones pueden, ellas mismas, o bien ser semánticamente primitivas, o bien ser compuestas a partir de primitivos semánticos por la aplicación de reglas sintácticas. Así, según Fodor (1987/1994:Apéndice) no sólo los objetos intencionales de los estados mentales pueden ser complejos, en el sentido de su carácter composicional —algo que, para Fodor, está supuesto ya en el realismo intencional; la hipótesis del lenguaje del pensamiento implica algo más que el mero realismo intencional, implica la afirmación de que *‘los estados mentales —y no sólo sus objetos proposicionales— tienen, de una forma característica, una estructura constitutiva’* (p. 192), esto es, que los estados mentales poseen una sintáctica apropiada al contenido que tienen. Así, la estructura sintáctica de los estados mentales refleja las relaciones semánticas *entre* sus objetos intencionales. El lenguaje del pensamiento es, entonces, la idea de que los estados mentales poseen una semántica combinatoria.

I.3. Procesos cognitivos.

Si una concepción representacional de la mente proporciona una manera de interpretar la naturaleza de los pensamientos individuales, no proporciona por sí misma una concepción comparable de la naturaleza de procesos mentales tales como razonar hacia una conclusión o formar y testear hipótesis, y por eso no proporciona las bases para una psicología de la cognición. Para una psicología de la cognición se necesita algo más: una teoría de los procesos mentales que use las propiedades de las representaciones mentales como base de una concepción causal de cómo un estado mental sigue a otro en un hilo de razonamiento. Suponga, por ejemplo, que uno desea explicar por qué Pérez ha cerrado la ventana. Una explicación bien podría darse según las siguientes líneas:

- (1) Pérez sintió frío.
- (2) Pérez notó que la ventana estaba abierta.
- (3) Pérez hipotetizó que había una corriente de aire helado soplando a través de la ventana.
- (4) Pérez hipotetizó que esta corriente de aire helado era la causa de su frío.
- (5) Pérez quería dejar de sentir frío.
- (6) Pérez hipotetizó que deteniendo el aire helado se detendría el frío.

Entonces, (7) Pérez se hizo el deseo de detener el aire helado.

- (8) Pérez hipotetizó que cerrando la ventana se detendría el aire helado.
Entonces, (9) Pérez se hizo el deseo de cerrar la ventana.
(10) Pérez cerró la ventana.

No estamos ante una serie aleatoria de pensamientos, sino ante una secuencia donde los últimos pensamientos se ven, plausiblemente, (a) como racionales, a la luz de los precedentes; y, (b) como consecuencias de esos estados previos —Pérez se hizo el deseo de cerrar la ventana *porque* pensó que al hacerlo se detendría el aire helado. Además, una teoría causal de la inferencia necesitaría establecer un vínculo estrecho entre las propiedades semánticas de los estados individuales y su rol en la producción de los estados subsecuentes. Son cambios en el contenido de las creencias y deseos de Pérez los que esperaríamos produzcan diferentes secuencias de pensamiento y diferentes conductas. Si Pérez hubiese notado el ventilador encendido, en lugar de notar una ventana abierta, esperaríamos que él manejara hipótesis diferentes, se hiciera deseos diferentes, y actuara de diferente manera, todo como consecuencia de cambiar el contenido de su creencia “la ventana está abierta” por “el ventilador está encendido”.

La TCM afirma que las propiedades semánticas e intencionales de los estados intencionales se identifican con las de las representaciones que involucran, pero cuando se pregunta cómo los estados intencionales pueden desempeñar un rol *causal* en la etiología de los procesos que involucran la generación de nuevos estados intencionales, la noción de representación, entendida sólo como el objeto intencional de un estado mental, resulta insuficiente para mostrar cómo el que Pérez esté en la relación R con una representación MP en el tiempo t puede desempeñar un rol causal en el que Pérez llegue a estar en la relación Q con una representación MP^* en $t+\delta$. En otras palabras, ¿en virtud de qué las representaciones pueden cumplir un rol causal en la producción de procesos cognitivos?

En orden a que una secuencia de representaciones estructure una línea de pensamiento racional, convincente, la cuestión de *cuáles* representaciones ocurrirán en la secuencia debe estar determinada por los significados de las representaciones precedentes. En orden a que la secuencia de representaciones tenga sentido, las representaciones posteriores necesitan

mantener relaciones semánticas apropiadas con las anteriores. Pero en orden a que la secuencia de representaciones sea una secuencia *causal*, la cuestión de qué representaciones ocurrirán después en la secuencia debe estar determinada por los poderes causales de las representaciones anteriores.

Debe resultar relativamente simple notar que este último aspecto de la TCM, relativo a la preservación de las relaciones semánticas de las cadenas de estados mentales que configuran un determinado proceso mental —lo que hacen vía su sintaxis— y la capacidad de un estado previo componente de tal cadena para causar un estado mental subsiguiente (o conducta), propiedades relevantes para la formulación de explicaciones (causales) psicológicas intencionales, nos remite a la discusión del capítulo precedente y evidencia la fuerza del vínculo entre la tesis de la individuación por poderes causales y la tesis computacional que supone la TCM. Las explicaciones intencionales seleccionan representaciones por su contenido —esto es, por sus propiedades semánticas. Pero, para ser explicaciones causales, deben seleccionar representaciones de un modo que las individúe según sus poderes causales. Esto sólo puede hacerse si los valores semánticos de las representaciones pueden vincularse a, o coordinarse con, los roles causales que éstas desempeñan en la producción de otras representaciones y de conducta. De no ser posible esto último, parecería que se bloquea la posibilidad de una psicología nomológico-causal, en tanto es notoriamente problemático ver las relaciones semánticas como relaciones causales o igualar razones y causas.¹⁸ Entonces, el

¹⁸ Hay una significativa historia de controversia sobre si la explicación intencional es una forma de explicación causal, y los temas que se mantienen entre, digamos, el positivismo del Círculo de Viena, por un lado, y Ryle, Wittgenstein, y la tradición de la *Verstehen*, por el otro. En particular, ha habido controversia sobre la naturaleza de la explicación intencional, notoriamente sobre si tales explicaciones proporcionan sólo *razones para* la acción, sólo *causas* de la acción, o ambas. Gran parte de la tradición continental en filosofía y las ciencias humanas ha sostenido que la explicación en ciencias humanas brinda razones para la acción y proporciona comprensión interpretativa (*Verstehen*), mientras la explicación en las ciencias naturales ofrece causas de eventos y proporciona explicaciones. En el mundo de habla inglesa, Wittgenstein, Ryle y otros filósofos del lenguaje ordinario han avanzado visiones similares a los efectos de que las explicaciones intencionales proporcionan razones para la acción y no designan causas. Por otro lado, buena parte de la tradición filosófica de habla inglesa ha seguido a los positivistas lógicos al asumir que el modelo de explicación nomológico-causal empleado en las

problema para convertir una teoría representacional de los estados mentales en una teoría psicológica de los procesos mentales requiere hallar una forma de vincular las propiedades semánticas de las representaciones mentales con los poderes causales de esas representaciones.

Precisamente en este punto el paradigma computacional empieza a ser de interés, pues los computadores se entienden como aparatos que almacenan y manipulan instancias de símbolos; y las manipulaciones que ejecutan dependen de qué representaciones ya estén presentes; también son de naturaleza mecánica e indiscutiblemente causal. La máquina de computación proporciona un paradigma general de los procesos de manipulación simbólica donde los símbolos ya presentes desempeñan un rol causal al determinar qué nuevos símbolos van a generarse. La TCM busca proporcionar una extensión de este paradigma a las representaciones mentales, y así proveer una concepción de los procesos cognitivos que pueda proporcionar una forma de discutir su etiología, respetando, al mismo tiempo, las relaciones semánticas entre las representaciones involucradas.

I.4. Formalización y computación.

Los defensores de la TCM creen que la computación de máquina proporciona un paradigma para comprender cómo un sistema de manipulación simbólica puede causar derivaciones de representaciones simbólicas de un modo que respete sus propiedades semánticas. Más específicamente, se cree que la computación de máquina proporciona respuesta a dos preguntas: (1) ¿Cómo pueden vincularse las propiedades semánticas de los símbolos con los poderes causales que permiten que la instancia de un símbolo s_1 en el tiempo t sea una causa parcial de la instancia de un segundo símbolo s_2 en el tiempo $t+\delta$?; y, (2) ¿cómo las leyes que gobiernan las regularidades causales pueden asegurar también que las operaciones que generan nuevas instancias de símbolos respetarán las relaciones semánticas entre los símbolos, tal que el proceso total resultante sea, en un sentido amplio, racional? Las respuestas a estas preguntas tienen un doble origen, a saber: la formalización en matemáticas y la operación de los computadores digitales.

ciencias naturales es normativo para la aplicación de la palabra ‘ciencia’, con la implicación de que la explicación psicológica debe bien ser causal y nomológica o bien no ser científica.

En la segunda mitad del siglo diecinueve, un problema importante en matemáticas fue la formalización de sistemas matemáticos, esto es, la eliminación en las reglas de deducción del sistema, de cualquier cosa dependiente del significado de los términos. La formalización deviene un tema importante en matemáticas después de que, independientemente, Gauss, Bolyai, Lobachevski y Riemann fundaran geometrías consistentes que negaban el postulado del paralelo de Euclídes. Esto condujo a un deseo de liberar los procedimientos empleados en las deducciones matemáticas de toda dependencia de las intuiciones semánticas del matemático (por ejemplo, sus intuiciones espaciales euclidianas). El proceso de formalización halló un portavoz definitivo en David Hilbert, cuyo libro sobre los fundamentos de la geometría, publicado en 1899, empleó un acercamiento a la axiomatización que involucró una abstracción completa de los significados de los símbolos. La formalización de la lógica, entretanto, había sido emprendida por Boole y luego por Frege, Whitehead y Russell; y la formalización de la aritmética por Peano.

Aun cuando hubo varias aproximaciones diferentes a la formalización en las matemáticas decimonónicas, la aproximación de “juego de símbolos” de Hilbert es de especial interés para los propósitos presentes. En esta aproximación, los símbolos usados en las pruebas son tratados como instanciaciones o piezas en un juego, cuyas “reglas” gobiernan la formación de expresiones y la validez de las deducciones en ese sistema. Las reglas empleadas en el juego de símbolos, sin embargo, se aplican a fórmulas sólo en tanto caen bajo particulares tipos sintácticos.

Este ideal de formalización en un dominio matemático requiere la habilidad para caracterizar completamente en términos notacionales (simbólicos y sintácticos): (a) las reglas para la correcta formación de símbolos, (b) las reglas para la correcta formación de las fórmulas, (c) los axiomas, y (d) las reglas que autorizan derivaciones. Así, por ejemplo, un lógico podría conocer una regla que autoriza una derivación al efecto de que, como sea que se hayan derivado fórmulas de la forma p y $p \supset q$, pueda válidamente derivar una fórmula de la forma q . Para aplicar esta regla, él no necesita conocer las interpretaciones de ninguna de las instancias de sustitución de p y q , o saber qué relación expresa \supset , sólo necesita ser capaz de reconocer

las estructuras de símbolos que tienen las formas sintácticas p y $p \supset q$. Como consecuencia, pueden efectuarse inferencias racionales, con sentido y que preserven la verdad, sin atender a —o incluso sin conocer— los significados de los términos.

El trabajo de formalización de símbolos en las matemáticas de los siglos diecinueve y veinte ha mostrado que, para dominios simbólicos interpretados sustanciales (aunque limitados; como es el caso de la geometría y el álgebra), pueden encontrarse formas de hacer derivaciones válidas de un modo que no depende de la intuición del matemático sobre los significados de los símbolos, en tanto (a) las distinciones semánticas entre los símbolos sean reflejadas por distinciones sintácticas; y, (b) pueda desarrollarse una serie de reglas, dependientes por completo de los rasgos sintácticos de las estructuras de los símbolos, que autorizarán aquellas deducciones que uno desearía estar autorizado a hacer sobre la base de los significados de los términos.

Un segundo tema que aparece en las matemáticas de cambio de siglo se refiere a la pregunta por cuáles funciones son “computables,” en el sentido de estar sujetas a evaluación mediante la aplicación de algún procedimiento mecánico o algoritmo. Los procedimientos para evaluar integrales son buenos ejemplos de algoritmos computacionales. Aprender integración es cuestión de aprender a identificar expresiones como miembros de clases particulares caracterizadas sintácticamente y aprender cómo producir las correspondientes expresiones que indican los valores de sus integrales. Uno aprende, por ejemplo, que integrales con la forma $\int x^n dx$ tienen soluciones de la forma $\frac{1}{n+1} x^{n+1}$.

Tales métodos computacionales son formales, en el sentido de que la habilidad de una persona para aplicar el método no requiere ninguna comprensión de los significados de los términos.¹⁹

¹⁹ Aunque puede decirse que los procedimientos computacionales (tales como las técnicas de integración y la suma de columnas) y los sistemas matemáticos formalizados (como la geometría de Hilbert) involucran “reglas formales”, dichas “reglas” son de diferentes clases. Los sistemas matemáticos formalizados, como la lógica formal y la geometría formal, contienen reglas que *autorizan derivaciones*, pero no *reglas de procedimiento*. Es decir, involucran reglas que gobiernan la *validez* de las derivaciones, pero no ofrecen ninguna clase de guía sobre

Para evaluar $\int 3x^2 dx$, por ejemplo, uno no necesita saber lo que la expresión indica —el área bajo la curva— sino sólo que es de un particular tipo sintáctico al que se aplica una particular regla para la integración. De modo similar, uno podría aplicar las técnicas aplicadas en la suma de columnas (otro procedimiento algorítmico) sin saber qué números está sumando. Por ejemplo, alguien podría aplicar el método sin mirar para saber qué números están representados, o los números podrían ser muy grandes para que los reconociera. Esa persona aún podría aprender las reglas para manipular dígitos sin que se le haya dicho que se usaban para representar números. El método de adición de columnas está diseñado de tal modo que los resultados no dependen de si la persona que ejecuta la computación conoce los significados de los términos. El procedimiento está diseñado para que, aplicado a las representaciones de dos números, A y B , confiablemente resulte la producción de una representación de un número C , tal que $A+B=C$.

Las técnicas de inferencia formal tienen un amplio alcance. Muchos problemas que caen fuera de las matemáticas e involucran información semántica altamente específica de un contexto pueden ser objeto de caracterización formal. Un juego como el ajedrez, por ejemplo, puede representarse mediante (1) un conjunto de símbolos que representan las piezas, (2) expresiones que representan estados posibles del tablero, (3) una expresión que selecciona el estado inicial del tablero, y (4) un conjunto de reglas que gobiernan la legalidad de las movidas mapeando expresiones que representan estados legales del tablero después de una movida m , al conjunto de expresiones que representan estados legales sucesores después de la movida $m+1$.

qué derivar. Por ejemplo, la regla de *modus ponens*, autoriza como válidas derivaciones de una expresión de la forma q dondequiera que uno haya derivado expresiones de la forma $p \supset q$ y p , pero no le dice al lógico que debe *tratar de derivar* q como parte de una prueba particular. Por otro lado, los algoritmos computacionales contienen reglas de procedimiento: reglas que especifican lo que debe hacerse bajo circunstancias particulares. El algoritmo para la suma de columnas no dice que uno *puede* poner un 7 en la próxima columna de la suma cuando encuentra un 3 y un 4 en la próxima columna de los dos sumandos, especifica que esto es lo que *debe* hacerse si uno va a aplicar el algoritmo. Pero tanto las reglas de procedimiento involucradas en los procedimientos algorítmicos, como las reglas de derivación involucradas en lógica formal, geometría, etc., puede decirse que son “formales” en el sentido de que la aplicabilidad de las reglas depende completamente de las formas sintácticas de las expresiones y no de sus valores semánticos.

Además de los juegos, también es posible representar los rasgos esenciales de muchos procesos del mundo en modelos formales de las clases empleadas por físicos, ingenieros y economistas. En general, un proceso puede modelarse si puede hallarse una manera adecuada de representar los objetos, las relaciones, y los eventos que estructuran el proceso, y crearse un conjunto de reglas de derivación que mapeen una representación R de un estado S de los procesos hacia una representación sucesora R^* de un estado S^* , sólo en el caso que el proceso es tal que S^* sería el estado sucesor para S . Como consecuencia, es posible crear sistemas representacionales en los que grandes cantidades de información semántica estén codificados sintácticamente, con el efecto de que la aplicación de técnicas de derivación puramente sintácticas pueden resultar en la producción de secuencias de representaciones que mantienen importantes relaciones semánticas: notablemente, secuencias que podrían contar como líneas de razonamiento racional, convincente.

La formalizabilidad de dominios simbólicos limitados muestra que las distinciones semánticas pueden preservarse sintácticamente y que la aplicación de reglas sintácticas de derivación puede resultar en una secuencia de representaciones semánticamente convincente. En términos gruesos, la formalización nos muestra cómo vincular la semántica con la sintaxis. Sin embargo, se requiere una forma de vincular las propiedades semánticas de representaciones con su habilidad para desempeñar un rol causal en la generación de nuevas representaciones con las que mantengan relaciones semánticas interesantes. En y por sí mismos, los métodos de prueba formal y los algoritmos formales no proporcionan tal vínculo, ya que dependen de las acciones del computador humano que las aplica. Es el paradigma de computación de máquina el que proporciona una forma de conectar los roles causales desempeñados por las representaciones con sus propiedades sintácticas, y así vincular indirectamente la semántica con el rol causal.

La transición crucial de las técnicas formales dependientes de un matemático humano a la computación mecánica llega con “On computable numbers”, de Alan Turing (1936). Este artículo buscó responder el problema matemático de hallar una caracterización general de las

clases de funciones que admiten soluciones computacionales (es decir, algorítmicas). La aproximación de Turing a este problema fue describir una máquina capaz de rastrear e imprimir símbolos en una cinta, máquina gobernada, en parte, por mecanismos internos y, en parte, por los símbolos específicos hallados en la cinta. Turing mostró que cualquier computación que pueda evaluarse mediante la aplicación de un algoritmo formal puede ser ejecutada por una máquina digital de la clase que él especificó. El objetivo original del artículo de Turing fue proporcionar una descripción general de todas las funciones computables: una función es computable sólo en el caso que pueda ser evaluada por una máquina de Turing. Pero al proporcionar esta respuesta para un problema en matemáticas, Turing también mostró algo mucho más interesante para los psicólogos y los filósofos, a saber: que es posible diseñar máquinas que no sólo almacenen símbolos pasivamente para uso humano, sino que también distingan activamente símbolos sobre la base de su forma y su orden sintáctico, y que, de hecho, operen de un modo parcialmente determinado por las propiedades sintácticas de los símbolos sobre los que operan. En resumen, Turing mostró que es posible vincular sintaxis y poderes causales en una máquina de computación.

Una máquina de computación es un aparato que posee varios rasgos distintivos: Primero, contiene medios en los que pueden almacenarse representaciones simbólicas. Estos símbolos, como símbolos escritos, pueden ordenarse en expresiones con estructura sintáctica y puede asignárseles interpretaciones mediante un esquema de interpretación. Segundo, un computador es capaz de diferenciar entre representaciones de un modo que corresponde a distinciones en sus “formas” sintácticas. Tercero, las regularidades causales que gobiernan qué nuevos símbolos causará el computador que se instancien, dependen de la forma sintáctica de los símbolos ya almacenados por la máquina.

Para tomar un ejemplo simple, suponga que un computador está programado para explorar dos lugares de almacenamiento, *A* y *B*, donde se almacenan las representaciones de los enteros, y para causar la instanciación de una representación en un tercer lugar *C*, de un modo tal que la representación instanciada en *C* será la representación de la suma de los dos números representados en *A* y *B*. Las representaciones halladas en *A*, *B* y *C* tienen estructura sintáctica:

Permítasenos asumir que cada representación es una serie de dígitos binarios (1s y 0s). También tienen interpretaciones semánticas, a saber: las asignadas por el esquema de interpretación empleado por el diseñador del programa. Cuando el computador ejecute el programa, causará la instanciación de una representación en *C*. Qué representación se instancie en *C* dependerá de qué representaciones se hallen en *A* y *B*. Más específicamente, dependerá del tipo sintáctico de las representaciones halladas en *A* y *B* —es decir, de qué secuencias de dígitos binarios estén presentes en esos lugares. Lo que el computador hace al ejecutar este programa es análogo a la aplicación de un algoritmo formal (como el empleado en la suma de columnas), que es sensible a las formas sintácticas de las representaciones en *A* y *B*. Si el programa ha sido diseñado adecuadamente, el proceso total también imitará ajustadamente la adición, en el sentido de que lo almacenado en *C* siempre será una representación de la suma de los dos números representados en *A* y *B*. Esto es, si el programa es diseñado adecuadamente, las operaciones sintácticamente dependientes ejecutadas por la máquina también asegurarán la producción de una representación en *C* que mantiene las relaciones semánticas deseadas con las representaciones en *A* y *B*. Las propiedades semánticas de las representaciones no desempeñan ningún rol causal en el proceso —son etiológicamente inertes. Pero dado que todas las distinciones semánticas son preservadas sintácticamente, y el tipo sintáctico determina la contribución causal de una representación, hay una correspondencia entre las propiedades semánticas de una representación y el rol causal que desempeña.

Este ejemplo ilustra tres puntos importantes. El primero es la comprensión, obtenida de la lógica formal y las matemáticas, de que al menos algunas relaciones semánticas pueden ser reflejadas o “rastreadas” por relaciones sintácticas. El segundo es la comprensión, obtenida de la ciencia computacional, de que pueden hacerse máquinas para operar sobre símbolos de forma tal que las propiedades sintácticas de los símbolos puedan reflejarse en sus roles causales. De hecho, para cualquier problema que pueda resolverse por la aplicación de un algoritmo formal *A*, es posible diseñar una máquina *M* que generará una serie de representaciones correspondientes a aquellas que serían producidas por la aplicación del algoritmo *A*. Estos dos puntos producen conjuntamente un tercero, a saber: que es posible para

las máquinas operar sobre símbolos de una forma que es, en palabras de Fodor, “sensible sólo a las propiedades sintácticas” de los símbolos y que está “confinada por completo a alterar sus formas,” en tanto

‘la máquina es creada de modo que transformará un símbolo en otro sí y sólo sí las proposiciones expresadas por los símbolos así transformados mantienen ciertas relaciones *semánticas* —e. g., la relación que las premisas sostienen con la conclusión de un argumento válido.’ (Fodor 1987:19)

En síntesis, “los computadores nos muestran cómo conectar las propiedades semánticas con las propiedades causales para *símbolos*” (Ibíd.) Y esto completa el deseado vínculo entre semántica y causalidad: para dominios que pueden formalizarse, las propiedades semánticas pueden vincularse a propiedades causales codificando las diferencias semánticas en sintaxis y diseñando una máquina manejada por los rasgos sintácticos de los símbolos.

Los computadores digitales son aparatos que almacenan y manipulan representaciones simbólicas, incluyendo tal manipulación la creación de nuevas instancias de símbolos. Las regularidades que gobiernan la generación de nuevas instancias simbólicas pueden exponerse en la forma de reglas que autorizan derivaciones basadas en los rasgos sintácticos de los símbolos ya instanciados en el sistema de almacenamiento del computador. Así, los símbolos cumplen roles causales en la generación de nuevos símbolos, y el rol causal que un símbolo puede desempeñar está determinado por su tipo sintáctico.

Resumiendo, la formalización muestra que (para dominios limitados) las propiedades semánticas de un conjunto de símbolos pueden ser “espejadas” por propiedades sintácticas; los computadores digitales ofrecen la prueba de que las propiedades sintácticas de los símbolos pueden ser determinantes causales en la generación de nuevos símbolos. En síntesis, el paradigma computacional muestra que uno puede coordinar las propiedades semánticas de las representaciones con los roles causales que estas pueden desempeñar codificando todas las distinciones semánticas en sintaxis.

I.5. Concepción computacional de los procesos cognitivos.

Vimos que la primera tesis constitutiva del la TCM era una concepción representacional de la naturaleza de los estados intencionales, a saber: que tales estados son relaciones a

representaciones mentales. La segunda tesis constitutiva de la TCM es una concepción computacional de la naturaleza de los procesos cognitivos, a saber: que los procesos cognitivos son computaciones sobre representaciones mentales, o “secuencias causales de instanciaciones de representaciones mentales” (Fodor 1987:17). Según Fodor:

‘Una cadena de pensamientos, por ejemplo, es una secuencia causal de instanciaciones de representaciones mentales que expresa las proposiciones que son los objetos de los pensamientos. En una primera aproximación, pensar ‘Va a llover; entonces me refugiaré’ es tener la instanciación de una representación mental que significa *Me refugiaré*, causada, de cierta forma, por una instanciación de una representación mental que significa *Va a llover*.’ (Ibid.)

Esta concepción puede descomponerse en varias afirmaciones constituyentes. Primera, los procesos cognitivos son secuencias de estados intencionales. Ahora, de acuerdo con la TCM, estar en un estado intencional particular es estar en una particular relación funcional con una representación mental. Así, si un organismo está realizando un proceso cognitivo, pasa por una secuencia de relaciones funcionales con representaciones mentales. Segunda, hay relaciones causales entre los estados intencionales que estructuran un proceso cognitivo. Estar en la relación R con una representación de tipo MP en el tiempo t (digamos, creer a las 12:00 del día que va a llover), puede ser una causa parcial de llegar a estar en la relación R^* con una representación de tipo MP^* en el tiempo $t+\partial$ (e. g., a las 12:01 llegar a la decisión de refugiarse). Tercera, la relación causal entre los estados seleccionados no es meramente incidental, sino que depende de modo regular de las propiedades sintácticas de las representaciones mentales. Es porque un organismo está en la relación R con una instanciación de tipo (sintáctico) MP en t que llega a estar en la relación R^* con una instanciación de tipo (sintáctico) MP^* en $t+\partial$, así como el programa de suma causa una particular representación a instanciarse en C porque representaciones con patrones sintácticos particulares están presentes en A y B . Tal como las representaciones en los computadores pueden desempeñar un rol causal en la generación de nuevas representaciones, y hacerlo en virtud de su forma sintáctica, así también ‘las representaciones mentales tienen sus roles causales en virtud de sus propiedades formales’ (Fodor 1981:26). Cuarta, como en el caso de un algoritmo formal o un programa de computador, cualesquiera diferencias semánticas entre representaciones mentales son reflejadas por distinciones sintácticas. Así, para cualesquiera dos representaciones MP y MP^*

con las que está relacionado un organismo simple, si MP y MP* difieren con respecto a sus propiedades semánticas, también deben ser de tipos sintácticos diferentes.

Ver los procesos mentales de esta forma es tratar la mente como si literalmente fuera un computador digital. Un computador es un aparato que ejecuta manipulaciones de símbolos sobre la base de los rasgos sintácticos de los símbolos, y puede hacerlo de un modo que respeta rasgos semánticos en tanto estén codificados en la sintaxis. De acuerdo a la TCM, los estados mentales involucran representaciones simbólicas de las cuales los primeros heredan sus propiedades semánticas. Todas las diferencias semánticas entre representaciones están codificadas sintácticamente, y la mente es un aparato cuyas regularidades causales están determinadas por las propiedades sintácticas de sus representaciones.

Esta concepción de la naturaleza de los procesos cognitivos posibilita que las explicaciones psicológicas intencionales subsuman propiedades intencionales correlacionadas con los poderes causales de los estados que componen dichos procesos. La explicación intencional selecciona estados intencionales por los valores semánticos de las representaciones que involucra, pero estos valores semánticos no son en sí mismos causalmente eficientes. De acuerdo con la TCM, las propiedades semánticas de las representaciones están correlacionadas con sus tipos sintácticos. Así, cuando se seleccionan representaciones por sus valores semánticos, su singular tipo sintáctico también es seleccionado, y el tipo sintáctico de una representación *es* un determinante del rol causal que ésta puede desempeñar al causar instancias de otras representaciones y en la etiología de la conducta. De este modo, las explicaciones psicológicas intencionales basadas en la TCM seleccionan *causas*. De hecho, las propiedades semánticas por medio de las cuales se seleccionan estados intencionales están correlacionadas con los roles causales que pueden desempeñar, porque las propiedades semánticas están correlacionadas con propiedades sintácticas, y las propiedades sintácticas determinan los poderes causales. Esto proporciona la posibilidad de dar cuenta de la causación mental de una manera que no requiere que las propiedades semánticas sean causalmente activas, correlacionando no obstante valor semántico y rol causal.

Tras esta presentación de aspectos fundamentales relativos a la TCM, que evidencian los vínculos entre teoría computacional e individualismo, pasamos a revisar la argumentación crítica de Wilson.

II. Wilson y el argumento computacional para el individualismo.²⁰

II.1. Computacionalismo en psicología.

El argumento que Wilson va a poner en discusión mantiene, según él, una semejanza *formal* con el argumento de los poderes causales discutido en el capítulo anterior, en el sentido de que éste afirmaba que la individuación de estados/procesos mentales debe ser individualista por la naturaleza *científica* de la psicología, mientras que el argumento para el individualismo sometido a consideración ahora dirá que debe ser individualista por la naturaleza *computacional* de la cognición. No obstante, existiría una diferencia entre ambos casos, a saber: aquí se trata de un argumento empírico para el individualismo, centrado en una afirmación *a posteriori* acerca de la naturaleza del procesamiento mental; su fundamento estaría en las prácticas explicativas particulares *en psicología*, antes que en una visión *a priori* sobre la naturaleza de la explicación científica o causal.

Wilson entiende al computacionalismo como una *ley de estructura cualitativa*, asumiendo lo planteado por la *hipótesis del sistema de símbolos físicos* —la idea de que ‘un sistema de símbolos físicos tiene los medios necesarios y suficientes para la acción general inteligente’ (Newell y Simon 1976:41)—, es decir, como una propuesta sobre la cognición que estructura un *paradigma* de investigación para las ciencias que estudian la cognición.²¹

Una “ley” de estructura cualitativa es un enunciado general cualitativo que determina el dominio y, en algunos casos, la metodología que abarca un campo de indagación, de forma

²⁰ El contenido de esta sección referido a Wilson, incluyendo las citas, corresponde a Wilson 1995:Cap. 3.

²¹ Al modo en que el conductismo habría constituido un paradigma de investigación para la psicología en la primera mitad del siglo pasado.

similar a los paradigmas kuhnianos.²² Según esto, el carácter “empírico” del computacionalismo debe entenderse significando que éste es ‘una afirmación que no puede ser refutada o confirmada por ningún experimento particular o conjunto finito de experimentos’ (p. 62).

Esta “precisión” de Wilson sobre el sentido del carácter empírico del computacionalismo constituye lo que denominará “compromisos mínimos del computacionalismo” (p. 63), compromisos a los que, declara, se acotará en lo que sigue y que, en su opinión, excluyen discusiones relativas a la relación entre la naturaleza de la psicología popular y el computacionalismo, así como discusiones respecto de las arquitecturas clásicas y conexionistas, es decir, a la concepción relativa al tipo de estado computacional que son los estados mentales. Por lo tanto, su interés primario estará puesto ‘en el argumento del computacionalismo al individualismo, no en el computacionalismo en sí’ (p. 62).

El interés primario de Wilson podría evidenciar cierta contradicción en el tratamiento que pretende hacer, pues discutir el argumento del computacionalismo al individualismo parecería suponer, al menos, el determinar si el computacionalismo entraña o no, individualismo. Esto implicaría, en cualquier caso, decir algo suficientemente sustantivo sobre el computacionalismo *en sí*, con independencia de que no se discutan otras o todas sus implicaciones. Entiendo que eso sustantivo debería referirse, en tanto concierne al individualismo, a la naturaleza de los estados mentales computacionales, que el individualista considera estados físicos bajo descripción funcional. El supuesto implícito de Wilson es que los “compromisos mínimos” del computacionalismo no exigen individualismo. Como cuestión de hecho, que los compromisos mínimos del computacionalismo puedan excluir la discusión sobre la relación entre computación y psicología popular es ya excluir un punto importante en la motivación para la concepción fodoriana de la TCM: la reivindicación de la psicología popular de deseos y creencias.

²² Ejemplos de leyes de estructura cualitativa en otros campos serían: la idea de que “el ladrillo básico de construcción de todos los organismos vivos es la célula”, en biología, y la afirmación geológica de que “la superficie del globo [terráqueo] es una colección de grandes placas” (Newell y Simon 1976:38-9).

II.2. El argumento computacional.

Wilson considera que la plausibilidad inicial de un argumento desde el computacionalismo al individualismo estriba, en parte, en que tal argumento proporciona fundamento teórico para intuiciones preteóricas que motivan a este último, ejemplo de las cuales sería la que señala que los *códigos* que nuestro cerebro almacena, no *lo que* codifican, son relevantes para la conducta cognitiva (McGinn 1982:210). Dado que los procesos mentales (como se dijo, cadenas de estados mentales) obedecen reglas formales, afectar el procesamiento cognitivo de un agente implica cambiar los códigos formales que constituyen su sistema cognitivo. Una segunda intuición individualista es que el organismo es una unidad *completa y natural*, a los propósitos de la teorización psicológica (Devitt 1990:388). Los estados computacionales, en tanto estados psicológicos individualistas, son intrínsecos al sistema en el que se instancian. Los sistemas computacionales, en tanto individualistas, son unidades completas y naturales.

La afirmación que hace Wilson en el sentido de que *los estados computacionales*, en tanto estados psicológicos individualistas, *son intrínsecos* al sistema en el que se instancian me parece descaminada, antes bien, y de acuerdo con lo revisado en el capítulo anterior, lo que un individualista sostiene es que los estados computacionales que describen estados mentales supervienen en los estados físicos intrínsecos del organismo que los instancia.

Para Wilson, el nexo entre computacionalismo e individualismo asume, como dijimos al inicio, la forma de un argumento al que se referirá como *argumento computacional* para el individualismo [en adelante AC] y que plantea lo siguiente:

Las ciencias de la cognición individúan taxonómicamente procesos mentales sólo en tanto procesos computacionales.

Los estados y procesos computacionales que un individuo instancia supervienen en los estados físicos intrínsecos de ese individuo.

En consecuencia,

Las ciencias de la cognición individúan sólo estados y procesos que supervienen en los estados físicos intrínsecos del individuo que instancia esos estados y procesos.

Antes de continuar, dos breves comentarios sobre AC. El primero para llamar la atención sobre la primera premisa, cuyo carácter totalitario considero tendencioso en exceso. Hasta donde puedo ver, tal premisa no parece hallar fundamento alguno en, o seguirse directamente de, la exigencia individualista de individuación por poderes causales. En principio, resulta ser por completo independiente del individualismo el que los individualistas crean que hay buenas razones para adoptar una concepción computacional de los procesos mentales. El segundo comentario se orienta hacia la segunda premisa, que parece conceder lo que señalé más arriba respecto de los alcances de la afirmación individualista sobre la superveniencia de los estados computacionales en estados físicos intrínsecos, antes que el carácter intrínseco de éstos.

En apoyo de mi primer comentario cito a Fodor, quien en su penúltimo libro escribe:

Por años, he escrito una cantidad de libros alabando la Teoría Computacional de la Mente (en adelante, TCM). En mi opinión, es lejos la mejor teoría de la cognición que hemos tenido; verdaderamente, la única que hemos tenido que merece la molestia de una discusión seria. Hay hechos sobre la mente de los que da cuenta y que estaríamos completamente perdidos para explicarlos sin ella; su idea central —que los procesos intencionales son operaciones sintácticas definidas sobre representaciones mentales— es sorprendentemente elegante. Hay, en resumen, razón para suponer que la teoría computacional es parte de la verdad sobre la cognición.

Pero no se me ha ocurrido que alguien pueda pensar que es una parte muy *grande* de la verdad; todavía menos, que está a kilómetros de ser la historia completa acerca de cómo trabaja la mente. ... Así, entonces, cuando escribí libros sobre lo buena que es la TCM, generalmente planteé un punto para incluir una sección diciendo que no supongo que pueda constituir más que un fragmento de una psicología cognitiva total y satisfactoria; y que los problemas más interesantes —ciertamente los más difíciles— sobre el pensar improbablemente van a ser muy iluminados por alguna clase de teoría computacional que seamos capaces de imaginar ahora. Conjeturo que en cierto modo di por concedido que incluso nosotros, fervientes admiradores de la psicología computacional, estábamos más o menos de acuerdo en eso. (Fodor 2000:1)

Me parece que Fodor es bastante elocuente respecto de su aprecio por la TCM, así como sobre su conciencia respecto de los límites de ésta, evidenciando una ponderación acerca de las posibilidades actuales de una psicología computacional que la primera premisa de evidentemente AC no contempla. Asimismo, creo que evidencia que las razones para ser

individualista no son equivalentes a las razones para ser computacionalista respecto de la cognición, aun cuando ambos grupos de razones no sean incompatibles.

El núcleo del cuestionamiento de Wilson a AC se organiza en torno a proponer y defender una alternativa que, en su opinión, habría sido pasada por alto o no tratada seriamente, a saber: el *computacionalismo amplio*, cuya “posibilidad” y “plausibilidad” para las ciencias de la cognición argumentará. Su ataque a AC se desarrolla en dos etapas dirigidas, respectivamente, a cada una de sus premisas. El punto de partida pretende socavar la segunda premisa, estableciendo la *posibilidad* del computacionalismo amplio.

II.3. Computacionalismo amplio.

II.3.1. Coherencia del computacionalismo amplio.

La idea básica en la propuesta del computacionalismo amplio dice que, aceptando que el procesamiento cognitivo es computacional, los estados/procesos sobre los que cuantifica una psicología computacional no necesitan ser individualistas porque el sistema cognitivo al que pertenecen puede ser un sistema *computacional amplio*. Tal sistema cognitivo puede trascender los límites del individuo e incluir partes de su ambiente, razón por la que al menos algunos de los estados computacionales que lo conforman no supervendrían en los estados físicos intrínsecos del individuo. Así, la psicología computacional resultante involucraría una referencia esencial al ambiente más allá del individuo y sus estados no podrían taxonomizarse de forma individualista. Si hay descripciones computacionales (formales) del ambiente de un organismo y de sus estados mentales, y si las transiciones causales desde las primeras a estos últimos pueden pensarse como computaciones, hay un proceso, que comienza en el ambiente y termina en el organismo, que puede pensarse como una computación, una computación amplia.

La descripción de la idea de Wilson sobre el computacionalismo amplio hace patente la semejanza entre su actual argumento y el argumento de los poderes causales [APC], discutido en el capítulo anterior. La exigencia de superveniencia, esta vez de los estados computacionales en las propiedades físicas intrínsecas del individuo, sirve también aquí como

factor crítico. Para el caso del argumento de los poderes causales, se dijo que el uso de una forma extensa y una restringida de “poder causal”, endosada ahí al individualista, incumplía la superveniencia de los poderes causales en los estados físicos intrínsecos. Aquí, Wilson introduce la dimensión “amplia” del computacionalismo que, como antes se dijo sucedía con el sentido extenso de poder causal, fracasará en supervenir sobre los estados físicos intrínsecos del individuo en quien se instancian contando, por eso, contra el individualismo.

Wilson busca apoyo para el computacionalismo amplio entre los propios individualistas, para lo cual selecciona argumentos ofrecidos por dos importantes filósofos: Gabriel Segal y Frances Egan. El primero de estos, discutiendo una afirmación de Martín Davies respecto a que “la psicología cognitiva trata los sistemas de procesamiento de información (módulos) y las criaturas en su totalidad *en tanto que* anidadas en sistemas mayores y finalmente en ambientes particulares” (Davies 1991:482), afirma que:

La base de superveniencia del contenido de una representación es algún sistema mayor en el que la representación está anidada. Esto puede ser: la criatura como un todo más su ambiente, la criatura como un todo, el mayor módulo en el que la representación ocurre, un sub-procesador de ese módulo, un sub-sub-procesador de ese módulo, un sub-sub-sub. . . . El individualismo es la tesis de que los estados representacionales de un sistema están determinados por propiedades intrínsecas de ese sistema. Parece probable que los sujetos como un todo (o cerebros totales) construyan sistemas computacionales mayores, integrados. Los sujetos como un todo más ambientes anidados no construyen sistemas computacionales integrados. Esta es una razón de por qué los individualistas trazan la línea donde lo hacen: el sujeto como un todo es el mayor candidato aceptable para la base de superveniencia porque es el mayor sistema integrado disponible. (Segal 1991:492, elipsis en el original)²³

En este texto de Segal, Wilson ve una concesión implícita a la coherencia del computacionalismo amplio situando, al mismo tiempo, el punto de divergencia entre un computacionalista *estrecho* y uno *amplio*. Este último rechaza precisamente la afirmación de que el “sujeto como un todo”, el individuo, sea “el mayor sistema físico integrado disponible”

²³ Al asumir que la base de superveniencia está constituida por las propiedades físicas intrínsecas *del individuo*, el individualismo iguala sistemas individuales y computacionales o asume que los últimos serán parte de los primeros, punto que bloquearía a priori la opción del computacionalismo amplio.

para la explicación psicológica computacional. Dado que Segal considera que la idea de que “los sujetos como un todo más ambientes anidados no construyen sistemas computacionales integrados” puede defenderse sobre bases a posteriori; el hecho de que no encontramos sistemas cognitivos computacionales mayores que el individuo implicaría, según Wilson, que aquél también debería admitir que la segunda premisa de AC tiene una justificación a posteriori, una que, permitiendo la mera posibilidad de sistemas computacionales amplios, muestra por qué *nuestros* sistemas cognitivos computacionales son individualistas.

Frances Egan (1992), por su parte, ha argumentado que las taxonomías computacionales son individualistas en su naturaleza: habría un rasgo general de la taxonomía en la psicología computacional, o quizás acerca de la teoría computacional más en general aún, que entraña que los estados y procesos computacionales sean individualistas. Esta visión implica que la segunda premisa de AC es verdadera, no de nosotros en tanto sistemas computacionales, sino de los sistemas computacionales *per se*. Egan comienza afirmando que la meta de las teorías computacionales de la cognición es ‘caracterizar los mecanismos subyacentes a nuestras varias capacidades cognitivas’ (pp. 444-5), y tales teorías ‘construyen procesos cognitivos como operaciones formales definidas sobre estructuras de símbolos’ (p. 446). Ahora:

Los símbolos sólo son objetos caracterizados funcionalmente, cuyas condiciones de individuación son especificadas por una *función de realización* f_R que mapea clases de equivalencia de rasgos físicos de un sistema en lo que podríamos llamar rasgos ‘simbólicos’. Las operaciones formales son precisamente esas operaciones físicas que son diferencialmente sensibles a los aspectos de las expresiones simbólicas que bajo la función de realización f_R se especifican como rasgos simbólicos. El mapeo f_R permite una secuencia causal de transiciones de estados físicos a ser interpretada como una *computación*.

Dado este método de individuar estados computacionales, dos sistemas ejecutando las mismas operaciones sobre las mismas estructuras de símbolos son computacionalmente indistinguibles. (p. 446)

De acuerdo con Egan, de lo anterior sigue que ‘si dos sistemas son duplicados moleculares entonces son duplicados computacionales. Las descripciones computacionales son individualistas —individuan estados por tipo sin referencia al ambiente o contexto social del sujeto’ (p.446).

Esta conclusión es cuestionada por Wilson, que la considera una petición de principio, en tanto afirma precisamente lo que el individualismo debería probar contra el computacionalismo amplio, a saber: la identidad entre sistemas computacionales e individuos. En contraste con dicho supuesto, el computacionalismo amplio avala el que pueda haber sistemas computacionales que se extiendan más allá de los límites del individuo.²⁴ Adicionalmente, Wilson cree que no hay nada en el método de individuación computacional que señala Egan, que implique que la clase de los rasgos físicos mapeados por una función de realización no pueda incluir miembros que sean parte del ambiente del individuo. Si esto es así, Egan no habría proporcionado un argumento de por qué el individualismo (respecto a la psicología computacional) se sigue de la naturaleza misma de la psicología computacional ni señalaría una incoherencia interna en la idea del computacionalismo amplio.

Comparto con Wilson que nada impide, en principio, que la clase de los rasgos físicos mapeados —creo que bien puede decirse *simbolizados*— por una función de realización cualquiera, pueda incluir miembros que sean parte del ambiente del individuo. No obstante, no es intuitivamente claro que tal sistema computacional amplio constituya un sistema cognitivo humano y, por eso, un objeto de la psicología. Difiero de Wilson en cuanto a aceptar que la cita de Egan en comentario afirma la identidad entre sistemas computacionales e individuos; en mi opinión, más bien lo que afirma es: (a) la identidad entre sistemas computacionales —su equivalencia funcional—, cuando ejecutan las mismas operaciones sobre las mismas estructuras de símbolos; y (b), que dos individuos molecularmente idénticos son, *ipso facto*, sistemas computacionales idénticos, afirmación que, por otra parte, se ajusta a la exigencia individualista, permitiendo la superveniencia de los estados computacionales en los estados físicos intrínsecos de tales individuos. Se trata de una afirmación más débil que la atribuida por Wilson, en cuanto a que se limita al caso de duplicados moleculares, antes que a individuos y sistemas computacionales en general.

²⁴ Planteado así, parece que implícitamente Wilson concede que existe algo como un “individuo” que tiene límites, más allá de los cuales se prolongan los potenciales sistemas computacionales amplios.

Continuando con su línea de argumentación, Wilson plantea una analogía entre computacionalismo amplio y funcionalismo amplio —la idea de que el rol conceptual de los estados mentales se extiende hacia el mundo (Harman 1987, 1988)—, pero señala una distinción entre ambos. En su opinión, el computacionalismo amplio sería más modesto y más radical que el funcionalismo amplio y, además, proporcionaría bases para un caso más fuerte contra el individualismo. Su modestia estribaría en conceder que el individualismo es verdadero de al menos algunos procesos mentales, rechazando sólo su carácter omnibarcativo. La radicalidad es consecuencia de negar que todo sistema formal o computacional se instancie en un individuo, punto que Wilson afirma los individualistas consideran concedido. Este último punto socavaría el argumento computacional de una manera fundamental y sería una objeción decisiva contra el individualismo, porque no sólo remueve al computacionalismo como uno de los mayores apoyos del individualismo, sino que también proporciona las bases para argumentar desde el computacionalismo una visión no-individualista de la psicología computacional como tal.

De esta forma, el desafío planteado a AC, antes que defender que los estados psicológicos requieren una construcción amplia, argumenta que los *sistemas computacionales* o *formales* en los que tales estados se instancian (o de los cuales son parte) se extienden más allá del individuo (cf. Kitcher 1985). La distinción entre un individuo y un sistema cognitivo computacional es central para la comprensión del computacionalismo amplio. Aun si muchos sistemas cognitivos computacionales se instancian por completo en un individuo, el computacionalismo amplio es una posibilidad porque los límites del individuo y el sistema computacional o cognitivo no necesitan ser idénticos.

Como ejemplo de un posible proceso computacional amplio, Wilson ofrece el caso del proceso de multiplicar (cf. Clark 1989:Cap. 4, 7; 1993:Cap. 6). Típicamente, uno multiplica números almacenando soluciones intermedias en alguna forma escrita, usualmente sobre papel, resolviendo el siguiente componente del problema, almacenando el resultado en un papel, y así sucesivamente. El proceso real que uno atraviesa al multiplicar números uno con otro involucra típicamente el almacenamiento de símbolos *en papel*. Una descripción del proceso

de multiplicación debe incluir una descripción de los símbolos matemáticos, y para la mayoría de los seres humanos tal descripción presupone una referencia a alguna cosa externa al organismo individual. Una parte crucial del proceso de multiplicación, a saber, el almacenamiento de símbolos matemáticos, se extiende más allá de los límites del individuo. En tanto multiplicadores, somos parte de sistemas computacionales amplios. Este caso ilustra un punto importante en la visión que Wilson apoya, la idea de que la resolución humana de problemas involucra *esencialmente* la explotación de representaciones del ambiente. Afirmación que se haría más plausible mientras más complejo sea el proceso computacional en el que participamos. Por ejemplo, pruebas de teoremas complejos en lógica cuatificacional raramente se efectúan por entero en la cabeza; al menos *algunos* de los símbolos se almacenan externamente.

Este ejemplo de proceso computacional amplio resulta en extremo desafiante para una comprensión individualista de los procesos mentales. Evidentemente, el caso implica al menos dos consecuencias anti-individualistas: (a) Borra el límite entre individuo y ambiente; la descripción computacional del proceso total abarca al individuo y al medio externo en el que almacena representaciones que forman parte de dicho proceso; (b) viola la idea de causalidad implicada por la superveniencia local (en el individuo) de los estados que constituyen el proceso computacional descrito en términos amplios.

Respecto a la primera consecuencia, no hay una propuesta sobre qué sería eso que ejecuta un proceso computacional amplio. La idea de interacción individuo-medio me parece inaplicable desde esta perspectiva, ya que el medio “externo” de almacenamiento es parte *esencial* del proceso total. Así, realmente no se trata de un individuo más parte del ambiente, sino de una entidad nueva e inespecificada. Esto mismo oscurece el tipo de causalidad en juego, la relación entre las capacidades causales de los estados computacionalmente descritos referidos a eso que es “ambiente” (en la descripción individualista) y eso que es “individuo” (también en la descripción individualista). Si hablamos de una entidad nueva, quizá siga siendo aplicable la superveniencia “local”, aunque de una forma muy contraintuitiva, lo que, por cierto, no podría contar muy sustantivamente en su contra. Por el contrario, la idea de que el

organismo sea el mayor sistema integrado disponible, a los efectos de implementar un sistema computacional, en tanto éste, como sistema autocontenido, proporciona una base de superveniencia local para los estados computacionales que describen estados mentales, brinda plausibilidad intuitiva al individualismo en este punto.

Dado que el computacionalismo amplio sería apropiado en los casos en que la interacción entre un individuo y algo extrínseco a él es parte crucial del proceso computacional a describirse como *explanans* en psicología, los procesos perceptivos y conductuales también deben ser asequibles a descripciones computacionales amplias. De acuerdo con Wilson, en el caso de la percepción es una parte intrínseca del proceso que el sistema acepta inputs del ambiente y los procesa, tal que ese input queda disponible para su ulterior procesamiento mental. El proceso perceptivo involucra una interacción ente un individuo y su ambiente. Dado que la percepción es un proceso que comienza con inputs ambientales, inputs que en sí mismos pueden tener una descripción formal y así ser accesibles a un sistema cognitivo computacional, todos los componentes del proceso perceptivo pueden describirse como parte de un sistema computacional amplio.

Un individualista puede objetar que esta caracterización del proceso de percepción simplemente supone la cuestión. Los objetos de percepción relevantes no son externos sino internos a un individuo; por ejemplo, lo son las imágenes retinales bidimensionales, no algún tipo de input ambiental. Una concepción computacional amplia de la percepción presupone una visión de la percepción que un individualista rechazará. Para Wilson, esta objeción concede un punto débil o negativo, a saber, que la naturaleza formal o computacional del procesamiento mental *en sí misma* no entraña individualismo. Para ello, se necesitaría, dice Wilson, hacer una afirmación sustancial acerca de, por ejemplo, los objetos de la percepción, en orden a derivar el individualismo del computacionalismo. Lo mismo sería verdadero para cualquier área de la cognición que se afirme es computacional. La formalidad de la cognición en sí no entraña individualismo, lo que permite la *posibilidad* del computacionalismo amplio.

No obstante, Wilson opina que puede hacerse y defenderse una afirmación aún más fuerte sobre el computacionalismo amplio, a saber, que los estados psicológicos son computacionales sólo en tanto son parte de un *sistema* formal implementado²⁵ y los sistemas formales de los que al menos algunos estados mentales son parte, no se instancian completamente en ningún individuo natural, esto es, en ningún organismo. Así, en tanto estados computacionales, tales estados psicológicos no se instancian en ningún individuo. Este argumento implicaría no sólo que una psicología computacional amplia es posible, sino también, asumiendo la verdad del computacionalismo, que al menos para algunos estados psicológicos tal psicología es *necesaria*. Lo contrario implicaría, cuando menos, una pérdida de potencia explicativa.

La evaluación de este argumento requiere examinar la noción de formalidad que Wilson emplea, examen que desarrollo a continuación.

II.3.2. La noción de formalidad.

El computacionalismo se expresa, a veces, como la idea de que, ya que la cognición es esencialmente *formal*, las ciencias cognitivas deben restringirse a postular y cuantificar sobre las propiedades formales de los estados mentales. Esta expresión de computacionalismo, que Fodor (1980:226-28) ha llamado *condición de formalidad*, puede hacer que el argumento del computacionalismo al individualismo parezca apremiante, pues las propiedades formales que tienen los estados mentales son pensadas con frecuencia como propiedades *intrínsecas* de los símbolos mentales, tales como su *forma* y *tamaño*²⁶. Además, esta concepción del computacionalismo considera las propiedades formales como una especie particular de *poderes causales* que tienen los estados mentales, propiedades que supervienen en las propiedades físicas intrínsecas del individuo. Esto es lo que haría tentador ver al computacionalismo como proporcionando un marco general para una especificación ulterior de la naturaleza de tales poderes. Para esta concepción de la formalidad, la tarea de la psicología computacional es descubrir las propiedades intrínsecas de las instanciaciones en el lenguaje del pensamiento.

²⁵ Véase la siguiente sección.

²⁶ Véase Fodor 1980:227; 1987:18.

Una prevención de hace Wilson respecto a la condición de formalidad consiste en indicarla como una *interpretación* del computacionalismo en psicología, una afirmación acerca de lo que el computacionalismo entraña o involucra, y no simplemente un enunciado del computacionalismo como tal.²⁷ Aunque la noción de formalidad se usa con frecuencia en la teoría computacional, Wilson considera ‘en ciertos aspectos descaminado’ (p. 72) hablar de *propiedades* como propiedades intrínsecas de los componentes individuales de sistemas computacionales. Al menos parte de lo descaminado de referirse a propiedades se derivaría del hecho de que la concepción de formalidad usada en lógica, matemáticas y ciencia computacional, disciplinas que proporcionan las bases últimas para el computacionalismo en psicología, es distinta de la expresada por la condición de formalidad. En estas disciplinas, dice Wilson, el foco está puesto en las propiedades y conducta de *sistemas* formales. Un sistema formal consta de primitivos, reglas de formación, fórmulas, axiomas y reglas de inferencia. Los fundamentos de la lógica están interesados, en parte, en la relación entre las nociones de sistema formal, procedimiento efectivo, algoritmo, computación y el conjunto de funciones recursivas. Para esta concepción de formalidad, que Wilson denomina *concepción sistémica de formalidad*, un sistema formal dado puede expresarse en notaciones alternativas y puede, en principio, ser realizado por una nación de personas relacionadas unas con otras según lo especifican las reglas del sistema (cf. Block 1978; Searle 1980). En este sentido, las propiedades físicas intrínsecas de los símbolos en un sistema formal son arbitrarias (véase Devitt 1990, 1991).

Para la concepción sistémica de formalidad que Wilson defiende, hay poco que decir de las propiedades formales que tienen los símbolos *particulares*; interesan más propiedades tales como estar cerrado bajo *modus ponens*, ser transitivo y ser compacto, que son propiedades de *sistemas formales* o, en forma derivada, propiedades de símbolos como elementos de sistemas formales. Los procesos computacionales, operaciones e instrucciones se piensan con

²⁷ Enunciados como este evidencian el sentido de la “precisión” de Wilson respecto a considerar el computacionalismo como una ley de estructura cualitativa. La generalidad de esta visión hace que cualquier especificación ulterior implique algo que puede afirmarse *más allá* de los compromisos mínimos del computacionalismo.

frecuencia como formales, pero esto sólo quiere decir que pueden describirse adecuadamente como el resultado de la aplicación de las reglas y algoritmos que constituyen el sistema al que pertenecen. En tanto símbolos particulares en un sistema formal, éstos tienen propiedades formales, pero no es claro si tales propiedades son intrínsecas o extrínsecas a los propios símbolos. Por ejemplo, una instancia del símbolo A conducirá a una instancia del símbolo B , y lo hará sólo en virtud de su “forma” en un sistema formal que contenga sólo la regla $A \rightarrow B$, pero lo hará sólo en un sistema formal con una regla de ese tipo. En consecuencia, Wilson recomienda ser precavido al identificar tales propiedades formales con poderes causales intrínsecos que posean los símbolos. Esas propiedades tienen el significado causal que tienen sólo en tanto los símbolos a los que se les atribuyen son parte de un sistema formal determinado.

Algunas puntualizaciones críticas son necesarias aquí. Cuando expusimos la TCM, destinamos una sección a tratar temas vinculados con la formalización y la computación de máquina. Señalamos allí que lo importante —para la TCM— de la relación entre formalización y computación era la posibilidad de: (1) vincular las propiedades semánticas de los símbolos con los poderes causales que permiten que la instanciación de un símbolo s_1 en el tiempo t sea una causa parcial de la instanciación de un segundo símbolo s_2 en el tiempo $t+\partial$; y, (2) que las leyes que gobiernan las regularidades causales pueden asegurar también que las operaciones que generan nuevas instanciaciones de símbolos respetarán las relaciones semánticas entre los símbolos, tal que el proceso total resultante sea, en un sentido amplio, racional. Luego, a partir de lo dicho debe ser claro que no son las propiedades formales, en tanto pertenecientes a *sistemas formales*, las que son concebidas como poderes causales; se dijo que es el paradigma de computación de máquina el que proporciona una forma de conectar los roles causales desempeñados por las representaciones con sus propiedades sintácticas, y así vincular indirectamente la semántica con el rol causal. Finalmente, la condición de formalidad nos remite al criterio de individuación solipsista metodológico,²⁸ es decir, a la doctrina que señala que los estados mentales se individualizan sin consideración por sus propiedades semánticas, en tanto estados mentales, lo que se pretende que pueda equivaler a afirmar que éstos se

²⁸ Cf. Capítulo 1.

individúan por sus propiedades “no-semánticas”, o sintácticas, o formales, o, como también se dice, por su “forma”. En consecuencia, evidentemente la condición de formalidad no funda la relación entre computacionalismo e individualismo. Sí podría hacerlo el que el sistema de símbolos físicos que *tiene* los medios necesarios y suficientes para la acción general inteligente supervenga en las propiedades físicas intrínsecas del organismo en el cual se instancia, haciendo posible y razonable que tales símbolos físicos, los estados mentales, tengan los poderes causales que tienen, y en virtud de los cuales habrán de individuarse.

De acuerdo con Wilson, la concepción sistémica de formalidad implica concebir las propiedades formales de los estados mentales como no-semánticas y no-físicas, a diferencia de la condición de formalidad fodoriana, que sólo considera el contraste entre propiedades formales y semánticas. Explica estos dos contrastes (formal/no-semántico; formal/no-físico) en la concepción de cognición de los computacionalistas (estrechos), examinando dos condiciones diferentes que debe satisfacer un nivel de descripción computacional de estados y procesos psicológicos.

En contraste con el nivel *físico* de descripción, dice Wilson, el nivel computacional es distinto de, e irreducible a, los niveles de descripción que caracterizan las realizaciones físicas de un sistema formal particular. El mismo programa computacional, el mismo sistema formal, puede instanciarse en muchas formas físicamente distintas. Dado el computacionalismo, este es el sentido en el cual la psicología es autónoma respecto de las ciencias físicas. Es esta autonomía, y por eso el contraste entre lo formal y lo físico, lo que Wilson piensa subyace a la primera premisa de AC. Por su parte, en contraste con el nivel *semántico* o representacional de descripción, el nivel computacional especifica las propiedades de los símbolos físicos y las reglas que constituyen el sistema formal del cual son parte esas instanciaciones mentales, sin referencia a lo que, si lo hay, esos símbolos representan. Un proponente de AC, en cambio, afirma que, a diferencia del nivel semántico de descripción, el nivel computacional especifica propiedades que son determinadas por los estados físicos intrínsecos del organismo. Este rasgo de las propiedades especificadas en el nivel formal de descripción hace plausible la segunda premisa del argumento computacional.

Antes de terminar esta sección, un apunte sobre el contraste entre propiedades formales y no-físicas en el nivel computacional de descripción. Wilson apoya su afirmación en la autonomía de la psicología, lo que parece sugerir que la psicología, de ser computacional, sería una ciencia de un dominio no-físico de propiedades. Sobre esto, me limito a comentar y suscribir lo que Fodor (1974) ha argumentado en términos de una autonomía *relativa* de las ciencias especiales, siendo la psicología un caso de éstas; el dualismo no es una opción aquí. Adicionalmente, entiendo que es efectivo que la condición de formalidad no considera que exista un contraste entre ser una propiedad “formal”, en el sentido ya indicado, y ser una propiedad física. Esto es, precisamente, lo que se pretende asegurar vía superveniencia. Las propiedades “formales” de los estados mentales son físicas y no-semánticas —o “sintácticas”, si lo prefiere.

II.3.3. Objeciones al computacionalismo amplio.

Para cerrar su argumentación en favor de la posibilidad del computacionalismo amplio, Wilson discute dos eventuales objeciones que éste podría enfrentar. La primera se refiere a lo que él considera una fortaleza *prima facie* del computacionalismo amplio, a saber, que está ingenuamente no comprometido en considerar el carácter computacional preciso de la cognición. Cita como ejemplo de esto, el hecho de que parecería ser compatible con las concepciones arquitectónicas clásica y conexionista del computacionalismo en psicología. Debido a esto, el computacionalismo amplio puede parecer problemático para alguien interesado en el problema del grado en que éste es una visión *real* de la psicología computacional, en oposición a una forma “débil” de considerar la metáfora computacional. La segunda objeción introduce fundamentos empíricos para dudar que *seamos* sistemas computacionales amplios, apelando para ello a la noción de modularidad. Si el éxito empírico ha llegado a la psicología computacional sólo o predominantemente con la presuposición de la modularidad, eso echa dudas sobre la idea de desarrollar una psicología computacional amplia empíricamente adecuada. Desarrollo ahora los aspectos sustantivos de estas objeciones y las respuestas que ofrece Wilson en cada caso.

La idea de que la mente de un individuo no es simplemente descrita o modelada por un programa de computadora es central para el paradigma computacional en psicología; la cognición es *guiada por reglas*, no solamente *regular* (Bennett 1964). El computacionalismo amplio sería, según esta perspectiva, posible sólo si uno se apoya en una lectura débil de la metáfora computacional, una lectura que no hace justicia a los compromisos computacionalistas en las ciencias cognitivas contemporáneas. Un sistema computacional estrecho o individualista —nuestros computadores personales, por ejemplo— sigue reglas instanciando o implementando programas. Un aparato físico es *capaz* de implementar un programa dado si sus estados físicos están configurados de tal modo que las transiciones entre estos estados son isomórficos a transiciones entre estados que el programa especifica, es decir, hay un mapeo desde clases de equivalencia de estados físicos a los estados simbólicos que constituyen el programa. Dado que el poder de implementación se caracteriza en términos de la noción matemática de isomorfismo, un aparato físico dado puede implementar un gran número de programas reales y un número infinito de programas posibles; qué programa sea el que de hecho implemente depende de cuál sea el programa que está codificado en el disco que insertamos o en cualquier otra cosa en la que almacenemos tal programa.

En opinión de Wilson, el carácter estrecho (individualista) del sistema no desempeña ningún rol significativo en la concepción de implementación; lo mismo puede decirse de sistemas computacionales *amplios*: un sistema computacional amplio tiene el poder de implementar sólo esos programas para los que hay un isomorfismo entre estados físicos del sistema y los estados simbólicos que especifica el programa. La concepción del poder de implementación real es una generalización de la del caso de sistemas computacionales estrechos: ‘un sistema computacional amplio implementa el “programa” almacenado físicamente en el ambiente con el que interactúa causalmente’. (p. 78) Determinar la descripción simbólica apropiada de los aspectos del ambiente de un organismo es con mucho un asunto a posteriori, tanto como lo es hacerlo respecto a la estructura *interna* de un organismo.

En el párrafo anterior, Wilson escribe “programa” entre comillas para destacar dos diferencias entre los programas que *corren* las computadoras y los que los computacionalistas (estrechos o amplios) afirman que corremos nosotros: (a) A diferencia de los programas que codificamos

en discos físicos, debemos descubrir qué interpretaciones simbólicas precisas podemos dar, bien de aspectos del ambiente de un organismo o de su estructura interna (o de ambos); (b) estas interpretaciones pueden resultar no ser lo bastante elaboradas como para garantizar la etiqueta “programa”. Entonces, (a) y (b) se supone que distinguen lo que sabemos de los organismos en tanto computadores. Simplemente no estamos, según Wilson, en una posición epistemológica apropiada para afirmar que nuestros cerebros o nuestros cerebros más nuestros ambientes instancian programas, precisamente en el sentido en que las computadoras lo hacen; he ahí, en parte, por qué el computacionalismo es una ley de estructura cualitativa *propuesta* para un programa de investigación y no una verdad familiar sobre la cognición. A la luz de las semejanzas y diferencias entre nosotros y las computadoras que emergen de la investigación empírica, seremos capaces de decidir si “programas” o “lenguajes” son categorías apropiadas con las que desarrollar explicaciones psicológicas. Nada de esto involucraría, nos dice Wilson, adoptar una comprensión débil de la metáfora computacional en psicología, sólo señala alguna precaución epistémica.

La cuestión de la implementación, como la plantea Wilson, merece una consideración más detenida. Un primer asunto requiere volver sobre la cuestión de saber qué clase de “sistema físico” sería el que implementa un programa computacional amplio, esto es, qué clase de “cosa” puede ser *un-(ex)individuo-más-una-u-otra-parte-de-su-(ex)ambiente*.²⁹ Visto así, el computacionalismo amplio rechaza el carácter mediador de la mente en la relación individuo y mundo; los pensamientos ya no “interactúan” con el mundo y “tener” una mente es una idea más difícil de concebir; de hecho lo que “piensa” en un sistema computacional amplio es, supongo, distinto de lo que se “comporta”. Recuérdese que una de las virtudes de la computación de máquina, según la TCM, radicaba en la posibilidad de conectar la semántica de los estados mentales con el rol causal de esos estados, vía su sintaxis. En otras palabras, la computación de máquina permite vincular y coordinar los aspectos informacional e informático de las representaciones mentales; ‘para hacer que la conducta sea sensible a la

²⁹ Uso “(ex)individuo” y “(ex)ambiente” para destacar que el problema que planteo se sigue, precisamente, del desdibujamiento de los límites respectivos en la relación individuo/ambiente en la concepción computacional amplia que defiende Wilson.

estructura del mundo, los pensamientos necesitan ser al mismo tiempo “cognitivos y ejecutivos” (Fodor 1997:99). La ventaja de esta concepción estriba en proporcionar una imagen (intuitivamente) coherente de lo que las ideas representan y lo que causan y que, al mismo tiempo, respeta su semántica. Desde esta visión, ‘las mismas propiedades de las representaciones mentales que determinan sus papeles informáticos son también portadoras de información acerca del mundo.’ (Íbid. p. 102)

Pasemos a la segunda objeción al computacionalismo amplio. Egan (1992:446, 457), citando ejemplos de investigación en la visión temprana y en el análisis sintáctico y morfológico en lingüística, señala que los procesos psicológicos para los que hay concepciones computacionales más satisfactorias son *modulares*: son de dominio específico y encapsulados respecto a la información.³⁰ Los mayores éxitos en la psicología computacional se han obtenido al explicar el carácter de los procesos psicológicos que funcionan con relativa independencia de la mayoría de las tareas *internas* del sistema cognitivo del individuo. Si el éxito empírico se vincula con el supuesto de la modularidad, surgen dudas sobre la idea de una psicología computacional *amplia* empíricamente adecuada.

Fodor (1983) trata más vivamente el punto sobre la relación entre la modularidad y la psicología computacional argumentando que ‘los sistemas globales son *per se* malos dominios para los modelos computacionales’ (p. 128). Específicamente, lo que él llama procesos *centrales*, tales como la resolución de problemas y la fijación de creencias, improbablemente tengan modelos computacionales precisamente porque son, en su opinión, no-modulares. La no-modularidad de los procesos centrales da una razón para ser escéptico sobre la posibilidad real (versus la mera posibilidad) de una psicología computacional adecuada para explicarlos. Y lo que es verdadero de los procesos centrales, que tienen acceso a una variedad de inputs representacionales, también lo sería de los procesos computacionales amplios, que dan acceso a representaciones de *fuera* del individuo.

³⁰ Un sistema de *dominio específico* es uno que opera sobre un tipo particular de información (un dominio); un sistema cognitivo que es *informacionalmente encapsulado* actúa como una función input-output para un conjunto específico de inputs y outputs y, así, es insensible a otra información.

Wilson responde inicialmente a este argumento en términos generales.³¹ Dice que, aún si se acuerda que, en general, los éxitos empíricos en las ciencias cognitivas hasta aquí han involucrado sistemas altamente modulares, tales como los empleados en la percepción visual y el reconocimiento de fonemas, esto podría ser así por una razón profunda, como que sólo los sistemas altamente modulares *son* computacionales; o, alternativamente, podría deberse a una razón trivial, como que sólo los procesos computacionales altamente modulares pueden ser comprendidos rápidamente como computacionales por los teóricos. En ambos casos, no habría nada que permita la extrapolación de un punto sobre los procesos centrales a los procesos computacionales amplios, ‘*ya que estos últimos también pueden ser modulares*’ (p. 80). La premisa de que los sistemas modulares son taxonomizados de forma individualista sería falsa porque los sistemas modulares bien pueden encapsular información que está en el *ambiente* del individuo, no en alguna parte del individuo; el módulo puede ser parte de un sistema computacional amplio.³²

Al cierre de esta sección, no deja de sorprender tanto esfuerzo destinado a la defensa sólo de la mera posibilidad del computacionalismo amplio, cuestión que difícilmente puede cuestionarse: en principio, *todo es posible*. Sobre la posibilidad real de sistemas computacionales amplios es de lo que trata la próxima sección, donde Wilson discute la *plausibilidad* del computacionalismo amplio.

II. 4. Plausibilidad del computacionalismo amplio en psicología cognitiva.

La posibilidad del computacionalismo amplio es, de acuerdo con Wilson, suficiente para mostrar que la segunda premisa de AC es falsa, pero el interés real en el argumento computacional descansa en la afirmación de que plausiblemente *somos* vistos como sistemas computacionales amplios. En tal sentido, Wilson piensa que el computacionalismo amplio se *hace* plausible con algo de la investigación computacional reciente en cognición humana y animal, lo que consolidaría y ampliaría su objeción a AC. Al efecto, discute dos ejemplos de

³¹ Una argumentación algo más detallada sobre el tema de la modularidad se revisará en el siguiente capítulo.

³² La pregunta sigue siendo, de *qué* cosa es un sistema computacional amplio, un sistema computacional.

investigación reciente en psicología cognitiva que mostrarían al computacionalismo amplio en acción. El primero corresponde a una aproximación pionera en el campo de la percepción de formas de Campbell y Robson (1968), conocida como *teoría de los canales espaciales múltiples*.

La idea básica de esta aproximación es que hay estímulos específicos a los que son sensibles conjuntos individuales de neuronas, estos estímulos se descomponen en *rejillas* [gratings] sinusoidales. Estas rejillas son relativamente simples, tienen sólo cuatro parámetros relevantes: frecuencia espacial, contraste, orientación y fase espacial. Una figura compuesta de estas rejillas es formalmente definible en términos de estos cuatro parámetros. La afirmación audaz y polémica de este programa de investigación es que *cualquier* escena natural en el ambiente de un organismo puede descomponerse en sus rejillas, y este hecho explica gran parte de la percepción humana de formas, incluyendo sus limitaciones.

De acuerdo con Wilson, para esta concepción de la percepción de formas, parte de la tarea del psicólogo de la percepción es identificar primitivos formales que describan adecuadamente el ambiente visual, y especificar algoritmos que se apliquen a estos primitivos para determinar escenas visuales completas. Para ver lo que esto significa, se emplea un caso más simple que la visión humana, el de un lente proyectando la imagen de un objeto en un pedazo de papel blanco. La *función de transferencia* grafica cómo el contraste es transferido a través del lente desde el objeto a la imagen y es definido sobre un rango de frecuencias espaciales. Como input, esta función toma el contraste en un objeto, produciendo como output contraste en la imagen. Del mismo modo, se define una *función de sensibilidad al contraste* para el sistema visual humano, que toma los mismos inputs desde el mundo para producir un output visual. Presuponiendo una equivalencia entre ambas funciones, Wilson afirma que el sistema formal que estudian los psicólogos de la percepción que trabajan dentro de este paradigma no es instanciado en ningún individuo: incluye, pero no está restringido a, las propiedades intrínsecas de un individuo. Esto se reflejaría en la metodología real empleada por tales psicólogos, que involucra un extenso y complejo análisis matemático de escenas naturales en sus primitivos computacionales. Tales análisis parecen ser una parte intrínseca del paradigma

de los canales espaciales múltiples, no simplemente algo preliminar para la psicología de la percepción.

Para considerar las afirmaciones de Wilson sobre el trabajo de Campbell y Robson, es necesario introducir unas mínimas distinciones, no obstante, me restringiré sólo a aquellos aspectos que me parecen cruciales para lo que pretendo plantear. En primer lugar, creo que la *función de transferencia* (FT) y la *función de sensibilidad al contraste* (FSC) no son equivalentes; considérese sólo el que la primera se refiere a un lente y la segunda al sistema visual humano involucrado en el fenómeno de la *percepción*, objeto de investigación de Campbell y Robson. La FSC corresponde a una *función de transferencia perceptiva* (FTP), que incluye tanto los componentes ópticos (lo que pasa en el ojo) —la parte que, quizá, puede homologarse al caso del lente— y los componentes neuronales; ambos componentes constituyen el *sistema visual* y hacen posible abordar el tema de la percepción visual humana. La sensibilidad del sistema visual se define en función del *umbral de contraste* requerido para detectar una determinada frecuencia espacial, el punto mínimo en el cual una rejilla determinada comienza a ser percibida. Midiendo los umbrales de contraste para una amplia gama de diferentes frecuencias espaciales, se deriva una curva que describe la sensibilidad total del sistema visual al contraste. Esta curva se denomina FSC, para distinguirla de FT. El término “sensibilidad” nos recuerda precisamente que estamos tratando con una propiedad del sistema visual, *con respecto* a una propiedad del estímulo.

Mi impresión sobre el análisis que Wilson hace de este caso, es que el análisis matemático (análisis de Fourier) que se efectúa del estímulo, su descomposición en un complejo de rejillas, sólo sirve a los efectos de correlacionar los parámetros “objetivos” característicos de cada rejilla con los umbrales perceptivos de individuos y no constituyen *per se* un sistema computacional amplio. Por supuesto, que no lo haga no implica, para concederle algo a Wilson, que no sea *posible* concebir eventualmente un sistema computacional amplio para describir lo que sucede.

El segundo ejemplo que Wilson nos invita a considerar corresponde al trabajo de Charles Gallistel (1989), quien reporta una investigación sobre las concepciones de espacio, tiempo y número en una variedad de animales, incluyendo abejas, ratas y hormigas. Una de las conclusiones primarias de Gallistel es que los modelos de base puramente sensorial de una gama de conducta animal son inadecuados. Más bien, la evidencia sugiere aplastantemente que estos animales construyen representaciones muy complejas de sus ambientes y las usan para guiar su conducta. Gallistel argumenta que tales representaciones son computacionales; que hay evidencia fuerte de que estos animales instancian módulos que son sensibles a la estructura formal (e. g., geométrica) de sus ambientes; y que esta sensibilidad es responsable de su navegación a través de sus ambientes físicos. Por ejemplo, en hormigas y abejas los procesos computacionales de *estimación precisa* [dead reckoning]³³ toma como inputs la orientación solar del animal, la velocidad de avance y una representación del acimut solar, produciendo como output una representación de la posición relativa de la criatura respecto de algún hito, tal como un nido. La *función efeméride*, que produce el tercero de estos inputs, toma como sus inputs una vista del sol y el tiempo de algún reloj endógeno. Wilson afirma que ‘en ambos casos, los procesos computacionales se extienden más allá de los límites del individuo.’ (p. 83)

No comparto la última afirmación de Wilson sobre una extensión *de hecho* de los procesos computacionales planteados por Gallistel, formulada simplemente a partir de la proposición de un isomorfismo entre descripciones formales de las propiedades de un ambiente y las de los procesos mentales. Según mi opinión, válida también para el caso de Campbell y Robson, que exista un análisis en términos de propiedades formales de aspectos del ambiente, y un isomorfismo entre esos modelos formales y los modelos formales de los procesos computacionales internos, no supone ni siquiera por asomo que exista en tales investigaciones la idea de un sistema computacional amplio que implemente a ambos sistemas formales.

³³ Proceso para determinar el cambio de posición (desde un punto de partida conocido) integrando velocidad con respecto a tiempo. Proporciona una representación continua de la posición propia en el mundo... (Gallistel 1989:161)

Gallistel llama a su idea una perspectiva *computacional representacional* para la cognición animal. Sobre la navegación animal, Gallistel dice:

Los movimientos rutinarios del animal son gobernados por un proceso de navegación cercanamente análogo a la práctica diaria de la marina. Esta práctica descansa sobre un isomorfismo extensivo entre la geometría del movimiento y la posición, y los procesos computacionales que subyacen a la navegación. En el nivel neurofisiológico de análisis, la hipótesis implica que la descripción matemática del proceso en el cerebro del animal que funciona durante la navegación animal es paralela a la descripción matemática de las computaciones que hace un sistema de navegación humano o computarizado. (1989:176-7)

Para Wilson, esta cita sugiere que Gallistel, así como aquellos que trabajan en el paradigma multi-canal en la percepción de formas, no ven nada misterioso en postular un isomorfismo extendido entre las propiedades formalmente descritas de un ambiente y las de los procesos mentales. Gallistel defiende una concepción general de aprendizaje que se enfoca en la afirmación de que un animal sintoniza su conducta con las ‘propiedades formales’ de su ambiente. Él dice:

Hay una rica correspondencia formal entre procesos y relaciones en el ambiente y las operaciones que ejecuta el cerebro. Los procesos y relaciones cerebrales recapitulan procesos y relaciones del mundo. La recapitulación no es fortuita. Para ajustar la conducta al ambiente, el cerebro crea modelos de los aspectos conductualmente relevantes del ambiente. Las propiedades formales de los procesos que realizan estos modelos reflejan las propiedades formales de la realidad externa correspondiente porque estos procesos han estado sujetos a evolución y selección natural. La selección se ha basado en la fidelidad de estos procesos a la realidad externa. La evolución por selección natural crea ricos isomorfismos de funcionamiento entre los procesos cerebrales y el ambiente, y el aprendizaje va a comprenderse en términos de estos isomorfismos. (1989b:27).

Así, el isomorfismo entre los procesos computacionales instanciados en la cabeza y ciertas “propiedades formales” en el ambiente es responsable de la exitosa conducta de navegación que exhiben muchos animales. Dos postulados centrales de estos programas de investigación son que el ambiente del organismo tiene cierta estructura formal y que es la sensibilidad del organismo a esta estructura la que explica partes medulares de su ejecución cognitiva.

Voy a resumir las afirmaciones de Gallistel consignadas en las últimas dos citas, como sigue: Primero, existe un isomorfismo “extensivo” entre los procesos computacionales de navegación animal y humana, con respecto a las descripciones formales ambientales; segundo, hay un

paralelismo entre la descripción matemática de la actividad cerebral animal y la de un sistema humano o computacional de navegación; tercero, la conducta animal no es aleatoria, por cuanto su cerebro refleja (representa) aspectos conductualmente relevantes de su ambiente; y, cuarto, la correspondencia entre los modelos internos y el ambiente evidencia un isomorfismo sustantivo, seleccionado por la naturaleza precisamente por su fidelidad representacional.

Si estas cuatro afirmaciones resumen bien lo que Gallistel señala, no evidencian en modo alguno, como quiere Wilson, la existencia *real* de un sistema computacional amplio de cualquier tipo en su propuesta, ni la *intención* de que lo haya.³⁴ Pienso que las cuatro afirmaciones pueden reducirse a dos: (1) Hay un isomorfismo entre sistemas de navegación animales/humanos y sistemas computacionales, lo que puede significar simplemente que los procesos cognitivos animales/humanos son computacionales; (2) Hay un isomorfismo básico mundo-mente o mundo-sistema cognitivo, que va en la línea de explicar por qué tener una mente es una buena idea.

Adicionalmente, varios comentarios que Gallistel hace sobre su proyecto apoyan una interpretación individualista. En primer lugar, Gallistel dice que su ‘agenda es reduccionista: comprender el aprendizaje como un fenómeno neuronal’ (1989b:24), llegando a decir que estudiar el sistema computacional total es simplemente un ‘preludio necesario para comprender lo que el sistema hace, en términos de lo que hacen sus elementos’ (1989b:24). Entonces, representar la estructura computacional del ambiente de un organismo, aunque metodológicamente necesario, es periférico para la comprensión del aprendizaje en sí. Además, en la conclusión de su libro, Gallistel dice que ‘la estructura de los mecanismos computacionales es dictada por la estructura formal de las representaciones a ser computadas y por los datos sensoriales o mnémicos a partir de los cuales son computadas’ (1989b:581), sugiriendo que él ve el sistema computacional de interés para el psicólogo como instanciado completamente en el organismo individual.

³⁴ Por cierto nada de esto puede contar contra la *posibilidad* de que lo haya.

Un rasgo general de la visión de Gallistel sobre la cognición animal señala que mucha de la conducta animal es gobernada por mapas internos y representaciones matemáticas del ambiente, antes que por input sensorial directo. Esta visión adscribe a un animal con un alto grado de autonomía de su ambiente y este aspecto de su concepción parece individualista: los animales navegan mediante *mapas internos*, no mediante rastreo sensorial, autodirección u otro método ambientalmente interactivo. Al caracterizar cómo navega un animal, abstraemos su ambiente real y nos concentramos en los rasgos intrínsecos de su mapa o modelo de ese ambiente.

Sobre el supuesto de que *puede* darse una interpretación computacional amplia de la perspectiva de Gallistel, Wilson funda la afirmación de ‘el computacionalismo en sí no entraña individualismo: los sistemas computacionales no necesitan ser individualistas’ (p. 86). Comparto la opinión de Wilson, con un carácter general. Lo que interesa es si los sistemas cognitivos computacionales que implementan leyes psicológicas son individualistas o no. He tratado de defender la idea de que, al menos los que Wilson discute, sí lo son.

II.5. Computacionalismo reconsiderado.

La primera premisa del argumento computacional dice que la naturaleza computacional del procesamiento psicológico agota el interés que las ciencias cognitivas tienen en los estados mentales. Wilson argumentará que esta afirmación es demasiado fuerte y que un debilitamiento apropiado de ésta es compatible con la aceptabilidad de explicaciones intencionales amplias en la psicología cognitiva. Aun si el computacionalismo amplio fuera falso o incoherente, uno tendría razón para rechazar AC. En lo que sigue, Wilson se concentra en la primera premisa de AC, con independencia de la discusión previa.

La primera premisa de AC establece una relación fuerte particular entre el computacionalismo y las ciencias cognitivas, a saber, que *toda* la investigación en las ciencias cognitivas presupone que los procesos mentales son computacionales (o que toda la investigación actual que merece proseguirse es computacional). En apoyo de esta afirmación “fuerte”, Wilson cita a Robert Cummins cuando escribe: ‘El computacionalismo es difícil de tragar... pero es igualmente difícil de escupir. Las únicas alternativas son el dualismo, que no es una teoría

alternativa sino la afirmación de que la teoría es imposible, y el ‘neuronismo’ —i.e., la doctrina de que las capacidades caracterizadas intencionalmente son realizables sólo como capacidades neurofisiológicas’ (1983:90; nota al pié omitida). De acuerdo con Wilson, no hay razón para pensar que “dualismo” y “neuronismo” sean las únicas alternativas al computacionalismo. Para él, ‘la falsedad del computacionalismo no entrañaría que los estudios científicos de la cognición, a parte de los proporcionados por las neurociencias, son imposibles, como no lo implica la falsedad del *conductismo*.’ (p. 88) La autonomía de la psicología cognitiva, respecto de las neurociencias, no dependería de que todos los estados psicológicos (o incluso algunos) sean computacionales. Sirva como ejemplo el que una proporción significativa de la investigación en psicología cognitiva experimental *del desarrollo*, no presupone computacionalismo.³⁵ Una razón, entonces, para rechazar la primera premisa es que la mejor investigación en algunas áreas de las ciencias cognitivas no es en absoluto computacional.

Dos puntos respecto de lo precedente. En primer término, ya indiqué mi desacuerdo con el carácter totalitario de la primera premisa de AC y con su eventual derivación del individualismo. Siendo así, esta premisa puede debilitarse tanto como se quiera, para que pueda incluir explicaciones intencionales amplias. Sobre esto último versa el segundo punto que quiero señalar. Hasta donde sé, no hay una equivalencia, ni siquiera remota, entre computación amplia y explicación intencional amplia. Entiendo a esta última como aquella explicación que cuantifica sobre contenidos individuados por sus propiedades semánticas (esto es, no de acuerdo con el solipsismo metodológico). El tema se vincula con la discusión sobre contenido amplio y estrecho en psicología. Al respecto, baste con señalar aquí el tratamiento que hace del tema Fodor (1994/1997), donde desarrolla una idea para hacer plausible la implementación computacional de leyes intencionales (amplias), sin recurrir al recurso del contenido estrecho. La solución propuesta no implica, en caso alguno, recurrir a sistemas computacionales amplios. Una nota final para remarcar que la propuesta de Fodor sigue girando sobre la existencia de una diferencia entre explicación intencional y explicación

³⁵ Véase Carey 1985; Keil 1989; Leslie 1987; Spelke 1990.

causal, la posibilidad de la segunda se apoya en el computacionalismo, porque la primera no puede hacerlo independientemente.

Al discutir la afirmación de que la cognición es formal, Wilson especificó dos contrastes: los procesos formales son *no-semánticos* y *no-físicos*. En su opinión, estos contrastes “sugieren” que los estados psicológicos pueden describirse en un nivel de descripción físico, computacional y semántico, pero que sólo el nivel computacional constituirá un estudio de la cognición desarrollado y propiamente científico. De acuerdo con Wilson, la motivación central para pensar que hay un nivel de descripción computacional en psicología, distinto al nivel de descripción físico, también da una razón para pensar lo mismo del nivel de descripción semántico. Además, las críticas individualistas de la relevancia causal de las propiedades semánticas amplias para la explicación psicológica, también se aplicarían a las propiedades especificadas en el nivel computacional. Estas dos afirmaciones proporcionan, según Wilson, razones para rechazar la visión de que *sólo* las propiedades físicas y computacionales de los estados psicológicos son relevantes para las ciencias cognitivas, no éstas y las propiedades semánticas amplias de los estados psicológicos.

Extrapolando el argumento de la *realización múltiple* de estados psicológicos propuesto por el funcionalismo, en cuanto que el nivel funcional es autónomo con relación a cualquier nivel de descripción físico, Wilson afirma que la misma relación se sostiene entre los niveles de descripción representacional y computacional, y por las mismas razones. Muchas funciones *representacionales* que los organismos instancian son realizadas diferentemente, tanto en diferentes individuos cuanto en diferentes especies, lo que ejemplifica con un caso de Sterelny (1990:98): Murciélagos y búhos instancian estados psicológicos con el contenido “ratón ahí abajo” y este estado perceptivo compartido explica sus consecuentes conductas comunes. Pero, ya que los murciélagos instancian este estado en un sistema de ecolocalización, mientras los búhos lo instancian en un sistema visual, es extremadamente improbable que haya una descripción *computacional* de su estado compartido. El hecho de que esta función se instancie, no precisamente *cómo* se instancie, es relevante para algunos aspectos de la explicación psicológica y debería contar en contra del carácter totalitario de la primera premisa en AC. Así

como la caracterización neurofisiológica de una capacidad psicológica es incompleta, la caracterización computacional de esa capacidad no agota el interés que los psicólogos cognitivos tienen en esa capacidad. Hay preguntas de interés *psicológico* que no se reducen a preguntas sobre cómo se realizan los estados psicológicos en estructuras computacionales o neurofisiológicas. Necesitamos ir más allá del nivel de descripción computacional, tanto como necesitamos ir *a* ese nivel, en orden a capturar generalizaciones que de otra forma perderíamos.

Wilson no desarrolla los argumentos que justifican su afirmación sobre la extrema improbabilidad de una descripción computacional del proceso por el que búhos y murciélagos llegan a tener en sus cabezas una representación cuyo contenido es “ratón ahí abajo”, asumo que éstos son los de Sterelny. Resumidamente, Sterelny (1990) piensa que la realización múltiple de estados intencionales es precisamente lo que bloquea su especificación computacional: hay innumerables formas en las que individuos pueden representarse una situación que haga semejante su comportamiento. Una primera respuesta podría señalar que, dado que los estados mentales son estados con contenido proposicional, lo que se instancia para todos estos individuos es una representación con un mismo contenido —para el ejemplo que comentamos, “ratón ahí abajo”. Sterelny piensa que esto es aún insuficiente. Primero, porque la solución no bloquea necesariamente la cuestión de la realización múltiple; ¿cómo sabemos que se ha instanciado ese estado representacional o cualquier otro que, no obstante, produce “el mismo” comportamiento? En segundo lugar, porque Sterelny cree que no podría haber generalizaciones intencionales verdaderas sobre bases funcionalistas “estrechas” (1990:101-4).

Sobre la imposibilidad de generalizaciones intencionales verdaderas *en* la psicología cognitiva computacional, remito una vez más a Fodor (1994/1997), ya comentado unos párrafos más arriba. Por otra parte, considerando nuestra somera discusión del trabajo de Gallistel sobre hormigas, abejas y ratas, por lo menos parece menos improbable hacer algo respecto de los búhos; quizá la situación del murciélago pueda requerir más tiempo. Por último, que se puedan desarrollar “explicaciones” de la conducta de estos animales sin conocer cómo se produce la

representación “ratón ahí abajo” no cuenta contra el computacionalismo. Uno puede suponer que hay tal representación; otra cosa es mostrar cómo sería posible que la haya y, por esa vía, aportar evidencia a favor de la teoría que asume tales representaciones. Reiterando una idea expuesta antes, la computación hace posible vincular y coordinar los aspectos informacional e informático de los procesos cognitivos que, evidentemente son distintos, pero deben aprender a convivir.

Una razón por la que aquellos interesados en desarrollar una psicología científica han sido cuidadosos al caracterizar estados mentales por su contenido (amplio), es que individuos instanciando diferentes contenidos amplios pueden comportarse idénticamente. Por esta razón, las diferencias en el contenido que dos individuos instancian son consideradas a veces como no haciendo *diferencia causal* para la psicología que un individuo posee; afirman que los contenidos amplios no son poderes causales.³⁶ Pero los sistemas formales en tanto sistemas *no-físicos* tienen constituyentes que *no* tienen poderes causales; los sistemas formales están constituidos por entidades abstractas, tales como números, conjuntos, axiomas y teoremas. En contraste, los sistemas formales *implementados* están constituidos por estados que sí tienen poderes causales y el computacionalismo en psicología obtiene mucha de su plausibilidad, desde una perspectiva individualista, porque el nivel computacional se concibe para describir los poderes causales que tienen los estados psicológicos.

Desde aquí, Wilson argumentará ahora que ‘los sistemas *semánticos* implementados, es decir, sistemas de estados caracterizados por su contenido, también tienen poderes causales.’ (p. 90) La idea se basa en la afirmación de correspondencia sistemática entre la sintaxis y la semántica del lenguaje del pensamiento.³⁷ Wilson postula que esta correlación entre sintaxis y semántica, más que correlación, es evidencia para postular una relación *instancial* —puede entenderse como co-instanciación— entre las dos, y afirma que esto explica cómo las propiedades semánticas podrían ser causalmente relevantes para las explicaciones de la conducta, puesto que la misma relación es la que se sostiene entre los niveles de descripción

³⁶ Sobre esto, véase Capítulo 1.

³⁷ Stich ha denominado a esta idea la *tesis de la correlación*. (1983:186)

computacional y físico, haciendo que las propiedades del primero sean causalmente eficaces. La consecuencia de lo anterior será que: ‘Si puede decirse que las propiedades computacionales van a ser causalmente eficaces en virtud de esta relación, entonces también pueden serlo las propiedades *semánticas*’ (p.91), junto con la conclusión ulterior de que los niveles de descripción computacional y representacional desempeñan los mismos roles explicativos dentro de las ciencias cognitivas.

Wilson afirma en estas últimas líneas que la co-instanciación sistemática de propiedades sintácticas y semánticas dota a estas últimas de, y explica el que puedan tener, relevancia causal, al modo en que la relación de superveniencia explica las propiedades causales de los estados computacionales. Uno podría querer conceder la relevancia causal de las propiedades semánticas, pero no podría hacerlo gratis. El precio implica estar en condiciones de especificar el mecanismo, que no sería la correlación, mediante el cual se asegura la co-instanciación sistemática de propiedades sintácticas y semánticas. Adicionalmente, un problema con esto puede emerger de lo que sabemos acerca de la sensibilidad de los computadores a la “forma” de las representaciones sobre las que operan, mientras son insensibles a sus contenidos, al menos es lo que el trabajo con sistemas formales nos enseñó y lo que los formalistas nos aseguran: ‘Si cuidas de la sintaxis, la semántica se cuidará sola’. (Haugeland 1985/1988:102)

III. Conclusiones.

Este capítulo estuvo dedicado a revisar algunos aspectos fundamentales de la Teoría Computacional de la Mente, con énfasis en su versión “fuerte”, para pasar luego a una línea de argumentación crítica dirigida contra un supuesto argumento desde el computacionalismo al individualismo. Básicamente, dicha argumentación crítica se fundó en demostrar la posibilidad y plausibilidad del computacionalismo amplio, como una visión computacional que especifica estados que no supervienen localmente en el individuo que los instancia. He desplegado con detenimiento cada una de las líneas de argumentación y he intentado señalar sus posibles dificultades.

La primera de estas dificultades la he situado en torno a la propia construcción de AC, específicamente su primera premisa, que considero no tiene bases individualistas que la justifiquen en su forma omni-abarcativa.

Respecto de la posibilidad del computacionalismo, considero que carece de real interés discutir su mera posibilidad. Las teorías posibles en todos los campos son seguramente infinitas, y la enorme mayoría de ellas probablemente no llegará jamás a concretarse. Al responder a Schiffer, en un contexto diferente, pero relacionado, Fodor señala: ‘Las teorías posibles no amenazan los argumentos para la mejor explicación; sólo teorías reales lo hacen’ (Fodor 1991b:310). Aun cuando no he defendido aquí la idea de que la visión individualista del computacionalismo en psicología constituya un argumento del tipo de la mejor explicación, pienso que se sostiene en un número de ámbitos que confirman su valor.

Por otra parte, la discusión de una eventual plausibilidad del computacionalismo amplio en prácticas psicológicas recientes resulta mucho más desafiante. Desde esta perspectiva, creo que los casos presentados por Wilson no apoyan contundentemente sus afirmaciones y no cumplen sus expectativas; su análisis gira básicamente sobre la posibilidad de una revisión que los convierta en casos reales de operación de sistemas computacionales amplios en psicología animal o humana; en otras palabras, no son casos de computacionalismo amplio, sino que podrían serlo.

En consecuencia, se puede seguir siendo individualista y computacionalista en psicología todavía.

Capítulo 4. Argumentos metodológicos

I. Introducción.

Este capítulo está destinado a exponer y discutir, desde la perspectiva crítica de Wilson,³⁸ dos argumentos metodológicos citados en apoyo del individualismo. La idea central es responder el siguiente par de preguntas: (a)¿Cuál es la influencia metodológica del individualismo? y (b)¿Hay razones metodológicas para adoptar el individualismo como constreñimiento para el estudio de la cognición? Las respuestas a estas preguntas implican evaluar la existencia de hechos acerca de cómo operan o deben desarrollarse las explicaciones psicológicas estrechas y amplias, hechos que supondrían, siguiendo a Wilson, la coherencia metodológica o la fertilidad empírica sólo de las primeras. Si fuera el caso responder afirmativamente (a) y (b), podría decirse que hay razón para pensar que la investigación en las ciencias cognitivas debe ser individualista. En este sentido, argumentos tales serían más fuertes que los argumentos a priori o los empíricos que hemos examinado en los capítulos precedentes; en tanto fuesen sólidos, no dejarían lugar para una apelación ulterior a la práctica explicativa en psicología. De acuerdo con Wilson, el individualismo en psicología ‘aún podría vindicarse si puede mostrarse que las mejores metodologías disponibles para investigar en psicología tratan a los estados y procesos psicológicos como si fueran individualistas o, respectivamente, que las metodologías que claramente son anti-individualistas padecen algún problema metodológico inherente’ (p. 94).

Así, la cuestión se orienta a determinar si podemos concebir al individualismo implicando una visión metodológica de la psicología, una visión acerca de cómo uno debe estudiar la cognición. De verificarse lo anterior, lo siguiente es determinar si hay o no diferencias metodológicas que distingan a individualistas y no-individualistas. Responder en forma afirmativa a estas cuestiones, presupone que empíricamente podemos identificar aproximaciones a la cognición —paradigmas de investigación en psicología— que expresen, respectivamente, perspectivas estrechas y amplias de teoría psicológica; esto, a su vez,

³⁸ Los argumentos son desarrollados en Wilson 1995:Cap. 4.

presupone que conceptualmente podemos distinguir dos de tales perspectivas en términos generales.

Al comparar las visiones individualista y amplia sobre la taxonomía y la explicación en la práctica psicológica, encontramos que difieren, al menos, en su visión de la importancia de la inserción del organismo en un ambiente físico y social para la construcción de teorías en psicología.³⁹ El individualista propone que debemos hacer abstracción del ambiente particular de un organismo; al ver al individuo como la mayor unidad psicológica completa, el individualista afirma que, para los propósitos de la psicología, el ambiente puede ponerse entre paréntesis respecto de los individuos localizados en él: toda la información estrictamente relevante para la psicología está contenida en las representaciones mentales de los individuos. El psicólogo amplio, en cambio, considera la inserción de un organismo en un ambiente específico algo intrínseco para el desarrollo de la taxonomía y la explicación psicológica.

Si también asumimos que individualistas y no-individualistas avalan el computacionalismo en psicología, el computacionalista estrecho afirma que el carácter del mundo es relevante sólo en tanto éste afecta a las propiedades *formales* de las representaciones mentales.⁴⁰ Para el computacionalista amplio, en cambio, la psicología computacional no está interesada solamente en lo que está “en la cabeza”, sino más bien en las causas de la conducta de un individuo, donde tanto la conducta cuanto sus causas se consideran como procesos computacionales en un contexto ambiental.

Wilson reconoce que mucha de la investigación hecha dentro del paradigma computacional, particularmente en áreas centrales de la inteligencia artificial, tales como planeamiento, representación de conocimiento y resolución de problemas, ha sido conducida dentro de un marco computacional individualista; pero, como se mostró en el capítulo anterior, también considera que tanto la aproximación multi-canales para la percepción de formas cuanto la perspectiva computacional-representacional del aprendizaje animal incorporan una perspectiva

³⁹ Véase Fodor 1980:229.

⁴⁰ *Íbid.* p. 231.

amplia en la investigación psicológica dentro del paradigma computacional. En síntesis, él piensa que computacionalismo amplio y estrecho son tipos de perspectivas de investigación activa en psicología. Otros ejemplos de ambos paradigmas serían: las investigaciones sobre el lenguaje natural en Chomsky (1957, 1965, 1980), en particular la gramática transformacional, típicamente muy individualista en su metodología; y la psicología perceptiva de J.J. Gibson (1979), su óptica ecológica, como caso paradigmático de psicología amplia.

Suponiendo que gruesamente podemos distinguir dos paradigmas de investigación en psicología, Wilson nos pide considerar una pregunta que Fodor (1980:233-4) ha planteado: ¿qué razón hay para pensar que no podemos aceptar *ambos* paradigmas —una psicología *racionalista*, que es individualista, y otra psicología, no-individualista o *naturalista*— como apropiados para la investigación en la disciplina? A pesar de la posibilidad de tal ecumenismo, posibilidad discutida en el capítulo anterior al examinar la mera posibilidad del computacionalismo amplio, habría razones metodológicas para pensar que la opción individualista es la única opción defendible en psicología. En defensa de esta afirmación, Wilson expone dos líneas de razonamiento que derivan, la primera, directamente de una formulación de Fodor; la segunda, se supone implícita en las afirmaciones de mismo Fodor sobre la modularidad de gran parte de la cognición.

II. Argumentos metodológicos para el individualismo.

II.1. Argumento contra el naturalismo en psicología.

En términos generales, el argumento de Fodor contra la posibilidad práctica de desarrollar una psicología naturalista, esto es, una psicología que afirma que uno puede hacer ciencia a partir de las relaciones entre un organismo y su ambiente, supone una estrategia de investigación que requiere la culminación de todas las demás ciencias, para que la psicología esté en condiciones de caracterizar adecuadamente su vocabulario básico. El argumento de Fodor para esta conclusión procede como sigue:⁴¹ La psicología naturalista especificará —de forma canónica— las relaciones que se sostienen entre (la mente de) un organismo y objetos de su ambiente; estas relaciones son las que aseguran que cuando alguien piensa en algo de su

⁴¹ Véase Fodor 1980:247-52.

ambiente, piensa *en* esa otra cosa; esto supone una relación causal entre quien piensa y la cosa en la que piensa, descrita de modo que dicha relación causal se obtiene cuando el “objeto” satisface tal descripción; se requiere, entonces, una descripción bajo la cual la relación entre ambos instancie una ley, lo que presupone que uno tiene acceso a descripciones proyectables de los *relata* relevantes. Pero las descripciones de las clases de cosas con las que los organismos se relacionan en un ambiente son el tema objeto de otras ciencias, no de la psicología. Por tanto, una psicología naturalista debe apoyarse en otras ciencias para obtener las descripciones que necesitan sus leyes. Esta relación de dependencia entre la psicología naturalista y las otras ciencias implica que la primera requiere una estrategia de investigación *holística*, que presupone la culminación de las otras ciencias antes que la psicología en sí misma pueda completarse. Como dice Fodor, ‘los psicólogos naturalistas heredarán la Tierra, pero sólo después de que todos hayan terminado con ella. Sin duda es correcto tener una estrategia de investigación que dice “espera un poco”. Pero, ¿quién quiere esperar *para siempre?*’ (1980:248).

Entonces, si determinar el nivel de descripción apropiado para una psicología naturalista requiere conocer las correspondientes descripciones en otras disciplinas, una estrategia de investigación naturalista es metodológicamente defectuosa. Sea cual sea la promesa que pudiese haber en una estrategia de investigación naturalista, por su dependencia metodológica de otras ciencias no hay esperanza práctica de completar una psicología naturalista y, citando una vez más a Fodor, ‘para la metodología, la esperanza práctica es *todo*’ (1980:252). De acuerdo con Wilson, la afirmación principal de Fodor es que caracterizar las relaciones organismo/ambiente requiere descripciones nomológicas de estados psicológicos que sólo pueden ser proporcionadas por otras disciplinas; de aquí la dependencia metodológica de una psicología naturalista del resto de las ciencias.

No obstante seguir, en general, la línea de la argumentación de Fodor, hay un aspecto que falta en la presentación que Wilson y que considero no menor. El punto de partida de la discusión donde Fodor plantea esta evaluación de una psicología naturalista, se sitúa en torno a las

consecuencias que se seguirían de asumir la condición de formalidad,⁴² una de las cuales consiste en que la psicología debe adoptar una forma de solipsismo metodológico.⁴³ En cambio, lo que una psicología naturalista supone, y no puede tener, son descripciones individuadas por tipo completamente transparentes de los objetos que el organismo tiene en mente —algo que una psicología estrecha logra, trabajando con descripciones opacas, con lo que hay *en* la mente. El problema a despejar es bajo qué descripción entran, en una psicología naturalista, los objetos de pensamiento en las generalizaciones científicas sobre las relaciones entre pensamientos y sus objetos. La respuesta de Fodor, en función del carácter nomológico de la relación requerida, es que debe ser bajo descripción física, pero que nadie sabe si ‘tales descripciones vayan a estar por llegar en, digamos, los *próximos* trescientos años’ (Ídem).

A modo de ejemplo, respecto de la práctica de los naturalistas en ausencia de las descripciones requeridas, Fodor señala que, como una de las alternativas:

Podemos evadirla: al modo en que lo hacen los teóricos del aprendizaje. Esto es, podemos “leer” la descripción del estímulo a partir del carácter de la respuesta del organismo. Como realidad histórica, esto ha conducido a una clase de psicología naturalista que es meramente una paráfrasis solemne de lo que sabe la abuela de cualquiera: por ejemplo, decir “los lápices son el estímulo discriminativo para la preferencia de ‘lápices’”, donde la abuelita habría dicho ‘lápiz’ se refiere a los lápices. Considero que la revisión por Chomsky (1959) de *Verbal Behavior* demostró, de una vez para siempre, la fatuidad de este camino. Lo que *sería* interesante —lo que habría sorprendido a la abuela— es una generalización de la forma Δ es el estímulo discriminativo para preferencias de ‘lápiz’, donde Δ es una descripción que selecciona lápices en un vocabulario proyectable (por ejemplo, en el vocabulario de la física) (Fodor 1980:252).

El hecho de que la relación entre el estímulo ambiental y la conducta no pueda establecerse en términos no-circulares, por ejemplo en un vocabulario físico, es lo que Fodor considera que bloquea el desarrollo de una psicología naturalista.

La siguiente sección desarrolla la contraargumentación de Wilson, en términos de rebatir, primero, el que una psicología naturalista o amplia necesite ser metodológicamente

⁴² Véase capítulo anterior.

⁴³ Esto es, que las clases psicológicas se individualizan con independencia de su semántica.

dependiente de otras ciencias en la forma en que Fodor lo afirma; segundo, que aun si fuera dependiente del modo señalado, esto no constituiría un problema metodológico profundo para las aproximaciones amplias a la psicología.

II.2. Crítica al argumento contra la psicología naturalista.

De acuerdo con Wilson, hay dos razones para sospechar que el argumento de Fodor contra una psicología amplia o naturalista debe estar equivocado. La primera sería una implicación del argumento fodoriano al extender su razonamiento a cualquier nueva ciencia que se refiera a objetos taxonomizados en alguna ciencia existente y que clasifique transversalmente estos objetos, pues esta nueva ciencia subsume estos objetos bajo clases nuevas. Incluso podría aplicarse a cualquier etapa en la *historia* de la ciencia, con la consecuencia de que la mayor parte de las ciencias actualmente existentes serían metodológicamente problemáticas en la forma que Fodor lo afirma. La segunda razón apela a ejemplos de investigación conducidos en la tradición naturalista *en psicología*, que no parecen padecer el problema que Fodor señala.

Según Wilson, un problema en el argumento de Fodor es que asume que hay una descripción física, nomológica, para cada cosa con la que un organismo entra en relación psicológica. Pero como cada disciplina define el nivel de descripción en el que se asientan sus generalizaciones, cualquier entidad puede ser clasificada de modo diverso por diferentes disciplinas; algo clasificado como un planeta en astronomía es sólo un cuerpo con masa en física. Así, no habría razón que impida a una psicología amplia construir descripciones taxonómicas propias en lugar de emplear las disponibles en otras ciencias. La idea de que las ciencias cognitivas deben asentarse en un *nivel computacional* de descripción presupone justamente el mismo punto. Resumiendo, Wilson piensa que la conclusión de Fodor está lejos de la práctica taxonómica real en la ciencia en general y la psicología en particular.

Para Wilson, la constatación de que los científicos usan confiable y exitosamente términos de las disciplinas de otros, y lo hacen sin saber todo lo que hay que saber sobre los referentes de esos términos (cf. Gauker 1991; Hardwig 1985), se explica por la existencia de una *confianza*

*mercenaria*⁴⁴ de éstos en la comprensión de esos otros, de los términos y el conocimiento en sus respectivas disciplinas, que se manifiesta en su confianza en los instrumentos, métodos y técnicas matemáticas de campos distintos al propio. Por lo tanto, las relaciones de dependencia epistémica entre disciplinas científicas no son una barrera para el desarrollo de taxonomías explicativas y conocimiento científico. La psicología no constituye una excepción a esto.

Entonces, de acuerdo con Wilson, el argumento de Fodor contra la psicología naturalista o amplia falla porque, aun si una psicología amplia requiere descripciones taxonómicas del organismo y el ambiente, estas descripciones probablemente van a ser desarrolladas *dentro* de teorías amplias particulares, no extraídas de otras disciplinas. Al respecto, una aproximación naturalista a la psicología *no* es metodológicamente dependiente del desarrollo de taxonomías en otras ciencias.

Creo que una evaluación de los argumentos presentados por Wilson debe partir por señalar que en ciencia, así como en otros ámbitos, las personas pueden hacer lo que les parezca. Así, el argumento de Fodor no supone una imposibilidad para la psicología amplia de proponer las taxonomías que le plazca. Por otra parte, tampoco se trata de negar que exista una relación entre nuestros estados mentales y el mundo, en virtud de la cual éstos tienen el contenido que tienen. Lo que Fodor cuestiona es la posibilidad práctica de una ciencia psicológica (u otra) cuyas generalizaciones instancian tal relación. La afirmación medular para lo que ahora se debate, me parece, tiene que ver con la opinión de Fodor en el sentido que ‘la verdad, la referencia y el resto de las nociones semánticas no son categorías psicológicas’ (Fodor 1980:252). Para eso se requiere disponer de una teoría del contenido, una que pueda hacer el trabajo requerido, una que los naturalistas no tienen —los individualistas tampoco.

II.3. Argumento de la modularidad al individualismo.

⁴⁴ Término de Trout (1992).

Con independencia del argumento considerado antes, una segunda apelación metodológica pueden aducirse para desarrollar las ciencias cognitivas de una forma individualista, por referencia al éxito relativo de las explicaciones computacionales de sistemas o facultades cognitivas *modulares*. Un argumento desde la modularidad al individualismo podría organizarse a partir de las siguientes premisas:

- (a) Nuestro mayor éxito empírico en la psicología computacional ha llegado al explicar procesos psicológicos modulares, procesos que funcionan con relativa independencia de las labores internas del sistema cognitivo;
- (b) La investigación de facultades cognitivas altamente modulares ha conducido a teorías robustas con confirmación empírica significativa;
- (c) Las taxonomías que emergen de tales investigaciones son individualistas *porque* las capacidades que individualizan son modulares;

Entonces, las taxonomías individualistas han tendido a reflejar éxito empírico, al menos dentro de la psicología cognitiva computacional y, aun si no son las únicas que hay, son nuestra mejor apuesta actual para el éxito empírico futuro en las ciencias cognitivas (véase Egan 1992; Fodor 1983).

Si bien el argumento implica varias dimensiones, a saber, qué éxito explicativo ha habido en la psicología cognitiva computacional, qué lugar tienen las explicaciones computacionales en psicología y qué relaciones mantienen con otras explicaciones psicológicas, y la significación de la modularidad de algunos procesos cognitivos; esta última dimensión, la de la modularidad, es la noción central de este argumento metodológico para el individualismo. Fodor (1983) ha caracterizado los sistemas modulares en la mente diciendo que son de dominio específico, inconscientes, obligatorios, rápidos, sus productos se refieren a aspectos superficiales, informacionalmente encapsulados, tienen localización neuronal fija, presentan pautas de deterioro fijas y específicas, y su ontogenia exhibe características de secuencia. Los conceptos más importantes entre estos son los de especificidad de dominio y encapsulamiento informacional.

Un sistema cognitivo de dominio específico es uno que opera sobre algún tipo particular de información, que constituye su dominio. Los sistemas de dominio específico se diferencian con respecto a los sistemas cognitivos de *propósito general*, que almacenan u operan sobre información de una variedad de fuentes, información que *no* constituye un dominio. Por su parte, un sistema cognitivo informacionalmente encapsulado (o autónomo) es uno cuyo funcionamiento no puede ser intervenido, ‘encapsula’ algún o algunos tipos de información y es insensible a otra información. Fodor considera que este rasgo es la esencia de la modularidad de un sistema cognitivo (1983:71).

De acuerdo con Wilson, el argumento metodológico desde la modularidad al individualismo hace dos afirmaciones, a saber: (1) que los éxitos empíricos en la psicología cognitiva computacional han tendido a seguir la presunción de la modularidad; y (2) que los sistemas modulares son taxonomizados de acuerdo con el individualismo.

Con respecto a la primera de estas afirmaciones, Wilson la valora como plausible y nos plantea una pregunta respecto de por qué debemos esperar este tipo de éxito explicativo. Según su opinión, especificar un dominio de operación de un sistema informacionalmente encapsulado facilita la tarea de proporcionar una concepción computacional de cómo ese sistema puede realizarse o implementarse. La razón de esta facilitación descansa en que suponer la modularidad constriñe el flujo de información que es plausible reconstruir en una explicación computacional. Por ejemplo, si suponemos que la percepción de forma es encapsulada respecto de la información contenida en las imágenes retinales de dos dimensiones (Marr 1982), entonces cualquier concepción computacional de los procesos de percepción de forma tendrá un rango circunscrito de inputs con los cuales trabajar, haciendo que el problema computacional a resolver probablemente sea más tratable.

Desde otra perspectiva, dado que típicamente se presume que los sistemas modulares son altamente localizados (localizaciones neuronales fijas), la investigación de sistemas modulares putativos se facilita mediante una consideración de las desviaciones clínicas o inducidas del funcionamiento normal en casos que involucran daño cortical particular (pautas de deterioro

fijas y específicas); por ejemplo, en personas con déficit cognitivos específicos como afasia y agnosia. Las concepciones computacionales particulares de sistemas modulares pueden recibir más rápidamente evidencia que las refute o confirme desde la neuropsicología que una concepción de otros sistemas.

En contraste con su acuerdo respecto vínculo entre el éxito empírico y la presunción de modularidad, Wilson piensa que la segunda afirmación en el argumento desde la modularidad al individualismo, la que dice que los sistemas modulares son taxonomizados en forma individualista, es falsa. La falsedad de esta afirmación se deriva del hecho que sistemas modulares pueden encapsular información que está en el *ambiente* del individuo, no en alguna parte en el individuo. Siendo esto correcto, para volver sobre la discusión del capítulo precedente, un módulo tal puede ser parte de un sistema computacional amplio.

La conclusión de Wilson es ninguno de los rasgos principales de los sistemas cognitivos modulares, su especificidad de dominio y su encapsulamiento informacional, implica que tales sistemas no puedan verse apropiadamente como partes de sistemas computacionales amplios. Ser amplio y ser modular serían propiedades compatibles que un sistema cognitivo puede poseer, de modo entonces que la modularidad no entrañaría individualismo, porque la localización de la información con respecto a la cual el sistema es encapsulado no afecta a su especificidad de dominio.

Coincido con Wilson en cuanto que la modularidad *per se* no entraña individualismo; efectivamente un módulo puede formar parte de un sistema computacional amplio, uno que va más allá de los límites del individuo. En tanto un módulo forma parte de un sistema cognitivo computacional, la cuestión vuelve a ser, como se intentó mostrar en el capítulo anterior, cómo individuar los poderes causales de un sistema cognitivo amplio. Que el sistema computacional quede incluido dentro de los límites del individuo, hace posible la individuación de poderes causales para la producción de otros estados cognitivos o conducta.

III. Conclusiones.

En este breve capítulo revisamos, como tercera opción, una defensa del individualismo sobre bases metodológicas: un argumento sobre la incoherencia de la psicología naturalista, y uno desde la modularidad al individualismo. La contra-argumentación de Wilson estuvo dirigida a señalar la falsedad de ambas líneas de argumentación.

Mi impresión es que el carácter apremiante que Wilson les asigna a ambos argumentos no va más allá de un efecto retórico del modo de presentarlos, en el sentido de magnificar su capacidad para imposibilitar algún tipo de práctica concreta en psicología. Particularmente evidente me parece en el caso del argumento de la modularidad al individualismo. Ningún éxito basado en el supuesto de la modularidad puede bastar para afirmar que con éste se agota toda posibilidad alternativa de indagación; asimismo, la naturaleza computacional de los módulos no entraña, no puede hacerlo, individualismo. En cuanto al argumento de incoherencia metodológica de la psicología naturalista, de nuevo lo cuestionado no es la existencia de taxonomías amplias, sino su esperanza práctica de disponer de las descripciones semánticas requeridas por sus generalizaciones. En este sentido, me parece que toda la confianza que uno pueda tener en las clases generadas en otras ciencias no garantiza que la psicología amplia tenga lo que requiere, como lo requiere.

Lo que en ningún caso se ha mostrado es la incoherencia o inconsistencia metodológica del individualismo.

Capítulo 5. Conclusiones.

Los cuatro capítulos que preceden a este apartado final, que se propone a modo de conclusiones sobre lo discutido, han seguido un itinerario que implica, por un lado, una exposición bastante exhaustiva de argumentos vinculados con el individualismo en psicología y, por otro, el despliegue de objeciones y contra-argumentos para los primeros. Los argumentos y contra-argumentos se ordenaron y presentaron organizados como clases, a saber: el argumento de los poderes causales; el argumento computacional; y argumentos metodológicos, siguiendo la propuesta de Wilson (1995). El principal objetivo de esta exposición de argumentaciones individualistas y contra-argumentaciones anti-individualistas en cada uno de estos tres ámbitos, se centró en generar las condiciones para realizar una evaluación sobre qué tan apremiantes son como razones para abandonar el individualismo, en favor de posiciones anti-individualistas.

Así, nuestro examen del argumento de los poderes causales nos advertía sobre la falsedad del individualismo global [IG], la afirmación de que la individuación en aquellas ciencias que buscan proporcionar explicaciones causales de los fenómenos que estudian debe ser por poderes causales, cuestionada a partir de la existencia de clases taxonómicas en algunas ciencias, clases que figurarían como *explanantia* en explicaciones causales formuladas por dichas ciencias, que no satisfarían el criterio de ser por poderes causales. En forma paralela, se afirmaba que no existe posibilidad de revisar tales clases en términos de una propiedad intrínseca y otra relacional o histórica, y se hacía también una apelación a la ciencia tal como es realizada, suponiendo que esto constituiría un criterio más válido que el apriorismo del constreñimiento individualista. Como un elemento crítico aún más decisivo, se afirmaba que en el argumento de los poderes causales, formulado por Wilson, el individualista incurriría en una contradicción, al emplear en forma ambigua el término ‘poder causal’.

Con respecto a lo afirmado en el párrafo que antecede indicábamos, al finalizar el Capítulo 2, que los casos ofrecidos como apoyo para falsar IG no pueden ser objeto de un análisis pormenorizado, lo que vale también para la presunta imposibilidad de revisión de las clases

involucradas en esos casos cruciales, imposibilidad que, no obstante, difícilmente pueda considerarse establecida en forma categórica. Por otra parte, apelar a la práctica científica sin más como criterio falla, pues en forma independiente esta apelación no es suficiente, dado que se requiere explicitar el o los criterios que posibilitarían su interpretación como práctica científica válida. Finalmente, con respecto a la afirmación de que existe un uso ambiguo de ‘poder causal’ en el argumento individualista, concluíamos que ésta es incorrecta, porque: (i) ‘poder causal’ y ‘lo que afecta a los poderes causales’ no son la misma cosa, siendo este último término el que admite propiedades relacionales (o históricas) como taxonómicamente válidas; (ii) la exigencia de superveniencia en la microestructura local se aplica sólo en el caso de poder causal —no en el de lo que afecta a los poderes causales; (iii) si lo anterior es verdadero, no hay contradicción en la afirmación individualista, porque no hay ambigüedad respecto de ‘poder causal’.

Una consideración adicional nos sugería que al menos parte de la discusión sobre el argumento de los poderes causales exige evidenciar los compromisos que se tienen con una u otra concepción de explicación causal en ciencia, lo que, en el caso de Fodor, implica un modelo del tipo nomológico-deductivo, con subsunción de los acontecimientos especificados en las generalizaciones de la ciencia en cuestión bajo leyes; para el caso de la psicología, leyes intencionales. Esto importa, porque el individualismo no exige en modo alguno que todas las taxonomías se hagan por, y sólo por, poderes causales, limitándose a exigirlo únicamente en aquellos casos en que las ciencias en cuestión se hayan dado como tarea el proporcionar explicaciones causales de los fenómenos de los que se ocupan, para lo cual es imprescindible comprometerse con alguna concepción explicativa causal.

La segunda línea argumentativa que revisamos se presentaba como un supuesto argumento desde el computacionalismo al individualismo, que era cuestionado intentando demostrar la posibilidad (lógica) y plausibilidad del computacionalismo amplio, concepción computacional que concibe estados computacionalmente relevantes para la psicología que no supervienen localmente en el individuo que los instancia. Una vez más, Wilson presenta los elementos básicos para su análisis en la forma de un argumento: el argumento computacional [AC].

Habiendo explicitado las premisas de AC y desarrollado los cuestionamientos de Wilson arribamos, al concluir su examen, a la idea de que la primera dificultad se situaba en torno a la misma construcción de AC, en particular de su primera premisa, que evaluamos como carente de bases individualistas para su forma omni-abarcadora. La discusión sobre la posibilidad del computacionalismo nos parece de escaso interés, en tanto consulta sólo la mera posibilidad. Por el contrario, mucho más sustantivo nos resultó discutir la eventual plausibilidad del computacionalismo amplio a partir del examen de prácticas psicológicas reales. Habiendo evaluado los casos que Wilson pone en análisis, correspondientes a diversos campos de investigación psicológica, concluimos que tales casos no apoyan de modo contundente sus afirmaciones, en razón de que básicamente giran sobre la posibilidad de revisarlos para hacer de ellos casos efectivos de operación de sistemas computacionales amplios: Como cuestión de hecho, éstos no constituyen casos efectivos de computacionalismo amplio, sino que podrían serlo, con lo que volvemos a la cuestión de la mera posibilidad.

Finalmente, revisamos una tercera opción, una supuesta defensa del individualismo sobre bases metodológicas, formulado como dos argumentos: un argumento sobre la incoherencia de la psicología naturalista, y otro desde la modularidad al individualismo. La contra-argumentación de Wilson se orientó a mostrar la falsedad de ambas líneas de argumentación.

Tras revisar los elementos expuestos por Wilson, concluíamos que el putativo carácter apremiante afirmado por Wilson no puede considerarse más que un efecto retórico del modo en que aquellos son presentados, que magnifica su capacidad para imposibilitar algún tipo de práctica concreta en psicología. Tal sobredimensión, se nos hacía particularmente evidente con respecto al argumento desde la modularidad al individualismo, puesto que ningún éxito pasado o futuro que se base en el supuesto de la modularidad podrá nunca bastar para afirmar que éste agota toda otra alternativa; la naturaleza computacional de los módulos no entraña, no hay forma en que pudiese hacerlo, individualismo. Para el caso del argumento de incoherencia metodológica de la psicología naturalista, el individualista no cuestiona la existencia de taxonomías amplias, limitándose a poner en entredicho su esperanza práctica de disponer de las descripciones semánticas requeridas por sus generalizaciones. Nuevamente, esto no puede

suponer que el individualista crea que se impide algún tipo de práctica particular de los científicos, tal omnipotencia está más allá —quizá muchísimo más allá— de la propuesta individualista. De cualquier forma, el análisis de los argumentos metodológicos no muestra tampoco incoherencia o inconsistencia en el individualismo.

¿Qué podemos decir tras este recorrido? Una primera impresión general es que los contra-argumentos ofrecidos por Wilson prometían más de lo que realmente pueden dar; ninguno de ellos es lo bastante contundente como para producir algún daño de consideración. Una posible razón de por qué esto es así, podría corresponder al hecho de que sólo desde la retórica wilsoniana las construcciones individualistas tienen el carácter totalitario y omni-abarcador que éstas le asignan. El individualismo aparece como una concepción bastante menos ambiciosa, aunque consistente, como espero haya podido apreciarse. Sin embargo, no es neutral respecto de cuestiones ontológicas y metodológicas. En eso radica, como para cualquier concepción alternativa, su posibilidad de ser sometida a análisis y crítica: se pueden compartir o no sus compromisos y, desde ahí, adoptar o no sus consecuencias.

Pero, si la argumentación de Wilson se propone una meta más modesta, esto es, evidenciar que el individualismo no puede copar y agotar el campo de visiones posibles sobre la taxonomía científica, el computacionalismo y las prácticas explicativas en psicología, despejando un lugar para alternativas, quizá no haya merecido el trabajo que Wilson se dio, por cuanto el propio individualismo implica desde siempre tal espacio, delimitado como he dicho por los compromisos que es necesario asumir cuando se es individualista, compromisos que no son neutrales y que, por cierto, no son ni los únicos ni necesariamente los correctos. El individualismo supone una opción, el individualista toma partido y asume las ventajas y desventajas del caso.

¿Qué ofrece Wilson como alternativa? Dado que finalmente acepta la validez parcial del individualismo o la imposibilidad de su refutación total, la oferta es pluralidad taxonómica, explicativa y metodológica en ciencia y, para el caso, en psicología. La heterogeneidad de la práctica taxonómica en ciencia haría improbable la existencia de una verdad general sobre el

tema, pero al rechazar el argumento de los poderes causales no se rechaza la idea de que, en algunos casos, sólo los poderes causales intrínsecos de una entidad se usan al taxonomizar esa entidad. Similarmente, al rechazar el argumento computacional para el individualismo no se sugiere que no hay capacidades cognitivas mejor explicadas como productos de sistemas computacionales estrechos. No se habría tratado de hacer un caso general para la explicación computacional amplia por sobre otras formas de explicación en psicología; la explicación computacional estrecha tiene un lugar en la psicología, correcto, pero no lo es todo. Del mismo modo, Wilson promueve una psicología metodológicamente pluralista; esto es, una que emplee tanto la perspectiva anidada cuanto la estrecha al desarrollar sus clases taxonómicas y sus explicaciones. Emplear sólo la perspectiva estrecha sería confundir éxitos explicativos parciales con una receta para el éxito explicativo; recíprocamente, emplear sólo la perspectiva anidada sería fracasar en reconocer algunos de los éxitos parciales que ha habido ya en la psicología. Lo contrario sería negar una simple pero importante verdad en la filosofía de la ciencia, a saber, ‘que *no hay recetas para el éxito explicativo*’ (Wilson 1995:113).

Todo lo anterior no ha sido jamás imposibilitado por el individualismo, no podría serlo; al contrario, como opción, el individualismo supone desde el inicio que no lo es todo, lo cual no constituye sino una razón adicional para seguir sosteniendo el individualismo en psicología.

Referencias.

- Bach, K. 1982: 'De Re Belief and Methodological Solipsism', en Woodfield 1982.
- Bennet, J. 1964: *Rationality*. London: Routledge and Kegan Paul; 1989: Indianapolis, IN: Hackett.
- Block, N. 1978: 'Troubles with Functionalism', en Savage. Reimpreso en Block 1980.
- Block, N. (Ed.) 1980: *Readings in the Philosophy of Psychology. Vol 1*. Cambridge, MA: Harvard University Press.
- Block, N. 1986: 'Advertisement for a Semantics for Psychology', en *Midwest Studies in Philosophy* 10, pp. 615-78.
- Block, N. y Fodor, J. A. 1972: 'What psychological States Are Not', en *Philosophical Review* 81, pp. 159-81.
- Burge, T. 1979: 'Individualism and the Mental', en *Midwest Studies in Philosophy* 5, pp. 73-122.
- Burge, T. 1986: 'Individualism and Psychology', en *Philosophical Review* 95, pp. 3-46.
- Burge, T. 1986: 'Intellectual Norms and Foundations of Mind', en *the Journal of Philosophy* 83, 12, pp. 697-720.
- Burge, T. 1988: 'Individualism and Self-Knowledge', en *Journal of Philosophy* 85, 11, pp. 649-63.
- Campbell, F. W. y Robson, J. G. (1968): 'Application of Fourier Analysis to the Visibility of Gratings', *Journal of Physiology* 197:151-66.
- Carey, S. 1985: *Conceptual Change in Childhood*, Cambridge, MA: MIT Press.
- Chomsky, N. 1957: *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. 1965: *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. 1980: *Rules and Representation*. New York: Columbia University Press.
- Clark, A. 1989: *Microcognition*. Cambridge, MA: MIT Press.
- Clark, A. 1993: *Associative Engines*. Cambridge, MA: MIT Press.
- Cummins, R. C. 1983: *The Nature of Psychological Explanation*. Cambridge, MA: MIT Press.
- Cummins, R. C. 1989: *Meaning and Mental Representation*. Cambridge, MA: MIT Press.

- Davies, M. 1986: 'Externality, Psychological Explanation and Narrow Content, Reply to Jerry Fodor's "Individualism and Supervenience"', trabajo enviado a la Joint Session of the Aristotelian Society and the Mind Association.
- Dennett, D.C. 1982: 'Beyond Belief', en *Thought and Object*, A. Woodfield (ed.), Clarendon Press, Oxford. Reimpreso en *The Intentional Stance*, Bradford/MIT Press, Cambridge, Mass., 1987.
- Dennett, D.C. 1983: 'Intentional Systems in Cognitive Ethology: The "Panglossian Paradigm" Defended', en *Behavioral and Brain Sciences* 6, pp. 343-90. Reimpreso en *The Intentional Stance*, Bradford/MIT Press, Cambridge, Mass., 1987.
- Devitt, M. 1990: 'A Narrow Representational Theory of the Mind', en *Mind and Cognition*, William Lycan (ed.), Oxford Basil Blackwell, 1990, pp. 371-398.
- Devitt, M. 1991: 'Why Fodor Can't Have It Both Ways', en *Meaning in Mind: Fodor and His Critics*, B. Loewer y G. Rey (ed.), Blackwell, 1991, pp. 95-118.
- Dretske, F. 1986: 'Misrepresentation', en *Belief*, R. Bogdan (ed.), OUP, Oxford.
- Dretske, F. 1988: *Explaining Behavior: Reasons in a World of Causes*, Bradford/MIT Press, Cambridge, Mass.
- Egan, M. F. 1992: 'Individualism, Computation and Perceptual Content', *Mind* 101:443-59.
- Fodor, J. 1968: *Psychological Explanation*, Random House, New York.
- Fodor, J. 1974: 'Special Sciences, or The Disunity of Science as a Working Hypothesis', en *Synthese* 28, pp.77-115. Kluwer Academic Publishers.
- Fodor, J. 1975: *The Language of Thought*. New York: Thomas Y. Crowell.
- Fodor, J. 1978: 'Propositional Attitudes', en *The Monist* 61, pp. 501-23. Reimpreso en Fodor 1981.
- Fodor, J. 1980: 'Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology', en *Behavioral and Brain Sciences* 3, pp.63-110. Reimpreso en Fodor 1981 y Rosenthal 1991.
- Fodor, J. 1981: *Representations*, Bradford/MIT Press, Cambridge, Mass.
- Fodor, J. 1984: 'Semantics, Wisconsin Style', en *Synthese* 59, pp. 1-20. Reimpreso en *A Theory of Content I & II*, Bradford/MIT Press, Cambridge, Mass., 1990.

- Fodor, J. 1985: 'Fodor's Guide to Mental Representation', en *Mind* 94, pp. 76-100. Reimpreso en *A Theory of Content I & II*, Bradford/MIT Press, Cambridge, Mass., 1990.
- Fodor, J. 1985: 'Banish Dis-Content', en *Mind and Cognition*, William Lycan (ed.), Oxford Basil Blackwell, 199, pp. 420-38.
- Fodor, J. 1987: *Psychosemantics*, Bradford/MIT Press, Cambridge, Mass. Traducción española de Oscar González-Castán, Madrid: Editorial Tecnos, 1994.
- Fodor, J. 1990: *A Theory of Content I & II*, Bradford/MIT Press, Cambridge, Mass.
- Fodor, J. 1991a: 'A Modal Argument for Narrow Content', en *Journal of Philosophy* 88, pp. 5-25. Reimpreso en Macdonald y Macdonald 1995.
- Fodor, J. 1991b: *Meaning in Mind. Fodor and His Critics*, Cambridge, MA: Basil Blackwell.
- Fodor, J. y Lepore, E. 1992: *Holism: A Shopper's Guide*, Cambridge, MA: Basil Blackwell.
- Fodor, J. 1994: *The Elm and The Expert. Mentalese and its Semantics*, Cambridge, MA: MIT Press. Traducción de Marco Aurelio Galmarini, Barcelona: Piados Ibérica, 1997.
- Fodor, J. 2000: *The Mind Doesn't Work That Way*, Cambridge, MA: Bradford Books/MIT Press.
- Gallistel, C. R. 1989a: 'Animal Cognition : The Representation of Space, Time and Number', *Psychology Annual Reviews* 40:155-89.
- Gallistel, C. R. 1989b: *The Organization of Learning*. Cambridge, MA: MIT Press.
- Gauker, C. 1991: 'Mental Content and the Division of Epistemic Labour', *Australasian Journal of Philosophy* 69:302-18.
- Gibson, J. J. 1979: *The Ecological Approach to Perception*. Boston: Houghton Mifflin.
- Hardwig, J. 1985: 'Epistemic Dependence', *Journal of Philosophy* 82:335-49.
- Harman, G. 1987: '(Nonsolipsistic) Conceptual Role Semantics', en E. Lepore (Ed.) *New Directions in Semantics*. London: Academic Press.
- Harman, G. 1988: 'Wide Functionalism', en S. Schiffer y D. Steele (Eds.), *Cognition and Representation*. Boulder, CO: Westview Press.
- Haugeland, J. (Ed.) 1981: *Mind Design*. Cambridge, MA: MIT Press.
- Haugeland, J. 1985: *Artificial Intelligence. The Very Idea*, Cambridge, MA: MIT Press. Traducción al español de Irene Tulli de Firmani, México D. F.: Siglo XXI, 1988.

- Hilbert, D. 1899: 'Grundlagen der Geometrie', en *Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmal in Göttingen*. Leipzig.
- Horst, S. 1996: *Symbols, Computation, and Intentionality*. Berkeley: University of California Press.
- Keil, F. 1989: *Concepts, Kinds, and Cognitive Development*, Cambridge, MA: MIT Press.
- Kim, J 1984: 'Supervenience and Supervenient Causation', en *The Southern Journal of Philosophy*, vol XXII Supplement, pp. 45-56.
- Kim, J 1984(?): 'Concepts of Supervenience', en *Philosophy and Phenomenological Research* 45, pp. 153-176.
- Kitcher, P. 1985: 'In Defence of Folk Psychology', *Journal of Philosophy* 81:89-106.
- Leslie, A. 1987: 'Pretense and Representation: The Origins of "Theory of Mind"', *Psychological Review* 94:412-26.
- Lycan, W. (Ed.) 1990: *Mind and Cognition*, Oxford: Basil Blackwell.
- Macdonald, C. Y Macdonald, G. 1995 (Eds.): *Philosophy of Psychology*, Blackwell.
- Marr, D. 1982: *Vision*, San Francisco: Freeman.
- McGinn, C. 1982: 'The Structure of Content', en Woodfield 1982.
- McGinn, C. 1989: *Mental Content*, Cambridge, MA: Basil Blackwell.
- Newell, A. y Simon, H. 1976: 'Computer Science as Empirical Inquiry: Symbols and Search', *Communications of the Association for Computing Machinery* 19:113-26. Reimpreso en Haugeland 1981.
- Putnam, H. 1973: 'Meaning and Reference', en *Journal of Philosophy* 70, pp. 699-711.
- Putnam, H. 1975: 'The Meaning of "Meaning"', en *Minnesota Studies in the Philosophy of Science*, vol. 7, K. Gunderson (ed.), University of Minnesota Press, Minneapolis. Reimpreso en *Mind, Language and Reality: Philosophical Papers Volume 2*, CUP, Cambridge, 1975.
- Pylyshyn, Z. 1980: 'Computation and Cognition: Issues in the Foundation of Cognitive Science'. *Behavioral and Brain Sciences* 3:111-132.
- Pylyshyn, Z. 1984: *Computation and Cognition: Toward a Foundation of Cognitive Science*. Cambridge, MA: MIT Press.
- Rosenthal, D. 1991 (Ed.): *Nature of Mind*, Oxford University Press.

- Searle, J. 1980: 'Minds, Brains and Programs', *Behavioral and Brain Sciences* 3:417-24.
Reimpreso en Haugeland 1981.
- Segal, G. 1989: 'Seeing What Is Not There', *Philosophical Review* 98:189-214.
- Segal, G. 1991: 'Defence of a Reasonable Individualism', *Mind* 100:485-94.
- Spelke, E. 1990: 'Principles of Object Perception', *Cognitive Science* 14:29-56.
- Stalnaker, R. C. 1989: 'On What's in the Head', en Tomberlin 1989. Reimpreso en Rosenthal 1991.
- Sterelny, K. 1990: *The representational Theory of Mind*. Oxford: Basil Blackwell.
- Stich, S. 1978: 'Autonomous Psychology and the Belief-Desire Thesis', en *The Monist* 61, pp. 573-91. Reimpreso en Rosenthal 1991.
- Stich, S. 1983: *From Folk Psychology to Cognitive Science*, Bradford/MIT Press, Cambridge, Mass.
- Tomberlin, J. 1978 (Ed): *Philosophical Perspectives*, Vol. 3: *The Philosophy of Mind and Action*. Atascadero, CA: Ridgeview.
- Vergara, P. 1999: 'Individualismo, Contenido Estrecho y Teoría Computacional Representacional de la Mente', *Anuario de Postgrado N°3*, Escuela de Postgrado de la Facultad de Filosofía y Humanidades de la Universidad de Chile, pp. 191-203.
- Wilson, R. 1995: *Cartesian Psychology and Physical Minds: Individualism and the Sciences of the Mind*, Cambridge University Press.
- Woodfield, A. (Ed.) 1982: *Thought and Object: Essays on Intentionality*. Oxford: Oxford University Press.