

**FUNDAMENTOS FILOSÓFICOS Y DETERMINANTES HISTÓRICOS DEL  
COGNITIVISMO, UN MODELO EN CIENCIAS COGNITIVAS**

**DANIEL ARBOLEDA ROJAS**

**Trabajo de grado para optar al título de Psicólogo**

**MAGÍSTER EN PSICOLOGÍA**

**JOHNY VILLADA ZAPATA**

**ASESOR ACADÉMICO**

**UNIVERSIDAD DE ANTIOQUIA**

**FACULTAD DE CIENCIAS SOCIALES Y HUMANAS**

**DEPARTAMENTO DE PSICOLOGÍA**

**MEDELLÍN**

**2019**

## **Resumen**

En este trabajo se realiza una revisión teórica sobre el cognitivismo, modelo que explica la cognición como el producto de procedimientos lógicos, algorítmicos, sobre representaciones mentales. Se analizan sus antecedentes, sus premisas básicas y algunas de las críticas que se le han hecho. Específicamente, se revisan algunas de las propuestas teóricas previas asociadas a la búsqueda de objetividad en psicología. Dichas propuestas son la psicología fisiológica de Wundt y el conductismo. Además, se desarrollan los argumentos esenciales del cognitivismo, a saber: el carácter representacional de la cognición y su interpretación lógico-formal. Finalmente, se exponen las críticas hecha al modelo cognitivista, especialmente aquellas que destacan la naturaleza semántica y contextual de la cognición humana. Toda esta revisión es hecha teniendo como elemento de análisis constante el problema del dualismo de sustancias, principal obstáculo en psicología para alcanzar el estatus científico.

**Palabras clave:** cognitivismo, funcionalismo, ciencias cognitivas, metáfora del ordenador, dualismo.

## **Abstract**

In this work, a theoretical review is carried out on cognitivism, a model that explains cognition as the product of logical, algorithmic procedures, on mental representations. Its antecedents, its basic premises and some of the criticisms that have been made are analysed. Specifically, some of the previous theoretical proposals associated with the search for objectivity in psychology are reviewed. These proposals are Wundt's physiological psychology and behaviorism. In addition, the essential arguments of cognitivism are developed, namely: the representational character of cognition and its logical-formal interpretation. Finally, critiques of the cognitivist model are presented, especially those that highlight the semantic and contextual nature of human cognition. All this revision is made having as an element of constant analysis the problem of dualism of substances, main obstacle in psychology to reach the scientific status.

**Keywords:** cognitivism, functionalism, cognitive sciences, computer metaphor, dualism.

## Contenido

Introducción.....	8
Capítulo 1: De Wundt a la computación simbólica.....	12
1. Objeto y método de la Psicología según Wundt.....	13
1.1. Aspectos históricos .....	13
1.2. Psicología Fisiológica.....	14
1.2.1. Objeto .....	14
1.2.2. Método.....	16
1.3. Problemas y limitaciones .....	20
2. Conductismo .....	23
2.1. Conductismo como programa .....	23
2.2. Skinner y el condicionamiento operante .....	25
2.3. Condicionamiento operante.....	27
2.4. Propósito, intención y errores categoriales .....	28
2.5. Críticas al conductismo.....	34
3. Cognitivismo .....	38
3.1. Principios generales.....	38
Capítulo 2: Antecedentes, principios y problemas del cognitivismo .....	44
1. Cognición .....	45
2. Funcionalismo.....	47
2.1. Argumento de las múltiples realizaciones .....	49
2.2. Funcionalismo y materialismo .....	51
3. De Turing a Putnam y de Putnam al Cognitivismo .....	55
4. El sujeto cognitivo .....	61
4.1. Sujeto natural-representacional.....	62
4.2. Sujeto abstracto-computacional.....	63
4.3. Sujeto computacional-representacional .....	66
5. Representación .....	68
5.1. El problema de la representación .....	70
5.2. Teoría de la causalidad o indicación.....	73
1. El conexionismo .....	80
2. Deducción y razonamiento semántico .....	89
3. Significado y Psicología popular .....	94

<b>Conclusiones</b> .....	100
<b>Referencias:</b> .....	104

## Índice de tablas

<b>Tabla 1. Tabla de instrucción para máquina de bebidas calientes: .....</b>	<b>53</b>
<b>Tabla 2. Tabla de instrucciones de una proposición: .....</b>	<b>55</b>

## Índice de figuras

<b>Figura 1. Algoritmo para hervir un huevo: .....</b>	<b>78</b>
<b>Figura 2. Conexión simple de nodos: .....</b>	<b>82</b>
<b>Figura 3. Red neuronal compleja: .....</b>	<b>82</b>
<b>Figura 4. Ejemplificación de modelo mental: .....</b>	<b>88</b>

## Introducción

Históricamente se han reconocido, sin mayores discrepancias, las Jornadas sobre Teoría de la Información realizadas en el MIT (Massachusetts Institute of Technology) en 1956 como el evento fundacional de aquella comunión interdisciplinaria que posteriormente se conocería bajo la denominación de *ciencias cognitivas* (Gardner, 2000).

La constitución de este campo interdisciplinario fue un producto directo de las desavenencias epistemológicas que, por aquellos días, se imponían ante los modelos vigentes de comprensión de la cognición. A saber, se trató de “una creciente conciencia, por parte de muchos científicos sensatos, de que la adhesión a los cánones conductistas estaba volviendo imposible el estudio científico de la mente” (Rivière, 1987, p. 28). Esto debido a la negación conductista de aquella ‘mente’ de la “psicología intencionalista de creencias-deseos” como un objeto susceptible de ser conocido bajo parámetros científicos (Rivière, 1991, p. 131).

Si bien al interior de las ciencias cognitivas cohabitan, y “no siempre de buena gana” (Rivière, 1991, p. 140), modelos explicativos parcial o completamente distintos, sí existen, por lo menos, dos rasgos fundamentales. El primero de ellos tiene que ver con la recuperación epistemológica, esto es, como objeto de conocimiento, de la mente representacional, o en otros términos, intencional. El segundo consiste en la aceptación de las computadoras como modelos de comprensión del sujeto cognitivo (Gardner, 2000).

Bächler (2014) señala que “en ciencia cognitiva existen tres enfoques utilizados para la comprensión de la mente —cognitivismo, conexionismo y cognición corporizada—”(p. 50). Además, añade que es el cognitivismo “el enfoque rector y guía de la investigación en este campo de conocimientos” (Bächler, 2014, p. 50). En atención a los propósitos de esta

investigación, se presentarán exclusivamente los aspectos fundamentales del enfoque cognitivista.

El cognitivismo, también llamado paradigma simbólico-computacional (Rivière, 1991) o perspectiva computacional clásica (Colom, 1993), posee dos premisas fundamentales: la primera destaca el carácter representacional de los estados mentales y la segunda la naturaleza computacional de los mismos (Restrepo, 2009).

Naturalmente, esta propuesta ha contado con numerosos simpatizantes, así como también ha sido víctima de múltiples críticas. Existen distintos trabajos en donde son presentadas tanto las premisas del modelo cognitivista, como también su críticas y replanteamientos. Uno de ellos, fundamental en la literatura cognitiva, es el ya citado libro de Gardner (2000) *La nueva ciencia de la mente*. Otro trabajo en donde es analizado este modelo es el texto *Conocer. Las ciencias cognitivas: tendencias y perspectivas. Cartografía de las ideas actuales*, publicado en 1990 por Francisco J. Varela. Además de estos, existen muchos más, como por ejemplo los de Rabossi (1995) y Bechtel (1988) en donde es abordado el cognitivismo desde la perspectiva de la filosofía de la mente.

Con lo anterior se pretende demostrar que en la literatura el cognitivismo no ha pasado desapercibido. En realidad, desde su aparición hasta la actualidad, ha tenido bastante atención. No obstante, los trabajos que se han ocupado de presentarlo, como los citados en el párrafo anterior, usualmente han tendido a concentrarse en aspectos colindantes al cognitivismo.

En concreto, el texto de Gardner (2000) no se concentra exclusivamente en el cognitivismo. Su visión es bastante amplia, abordando asuntos asociados entre sí en el contexto de la ciencia cognitiva, pero con cierta independencia conceptual, como lo son la

lingüística, antropología, neurociencia, etc. A su vez, Varela (1990) centra su interés en la enacción o cognición corporizada, y tanto Rabossi (1995) como Bechtel (1988) hacen énfasis en aspectos primordialmente filosóficos.

Así pues, si bien existe un centenar de bibliografía asociada al cognitivismo, una de las particularidades del siguiente trabajo consistirá en enfocarse estrictamente en dicho modelo, sin servirse de él para introducir otros aspectos o campos que sobrepasan sus límites de competencia directa.

En concreto, durante este trabajo se intentará alcanzar un objetivo claro: examinar el llamado modelo cognitivista o cognitivismo, algunos de sus antecedentes teóricos, sus premisas elementales y las objeciones que se le han hecho. Además de esto, se hará énfasis recurrentemente en dos puntos fundamentales para la comprensión del cognitivismo: el intento de alcanzar un carácter admisiblemente científico para la psicología y la superación del dualismo de sustancias como principal obstáculo para hacerlo. Para lograr esto, el trabajo estará dividido en tres capítulos, uno para cada elemento enunciado.

De esta manera, en el primer capítulo serán presentados algunos aspectos históricos asociados al cognitivismo, específicamente los principales intentos de científicidad en psicología, a saber: la psicología fisiológica de Wundt y el conductismo. En el segundo capítulo se abordarán las características básicas del cognitivismo, haciendo énfasis en su visión lógico-representacional de la cognición. Finalmente, en el tercer capítulo, se expondrán algunas de las críticas hechas al cognitivismo y las nuevas propuestas explicativas de la cognición derivadas de dichas críticas.

Como ya fue mencionado, en ciencias cognitivas, actualmente el enfoque cognitivista es el modelo de comprensión rector (Bächler, 2014). En esta medida, toda reflexión que se

proponga examinar las premisas básicas de este paradigma, bien sea teniendo como objetivo su refutación, afirmación o descripción, contribuirá a la discusión actual relativa al núcleo conceptual de las ciencias cognitivas.

El principal mérito de este trabajo será actuar como presentación general de un tema determinante para la consolidación actual de la psicología. En efecto, aunque se trate de un tema ampliamente criticado, está asociado a nociones tan nucleares en la psicología contemporánea como las de *procesamiento de la información* o *representación*, por mencionar solo dos. Así, la comprensión del significado histórico-filosófico de dichos conceptos, por ejemplo, permitirá al lector tener una mayor conciencia de las características actuales de la psicología. Este hecho convierte el siguiente trabajo, plenamente monográfico, en un esfuerzo justificable.

## **Capítulo 1:**

### **De Wundt a la computación simbólica**

Durante las siguientes páginas presentaré, en sentido amplio, el trayecto que ha atravesado la psicología en su búsqueda de objetividad y rigurosidad científica. Elegí específicamente tres propuestas, hechas en distintas circunstancias históricas, porque se solapan entre sí y permiten dar cuenta de una especie de continuidad en el trayecto señalado. Estas propuestas son, en orden cronológico, la psicología fisiológica de Wundt, el conductismo y el cognitivismo.

En lo relativo a la propuesta hecha por Wundt, destacaré principalmente sus esfuerzos por hacer de la observación interna un método científicamente válido y semejante a la observación externa. Así mismo, describiré cuál es el objeto de estudio que a su juicio deberá tener la psicología. Finalmente, comentaré la forma en que Wundt enfrenta el problema del dualismo de sustancias para acabar mencionando algunas de las limitaciones de su propuesta.

En cuanto al conductismo, comenzaré presentando sus aspiraciones principales como filosofía y modelo de la psicología. Luego describiré en términos generales la propuesta hecha por Skinner, concentrándome particularmente en la redefinición de los estados mentales como conductas operantes, soportada en la concepción de aquellos como errores categoriales. Por último, mencionaré las principales críticas hechas al conductismo.

Para finalizar, señalaré los aspectos esenciales del modelo cognitivista, haciendo especial énfasis en la conceptualización de la cognición como computación simbólica característica de este paradigma. Este apartado será tan sólo un abrebocas para el siguiente capítulo; en éste presentaré con mayor detenimiento y atención los supuestos nucleares del Cognitivismo.

## **1. Objeto y método de la Psicología según Wundt**

### **1.1. Aspectos históricos**

Iniciaré este apartado dando cuenta de algunos aspectos biográficos generales de Wundt. Particularmente, me interesa señalar dos puntos: uno, relativo a su formación académica, y otro asociado a su actividad profesional. Considero estos puntos fundamentales ya que determinaron en gran medida su propuesta.

Ante todo, debe ser señalado que la formación de Wundt fue en Medicina en la Universidad de Heidelberg (Tortosa & Civera, 2006). En 1856 obtuvo su doctorado y al año siguiente comenzó a trabajar como docente ayudante en el laboratorio de Fisiología experimental de la misma universidad. Allí trabajó hasta 1864 bajo la supervisión de Helmholtz, fecha en la que este se traslada a Berlín (Tortosa & Civera, 2006). En Heidelberg, Wundt permaneció como docente hasta 1874. Durante este período publicó algunas de sus obras más importantes, entre ellas los *Elementos de Psicología fisiológica* (1873-1874) “en la que presentaba la primera versión de su sistema teórico, definiendo el objeto y método de una nueva disciplina científica que bautizó con el nombre de Psicología fisiológica” (Tortosa & Civera, 2006, p. 82). Debido a todo esto, no es atrevido afirmar que los antecedentes formativos y profesionales de Wundt influenciaron la propuesta que desarrollaría durante años.

En 1875, luego de haber obtenido una cátedra en la Universidad de Leipzig, fundó allí un “pequeño laboratorio para su propio trabajo de investigación y para demostraciones prácticas” (Tortosa & Civera, 2006, p. 82). No obstante, es sólo hasta 1883 que “la Universidad reconoció aquel Laboratorio como Instituto de Psicología Experimental y lo financió” (Tortosa & Civera, 2006, p. 82). Este laboratorio pasaría a ser reconocido históricamente como el primer laboratorio de psicología.

Además de la fundación de este laboratorio, en 1881 Wundt creó una revista que contenía, justamente, publicaciones relativas a las investigaciones hechas en su laboratorio (Martínez, 2007). Esta revista tenía por nombre *Philosophische Studien* y fue la primera revista de divulgación sobre psicología (Danziger, 1980).

Ahora, si bien antes de Wundt existieron algunos autores cuya aspiración era dar un trato científico a problemas psicológicos (Fechner y Helmholtz, por ejemplo) (Gardner, 2000), es a él a quien se reconoce como fundador de la psicología, y aún más, de la psicología experimental. Esto se debe a que la constitución de una ciencia no es sólo un esfuerzo académico-epistemológico, sino además un esfuerzo institucional (Gardner, 2000). La fundación de un laboratorio dedicado exclusivamente a la investigación experimental en psicología, y la creación de una revista dedicada a la divulgación de los productos de la investigación misma, son justamente esfuerzos institucionales.

## **1.2. Psicología Fisiológica**

### **1.2.1. Objeto**

La Psicología, en el sentido en que Wundt la pensó, debe ocuparse de la experiencia inmediata (Quiñones & Ato, 1983). Wundt, señalan Quiñones y Ato (1983), “(...) divide la experiencia en mediata (experiencia utilizada como medio para conocer algo distinto de la experiencia misma) e inmediata, que se refiere a la experiencia misma y es objeto de la psicología” (p. 134). Este tipo de experiencia, en oposición a la experiencia mediata, se caracteriza por no prescindir del sujeto cognoscente (Tortosa & Civera, 2006). De hecho, como veremos, la experiencia inmediata es justamente la experiencia del sujeto que conoce en tanto tal.

La división entre experiencia inmediata y experiencia mediata es producto de la descomposición de toda experiencia en dos elementos: contenido y conocimiento del

contenido, o lo que es igual, objeto de la experiencia y sujeto cognoscente (Wundt, 1902).

La consecuencia de esta escisión de la experiencia, según Wundt (1902), es la siguiente:

De aquí dos caminos que se abren para el estudio de la experiencia: uno es el de la ciencia natural, que considera los objetos de la experiencia en su naturaleza, pensada independientemente del sujeto; el otro es el de la psicológica, por el cual se marcha a la investigación del contenido total de la experiencia, en su relación con el sujeto y de las cualidades que éste atribuye inmediatamente a dicho contenido. (p. 2)

Es conveniente destacar, en relación a lo anterior, que la división de la experiencia en inmediata y mediata no significa que existan dos tipos de experiencias sustancialmente diferentes. Esta confusión se presenta con mayor facilidad cuando los conceptos inmediato y mediato son reemplazados, respectivamente, por interno y externo. Sin embargo, “las expresiones experiencia interna y experiencia externa, no indican dos cosas diferentes, sino solamente dos diversos puntos de vista que usamos en el conocimiento y en la exposición científica de la experiencia en sí única” (Wundt, 1902, p. 1).

En este sentido, que la psicología y las ciencias naturales no se ocupen de objetos fundamentalmente distintos, sino de puntos de vista diferentes de la experiencia, que es una sola, implica que se pueden complementar entre sí. Al respecto señala Wundt (1902):

Si la psicología debe, en la conexión de los procesos psíquicos, en cuanto son datos inmediatos de la experiencia, evitar la ayuda de hipótesis metafísicas, puede con todo—puesto que experiencia interna y externa son dos puntos de vista que se completan mutuamente de una sola e idéntica experiencia—, volver sobre todo donde la conexión de los fenómenos psíquicos presenta vacíos, a considerar físicamente los mismos procesos, para ver si, mediante este nuevo

punto de vista de la ciencia natural, se puede restablecer la continuidad que se creía faltaba. Lo propio cabe decir, aunque en sentido inverso, respecto de los vacíos que se presentan en la cadena de nuestros conocimientos fisiológicos, que puede completarse con anillos suministrados por una exposición de la experiencia desde el punto de vista puramente psicológico. (p. 5)

Esta especie de maridaje entre la psicología y las ciencias naturales, particularmente la fisiología, no se reduce exclusivamente a una consecuencia directa de que ambas tengan por objeto la experiencia. Esta comunión también tiene consecuencias puramente metodológicas, como lo indicaré más adelante.

Hasta ahora he dicho que, para Wundt, “la psicología tiene, (...) como objeto propio, el total contenido de la conciencia en su constitución inmediata” (Wundt, 1902, p. 2). En palabras de Gardner (2000) “según Wundt (...) la psicología consiste en el estudio de la experiencia consciente como tal” (p. 120). Me parece conveniente recurrir a estos pasajes en este punto dado que el hecho de que se llame experiencia consciente a lo que hasta ahora hemos llamado experiencia inmediata permite establecer la transición entre objeto y método en la psicología de Wundt.

### **1.2.2. Método**

Metodológicamente, las ciencias naturales se fundamentan sobre el experimento y la observación (Wundt, 1902). El experimento “consiste en una observación en la cual los fenómenos observables surgen y se desarrollan por la acción voluntaria del observador” (Wundt, 1902, p. 9). Por otro lado, “la observación, en sentido estricto, estudia los fenómenos sin semejante intervención, tal como se presentan al observador en la continuidad de la experiencia” (Wundt, 1902, p. 9).

Si la experiencia puede dividirse en inmediata y mediata (interna y externa), la observación de la experiencia también es susceptible de ser dividida de la misma forma. En este sentido, la observación externa se corresponde ante todo con la experiencia mediata (la experiencia del contenido de la experiencia), en tanto existe una independencia entre sujeto y objeto (Danziger, 1980). De esto se sigue que la observación interna se asocie a la experiencia inmediata. Ahora, la separación entre el sujeto y el objeto, característica de la observación externa, es difícil de reproducir en la observación interna (Danziger, 1980). Lo anterior dado que justamente la experiencia inmediata es la experiencia de la consciencia, en donde sujeto y objeto fácilmente se mezclan entre sí (Danziger, 1980). En palabras de Wundt (1902):

Consistiendo uno de los principales objetos de la psicología en la exacta investigación del modo de surgir y de desarrollarse de los procesos subjetivos, es fácil comprender cómo, en este punto, la intención del observador altera sustancialmente los hechos observables ó ella misma se suprime en todo. (p. 10)

Tradicionalmente, se ha nombrado el método usado por Wundt como *introspección*, no obstante, el verdadero método desarrollado y promovido por él es la observación interna (Danziger, 1980). ¿Qué diferencias existen entre ambos métodos? Antes que nada, habrá que señalar que la introspección en sentido estricto también puede ser llamada autoobservación (Danziger, 1980). Ahora, la autoobservación, en principio, se invalida metodológicamente en tanto “la observación científica exige la clase de independencia entre sujeto y objeto que la introspección (autoobservación) justamente destruye” (Danziger, 1980, p. 6). Por su parte, la observación interna, como de alguna forma ya fue señalado, también puede invalidarse con facilidad en el mismo sentido, convirtiéndose en autoobservación (Danziger, 1980).

A pesar de esto, la observación interna no es inválida intrínsecamente, como sí lo es la autoobservación, de ahí que pueda convertirse en un método confiable como la observación externa. De esta manera, “para una psicología científica sería necesario manipular las condiciones de la percepción [observación] interna de modo de aproximarlas a las condiciones de la percepción externa” (Danziger, 1980, p. 6). ¿Qué condiciones? Fundamentalmente, la independencia entre el sujeto y el objeto. Esto se lograría a través del experimento, el cual estaría asociado a la manipulación fisiológica.

El experimento, en tanto consiste, como vimos, en la manipulación controlada del flujo de los procesos investigados, permite supeditar el desarrollo de los mismos a factores externos, privándolos de la influencia directa del sujeto. Tal como lo indica Wundt (1902):

En efecto, el contenido de la psicología consiste en procesos y no en objetos persistentes. Para indagar la aparición y el curso exacto de estos procesos, su composición y las recíprocas relaciones de sus diversas partes, tenemos, antes de nada, que producir a nuestra voluntad aquellas apariciones y poder variar las condiciones según nuestros propósitos, lo que únicamente es posible mediante el experimento y no por la mera observación. (p. 10)

Así pues, concretamente, ¿a qué tipo de modificaciones debe ser sometida la observación interna para alcanzar, en alguna medida, la rigurosidad de la observación externa? En primer lugar, se deberá procurar que el tiempo transcurrido entre la experiencia inmediata, la observación de ésta y el reporte, sea mínimo (Danziger, 1980). De esta manera, “si la observación y el informe podían seguir inmediatamente a la percepción original, sin tiempo de reflexión y sin autoconciencia, las condiciones de la observación psicológica podían aproximarse a las condiciones de la observación externa” (Danziger, 1980, p. 6).

En segundo lugar, está el factor de repetición (Danziger, 1980). En efecto, “si partimos del hecho de que estímulos externos idénticos producen experiencias subjetivas idénticas o muy similares, se pueden iniciar tales experiencias a voluntad y repetirlas tanto como se desee” (Danziger, 1980, p. 8). En este punto cobra especial relevancia la fisiología, dado que la repetición tal como fue señalada sólo es posible a través de la manipulación y presentación deliberada de estímulos externos (Danziger, 1980). De ahí que Wundt la haya considerado fundamental para convertir la psicología en ciencia (Martínez, 2007).

Sólo bajo estas condiciones la observación interna podrá asemejarse a la observación externa, y sería válida metodológicamente para la constitución de la psicología como ciencia. No obstante, esto implica que sólo aquellos fenómenos que sean proporcionalmente equivalentes, a los estímulos externos evocadores, son susceptibles de ser estudiados mediante la observación interna. Concretamente, “la sensación y la percepción eran las áreas que más claramente cumplían los requerimientos fundamentales (...), y persistieron siempre como las áreas más investigadas en el laboratorio de Wundt” (Danziger, 1980, p. 8).

Lo anterior tiene una consecuencia importante, a saber, que el método propuesto excluya fenómenos que hacen parte del amplio campo de la psicología. En otros términos, la observación interna deja por fuera aspectos importantes de la experiencia inmediata, es decir, la experiencia consciente. En relación a este punto, Danziger (1980) señala lo siguiente:

La capacidad del experimento para la solución de los variados problemas psicológicos resulta ser un problema de grado. En un extremo están los problemas para los cuales el método experimental provee una fuente excelente de datos válidos; en el otro extremo están los problemas que son casi imposibles para la investigación experimental. En esa escala Wundt consistentemente asigna el extremo más apto a los problemas del área de la sensación y la percepción,

mientras que los problemas del pensamiento, el afecto, la actividad voluntaria, y la psicología social se ubican en el punto más bajo. (p. 8)

Este hecho condujo a Wundt a una situación particular, a saber: la no correspondencia entre teoría y método.

### **1.3. Problemas y limitaciones**

Son varias las limitaciones que presenta el sistema propuesto por Wundt, particularmente debido a su método. A continuación, se mencionarán tan sólo dos, haciendo especial énfasis en una de éstas debido a los propósitos generales de este trabajo.

En primer lugar, está el hecho de que el método propuesto por Wundt, para no sacrificar la objetividad lograda, sólo permitía hacer observaciones empíricas relativas al estímulo. Esto dado que, en tanto el carácter científico de la observación interna depende en gran medida de la manipulación de estímulos, “el informe introspectivo en el laboratorio estaba marcadamente limitado a juicios acerca de la magnitud, intensidad, y duración del estímulo físico, complementados a veces por juicios sobre su simultaneidad y sucesión” (Danziger, 1980, p. 9).

Ahora, en segundo lugar, está lo mencionado al final del apartado anterior. En tanto la observación interna científica sólo es posible en aquellos casos en donde la experiencia inmediata es proporcional a la mediata, la psicología fisiológica de Wundt se ve obligada a dejar por fuera de la investigación experimental procesos indiscutiblemente psicológicos como el pensamiento (Danziger, 1980). La consecuencia final de esta limitación es el hecho de que “los conceptos explicativos y los principios de su psicología no estaban fundados en la percepción interna, sino que eran esencialmente construcciones hipotéticas” (Danziger, 1980, p. 12).

Esta limitación se asocia, en términos generales, a uno de los principales problemas que ha tenido que enfrentar la filosofía y la psicología: dar cuenta de la relación existente entre *la mente y el cuerpo* (Araujo, 2007). El hecho de que no se tenga claridad acerca de cómo investigar experimentalmente aquellos procesos psicológicos no determinados directamente por la manipulación externa de estímulos, revela un problema más profundo: desconocer de qué manera dichos procesos interactúan con las condiciones fisiológicas.

En efecto, en relación a este punto, “Wundt postuló que la mente y el cuerpo eran sistemas paralelos y no interactuantes” (Quiñones & Ato, 1983, p. 134). Esta posición se conoce como paralelismo psicofísico y está constituida por dos factores (Araujo, 2007). Uno de ellos puede ser llamado el factor negativo y consiste en la renuncia a la “tradicional idea de causalidad psicofísica, según la cual el cerebro puede ser visto como causa de la mente” (Araujo, 2007, p. 104). La renuncia a esta idea tiene por consecuencia la superación, a su vez, de “la suposición de que el conocimiento neurofisiológico puede reemplazar al psicológico y ofrecer una explicación suficiente de los fenómenos mentales” (Araujo, 2007, p. 104).

El otro factor, el positivo, consiste en el “reemplazo de la hipótesis de la causalidad psicofísica a través de la suposición de una simple coexistencia entre los fenómenos mentales y cerebrales, que se da en la misma experiencia” (Araujo, 2007, p. 104). Lo que subyace a estas dos premisas, en último término, es algo que ya fue señalado páginas atrás, a saber: la fragmentación de la experiencia, no en partes esencialmente diferentes, sino en dos perspectivas de una misma.

Todo esto tiene una consecuencia fundamental: si bien el paralelismo psicofísico conserva de alguna manera una especie de dualismo, no se trata de un dualismo de sustancias, se trata de un dualismo epistemológico, en tanto se deriva de la división de la experiencia en

inmediata (sujeto cognoscente) y mediata (objeto conocido). Al respecto, Araujo (2007) señala lo siguiente:

Solo en este sentido empírico o heurístico debemos entender el paralelismo psicofísico en Wundt. Este es sobre todo una consecuencia directa de sus principios epistemológicos de base, en la medida en que remite a las bases originarias de la experiencia. Ahora bien, no se trata de alguna tesis metafísica, ni de una idea de alguna sustancia escondida detrás de los fenómenos (...) Más bien, este postulado es un principio heurístico para la investigación de la experiencia habitual. (p. 105)

Realizar este matiz es, a mi juicio, fundamental. Lo es en la medida en que, para lograr la conversión de la psicología en una ciencia, principal objetivo de Wundt, es necesario diferenciarla de su larga tradición filosófica (Danziger, 1980). La solución al problema de la interacción entre la mente y el cuerpo es el principal desafío que permitiría tal diferenciación (Danziger, 1980).

Ahora bien, el hecho de que Wundt confiera al problema del dualismo un carácter epistemológico, y no metafísico, es un avance significativo en relación a la privación de las reminiscencias filosóficas presentes en la psicología. No obstante, en tanto el problema del dualismo se deriva, en este caso, de la no correspondencia entre objeto y método, sigue siendo un factor que entorpece la consecución de un estatuto científico para la psicología. Esto dado que toda ciencia deberá procurar, justamente, la congruencia entre su objeto de estudio y el método usado para dar cuenta del mismo. Es éste el objetivo buscado por el conductismo.

## **2. Conductismo**

### **2.1. Conductismo como programa**

Hasta ahora se ha reconocido, como en todos los manuales de historia de la psicología, el sistema propuesto por Wundt, como el primer intento explícito de alcanzar un estatuto científico en la investigación psicológica. Sin embargo, sabemos que tal proyecto tropezó con uno de los grandes líos (tal vez el mayor) que ha tenido que enfrentar la psicología: el dualismo mente-cuerpo. Si bien en Wundt este problema es superado en cierta medida, en tanto se pasa de un dualismo de sustancias a uno epistemológico, en el sentido descrito en el apartado anterior, esta nueva forma de dualismo se deriva de una incongruencia, fatal para cualquier proyecto científico, propia del sistema wundtiano: la no correspondencia entre objeto y método.

Watson, que tal como lo indica Yela (1996) comenzó a trabajar según el modelo general propuesto por Wundt, pero adaptado al campo de la psicología animal, en donde además del control de los estímulos y las circunstancias externas (cajas experimentales, laberintos, etc.) también se hace “un análisis de la experiencia subjetiva del animal, de las sensaciones y afectos que en su conciencia acontecen” (Yela, 1996, p. 168), finalmente optó por rechazar dicho análisis mentalista admitiendo exclusivamente “lo que queda en estos estudios cuando se prescinde de esa conciencia: la conducta observable” (Yela, 1996, p. 169). En términos generales, esta renuncia marcó el inicio de lo que pasaría a conocerse en la historia de la psicología como conductismo.

En efecto, “las grandes dificultades que planteaba la tarea de compaginar el objeto de estudio (la conciencia) con la metodología científica abonaron el terreno para la aparición de este movimiento” (Zumalabe, 2012, p. 91). Es con la intención de superar dicha dificultad, esto es, la imposibilidad de articular la conciencia como objeto de estudio al método

científico, que Watson “apostó decididamente por el segundo eliminando para ello al primero” (Zumalabe, 2012, p. 91). De hecho, es justamente dicha negación de la conciencia, o en sentido amplio de lo mental, a favor de la rigurosidad metodológica, lo que finalmente define a este nuevo modelo. En palabras de Yela (1996) “el conductismo teórico de Watson (...) termina por consistir en una psicología caracterizada exclusivamente por su método, el *objetivismo antimentalista*” (p. 170).

Ahora bien, ¿qué ventajas ofrece la conducta para conservarse como objeto de estudio? Pues en tanto lo que se pretendía era ajustar la psicología a la epistemología de la ciencia natural (Zumalabe, 2012), la conducta como objeto de conocimiento permitió “hacer una psicología cuyos enunciados teóricos y empíricos fueran potencialmente exteriorizables y por tanto susceptibles de ser observados por una tercera persona (...)” (Zumalabe, 2012, p. 92). Esto supuso, a su vez, la reducción de los “procesos psicológicos a simples mecanismos de estímulo-respuesta” (Zumalabe, 2012, p. 92).

No está de más preguntarse, en este punto, además del *objetivismo antimentalista*, cuáles son las características generales del conductismo como nuevo modelo de comprensión en psicología. Para dar cuenta de dichas características, recurriré, una vez más, a las palabras de Yela (1996):

Para ello [para alcanzar la objetividad de las ciencias naturales y predecir y controlar la conducta] la psicología habrá de ser reduccionista y fisicalista (el psiquismo se reduce a la conducta, y ésta a movimientos físicos), elementalista, asociacionista y mecanicista (la conducta es un conjunto de elementos, energías físicas y movimientos musculares y glandulares, que se asocian mecánicamente en un organismo reactivo y pasivo), perifera y ambientalista (todo acontece en el ambiente, en él se observan los estímulos y las respuestas; las leyes de su

conexión son las que son, independientemente de lo que pase entre medias, dentro del organismo, que a efectos psicológicos, es algo vacío, una *black box*).  
(p. 170)

Tal es, en sentido amplio, la propuesta hecha por Watson, propuesta que sería adoptada, en ocasiones de manera radicalmente distinta, por autores como Hull, Tolman, Guthrie y Skinner (Yela, 1996). Todos conservan el objetivismo metodológico y prescinden de toda referencia mentalista, excluyen a la mente o la reducen a conceptos operacionales, juzgan como científicamente válido sólo lo comprobado empíricamente, etc. (Yela, 1996). No obstante, aunque todos investigan el aprendizaje animal en situaciones experimentales, discrepan en relación a lo que es aprendido por el animal (respuestas, asociaciones entre estímulos y respuestas, asociaciones entre estímulos, expectativas, etc.), el mecanismo que permite el aprendizaje (continuidad, reforzamiento, aprendizaje vicario, etc.) y, como si con eso no bastara, discrepan en la conceptualización de dicho mecanismo (Yela, 1996).

En el apartado siguiente me ocuparé de presentar algunos aspectos generales del sistema propuesto por Skinner. Tomaré exclusivamente este autor y su propuesta debido al impacto histórico que tuvo en la psicología, y en esto creo que no es necesario dar mayores justificaciones.

## **2.2. Skinner y el condicionamiento operante**

Iniciaré señalando, una vez más, que para el conductismo el objeto de estudio de la psicología deber ser la conducta. Esto no implica que por primera vez aparezca la conducta como fuente de información valiosa para la psicología. De hecho, “los referentes empíricos de las investigaciones psicológicas, conductistas o no conductistas, son [y han sido] la conducta” (Peña, p. 127). Lo que sí implica es que la conducta, con la negación de cualquier constructo mentalista, pasó a tener un valor central en lugar de periférico.

Ahora: ¿qué problema supone la conducta? Pues bien, dado “que tiene sentido hablar de ciencia sólo si lo que se nos ofrece son explicaciones de orden causal” (Tomasini, 1990, p. 54), la psicología, como ciencia de la conducta, deberá dar cuenta, ante todo, de las causas de ésta. No en vano en el primer capítulo de *Sobre el Conductismo* (1994), Skinner inicia preguntándose justamente a qué se debe el hecho de que las personas se comporten tal como lo hacen. Esta pregunta, tal como él lo señala, en un primer momento debió ser de carácter práctico (en tanto su solución permitiría predecir y controlar la conducta), no obstante, “eventualmente se convirtió en el problema de entender y explicar el comportamiento” (Skinner, 1994, p. 12).

Tradicionalmente, la conducta se ha considerado un efecto de la actividad mental. Esto es así, entre otras cosas, dado que la causalidad ha sido entendida como una sucesión directa de eventos (Skinner, 1994). De ahí que sea fácil concluir que lo experimentado antes de una conducta en particular sea la causa de ésta última. Así, “si se nos pregunta por qué hemos hablado hostilmente a un amigo, podemos responder «Porque sentí ira»” (Skinner, 1994, p. 12).

No obstante, este tipo de explicación, que por lo general aboga a constructos mentalistas como causas de la conducta, debe inevitablemente enfrentar los problemas asociados a toda forma de mentalismo. Dichos problemas son, fundamentalmente, la explicación sustancial de los estados mentales y la solución al problema de la interacción entre sustancias. En relación a este punto, Skinner (1994) indica lo siguiente:

Pero ¿dónde están esos sentimientos y estados de la mente? ¿De qué están hechos? La respuesta tradicional es la de que están localizados en un mundo de dimensiones que no son físicas llamado mente, y que son mentales. Pero entonces surge otra pregunta: ¿cómo puede un hecho mental causar o ser

causado por un hecho físico? Si queremos predecir lo que hará una persona, ¿cómo podemos descubrir las causas mentales de su comportamiento, y cómo podemos producir los sentimientos y estados de la mente que la inducirán a comportarse de una manera determinada? (p. 13)

Se sabe ya, en atención a lo que he venido desarrollando hasta ahora, que enfrentar este tipo de problemas complica significativamente la consecución de aquella rigurosidad científica deseada para la psicología. Es debido a esto que el conductismo opta por dirigir su atención exclusivamente a la conducta y, en relación a ésta, a las circunstancias ambientales. Ahora, no está de más indicar que esta decisión no fue tomada por capricho; en realidad, existe toda una reconceptualización de *lo mental*, por lo menos en Skinner, como conducta en último término. Tal reconceptualización será presentada, en lo que sigue, luego de indicar a juicio de Skinner cuáles son las causas de la conducta.

### **2.3. Condicionamiento operante**

Ante los problemas vinculados a una explicación mentalista de la conducta, el conductismo sugiere buscar las causas de aquella en la historia ambiental previa del organismo (Skinner, 1994). Así, “si todos los vínculos son válidos [entre la conducta y la historia ambiental], nada se pierde al descuidar los supuestos vínculos que no son físicos” (Skinner, 1994).

En este punto es necesario introducir una aclaración. Con ‘historia ambiental previa’, como veremos, no se está haciendo referencia a los estímulos ambientales que preceden, inmediatamente, a la conducta. Este sería el tipo de causalidad defendida por las explicaciones mentalistas de la conducta. En su lugar, se hace referencia a las contingencias de refuerzo que han permitido la conservación de una conducta en particular (Skinner, 1994). Este proceso es conocido como condicionamiento operante. En efecto, “a través del proceso

de condicionamiento operante, el comportamiento que tiene esta clase de consecuencia llega a tener mayor probabilidad de ocurrencia” (Skinner, 1994, p. 40).

Ahora bien, ¿qué tipo de consecuencias son reforzantes? Iniciemos señalando que a juicio de Skinner el condicionamiento operante “complementa la selección natural” (Skinner, 1994, p. 46). Lo anterior en tanto “las consecuencias importantes del comportamiento que no pudieron desempeñar un papel en la evolución por no ser rasgos del ambiente suficientemente estables, se hacen efectivas a través del condicionamiento operante durante la vida del individuo (...)” (Skinner, 1994, p. 46). Esto nos permite señalar que toda consecuencia reforzante lo es en la medida en que tenga un valor adaptativo, de ahí que la conducta que la produce sea conservada tal como son conservadas, por la selección natural, algunas modificaciones genéticas propias de una especie en particular. En palabras de Skinner (1994), “la susceptibilidad de refuerzo se debe a su valor de supervivencia (...)” (p. 46).

Todo lo anterior permite afirmar que el tipo de causalidad expresada por las contingencias de refuerzo, y las de supervivencia, no depende de los antecedentes sino de las consecuencias. En efecto, “ambas tratan de explicar el propósito después del hecho (...)” (Skinner, 1994, p. 40). Este punto es más importante de lo que parece a simple impresión. Lo es en la medida en que “este proceso [el condicionamiento operante] y sus efectos han dado lugar a una amplia gama de conceptos mentalistas” (Skinner, 1994, p. 40), conceptos que, tal como anticipé hace unos párrafos, Skinner redefine como conductas. Tomemos como ejemplo dos conceptos mentalistas por excelencia: propósito e intención.

#### **2.4. Propósito, intención y errores categoriales**

Antes que nada, preguntémonos lo siguiente: ¿por qué la conducta operante ha dado lugar a conceptos mentalistas, tal como lo indica Skinner en el último pasaje citado? La razón parece hallarse en el tipo de causalidad que supone el condicionamiento operante: una

causalidad determinada por las consecuencias. Así, “la aparente falta de una causa inmediata en el comportamiento operante ha llevado a la invención de un hecho iniciador” (Skinner, 1994, p. 52). Esto dado que, como indiqué hace unos párrafos, se ha tendido a suponer como causa de un evento el fenómeno que le antecede inmediatamente. Esta suposición, en lo que concierne a la psicología, y para resolver la pregunta con la que inicié, ha facilitado la atribución de propiedades causales a los estados *mentales internos*, así como la promulgación de los mismos. No obstante, dichos estados, como intentaré demostrar, no son de una naturaleza diferente a lo material y tampoco son internos. De ahí que resalte parte de la expresión en cursiva.

Ahora, ¿qué relación guarda todo lo anterior con los propósitos y la intención? La relación radica en que estos conceptos, mentalistas sin lugar a dudas, son justamente dos de los tantos que se han sugerido como causas de la conducta. Suele creerse que la conducta intencional está dirigida *por* un propósito. Esta expresión parece suponer que existe un *fenómeno mental* llamado propósito, que es discreto y diferenciable de otros fenómenos mentales, y que además posee propiedades causales. Sin embargo, lo que quisiera argumentar, siguiendo las reflexiones de Skinner (1994), es que en realidad la conducta intencional está dirigida *hacia* un propósito, lo cual es radicalmente distinto, incluso en sentido ontológico.

Imaginemos la siguiente situación. Por alguna razón, irrelevante para el desarrollo del argumento, estamos perdidos en un bosque. Nuestra principal urgencia es conseguir agua, y a pesar de haber caminado ya durante un par de horas, no hemos encontrado ningún río o arroyo. Ante esta situación, decidimos tomar un pedazo de rama y comenzar a excavar algunos centímetros con el fin de hallar algún tipo de corriente subterránea. Sabemos que

este resultado es posible dado que antes, en situaciones menos riesgosas, hemos llevado a cabo el mismo procedimiento.

Si luego, al contar la experiencia a alguien más, esta persona indica que hemos actuado tal como lo hicimos porque antes de iniciar a excavar nos planteamos como propósito hallar agua, estaremos tentados a aceptar dicho razonamiento. Sin embargo, según Skinner, en realidad actuamos de tal manera porque en circunstancias pasadas el tomar un pedazo de rama, y remover con ella alguna cantidad de tierra, ha sido reforzado por el hallazgo de agua. Es decir, el hallar agua, como propósito, no precede a la conducta; de hecho, como se supone en este caso, es una consecuencia de la misma. De ahí que Skinner (1994) señale que, en realidad, “los motivos y propósitos son los efectos de los refuerzos” (p. 55). En este sentido, “cuando la persona es «consciente de su propósito», está sintiendo u observando introspectivamente una condición producida por el refuerzo” (Skinner, 1994, p. 55), por ejemplo, la supresión de la sed en la situación propuesta. En síntesis, el propósito, o intención de una conducta, es una expectativa de resultado (Skinner, 1994).

Se sabe ya que el propósito o la intención de una conducta en particular en realidad no antecede a dicha conducta. El punto está en que lo que llamamos propósito o intención no es más que la anticipación de un escenario específico que sucede a una conducta también específica. Ahora, si bien la creación de constructos mentalistas está asociada a la comprensión de la causalidad según los antecedentes (como se puede concluir de lo expuesto hasta aquí), también está asociada a la conservación de errores categoriales en el lenguaje. Veamos qué significa esto.

Primero lo primero: ¿qué es un error categorial? Un error categorial es, en palabras de Freixa (2003), “un proceso erróneo de atribución [y de no atribución, añadiría] de un elemento a una categoría” (p. 597). Existen algunos ejemplos dentro de la psicología misma.

La comprensión general que se tiene de la conducta es uno de ellos (Freixa, 2003). Tradicionalmente, se ha entendido por conducta “el movimiento visible de un ser vivo o de una de sus partes” (Freixa, 2003, p. 597). Esta conceptualización de la conducta excluye operaciones como el cálculo *mental* en la medida en que, justamente, es *mental*. No obstante, dicho procedimiento antes de *ser mental* (filogenética y ontogenéticamente) fue público, es decir, fue conducta (Freixa, 2003). Esto dado que “cuando se está en fase de aprendizaje del cálculo, uno se ayuda (...) de los elementos externos que tiene a mano (...), elementos que pueden ser contados y manipulados a voluntad (...)” (Freixa, 2003, p. 599). De ahí que calcular *mentalmente* sea, en primera instancia, “una conducta manual, manifiesta, motora y pública (...)” (Freixa, 2003, p. 599).

Naturalmente, como consecuencia de la repetición constante de una conducta (en sentido público), “la conducta se automatiza y se vuelve cada vez menos dependiente de su soporte manipulativo” (Freixa, 2003, p. 600). Esto no implica que la conducta se *vuelva mental* y por lo tanto pase a pertenecer a un nuevo orden ontológico. Es en este punto en donde a la conducta encubierta le es negado su carácter conductual, en tanto no es pública, lo cual constituye un error categorial. Ahora, lo más grave de todo, tal vez, es que esta automatización parece suponer (erróneamente, como acabo de señalar) que la conducta externa al automatizarse cambia sustancialmente, es decir, en su esencia. A propósito de este punto Freixa (2003) señala lo siguiente:

Una simple diferencia de accesibilidad a un fenómeno [que sea público o encubierto] nunca tuvo el poder de cambiar ni la naturaleza ni el estatus del fenómeno en cuestión, que es independiente del hecho de que se pueda acceder hasta él más o menos fácilmente. En otras palabras, la diferencia de accesibilidad concierne al observador, no al fenómeno. El fenómeno es lo que

es, independiente de su accesibilidad, que es una característica dependiente del observador. (p. 600)

Todo esto nos permite comprender con mayor precisión en qué sentido el conductismo redefine *lo mental*. Sin embargo, a mi parecer, sigue quedando un problema por resolver. Sabemos ya que *lo mental* es en realidad una serie de fenómenos externos, no obstante, el sentido común parece resistirse a aceptar este hecho. Nos cuesta abandonar la concepción tradicional de *lo mental* como un ‘reino’ diferente a lo material, de ahí que nos veamos tentados a seguir suponiendo que las creencias, los propósitos o intenciones y los sentimientos (por mencionar tan sólo algunos ejemplos) son *cosas* internas, *mentales*. El problema está en que dicha suposición no es más que una corazonada que se deriva de la función reificadora del lenguaje (Ramos, s.f.).

Gramaticalmente, sabemos que los sustantivos, en términos generales, representan cosas. La palabra árbol, como palabra, semánticamente se remite a un hecho externo. Asimismo, “la existencia en nuestro lenguaje de sustantivos como ‘mente’, ‘inteligencia’ y ‘pensamiento’ dan lugar a la creencia de que debe haber algo, material o inmaterial, que corresponda a tales expresiones” (Ramos, s.f., p. 204). Sin embargo, la correspondencia de estos sustantivos con hechos externos es bastante confusa, de ahí que sea más fácil suponer una correspondencia con hechos internos. No obstante, tal supuesta correspondencia reposa en el efecto sustancializador del lenguaje, esto es, la suposición de que a todo sustantivo le corresponde *algo*.

Veamos un ejemplo de lo anterior. La oración ‘tengo una moneda’ es semejante gramaticalmente a la oración ‘tengo un dolor’. En palabras de Moreno (2011), “las dos parecen hablar de objetos que son poseídos (dolor, moneda) y ambas sugieren la existencia de un espacio donde tales objetos se hallan en nuestra posesión (el bolsillo, la mente)” (p.

191). Así, dado que “una de las proposiciones describe un hecho en el mundo material, así mismo pensamos que la otra proposición también describe la posesión de un objeto en un mundo diferente, el mundo mental” (Moreno, 2011, p. 191).

Ciertamente, establecer tal equivalencia semántica es un error categorial bastante grave. Esto último dado que suponer la existencia de un *mundo mental* implica conservar el viejo dualismo de sustancias. En síntesis, considerar equivalentes semánticamente las proposiciones relativas a fenómenos externos y las proposiciones mentalistas “nos ofrece la falsa imagen de la mente como un tipo de mundo que puede ser descrito, un espacio, que al igual que el espacio físico se encuentra lleno de estados (dolores), procesos (pensamientos, inferencias) y objetos (ideas)” (Moreno, 2003, p. 192). Esto dado que, en palabras de Wittgenstein (2008), “donde nuestro lenguaje hace presumir un cuerpo y no hay un cuerpo, allí, quisiéramos decir, hay un espíritu” (p. 14).

Bastará señalar un último punto para pasar al siguiente apartado. El hecho de que el conductismo redefine en términos de conducta lo que tradicionalmente se ha considerado como mental, y por lo tanto diferente esencialmente de las situaciones y fenómenos externos, le ha permitido evitar toparse con los problemas del dualismo de sustancias. Tal como indica Novoa (2002):

Se puede decir que la dualidad cartesiana intentó resolverse con los planteamientos del conductismo radical. La incorporación del monismo conductual, referente a la consideración de la persona como una unidad o un todo, en el cual es indivisible biología de conducta, o sentimiento de conducta; y la remisión del estudio del mundo debajo de la piel, al estudio de las relaciones funcionales entre ambiente y comportamiento deja sin piso el debate de si la

mente es una sustancia o una estructura funcionalmente distinta para dar paso al estudio del comportamiento como interacción. (p. 77)

Llegados a este punto, queda la sensación de que el conductismo solucionó todos los líos, conceptuales y ontológicos, que han entorpecido el camino de la psicología hacia la rigurosidad científica. Lo cierto es que, en realidad, ha sido fuertemente criticado. Dichas críticas han estado asociadas particularmente a su objeto.

## **2.5. Críticas al conductismo**

Partamos de lo siguiente: el principal punto de diferenciación entre el conductismo y la psicología fisiológica de Wundt consiste en la redefinición de los constructos mentalistas como conductas por parte de aquél. Esta diferencia es fundamentalmente objetiva, es decir, relativa al objeto. Advierto esto último dado que en términos metodológicos los principios, en sentido general, continuaron siendo los mismos: experimentación y observación científica.

Según esto, lo que realmente hace el conductismo como programa o modelo, es redefinir los criterios de justificación para los objetos de la psicología. Si antes dichos objetos eran internos e intencionales (contenían información relativa a algo), con el conductismo pasan a ser fenómenos observables que deben su génesis y transformación a procesos asociativos de aprendizaje y condicionamiento (Rivière, 1991). Es justamente tal reducción de lo mental a conductas observables, sin admitir el valor de algunos procesos mediadores que determinan la interacción entre el organismo y el ambiente, así como el reconocimiento del condicionamiento clásico y operante como únicos mecanismos de aprendizaje, lo que frenó, paulatinamente, el avance del conductismo a mediados del siglo XX. Pasemos ahora a las críticas.

En 1948 tuvo lugar, en el Instituto de Tecnología de California, el famoso simposio de Hixon. Tal encuentro tuvo por nombre “Los mecanismos cerebrales de la conducta”

(Gardner, 2000). A él asistieron grandes figuras académicas. Es importante detenerse un momento en este episodio histórico dado que puede advertirse en él un marcado inconformismo con las tesis conductistas. Coincidieron en el simposio matemáticos como John von Neumann y Warren McCulloch, quienes compartían un interés particular: “explorar ciertos paralelismos entre el sistema nervioso y los procedimientos lógicos (...)” (Gardner, 2000, pág. 26). Señalo este punto en particular porque es un antecedente importante a lo que pasaría a ser la ciencia cognitiva, no obstante, me interesa especialmente las críticas hechas al conductismo en este encuentro. Dichas críticas estuvieron en manos, fundamentalmente, del psicólogo Karl Lashley (Gardner, 2000).

La charla de Lashley, titulada “El problema del orden serial de la conducta”, sintetizó bastante bien el inconformismo que por esos días generaban las premisas conductistas. En ella sugirió que el modelo explicativo propuesto por el conductismo no podía dar cuenta de conductas complejas ordenadas en forma de secuencia, dado que en este tipo de conductas (como tocar piano) no es posible que una de las acciones implicadas constituya la causa de la conducta que le sigue inmediatamente (debido a la rapidez y continuidad con la que suceden) (Gardner, 2000, pág. 29). Tal rapidez no deja lugar a procesos de aprendizaje como el condicionamiento operante o clásico, así como tampoco a cualquier forma de retroalimentación como “para que un tono dependa del anterior o refleje de algún modo su efecto” (Gardner, 2000, p. 29). Este tipo de conductas “desafían toda explicación en términos de cadenas lineales del tipo A produce B” (Gardner, 2000, p. 29). Dentro de dichas conductas se halla el lenguaje, reducido por el conductismo, en el mejor de los casos, a “conductas más simples, de naturaleza asociativa (...)” (Rivière, 1991, pág. 141).

En este sentido, es posible entrever una desazón relativa al modelo explicativo de “izquierda a derecha” (Rivière, 1991, pág. 141), es decir, estímulo-respuesta (o respuesta-

consecuencia en el caso del condicionamiento operante), propuesto por el conductismo. Lo que propone Lashley es reemplazar dicho modelo por “mecanismos de determinación <de arriba abajo> y (...) sistemas jerárquicos de organización de la conducta (...)” (Rivière, 1991, pág. 141). En este sentido, Gardner (2000) señala lo siguiente:

(...) hay planes globales muy amplios, dentro de los cuales se orquestan secuencias de acciones cada vez más densas y tupidas (...) así, en el caso del lenguaje, los núcleos superiores de la jerarquía abarcan la intención global que instó a la preferencia, en tanto que la elección de la sintaxis y la producción efectiva de los sonidos ocupan núcleos inferiores de la jerarquía (...) la forma precede y determina toda conducta específica: la organización no es impuesta desde afuera sino que emana del interior del organismo. (pág. 29)

En suma, el simposio de Hixon, particularmente la conferencia dictada por Lashley, expresa la necesidad de una nueva forma de concebir la conducta (y lo psicológico) ante la insuficiencia explicativa del modelo conductista. Esto condujo, años después, a “la recuperación de los viejos conceptos de la psicología natural mentalista, de los planes y propósitos (...), las estrategias activas de elaboración del conocimiento (...) los símbolos mentales, como objetos portadores de conocimiento (...)” (Rivière, 1991, pág. 141), etc.

Otro aspecto importante que ha sido frecuentemente criticado al conductismo es la consideración del condicionamiento clásico y operante como únicas formas de aprendizaje. Los avances en la investigación de algunos psicólogos, como por ejemplo Bandura, han demostrado “la importancia de otro tipo de condicionamiento, el llamado vicario o por observación, ya insinuado por Tolman, en el que el sujeto aprende sin dar ninguna respuesta manifiesta y sin que ninguna pueda, por consiguiente, ser reforzada (...)” (Yela, 1996, p. 176). Este tipo de aprendizaje supone “la intervención de procesos cognitivos, como la

asociación e integración de experiencias sensoriales, imágenes y recuerdos, la codificación de señales y, sobre todo, cuando el sujeto dispone del lenguaje, la codificación y la comprensión verbales” (Yela, 1996, p. 176). Ciertamente, la suposición de constructos mediadores en el aprendizaje, necesarios para la comprensión del condicionamiento vicario, resultaría escandalosa para cualquier conductista.

Es este último punto el aspecto final que señalaré en este apartado, a saber: el radicalismo de los conductistas, entre ellos, encabezándolos, Skinner. Este radicalismo puede notarse en las constantes generalizaciones hechas por Skinner y los supuestos que subyacen a las mismas. En palabras de Yela (1996), dichos supuestos son los siguientes:

El supuesto de la generalización ambiental: la caja de Skinner es representativa de todos los ambientes; el supuesto de la generalidad específica: la rata y la paloma, son representativas de todas las especies de seres vivos; y el supuesto de la generalidad comportamental: las operantes, estímulos y refuerzos empleados por Skinner, y la tasa de respuestas, como variable dependiente, son representativos de los aspectos importantes de todo comportamiento. Ninguno de todos estos supuestos (...) encuentra justificación en los resultados experimentales. (p. 179)

La conjunción de todo lo anterior produjo una sensación en muchos investigadores “(...) de que la adhesión a los cánones conductistas estaba volviendo imposible el estudio científico de la mente” (Gardner, 2000, p. 28). Es así como de la mano de los avances en neurociencias, computación, lingüística, filosofía y otros campos se va consolidando la ciencia cognitiva, un nuevo intento en la historia de la investigación psicológica por alcanzar la tan anhelada rigurosidad científica.

### **3. Cognitivismo**

#### **3.1. Principios generales**

Durante la primera mitad del siglo XX, la primacía del modelo conductista derivó en lo que Rivière (1991) llama una “anomalía histórica” (p. 130). Esta afirmación cobra sentido si se tiene en cuenta que el conductismo rompió con la tradición epistemológica de la psicología. En otros términos, la reflexión psicológica históricamente ha tenido por objeto de estudio entidades de carácter epistémico, intencional, excepto durante el breve periodo en que el conductismo se impuso como filosofía imperante (Rivière, 1991). Esto es así dado que “el problema fundacional del conductismo fue, como bien sabe cualquier psicólogo, el de la objetividad” (Rivière, 1991, p.132). No obstante, la búsqueda de ésta condujo a la negación de los estados mentales como objetos susceptibles de ser conocidos científicamente (Rivière, 1991). En atención a esto los criterios de justificación para los objetos psicológicos fueron redefinidos; debían ser 1) fenómenos observables 2) que debieran su génesis y transformación a procesos asociativos de aprendizaje y condicionamiento (Rivière, 1991).

Dentro de este contexto, el cognitivismo, principal modelo de investigación en ciencia cognitiva (Bächler, 2014), surge como alternativa teórica ante la insuficiencia explicativa del conductismo (Gardner, 2000). Sin embargo, no se desvincula completamente de éste. En términos generales, el cognitivismo hereda el interés objetivista del conductismo, pero se diferencia de él en que incluye en su vocabulario conceptos de carácter estrictamente representacional (Rivière, 1991).

Todo esto, leído a la ligera, pareciera indicar que el cognitivismo es una recuperación de la propuesta hecha por Wundt. Es fácil llegar a esta conclusión si se tiene en cuenta que, según se ha comentado, el cognitivismo pretende dar cuenta de lo mental sirviéndose de constructos representacionales sin renunciar a la rigurosidad metodológica. Si bien, en efecto,

la objetividad del método puede ser pensada como un denominador común del cognitivismo y la psicología de Wundt (así como del conductismo, como señalé en el párrafo anterior), lo que realmente diferencia a estas propuestas es la forma en que se concibe el objeto de estudio.

Con la llegada del cognitivismo, la Psicología no se vuelve cognitiva por primera vez, siempre lo ha sido (excepto durante los años de primacía conductista, como indiqué hace un momento), dado que por regla general sus objetos de estudio se han caracterizado por poseer una naturaleza representacional, es decir, por tener contenido (Rivière, 1991). Esta característica se cumple tanto para la psicología de Wundt, como para la propuesta cognitivista y la larga tradición filosófica que precede la formalización de la psicología (Rivière, 1991). No obstante, y este es el punto de diferenciación entre la psicología fisiológica y el cognitivismo, *lo psicológico* comienza a pensarse de una manera radicalmente distinta, sin renunciar a sus propiedades representacionales, a partir del paradigma cognitivista (Rivière, 1991).

La aceptación del cognitivismo derivó en la redefinición, una vez más, de los criterios de justificación de los objetos psicológicos. Ahora eran concebidos como “objetos computables o como productos de algoritmos de cómputo” (Rivière, 1991, pág. 136). Esto permite, por un lado, privarlos del carácter fenoménico meramente accesible a través de la introspección y, por otro, “simular su funcionamiento mediante algoritmos rigurosos que dieran cuenta de sus <productos extensionales>” (Rivière, 1991, pág. 136). La nueva psicología cognitiva pretende, entonces, dar cuenta de lo mental en términos intencionales sin echar a un lado la aspiración objetivista del conductismo. Según Rivière (1991), “es eso, por ejemplo, lo que hace el psicólogo cognitivo que fundamenta sus conceptos de proposición, esquema, rasgo semántico, en los tiempos de respuesta de sus sujetos en tal o cual tarea, o en los errores que cometen, etc.” (p. 134).

Según Bächler (2014), el núcleo filosófico del Cognitivism está compuesto por la teoría representacional de la mente (TRM) y la hipótesis del sistema de símbolos físicos. Por un lado, la TRM, en respuesta a las premisas conductistas, “retoma el valor explicativo de los estados mentales —creencias, deseos, etc. —, pero lo hace desde un punto de vista radicalmente distinto” (Bächler, 2014, p. 51). Aquello distinto no es más que la redefinición de los estados mentales en términos computacionales (Rivière, 1991). La importancia de esta nueva concepción de los estados mentales radica en que “lleva implícita una visión física y no subjetiva de la noción de mente” (Bächler, 2014, p. 51). Dicha índole física obedece, como veremos a continuación, al lugar que ocupan los símbolos en la definición computacional de los estados mentales.

La hipótesis del sistema de símbolos físicos, a su vez, “sostiene que un sistema de símbolos físicos tiene los medios necesarios y suficientes para la acción inteligente general” (Bächler, 2014, p. 52). Esto significa que todo sistema que posea inteligencia general, entendida como la ejecución de un “comportamiento apropiado a los fines del sistema y adaptativo en relación con las demandas del ambiente” (Bächler, 2014, p. 52), deberá tener, indispensablemente, determinado sistema de símbolos.

Ahora, ¿en qué consiste el carácter físico atribuido a los símbolos en esta hipótesis? Consiste, en primer lugar, en el tipo de relación que media entre los símbolos de dicho sistema (Bächler, 2014). Según el cognitivism, los símbolos interactúan entre sí en virtud de su forma y no de su contenido; en este sentido, este modelo puede ser descrito como un mecanicismo abstracto (Rivière, 1991). En segundo lugar, el carácter físico de los símbolos hace referencia a la implementación física necesaria para que tal sistema de símbolos sea operativo (Bächler, 2014). En otros términos, se trata de la materialización, biológica o artificial, de dicho sistema. Esta aparente indiferencia en relación a las condiciones físicas

que soportan el sistema de símbolos ha sido nombrada como funcionalismo (Gardner, 2000). En el siguiente capítulo comentaré con mayor detenimiento este punto.

Ahora bien, para el cognitivismo se debe entender por cognición, como ya he señalado de alguna forma, la “computación de representaciones simbólicas” (Bächler, 2014, p. 50). En palabras de Varela (1990) “la cognición consiste en actuar sobre la base de representaciones que adquieren realidad física con la forma de un código simbólico en el cerebro o en una máquina” (p. 39). Esta definición es una buena síntesis de la concepción general de este modelo, dado que resalta el papel que cumplen los símbolos y las representaciones al interior del mismo. Además de esto, deberá ser señalado que lo realmente característico del enfoque cognitivista es su aceptación de la “versión fuerte y literal de la metáfora del ordenador” (Rivière, 1991, p. 140). Esto significa que a juicio del Cognitivismo el computador es un sistema funcionalmente equivalente al sujeto cognitivo, y viceversa. Tal es, en términos generales, la famosa metáfora del ordenador, cuya primera formulación puede ser rastreada hasta las ideas de Turing (Rivière, 1991).

Asumir tal equivalencia tuvo un efecto definatorio para el modelo cognitivista. Que el sujeto cognitivo sea funcionalmente semejante al software de un computador (porque es aquí donde se establece la relación) implica que la mente actúa a partir de determinado conjunto de símbolos, como ya ha sido señalado. Como ahora sabemos, éstos se “amalgaman, transforman o contrastan entre sí” (Gardner, 1987, p. 55) en virtud de principios puramente formales, sintácticos. Así, un sistema constituido simbólicamente (como un software o la mente cognitivista) “interactúa sólo con la forma de los símbolos (sus atributos físicos), no su significado” (Varela, 1990, p. 44). Es este el rasgo esencial del modelo en cuestión.

Para Rivière (2003) este énfasis formalista hace parte de un “ideal algorítmico formulado por Raimundo Lull, Descartes y Leibniz. Un ideal que implicaba, entre otras

cosas, la traducción de representaciones continuas o densas a códigos analíticos (...)” (p. 24). Tal ‘traducción’ corresponde en el Cognitivismo a la conceptualización de los estados mentales como actitudes proposicionales, constituidas, a su vez, por símbolos, es decir, “entidades con una estructura formal, sintáctica (...)” (Restrepo, 2009, p. 82). En síntesis:

En la dialéctica que siempre ha caracterizado a este enfoque, que es la dialéctica entre sintaxis y semántica, entre reglas y contenidos, entre forma e intencionalidad, fueron los tres primeros miembros de estos pares -sintaxis, reglas y formas- los que impusieron una abrumadora presencia en los años sesenta, constituyendo una imagen del sujeto psicológico como sujeto esencialmente racional y racionalmente independiente de los contenidos procesados. (Rivière, 1991, p. 145).

Este intento cognitivista de “construir un lenguaje formal, carente de contenido o semánticamente transparente” (Bächler, 2014, p. 52) ha sido ampliamente criticado, entre otras cosas, debido a la “distancia existente entre la definición sintáctica de la inteligencia y la realidad semántica de la cognición humana” (Bächler, 2014, p. 53). Es justamente de esta naturaleza la crítica hecha por Searle (1994, por ejemplo) al cognitivismo. No obstante, la presentación de dicha crítica será hecha en el capítulo número tres en donde se señalan, justamente, las objeciones hechas al modelo cognitivista.

Si se toma como ejemplo específico del modelo cognitivista la propuesta hecha por Fodor (1975), podrá notarse que estas premisas están situadas en el núcleo conceptual de la famosa hipótesis fodoriana del lenguaje del pensamiento. Dicho lenguaje, en efecto, se caracteriza por su naturaleza representacional, entendiendo por representación “el resultado de una combinación de estructuras formales (sintácticas) y significativas (semánticas)” (Colom, 1993, p. 6). Así mismo, según esta propuesta, “las transformaciones de los estados

mentales (operaciones) estarían regidas por las restricciones estructurales (sintácticas) de las representaciones” (1993, p. 6). Esto significa, en otros términos, que los estados mentales se transforman en virtud de las características sintácticas que poseen las representaciones que los constituyen. Es fácil reconocer en esta caracterización de la propuesta fodoriana, respectivamente, las dos premisas esenciales del cognitivismo.

Durante este último apartado he intentado dar cuenta, en sentido amplio, de los supuestos principales del modelo cognitivista. Mi interés ahora será comentar, con mayor detenimiento, algunos aspectos que han sido insinuados y dejados deliberadamente para el siguiente capítulo.

## Capítulo 2:

### Antecedentes, principios y problemas del cognitivismo

En términos generales, hasta ahora han sido presentadas las premisas fundamentales del llamado modelo cognitivista. Como características básicas de dicho modelo, se propuso la comprensión de los estados mentales como estados representacionales y la constitución simbólica de los mismos, dejando por fuera una característica más: la causalidad de los estados mentales en relación a la conducta.

La atribución de un papel causal a los estados mentales se conoce en la literatura como teoría causal de la mente. Ésta supone que “la conducta es efecto de causas mentales y que tales causas son procesos internos (...)” (Martínez-Freire, 1995, p. 77). Que los estados mentales sean internos y además causen la conducta implica que dichos estados son fundamentalmente distintos de la conducta misma (Martínez-Freire, 1995). Esto supone una diferenciación radical en relación a las premisas del conductismo.

Naturalmente, existen varias concepciones asociadas a la teoría causal de la mente. En otras palabras, la comprensión de los estados mentales como estados internos que causan el comportamiento puede hacerse desde perspectivas, incluso, radicalmente diferentes. Usualmente, se reconocen por lo menos dos modelos explicativos: la teoría de la identidad mente-cerebro, defendida por autores como David Armstrong y David Lewis, y el funcionalismo, propuesto por Hilary Putnam y adoptado, entre otros, por Jerry Fodor (Martínez-Freire, 1995). Debido a los intereses particulares de este trabajo, sólo será expuesto el modelo funcionalista por considerársele un antecedente directo del cognitivismo.

Primero se expondrá, en sentido amplio, lo que se entiende por cognición dentro de la psicología cognitiva en general. Luego, se presentará el modelo funcionalista, pretendiendo trazar con mayor exactitud el camino que condujo, a partir de la convergencia

de múltiples esfuerzos interdisciplinarios, al establecimiento del paralelismo mente/software (supuesto fundamental del modelo cognitivista). Después, se pasará a exponer la interpretación logicista del sujeto propia del cognitivismo. Finalmente, se presentará el problema de la representación y la explicación que de él dan algunos teóricos.

## **1. Cognición**

En algunos apartados de este trabajo ya se ha hecho mención a lo que se entiende por cognición dentro del cognitivismo. Se ha dicho que la cognición, desde este modelo, es entendida como la manipulación de representaciones simbólicas (Bächler, 2014). No obstante, aunque ya exista tal mención, y aunque seguramente vuelva a hacerse algún comentario en otros apartados, definiré en un par de párrafos lo que se entiende por cognición dentro de la psicología cognitiva en general.

Hablar de psicología cognitiva implica hablar de una disciplina encargada de estudiar el comportamiento humano y las variables determinantes del mismo, que cuenta con algunas características elementales: 1) posee una imagen activa del sujeto, un sujeto que no es solamente reactivo sino que procesa información y organiza su comportamiento en función de planes y estrategias; 2) renuncia a las explicaciones lineales y opta por explicaciones jerárquicas del comportamiento; 3) abandona la explicación E-R e introduce variables mediadoras como estrategias, planes, procesos, representaciones, esquemas, etc. (Rivière, 1987). Todo esto puede ser condensado con la idea del *procesamiento de información* (Villar, 2001).

La perspectiva del procesamiento de información, que es la condición más general que todo modelo que pretenda ser cognitivo debe aceptar, consiste en pensar el sujeto como un agente que procesa (selecciona, organiza, controla, almacena, recupera, etc.) información procedente del entorno (Villar, 2001). Esta condición deriva en otras características.

La primera de ellas consiste en suponer una semejanza entre el procesamiento de información humana y el procesamiento de información artificial en tanto, entre otras cosas, ambos operan sobre representaciones. La segunda se refiere a la idea de que toda actividad cognitiva, por más compleja que sea, puede ser reducida a un puñado de procesos fundamentales. La tercera indica que dichos procesos elementales funcionan de manera conjunta para producir mayor complejidad y rendimiento en distintas tareas. La cuarta y última, consiste en el reconocimiento de las limitaciones propias de la cognición humana (cantidades limitadas de datos que pueden ser procesados, ejecución secuencial de procesos, ‘sobrecargas’ del sistema, etc.) (Villar, 2001). No se desarrollará cada una de estas características dado que el interés particular de este apartado es dar cuenta del concepto de cognición latente en estos rasgos.

Según Neisser (1967), citado por Martínez-Freire (2005), “el término cognición se refiere a todos los procesos mediante los cuales la entrada sensorial es transformada, reducida, elaborada, almacenada, recobrada o utilizada (...)” (p. 119-120). Ciertamente, esta definición de cognición puede ser respaldada por la perspectiva del procesamiento de la información.

Así, juntando el planteamiento de Neisser con las características de la perspectiva del procesamiento de la información, se puede decir que, por cognición, se entiende una serie de procesos diferenciables, pero no independientes, que actúan en conjunto sobre representaciones, posibilitando la organización y utilización de la información.

Ciertamente, esta sigue siendo una conceptualización vaga y abstracta. Es por esto que existen tantos modelos al interior de la psicología cognitiva; cada uno se ajusta, en mayor o menor medida, con mayor y menor énfasis, a las características señaladas. Por tal razón es posible dividir dichos modelos en dos áreas generales: los *modelos blandos*, que no asumen

el paralelismo entre la cognición humana y artificial como una relación de identidad, y los *modelos duros*, como el cognitivismo, que sí lo asumen (Villar, 2001).

Se verá ahora en virtud de qué principios y antecedentes el cognitivismo puede ser presentado como un *modelo duro*, es decir, un modelo que concibe la cognición artificial como un ejemplo transparente de la cognición humana.

## **2. Funcionalismo**

Según Block (1996), “las teorías de la mente previas al funcionalismo han estado ocupadas de (1) qué *es* y (2) qué es lo que da a cada tipo de estado mental su identidad propia (...)” (p. 4). Respectivamente, dichos problemas son de carácter ontológico y metafísico (Block, 1996). El funcionalismo, como toda teoría de lo mental, deberá intentar responder a dichos problemas. En lo siguiente, entiéndase por metafísica el estudio del *ser del ente*, es decir, lo fundamental del ente, y por ontología el estudio del *ente* en cuanto tal (Guerra, 1996).

Si bien a juicio de Block (1996) el funcionalismo responde la cuestión metafísica sin responder la ontológica, señalando que lo fundamental a cada tipo de estado mental es su función (cuestión metafísica), otros autores sí admiten el funcionalismo como una respuesta al problema ontológico. Así pues, según Moya (2004) “el funcionalismo, en relación con el problema ontológico de la naturaleza de la mente y su relación con el cuerpo, es la tesis según la cual las propiedades mentales son propiedades funcionales” (p. 99).

A simple impresión no parece quedar clara la diferencia entre ambas respuestas. Pongámoslo en estos términos: 1) los estados mentales *son* estados funcionales (cuestión ontológica), esto implica que no *son* reductibles a conductas, estados fisiológicos e incluso tampoco lo son a las intenciones conscientes o contenidos mentales que la psicología popular reconoce como causa del comportamiento propio y ajeno (Rivière, 1987); 2) lo que identifica

a cada conjunto específico de estados mentales (cuestión metafísica) es su función también específica, así pues, “lo que los dolores tienen en común –lo que los hace dolores– es su función [*doler*]” (Block, 1996, p. 4).

Dado que ambas cuestiones (metafísica y ontológica) están asociadas entre sí, se revisará específicamente las implicaciones de ésta última. Partamos de la siguiente pregunta: ¿qué significa decir que los estados mentales son en realidad estados funcionales? Pues bien, antes que todo, parece sensato definir qué es un estado funcional. Por estado funcional debe entenderse un estado definido por su relación con inputs específicos, otros estados funcionales y la salida de outputs también específicos. En palabras de Block (1996):

De acuerdo con el funcionalismo, la naturaleza de un estado mental es como la naturaleza de un estado de autómeta: constituido por sus relaciones con otros estados y con los *inputs* y *outputs*. Si no es más que estar [en un estado] tal que recibiendo un *input* de ‘1’ ocasiona tal y tal cosa, etc. De acuerdo con el funcionalismo, tener dolor no es más que disponerte a decir ‘ouch’, preguntarte si estás herido, distraer tu atención, etc. (p. 2).

Así pues, en términos proposicionales, si se tiene la creencia *si p entonces q* (siendo p “llover” y q “abrigarse”, por ejemplo) y luego se tiene la evidencia (input) de que *no p*, se pasará entonces al estado mental *no q*, produciendo su output correspondiente (en efecto, no abrigarse). Tal es la concepción funcionalista de los estados mentales, a saber, estados caracterizados fundamentalmente por sus propiedades operacionales. Otros ejemplos de propiedades funcionales son el ser un reloj, ser un carburador o ser dinero (Moya, 2004). Ciertamente, “un objeto es un carburador en virtud de su aptitud para desempeñar cierto papel causal (mezclar aire y gasolina) en un motor” (Moya, 2004, p. 99).

Otro caso posible de un objeto definido funcionalmente es el de las calculadoras. Que la calculadora de un *smartphone* sea capaz de arrojar un resultado (output), ante una función particular (por ejemplo, la función *suma*), depende de los inputs recibidos (la digitación de las unidades que deberán ser sumadas) y de la ejecución lógica de la función *suma*. Este ejemplo es sugerente en tanto permitirá en el siguiente apartado asociar la concepción funcionalista de los estados mentales con la definición computacional del sujeto cognitivo. Además, también lo es dado que posibilita introducir la consecuencia esencial del funcionalismo en relación al problema ontológico, a saber: el principio de las *múltiples realizaciones*.

### **2.1. Argumento de las múltiples realizaciones**

Es importante recordar lo que fue señalado hace unos párrafos: la concepción funcionalista de los estados mentales no admite una identificación entre éstos y las conductas manifiestas, las disposiciones fisiológicas o los estados intencionales. Esta característica es nombrada por Rivière (1987) como un nivel de discurso propio o, tal vez con mayor claridad, una suerte de autonomía funcional.

Dicha autonomía funcional parece sugerir que lo verdaderamente relevante para la definición de los estados mentales no son las condiciones físicas o comportamentales que se le asocien; lo verdaderamente relevante son sus relaciones funcionales con inputs, outputs y otros estados (sin prescindir de las condiciones físicas, como veremos). La consecuencia de esto es la aparente indiferencia que para los estados mentales implica ser *realizados* en X o Y condición material. En términos de Moya (2004):

La clasificación funcional de las cosas es más abstracta que la clasificación en términos de propiedades físicas, y no hay razones para pensar que ambas clasificaciones guarden entre sí una correspondencia precisa. Así, un reloj de

arena, un reloj de pulsera analógico, un reloj digital, etc., son muy diferentes desde el punto de vista físico. La física los clasificaría bajo categorías muy distintas. Sin embargo, todos ellos son relojes, es decir, son clasificados bajo la misma categoría desde el punto de vista funcional. (p. 100).

Que la función de un reloj siga siendo la misma independientemente de las condiciones materiales en que esté realizado, o que una calculadora digital sea funcionalmente equivalente a un ábaco, sin ánimo de llevar la comparación al extremo, es lo que se conoce como principio de múltiples realizaciones. Según este principio, las propiedades funcionales tienen realizaciones físicas, es decir, *supervienen* a disposiciones materiales específicas; dichas disposiciones pueden ser llamadas “bases de superveniencia” (Moya, 2004, p. 101). Esto implica que si bien las propiedades funcionales, como lo son las mentales, son operacionalmente autónomas, necesitan *alguna* condición material sobre la cual realizarse. No obstante, y esto puede resultar confuso, las propiedades funcionales no necesitan una condición material *específica*.

Si tomamos como ejemplo la calculadora, notaremos que la función *sumar* puede ser eficazmente realizada tanto por una calculadora digital como por un ábaco. Aunque ambas bases de superveniencia son fundamentalmente distintas, funcionalmente son equivalentes. Lo relevante aquí es señalar que, si bien la función calcular *necesita* una base de superveniencia, dicha base no *tiene* que ser una en particular. En pocas palabras, para que un objeto cumpla con la función calcular es *suficiente* con tener un circuito eléctrico específico, una pantalla LED, un teclado con dígitos del 0 al 9, etc., pero no es *necesario*, es decir, existen más posibilidades (Moya, 2004). Tal es el principio de múltiples realizaciones.

Todo esto tiene consecuencias que, si bien están insinuadas en lo anteriormente descrito, merecen ser declaradas explícitamente. La primera de ellas es que toda propiedad

funcional, incluidos los estados mentales, depende de propiedades físicas. La segunda consiste en que dos objetos pueden ser materialmente diferentes, pero funcionalmente semejantes. Veamos la explicación que Moya (2004) da a ambas consecuencias:

Las propiedades funcionales no son idénticas a propiedades físicas. (...) Sin embargo, hay un sentido en el que se puede decir, desde una orientación materialista, que las propiedades funcionales dependen de las propiedades físicas. El sentido es el siguiente: si dos objetos difieren en sus propiedades funcionales, han de diferir en sus propiedades físicas. (...) Sin embargo, lo inverso no es cierto. Dos objetos que difieren en sus propiedades físicas no difieren necesariamente en sus propiedades funcionales. Dos objetos pueden ser físicamente distintos y funcionalmente equivalentes. (p. 100).

Resulta conveniente señalar estas consecuencias dado que de ellas se derivan asuntos fundamentales para la comprensión del funcionalismo y, por extensión, del cognitivismo. Concretamente, la primera consecuencia permite señalar en qué medida el funcionalismo es una postura materialista (monista) y no dualista en lo concerniente al problema de la relación mente-cuerpo. La segunda, permite establecer una conexión directa con un punto esencial del cognitivismo: la comprensión de la cognición como computación de representaciones simbólicas, equivalente a la cognición artificial. A continuación serán presentados, respectivamente, cada uno de estos asuntos.

## **2.2.Funcionalismo y materialismo**

De una u otra forma, se ha insistido hasta ahora en presentar el funcionalismo como una solución, o alternativa por lo menos, al principal problema filosófico que ha debido enfrentar la psicología: la relación mente-cuerpo. Históricamente, las posturas relativas a este problema se han ubicado en dos polos: posturas monistas o dualistas. Recordemos el

caso de la psicología fisiológica de Wundt, cuyas premisas la condujeron, paradójicamente, a un tipo de dualismo. Por otro lado, la conceptualización de los estados mentales como conductas por parte del conductismo, lo posiciona dentro de las posturas monistas. Al funcionalismo se le ha reprochado ser una especie de dualismo enmascarado, no obstante, a continuación se pretenderá señalar en qué medida el funcionalismo es simpatizante de las teorías materialistas.

Para efecto de los siguientes argumentos, convéngase en replantear conceptualmente los elementos involucrados en el problema ontológico. Desde ahora, por *mente* se entenderá *estados funcionales*; a su vez, por *cuerpo* se hará referencia a la *base de superveniencia* de dichos estados. Ahora bien, se sabe ya que los estados funcionales necesitan *alguna* base de superveniencia, lo cual no implica que dicha base *deba* ser una en particular. Efectivamente, es este el principio de las múltiples realizaciones. Además de esto, que los estados funcionales no dependan de una base de superveniencia particular, por ejemplo el cerebro, es lo que párrafos atrás era llamado ‘autonomía funcional’. De todo esto se puede concluir que la relación estados funcionales-base de superveniencia no es una relación de identidad; en realidad, es una relación que podría ser llamada de *dependencia inespecífica*. Con lo anterior se quiere destacar que los estados funcionales sí dependen de alguna base de superveniencia capaz de realizarlos, pero que esta base es inespecífica, es decir, puede ser de cualquier tipo.

Es justamente esta inespecificidad la que coloca en riesgo el carácter materialista del funcionalismo. Esto es así dado que dentro de la irrelevancia concreta de las bases de superveniencia, pueden entrar posibilidades tanto físicas como no-físicas (Block, 1996). Lo fundamental para que una base de superveniencia pueda realizar un estado funcional no es

su constitución esencial (ser física, por ejemplo). Lo fundamental es ser causalmente capaz de *ejecutar* el estado funcional. En palabras de González (2011):

(...) la caracterización de los estados mentales mediante estados funcionales también permitiría la existencia de entidades conscientes no-físicas, con perfiles causales compatibles al humano o animal. Desde un punto de vista conceptual al menos, es posible plantear que existen entidades conscientes no-físicas con perfiles causales adecuados similares al humano, pero sin ninguna realización física. Por ejemplo, resulta al menos concebible la existencia de un estado mental como el dolor en seres ectoplásmicos o en ángeles, en el sentido de que estos seres podrían tener un perfil causal similar al de humanos o animales y una implementación que no fuera física. (p. 194-195)

A primera impresión, todo esto puede parecer una crítica caprichosa, e incluso caricaturesca, en la medida en que se deriva de una sutileza conceptual. No obstante, se trata de una posibilidad real, por lo menos conceptualmente, coherente con los principios del funcionalismo. Es decir, las premisas funcionalistas permiten tener tal sospecha de su carácter materialista. Esto es así dado que, tal como indica Block (1996), lo que es común a los estados mentales como el dolor es su función específica, no obstante, “el funcionalismo no nos dice si los seres que tienen dolores tienen alguna parte no-física” (p.p. 4-6). En otros términos, el funcionalismo no nos dice que los seres que tienen estados funcionales necesitan tener bases de superveniencia estrictamente físicas.

Ahora, si bien como se ha insistido todo lo anterior es plenamente justificable, lo cierto es que el funcionalismo nace “(...) como una clase de fisicismo y, por tanto, como una teoría auténticamente materialista” (González, 2011, p. 196). Ciertamente no se trata de un materialismo semejante al de la teoría de la identidad mente-cerebro, por ejemplo. De hecho,

en la medida en que funcionalismo supone que una propiedad funcional tiene múltiples realizaciones, demuestra que la teoría de la identidad es falsa. Esto es así dado que “si una criatura sin cerebro puede pensar [el cerebro no es la única base de superveniencia posible para el pensamiento], el pensamiento no puede ser un estado cerebral” (Block, 1996, p. 3). Sea como sea, el hecho es que el funcionalismo tuvo en su génesis aspiraciones fisicistas y, en realidad, es un tipo particular de materialismo. Veamos esto con más detalle.

Primero, es sensato preguntarse qué se entiende por materialismo. A juicio de Moya (2004) el materialismo es “la tesis según la cual toda entidad o evento particular es una entidad o evento físico, en unión con la tesis según la cual todas las propiedades de una entidad o evento particular sobrevienen a sus propiedades físicas” (p.p. 103-104). ¿Cumple el funcionalismo estas condiciones? Pues bien, “por sí mismo, el funcionalismo no implica lógicamente el materialismo” (Moya, 2004, p. 104).

Tal como se ha descrito, el funcionalismo en ningún momento condiciona las bases de superveniencia a una naturaleza exclusivamente física. Es esta sutileza la que soporta las interpretaciones dualistas del funcionalismo. No obstante, el funcionalismo puede ser considerado una tesis materialista en la medida en que se le añada una premisa más: “la afirmación según la cual todas las causas y efectos son causas y efectos físicos” (Moya, 2004, p. 104). Así, “en unión con este supuesto, el funcionalismo implica que todo suceso mental particular es un suceso físico particular, lo cual, junto con la tesis de la superveniencia, hace del funcionalismo una teoría materialista” (Moya, 2004, p. 104).

Según lo anterior, el funcionalismo se convertiría en una teoría materialista en tanto juzgaría como necesario a toda propiedad funcional una base de superveniencia de carácter estrictamente físico, sin importar sus características específicas. Este tipo de materialismo puede ser llamado como materialismo de casos o particulares, por oposición de un

materialismo de propiedades o tipos (Block, 1996; Moya, 2004). Se trata de un materialismo que reconoce las propiedades mentales como propiedades funcionales (no reductibles a disposiciones físicas), pero que admite que cada caso particular de estas propiedades sobreviene a una propiedad física también particular. En palabras de Moya (2004):

Es importante advertir que la identidad psicofísica que defiende el funcionalismo es una identidad de particulares mentales y físicos, no una identidad de propiedades mentales y físicas. Recordemos que, para el funcionalismo, las propiedades mentales no son propiedades físicas, sino propiedades funcionales. Aunque éstas sobrevengan a las propiedades físicas, no son idénticas a ellas. (p. 104)

Así las cosas, la perspectiva materialista del funcionalismo pareciera ser más prometedora y menos reduccionista que la teoría de la identidad, por ejemplo. Ahora bien, si al materialismo funcionalista además se le asocia el mecanicismo abstracto de las hipótesis cognitivistas, se tendría por resultado un funcionalismo computacional. Este funcionalismo está directamente asociado, tal como la segunda consecuencia señalada hace unas páginas (la concepción de la cognición como fundamentalmente simbólica), a las ideas de las cuales se derivó en un primer momento el cognitivism, a saber: el llamado desafío de Turing.

### **3. De Turing a Putnam y de Putnam al Cognitivism**

Si bien la formalización del funcionalismo es atribuible a los trabajos de Hilary Putnam, particularmente a su artículo *Psychological Predicates* (1967), sus premisas fundamentales ya estaban insinuadas en los trabajos de Turing (*Computing machinery and intelligence*, 1950). No en vano, Putnam toma como punto de partida para su conceptualización del funcionalismo las llamadas máquinas de Turing.

Una máquina de Turing es, en sentido amplio, una “máquina abstracta muy simple, capaz de simular cualquier actividad que pudiera descomponerse en un procedimiento efectivo, es decir, en un algoritmo consistente en una serie de transformaciones discretas de cadenas, también discretas, de símbolos” (Rivière, 1987, p. 136). Dado que la definición estricta de una máquina de Turing o computador digital, según los términos del autor, es algo que demandaría presentar con detalle cada uno de los elementos fundamentales de dicha máquina (unidad de almacenaje, unidad de ejecución y unidad de control), solo se hará énfasis en la llamada tabla de instrucciones de todo computador digital (Turing, 1950).

Por tabla de instrucciones o tabla de máquina se entiende el conjunto de reglas según las cuales se organiza funcionalmente el sistema (Villena, 2017). Se trata de los procedimientos que determinan la relación de la máquina con los inputs, sus distintos estados internos y la emisión de outputs específicos. Pensemos la siguiente máquina de bebidas calientes para luego presentar su tabla de instrucciones. Esta máquina contará tan solo con cuatro inputs posibles (I), cinco estados internos (E) y cinco outputs posibles (O). Se trata de una máquina que puede preparar tres tipos de bebidas (expreso [E3], cappuccino [E4] y chocolate [E5]). A cada una de estas posibilidades le corresponde un input (I2, I3, I4, respectivamente) y un output (O3, O4, O5). Además de esto, la máquina cuenta con dos estados internos más: E1, en donde la máquina no ha recibido dinero y arroja como output (O1) la frase “Inserte dinero”, y E2, en donde la máquina recibe una cantidad de 1,000\$ (sólo puede recibir exactamente esta cantidad y cada bebida tiene el mismo precio), y arroja por output (O2) la frase “Seleccione bebida”, haciendo referencia a los inputs I2, I3 e I4. Veamos ahora la tabla de instrucciones de dicha máquina:

Tabla 1

*Tabla de instrucciones para máquina de bebidas calientes*

<b>Input</b>	<b>Estado</b>	<b>Output</b>
	E1	(O1) "Inserte dinero"
(I1) 1,000\$	E2	(O2) "Seleccione bebida"
(I2) Opción "Expreso"	E3	(O3) Bebida Expreso (E1)
(I3) Opción "Cappuccino"	E4	(O4) Bebida Cappuccino (E1)
(I4) Opción "Chocolate"	E5	(O5) Bebida Chocolate (E1)

Fuente: Elaboración propia

Ahora, si bien la máquina del ejemplo propuesto *precisa* una serie de condiciones mecánicas que permitan, por ejemplo, la preparación de las bebidas (O3, O4 y O5), esta tabla de instrucciones describe con exactitud las relaciones *funcionales* de los inputs, outputs y estados internos del sistema. Este punto es importante de resaltar. La tabla de instrucciones de una máquina de Turing no es más que la *formalización* de funciones complejas sin necesidad de recurrir, y esto es lo esencial, a descripciones materiales. De hecho, el que la tabla de instrucciones sea una formalización indeterminada materialmente es un antecedente directo del principio de múltiples realizaciones del funcionalismo. Es por esta condición que Rivière (1991) llama a las máquinas de Turing ‘máquinas abstractas’. Debido a esto, es posible concluir que la idea de máquina de Turing lleva implícitos los dos puntos

fundamentales del funcionalismo, a saber: la descripción funcional de procesos complejos y las múltiples realizaciones posibles para dichos procesos.

Hace un par de párrafos se dijo que fue Putnam (1967) quien conceptualizó explícitamente los principios del Funcionalismo a partir de la noción de máquina de Turing. Según Bechtel (1981), el interés de Putnam por las ideas de Turing se debe a lo siguiente:

Putnam se interesó inicialmente en Máquinas de Turing porque la relación del programa que gobierna la Máquina de Turing [tabla de instrucciones] parecía estar en una relación con el dispositivo físico que era, con mucho, igual a aquella en la que está la mente con el cerebro. Pensó que apelando a esta analogía podría disipar gran parte de las preocupaciones sobre el status ontológico de la mente, puesto que parece que no hay razón alguna para ser dualista respecto de una Máquina de Turing y los casos parecen ser completamente comparables. (p. 154)

Inicialmente, el objetivo de Putnam (1967) consistió en determinar si el dolor era o no un estado cerebral. Así, luego de un análisis que sería innecesario reproducir aquí, Putnam concluyó que el dolor no era un estado cerebral; en realidad, se trataba de un estado funcional (ya en varias ocasiones se ha señalado en este mismo trabajo qué significa que el dolor, y los estados mentales por extensión, sea un estado funcional). Sumado a esto, el tener presente las posibilidades que ofrecía el concepto de Máquina de Turing para el análisis de la relación mente-cuerpo, más los avances en computación durante la segunda mitad del siglo XX, permitió consolidar el paralelismo mente/software ya insinuado desde Turing (Rivière, 1991).

En general, la idea es que la descripción funcional dada a procesos mecánicos es posible de ser extendida a procesos mentales. Tómese como ejemplo, una vez más, la

proposición “si llueve (P) entonces me abrigo (Q)”. Una posible descripción funcional de esta proposición podría darse tal como se indica en la siguiente tabla de instrucciones:

Tabla 2

*Tabla de instrucciones de una proposición*

Input	Estado	Output
***	(E1) Si P entonces Q (si llueve [P] entonces me abrigo [Q])	***
(I1) Comprobación sensorial de que llueve (P)	(E2) P entonces Q ( <i>modus ponendo ponens</i> )	(O1) Abrigarse (Q)
(I2) Comprobación sensorial de que no llueve (no P)	(E3) No P entonces no Q ( <i>modus tollendo tollens</i> )	(O2) No abrigarse (no Q)

Fuente: Elaboración propia.

Esta misma caracterización es posible de ser dada al dolor, por ejemplo, y otros estados mentales. Ahora bien, el hecho de que el ejemplo propuesto sea de tipo proposicional permite introducir un elemento importante que permitirá establecer la conexión directa entre el funcionalismo y el cognitivismo.

Hasta ahora se han indicado los aspectos esenciales del funcionalismo sin hacer referencia directa a la relación que guarda con el cognitivismo. Recuérdese que la idea fundamental de la cual se deriva este apartado es la relación que tiene el principio de múltiples realizaciones con la concepción simbólica de la cognición característica del cognitivismo.

Sabemos ya que la idea de máquina de Turing supone el principio de múltiples realizaciones del Funcionalismo, principio soportado en la concepción funcionalista de los estados mentales. Si a esto se le suma el hecho de que la tabla de máquina no es más que la *formalización* funcional de procesos complejos, se tendría como resultado el llamado funcionalismo computacional o de IA (Bechtel, 1981).

El funcionalismo computacional asume el funcionalismo de tabla de máquina, es decir, acepta la descripción de los estados mentales como estados funcionales. No obstante, se anima a “caracterizar las actividades mentales en términos de símbolos y reglas para manipular esos símbolos” (Bechtel, 1981, p. 157). Este paso entre lo que podría llamarse una concepción *funcionalista simple* a una concepción *lógico-funcionalista* está determinado por múltiples factores. Uno de ellos es el llamado desafío de Turing.

Ya fue mencionada una parte de la historia, a saber: la propuesta de Turing de pensar las operaciones de un sistema como determinadas por la descripción funcional de las mismas, entendiendo por descripción funcional las relaciones entre inputs, outputs y estados internos del sistema. En conjunto con esta propuesta, Turing (1950) plantea que “un computador es comparable en inteligencia a un ser humano si los seres humanos no pueden distinguir las realizaciones de un computador de las de un ser humano” (Bechtel, 1981, p.p. 157-158). En síntesis, para Turing, dos sistemas son equivalentes si emiten un mismo output a partir del mismo input (Bechtel, 1981).

Es cierto, sin embargo, que “la equivalencia de Turing no establece (...) que los dos sistemas funcionan de la misma manera, puesto que considera sólo los outputs de conducta y no si se emplean los mismos procedimientos internos (reglas y símbolos)” (Bechtel, 1981, p. 158). No obstante, diríase, el paralelismo queda establecido, y los avances en computación

permitirán pasar de la simple equivalencia de outputs entre sistemas cognitivos naturales y artificiales, a la compleja equivalencia funcional (interna) entre ambos.

Así pues, en la medida en que la computación ofrece una solución sensata al problema de la relación mente-cuerpo (ver cita de Bechtel, p.47), en tanto da cuenta de cómo es posible que estados inmateriales (software) determinen causalmente estados materiales (hardware), y conforme avanzan los desarrollos en inteligencia artificial, la psicología va imponiendo una explicación logicista, simbólica, sintáctica, a la cognición humana. Tal es el propósito del funcionalismo computacional, el cual no es más que el antecedente filosófico del cognitivismo.

Es así como con el cognitivismo asume una “versión fuerte y literal de la metáfora del ordenador” (Rivière, 1987, p. 140), entendiendo por tal la concepción metafórica de la cognición humana en términos artificiales. Con esto se asume también una interpretación logicista del sujeto cognitivo; veamos ahora dicha interpretación.

#### **4. El sujeto cognitivo**

Si lo que define al cognitivismo es su reconceptualización de la cognición, entendiéndola como computación de representaciones en virtud de principios lógico-sintácticos, el sujeto cognitivo, esto es, el agente de la cognición, será justamente un sujeto computacional-representacional. Ahora, este sujeto es el resultado de la conjunción de varias nociones más de sujeto. Específicamente, el sujeto computacional-representacional es un derivado de la conjunción del sujeto natural-representacional y el sujeto abstracto-computacional. A continuación, presentaré cada una de estas nociones para finalmente definir el sujeto propio del cognitivismo.

#### **4.1.Sujeto natural-representacional**

Comúnmente, los humanos atribuyen a otros humanos estados internos, mentales, cuando los ven comportarse de una manera en particular. La certeza de los propios estados internos se extiende hacia los demás. En efecto, “desde nuestro sentido común, esto es, desde nuestras concepciones compartidas, todos pensamos que además de cada uno de nosotros existen otros sujetos (...)” (Martínez-Freire, 2005, p. 82). Estos sujetos, tal como nosotros, actúan en virtud de creencias, deseos, sentimientos, etc. Es este el núcleo del sujeto natural de la psicología del sentido común, a saber, “la atribución recíproca de creencias y deseos” (Rivière, 1991, p. 130).

Ahora bien, dichos estados, además de tener, presumiblemente, propiedades causales en relación a la conducta, tienen un contenido específico. Las creencias son sobre algo, así como los deseos y sentimientos, por mencionar solo algunos casos, también lo son. En pocas palabras, esto significa que los pensamientos son representacionales. Así las cosas, el sujeto natural se convierte en un sujeto representacional, dado que no sólo cuenta con estados internos asociados a su conducta, sino que dichos estados tienen un contenido, representan algo. Esta ‘propiedad’, por nombrar el asunto de alguna manera, en la tradición filosófica ha sido conocida como intencionalidad (Crane, 2008).

Es bien sabido que el concepto de intencionalidad, y la identificación de ésta como la característica fundamental de los estados mentales, es obra del filósofo alemán Franz Brentano (Martínez-Freire, 2005). Ahora, ¿qué significa ‘intencionalidad’?, ¿qué designa? Pues bien, según Brentano, con intencionalidad se designa la inexistencia intencional de los estados mentales, o lo que es igual, la objetividad inmanente de estos (Martínez-Freire, 2005). Ciertamente, se ha pasado de una denominación extraña a dos que lo son aún más. Sin

embargo, el asunto puede ser resuelto con un poco de etimología. Para ello, se recurrirá a las palabras de Martínez-Freire (2005):

El término inexistencia puede llevar a malentendidos si se entiende literalmente, esto es, como no-existencia, pero en realidad debe entenderse en la tradición escolástica del inesse, como existencia en algo. Así pues, la inexistencia intencional, característica de todos los fenómenos mentales, sería la propiedad que tiene cada fenómeno mental de incluir dentro de él un objeto. En esta línea de la inexistencia intencional, la intencionalidad de los procesos mentales no indica sino que sus objetos son inmanentes. (p. 84)

Claramente, la intencionalidad de los estados mentales no es más que el reconocimiento de estos como ‘entidades’ con contenido, tal como lo señalaba al inicio de esta sección. Dicho ligeramente, los estados mentales son representaciones de algo más. Si bien el problema de la representación cobrará mayor importancia en el siguiente apartado de este capítulo, quise por lo pronto introducir el concepto de intencionalidad. Esto dado que permite comprender con mayor claridad qué significa decir que los estados mentales son representacionales. Veamos ahora la descripción de la otra concepción del sujeto implicada en la constitución del sujeto cognitivista.

#### **4.2.Sujeto abstracto-computacional**

En sentido general, las distintas disciplinas que constituyen las ciencias cognitivas, han tenido como supuesto común el del sujeto como un sistema que procesa información (SPI) (Martínez-Freire, 2005). Esta concepción es aplicable tanto para los sistemas cognitivos naturales, como para los sistemas cognitivos artificiales (Martínez-Freire, 2005). Esto implica que la noción del sujeto cognitivo como SPI es, en realidad, bastante abstracta y general.

Un SPI es, según la definición dada por Martínez-Freire (2005) basada en el texto de Newell y Simon (1972), “un sistema que consta de una memoria [a largo plazo o MLP] que contiene estructuras de símbolos, de un procesador, efectores y receptores” (p. 116). Debe ser señalado, además, que dentro del procesador existen varios elementos. Primero, un conjunto de procedimientos que posibilitan el procesamiento de la información; segundo, una memoria a corto plazo (MCP) que conserva las estructuras de símbolos de entrada y salida, y finalmente un ‘interprete’ que determina los procedimientos de procesamiento en función de las estructuras de símbolos presentes (Martínez-Freire, 2005).

Ciertamente, se trata de una concepción del sujeto bastante abstracta, lo cual implica, como ya fue señalado y como reconocen Newell y Simon, admitir que además del ser humano “un computador es también un caso de procesador de información” (Martínez-Freire, 2005, p. 117). Veamos ahora el lugar que ocupan los símbolos en esta descripción del sujeto. Justamente a partir de ellos será posible, afianzar más el paralelismo sujeto natural/abstracto que ya comienza a insinuarse, para así llegar finalmente al sujeto computacional-representacional motivo de estas disertaciones.

Convéngase, antes que nada, que un símbolo es cualquier “ítem de carácter representacional” (Martínez-Freire, 2005, p. 117). Esto implica que los símbolos pueden contener en ‘su interior’, por decirlo de alguna forma, un significado. Tómese como ejemplo la multiplicación  $Y * Q = Z$ , en donde Y, Q y Z son símbolos que pueden representar potencialmente cualquier número. Ahora, la función ‘multiplicación’, entendiendo por función una relación específica entre elementos (Crane, 2008), permite indicar que además de representar, los símbolos se relacionan entre sí (están determinados por funciones específicas). Dicha relación es ‘ejecutada’ mediante algoritmos, es decir, métodos o

procedimientos con etapas definidas “para encontrar el valor de una función” (Crane, 2008, p. 148).

Ahora un ejemplo. Tomemos la función utilizada en el apartado anterior y reemplacemos sus valores arbitrariamente:  $51 * 32 = 1.632$ , en donde Y es 51, Q es 32 y Z es 1.632. Esta operación o función puede ser ejecutada a partir de varios métodos o algoritmos. Por ejemplo, podría sumarse la variable Y la cantidad de veces indicada por Q, o podría llevarse a cabo una multiplicación larga en donde primero se multiplique 2 por Y, luego 3 por Y y finalmente se sumen ambos productos (Crane, 2008).

Ahora bien, ¿qué sentido tiene hablar de símbolos, funciones y algoritmos en este apartado? Fundamentalmente, el motivo por el cual rescato estos conceptos es que posibilitan pasar del sujeto como SPI al sujeto computacional. Hace unos párrafos decía que un SPI posee una MLP en donde se almacena símbolos, y que además posee una MCP en donde se conservan aquellas estructuras simbólicas que están siendo procesadas en un momento específico. Pues bien, según esto, y teniendo en consideración los conceptos recientemente introducidos, puede afirmarse que aquellos procedimientos de procesamiento incluidos en el procesador de todo SPI no son más que algoritmos, es decir, métodos que permiten la ‘manipulación’ de estructuras simbólicas.

En este sentido, si entendemos por procesamiento de información, la manipulación algorítmica de símbolos, estamos bastante cerca de la definición común dada al *software* o programa de un computador. En efecto, según Crane (2008), “una computadora es un dispositivo que procesa representaciones de una manera sistemática” (p. 145). Con esto quiere decir que una computadora, o su *software* para ser más preciso, lo que hace fundamentalmente es procesar sistemáticamente, mediante algoritmos, estructuras simbólicas. En palabras de Crane (2008):

Algunas veces las computadoras se llaman “procesadores de información”. Otras “manipulaciones simbólicas”. En mi terminología esto es lo mismo que decir que las computadoras procesan representaciones. Las representaciones llevan información en el sentido de que “dicen” algo o se les interpreta como “si dijeran” algo. Esto es lo que los procesos de computadora manipulan. Cómo procesan o manipulan es realizando procedimientos eficaces. (p. 169)

Así pues, hemos pasado de una noción de sujeto más amplia a una más específica. Primero, se presentó el sujeto como un SPI, constituido por varios elementos, cada uno de ellos funcionalmente valioso. Después, se presentó al sujeto como un sujeto computacional, es decir, un sujeto definido fundamentalmente por el procesamiento algorítmico de símbolos. No obstante, ambas concepciones no son excluyentes entre sí; en realidad, puede decirse que la segunda se trata de una focalización o especificación de la primera.

En este punto estamos bastante cerca de la caracterización final del sujeto como sujeto computacional-representacional. De hecho, solo basta formalizar y establecer sin timidez la equivalencia entre el sujeto abstracto-computacional y el sujeto natural-representacional.

#### **4.3.Sujeto computacional-representacional**

Antes de pasar al ‘ensamblaje’ de las dos nociones de sujeto descritas hasta acá, quiero recurrir, una vez más, a la definición de cognición dada por Neisser (1967). Hace unas páginas señalé que, para este autor, la cognición consiste en la transformación, reducción, elaboración, almacenamiento y utilización de la información sensorial.

Si recordamos la caracterización dada hace unos párrafos de todo SPI, resultará sencillo comparar la noción de procesamiento con la noción de cognición dada por Neisser (1967). En este sentido, puede afirmarse que la cognición no es más que el procesamiento de

información externa. Justamente por este motivo puede indicarse que los computadores, sistemas que procesan información, son sujetos cognitivos en sentido general.

Sabemos entonces que tanto los humanos como los computadores son agentes de cognición, naturales y artificiales respectivamente. No obstante, en relación a los sujetos artificiales, conocemos más sobre las particularidades de su cognición. Sabemos que la forma en que procesan información es fundamentalmente computacional, según lo que decíamos en el apartado precedente. Ahora, también sabemos que los símbolos son “entidades con significado (semántica) y forma (sintaxis)” (Restrepo, 2009). Esto implica que los símbolos, además de tener una estructura, tienen un contenido, es decir, representan cosas.

Pensemos ahora las características de la cognición humana, del sujeto natural-representacional. Decíamos que este sujeto se define, fundamentalmente, por la naturaleza representacional de sus estados internos. Entonces, si tanto los humanos como los computadores tienen propiedades representacionales, y se sabe que dichas propiedades obedecen, en el caso de los computadores, a la existencia de estructuras simbólicas con forma y contenido, ¿por qué no suponer esta misma constitución simbólica a la cognición humana? De esta comparación entre cognición artificial y humana se deriva la noción de sujeto computacional-representacional, entendiendo por tal un sujeto que procesa representaciones simbólicas a partir de un sistema de símbolos que opera fundamentalmente en virtud de principios lógico-sintácticos (algorítmicos).

El apartado precedente, en donde fue presentado el funcionalismo, permite comprender con mayor detalle qué implica decir que la cognición es algorítmica. No obstante, aunque se ha hecho énfasis en la naturaleza representacional tanto de la cognición humana como de la artificial, no se ha indicado qué significa esto. Tal es el propósito del siguiente apartado.

## 5. Representación

Durante este capítulo se ha usado reiterativamente el concepto de ‘representación’ para caracterizar el núcleo del cognitivismo. Se ha dicho que la cognición consiste en la computación de representaciones simbólicas, entendiendo por esto el relacionamiento lógico de estados con contenido. No obstante, en ningún momento se ha dado una definición del concepto de ‘representación’ y del problema que supone. Tal es el propósito de este apartado.

En efecto, el concepto de representación juega un papel importante dentro de la propuesta cognitivista. El problema está en que debido a su ‘naturalidad’, es decir, a su naturaleza intuitiva, se escapa fácilmente de la valoración crítica. Lo cierto es que la representación es un asunto bastante enigmático, y uno de los principales problemas de la ciencia cognitiva.

Primero lo evidente: “una representación es algo que representa a otra cosa” (Crane, 2008, p. 35). Así, si una representación es ‘algo’ que representa a ‘otra cosa’, entonces puede decirse que existen, por lo menos, dos elementos involucrados: la representación propiamente y el objeto representado. Esto abre el camino, según Crane (2008), a dos cuestiones: 1) ¿qué clase de cosas pueden ser representaciones? y 2) ¿qué clase de cosas pueden ser objetos de representación? En otras palabras, ¿qué es una representación (R) y qué objetos puede representar (O)?

Sobre ambos la respuesta es tan amplia como para resultar insatisfactoria. En primer lugar, son muchas las cosas que pueden representar: un semblante (R) puede representar un estado de ánimo (O), los anillos de un árbol (R) representan la edad del mismo (O), las flores (R) pueden representar un duelo (O), etc. (Crane, 2008). En segundo lugar, pueden ser representados objetos físicos concretos, rasgos o propiedades de dichos objetos, estados de cosas o hechos, objetos no físicos, y un largo etc. (Crane, 2008).

Así, si bien estos problemas no cuentan con una solución clara, sí permiten aproximarse más a la definición de lo que es la representación. Hasta ahora se sabe que en lo concerniente a ésta se colocan en juego, por lo menos, dos cosas: la representación en sí misma y el objeto de representación. Para efectos de claridad, y siguiendo la terminología de Perner (1994), llámese a la primera *medio representacional* y al segundo *contenido representacional*.

Existe aún un elemento dejado por fuera: la *relación de representación* (Perner, 1994). Por tal se entiende el vínculo establecido entre el medio y el contenido representacional. Así, si se toma como ejemplo la fotografía de un perro en particular, el medio será la fotografía, el contenido el perro en concreto, y la relación será el vínculo establecido entre fotografía y perro.

Es posible que sea más inteligible la idea de medio y contenido que la idea de relación. Esto es así dado que el sentido común da por hecho que la imagen de un perro representa a un perro en concreto, sin detenerse a pensar que entre ambos existe una relación representacional. Esta relación es tan natural que fácilmente pasa inadvertida. Sin embargo, esto no significa que sea un desperdicio preguntarse por ella. En realidad, es este el problema fundamental de la representación, a saber: “*cómo* alguna clase de representación –imágenes, palabras o lo que sea- consigue representar” (Crane, 2008, p.p. 37-38). Antes de pasar a determinar por qué esto es un problema, sigamos con la caracterización del concepto de representación.

Se ha dicho hasta ahora que la representación no se reduce al medio representacional. También abarca el contenido representacional y la relación de representación. Para efectos descriptivos, formalicemos el proceso de representación en los siguientes términos:  $X$  r  $Z$ , siendo  $X$  el medio,  $r$  la relación de representación y  $Z$  el contenido. Ahora bien, además de

estos elementos, según Perner (1994) la representación tiene una serie de características adicionales:

- Asimetría:  $X r Z$  y no al contrario, es decir,  $Z r X$ . En otros términos, el medio representa el contenido pero el contenido no representa el medio.
- Singularidad:  $X r Z1$  y no a  $Z2$  o  $Zn$  (otro contenido semejante a  $Z1$ ). En otros términos, una fotografía de un perro en concreto representa a dicho perro y solo a él aunque sea semejante a otros perros.
- Representación errónea:  $X r Z$  aunque  $X$  no sea idéntica a  $Z$ . En otros términos, una foto representa a alguien aunque en la foto la persona tenga los ojos rojos por el flash.
- No existencia:  $X r Z$  aunque  $Z$  no exista. Por ejemplo, la representación de un centauro representa a un centauro aunque éste no exista.

Se sabe ahora qué debe entenderse por representación y cuáles son sus características elementales. Es tiempo entonces de comentar por qué la relación de representación supone un problema.

### **5.1.El problema de la representación**

Crane (2008) utiliza el caso del Pioneer 10 para dar cuenta del problema de la representación. El Pioneer 10 fue una nave lanzada en 1972 por la NASA con el propósito de explorar el sistema solar. En su estructura fue colocada una placa de metal con varias imágenes y signos grabados. La imagen más grande era la de un hombre y una mujer desnudos. El hombre está alzando la mano en gesto de saludo. Además, en la placa figuraban dos diagramas más: uno, de un átomo de hidrógeno, y otro con los tamaños proporcionales de los planetas de nuestro sistema solar, señalando el planeta desde donde procedía el Pioneer 10.

Lo que se buscó con esto fue que una vez la nave saliera de nuestro sistema solar, si era hallada por alguna civilización extraterrestre, dicha civilización entendiera su lugar de procedencia, así como las características del planeta y de sus habitantes. No obstante, la posibilidad de que tal civilización ‘entendiera’ lo que *representan* los símbolos de la placa es, a juicio de Crane (2008), incluso más improbable que la llegada de la nave a un planeta con vida inteligente. En sus palabras:

Piénsese en algunas de las cosas que deberían comprender. Tendrían que comprender que los símbolos sobre la placa eran símbolos, que aspiraban a representar cosas y no eran sencillamente raspaduras en la placa, o simple decorado. Una vez que los alienígenas supieran que eran símbolos, tendrían que comprender de qué clase de símbolos se trataba (...) Entonces deberían tener alguna idea de qué clase de cosas representaban los símbolos: que el dibujo del hombre y la mujer simbolizan formas vitales en lugar de elementos químicos, que el diagrama del sistema solar simbolizaba nuestra parte del universo antes que la forma de quienes diseñaron el vehículo espacial. Y –quizá lo más absurdo de todo-, aun cuando se hiciesen una idea de lo que el hombre y la mujer eran, tendrían que reconocer que la mano alzada era un signo pacífico y no de impaciencia, de agresión o desdén, o sencillamente que era la posición normal de esta parte del cuerpo. (Crane, 2008, p. 32)

De todo esto, cuando menos, se puede sacar una conclusión: los símbolos no tienen un significado intrínseco (Crane, 2008). Como ya se ha señalado, para el sentido común la representación es tan natural como lo es el concepto de tiempo. Sin embargo, “lo que parece natural y evidente en nuestras vidas se torna, al reflexionar, hondamente misterioso” (Crane, 2008, p. 34).

Se tiene entonces que el problema de la representación consiste en determinar cómo representan los medios a los contenidos. En otras palabras, se busca determinar cuáles son las condiciones necesarias y suficientes para que la relación de representación sea operativa (Crane, 2008).

Este problema conduce a uno más. Dado que existen distintos medios de representación, habrá que preguntarse si todos suponen el mismo tipo de relación de representación. Por ejemplo, asúmase que las imágenes representan por semejanza (como veremos, esto es bastante cuestionable). Entonces, si las imágenes representan por semejanza, habrá que preguntarse si las palabras también lo hacen o si su relación de representación es de otra naturaleza. Este problema derivará, a juicio de Crane (2008), en la pregunta por si existen representaciones más fundamentales que otras. Para el autor dichas representaciones son las mentales (Crane, 2008).

A continuación será presentada una de las explicaciones existentes a propósito del problema principal de la representación: la teoría de la causalidad o indicación. Para su análisis se juzgará si puede o no dar cuenta de las características previamente descritas de la representación (asimetría, singularidad, representación errónea y no existencia). Se ha seleccionado esta alternativa por ser la respuesta aceptada, en mayor o menor medida, por cognitivistas como Fodor (Skidelsky, 1999). Además, por ser la principal propuesta enmarcada dentro de las respuestas naturalistas (Perner, 1994).

Para Skidelsky (2003) el calificativo naturalista dentro de este contexto alude a un requisito, fijado por los filósofos, que toda teoría de la representación debe satisfacer. Dicho requisito consiste en la no apelación “a términos semánticos/intencionales para explicar la noción de representación” (Skidelsky, 2003, p. 99). Con esto lo que se pretende es evitar explicaciones tautológicas, en donde se recurra a conceptos representacionales para explicar

la representación. Así, según Crane (2008) la explicación de la representación deberá ser buscada en los hechos naturales que la sustentan.

Finalmente, antes de iniciar con la presentación de la teoría de la causalidad o indicación, considero importante resaltar que para evitar el problema de los distintos tipos de medios, y por lo tanto distintos tipos de relación de representación, la teoría señalada se asumirá como una explicación válida para cualquier tipo de medio representacional (imágenes, lenguaje, etc.).

## **5.2. Teoría de la causalidad o indicación**

Según esta teoría, “la representación es a fin de cuentas una relación causal o, más precisamente, se basa en ciertas relaciones causales” (Crane, 2008, p. 277). Para contextualizar esta idea, tómese como ejemplo las manchas características del sarampión. En este caso, “las manchas son (...) signos naturales del sarampión: representan la presencia de sarampión” (Crane, 2008, p. 277). En tal sentido es posible hablar de la representación en términos causales, a saber, tomando el medio representacional (las manchas) como un efecto del contenido representacional (el sarampión). Esto puede ser llamado *significado natural*, entendiendo por tal una correlación entre causas y efectos (Crane, 2008).

La idea del significado natural puede ser más clara en relación a casos como el sarampión, el humo y el fuego o ciertas nubes y una tormenta. No obstante, “algunas teorías causales de la representación mental opinan que las correlaciones causales entre los pensamientos y las cosas que representan pueden formar la base natural de la representación” (Crane, 2008, p. 278). Es decir, el mismo principio de correlación causal es extensible a las representaciones mentales. Veamos de qué manera es esto posible.

Primero, para avanzar en la argumentación, es necesario introducir un concepto: la *indicación confiable*. Con esto se intenta denominar la regularidad natural de toda correlación

causal. Así, “X indica de manera confiable a Y cuando hay un enlace causal confiable entre X y Y” (Crane, 2008, p. 279). En este sentido, el humo, las nubes y las manchas indican (representan) confiablemente el fuego, una tormenta y el sarampión (Crane, 2008). Si se aplica lo anterior al caso de los estados mentales, se tendría que “un estado mental representa Y si y solo si hay una correlación causal entre este tipo de estado mental y Y” (Crane, 2008, p. 280), entendiendo por correlación causal la indicación confiable del estado mental en relación al contenido representado.

De todo esto puede quedar un sin sabor. Esto es así dado que la idea de las representaciones mentales conectadas causalmente con los hechos que representan es difícil de digerir. Es sencillo preguntarse cómo es posible que un objeto externo (contenido representacional) este correlacionado con un “objeto interno” (medio representacional), si entre ambos elementos no existen intercambios mecánicos, por ejemplo. Una vez más se cruza en el camino de la psicología el problema del dualismo. No obstante, como se verá a continuación, esta dificultad puede ser resuelta por la idea de *covariación de propiedades*.

Se entiende por covariación de propiedades la relación nómica (en virtud de leyes naturales) dada entre las propiedades de ciertos objetos o sucesos con las propiedades de otros objetos o sucesos (Moya, 2004). Según Moya, un ejemplo de esto puede ser el siguiente (2004):

Podemos concebir nuestros órganos de los sentidos como sistemas cuyas propiedades covarían nómicamente con los cambios en ciertas propiedades del entorno. Por ejemplo, los estados de nuestro aparato visual varían nómicamente, en condiciones normales, con las variaciones de la luz reflejada por los objetos. De este modo, el estado de estos órganos visuales, en circunstancias normales, contiene información sobre ciertas propiedades del entorno. (p. 180)

De esta manera, tal como una fotografía de un caballo negro representa al caballo con su propiedad de *ser negro*, así mismo, un estado mental sobre dicho caballo representa su propiedad de *ser negro* en la medida en que dicha propiedad covaría con las propiedades perceptivas del aparato visual (Perner, 1994). Es cierto que este ejemplo supone un paso de la percepción a la representación mental, algo que resulta tan enigmático como la representación misma; no obstante, para efectos del argumento, admitamos dicha transición.

Así las cosas, la idea de covariación de propiedades (dicho sea de paso, formulada por Dretske, 1981), contiene “el germen de una concepción del contenido perceptivo en términos no intencionales, en términos de dependencia nómica entre propiedades y de la información consiguiente” (Moya, 2004, p.180). En otros términos, la covariación de propiedades permite explicar cómo es posible el paso factores externos a internos, reduciendo las características de toda representación a las características propias del objeto representado.

Se ha dicho hasta ahora que la teoría causal de la representación explica la relación de representación como un producto de la correlación causal entre X y Y (siendo X el medio y Y el contenido). Así mismo, se señaló que dicha correlación causal puede ser entendida como indicación confiable, es decir, X indica confiablemente a Y dado que Y es su causa. Por último, se complementó esto señalando que la relación de causalidad entre X y Y es posible dado que las propiedades de Y covarían con las propiedades de X. Concretamente, la línea argumentativa de esta propuesta puede ser la siguiente:

1. En la naturaleza, hay relaciones de representación: el humo representa fuego.
2. Estas relaciones obedecen a relaciones causales: el humo es causado por el fuego.
3. Las relaciones causales indican confiablemente: si hay humo es porque hay fuego.
4. Las propiedades de dos cosas covarían entre sí: la intensidad del humo depende de las propiedades del objeto quemado.

5. Todo esto es aplicable a las representaciones mentales: la creencia de que lloverá depende de la percepción de nubes con propiedades específicas, cuyas propiedades indican confiablemente la probabilidad de una tormenta.

Estando hechas estas claridades, es momento de analizar en qué medida esta teoría da cuenta de las características de la representación (asimetría, singularidad, representación errónea y no existencia). En relación a la asimetría y la singularidad, Perner (1994) señala que esta teoría es capaz de dar cuenta de ambas características.

Sobre la asimetría, dado que el medio es un efecto del contenido, quien es su causa, es difícil que sean confundidos (Perner, 1994). Así, ya que la fotografía de un caballo negro es el efecto de un caballo negro, no es posible confundir el caballo negro como una representación de su fotografía. Sus distintos lugares en la relación de causalidad les confiere la asimetría propia de la representación. Acerca de la singularidad, puede decirse que en tanto la fotografía de un caballo negro fue causada por un caballo negro en particular, aun cuando dicho caballo sea semejante a otros, él fue quien causó la fotografía. En esta medida, la fotografía solo lo representa a él (Perner, 1994).

El verdadero problema viene al momento de intentar ajustar esta teoría tanto a la representación errónea como a la no existencia. Para Crane (2008), estos dos problemas pueden ser llamados en conjunto el *problema del error*. Con esto se pretende nombrar “el hecho de que las teorías causales de la representación pasen un mal rato dando razón de errores en el pensamiento (Crane, 2008, p. 281). En efecto, el que podamos representar cosas que no existen o representar mal algunas otras puede ser entendido como errores de la representación. Veamos ahora la explicación que Crane (2008) a cada problema.

Sobre la no existencia, Crane (2008) señala lo siguiente: “si se supone que la indicación confiable es una condición necesaria de la representación, entonces X no puede

representar Y en ausencia de Y. Si es una condición necesaria que algunas manchas representen sarampión o indiquen sarampión, entonces las manchas no pueden representar sarampión en ausencia de sarampión”. (p. 285)

Esto es problemático porque, a decir verdad, sí podemos representar Y con X aun cuando no exista Y. Es decir, podemos tomar cierto tipo de manchas (X) como representación de Y (sarampión). No obstante, puede darse el caso de que estas manchas (X) en realidad representen Q (varicela) (Crane, 2008). El problema radica en que si la indicación confiable es el fundamento de la representación, esto no tendría por qué suceder, dado que X (manchas) no indica confiablemente Y (sarampión) sino Q (varicela), pero es tomada como una representación de Y.

Ahora bien, a propósito de la representación errónea, Crane (2008) indica:

Supóngase que soy capaz de reconocer borregos, soy capaz de percibir borregos cuando hay borregos a la vista. Mis percepciones de borregos son representaciones de algún género –llamémoslo representación S para abreviar- y son indicadores confiables de borregos, y por lo tanto la teoría dice que representan borregos. (...) Supóngase también que, en ciertas circunstancias (...) soy incapaz de distinguir borregos y cabras. Y supóngase que esta conexión es plenamente sistemática: hay una conexión confiable entre cabras-en-ciertas-circunstancias y percepciones de borregos. Tengo una representación S cuando veo una cabra. Este parece ser un caso de representación errónea. (p. 283)

En este caso, el énfasis del problema no está tanto en si existe o no un contenido representado. El énfasis está en la confusión entre dos contenidos y un mismo medio. Ciertamente, ambos problemas están unidos el uno al otro; de hecho, pueden ser pensados como dos caras de un solo problema (Crane, 2008). Sea como sea, lo importante es que en la

práctica es posible tener representaciones erróneas; en la teoría, sin embargo, es algo inadmisibles. Esto es así dado que algo representa a algo si y solo si existe una relación de causalidad entre ambas cosas. Si no hay tal relación no debería ser posible la representación, pero, como se ha indicado, sí lo es.

En síntesis, podría resumirse el problema así: si se supone que el medio representacional representa al contenido porque hay una relación causal entre ambos (indicación confiable y covariación de propiedades), ¿cómo es posible que una cosa represente a otra cosa con la cual no está unida causalmente? La respuesta de Fodor a este problema consistirá en introducir la noción de *dependencia asimétrica* (Skidelsky, 1999).

La idea de dependencia asimétrica es, en realidad, bastante simple. Se ha dicho que un medio X representa a Y porque existe entre ambos una relación de causalidad y covariación de propiedades. Esto implica que cualquier modificación en Y supondrá una modificación en X. Esto puede ser llamado dependencia simétrica o directa. De esta manera, la dependencia asimétrica o indirecta consistirá en la relación que tiene un contenido Q con un medio X sólo porque X representa a Y. Es decir, la relación entre X y Q está subordinada a la relación original entre X y Y.

Recurramos a uno de los ejemplos ya usados para alcanzar mayor claridad. Se ha dicho que en ciertas circunstancias las cabras pueden causar representaciones de borregos (representaciones S). Lo que propone Fodor es que “si hay condiciones en las cuales las cabras-en-ciertas-circunstancias también causan que tengamos representaciones S, tiene sentido suponer que las cabras hacen esto porque los borregos ya causan representaciones S” (Crane, 2008, p. 290). Entonces, el hecho de que una cabra (en circunstancias que la hagan semejante a un borrego) cause una representación de borrego, depende del hecho inicial de que los borregos causen representaciones de borrego. Para Skidelsky (1999), esto puede

resumirse así: “Para todo Y no igual a X, el que los Y que [en tanto] Y causen realmente <X> [representación de X] es asimétricamente dependiente de que los X causen <X>” (p. 443).

Esta alternativa soluciona el problema del error en tanto explica cómo es posible que algo represente a algo aunque no haya alguna relación de causalidad que medie entre ambos. No obstante, como el mismo Fodor reconoce, la dependencia asimétrica es sólo una condición suficiente para la representación mental (Crane, 2008). Es decir, “no afirma que cualquier posible clase de representación mental deba exhibir la estructura de dependencia asimétrica, sino que si algo exhibe de hecho esta estructura, entonces es una representación mental” (Crane, 2008, p. 292).

En este sentido, dado que una explicación además de condiciones suficientes debiera procurar condiciones necesarias, es decir, rasgos fundamentales, no queda claro cómo todo esto pueda suponer un avance significativo en lo relativo al problema de la representación. De ahí que el asunto siga siendo problemático y objeto de reflexión y discordia tanto en psicólogos como en filósofos.

### **Capítulo 3: Críticas al cognitivismo**

Durante los capítulos anteriores ha sido presentada tanto la historia que condujo al cognitivismo, como sus premisas y supuestos fundamentales. En este capítulo, se presentarán algunas de las críticas planteadas a dicho modelo, y las propuestas hechas para superar dichas críticas.

En primer lugar, se indicarán los señalamientos hechos desde el conexionismo, modelo igualmente computacional pero preocupado por tener una mayor validez fisiológica. Luego, se caracterizará la teoría de los modelos mentales, que aspira describir procesos aparentemente formales, como la deducción, desde una perspectiva semántica. Finalmente, se indicará la desvinculación contextual que supone el cognitivismo y la urgencia de una psicología cultural, centrada en el significado.

#### **1. El conexionismo**

El Cognitivismo, concentrado en describir la cognición humana en términos puramente sintácticos, formales, poco a poco fue tropezando en su proyecto. La comparación entre sistemas cognitivos artificiales y naturales, y la imposición de los mecanismos de funcionamiento de los primeros a los segundos, pasó de ser una solución prometedora para el problema de la cientificidad de la psicología, a una aspiración excesivamente abstracta y apartada de la realidad (Fernández, s.f.).

En efecto, el principal problema del cognitivismo consistió en no hallar una correspondencia entre sus premisas y los alcances diferenciales de la cognición humana y artificial. En otros términos, aunque en principio la comparación entre ambos sistemas resultara admisible, fue inevitable no advertir la amplia brecha que separaba las capacidades

de éstos. Así, si bien los computadores demostraron tener una competencia superior en actividades como el procesamiento numérico, son remotamente comparables con nuestra efectividad en tareas como la comprensión del habla, la planificación y el aprendizaje (Tienson, 1987).

Progresivamente, esto condujo a una sospecha, a saber: “que los cerebros humanos no hacen cosas de la manera convencional en la que los computadores lo hacen, que nuestra inteligencia es de una clase diferente” (Tienson, 1987, p. 361). Además de esta limitación comparativa, los precursores y simpatizantes del conexionismo (modelo con el cual se pretendió subsanar las fallas cognitivistas), recurren a dos objeciones más para argumentar la transición de un modelo a otro. Dichas objeciones son la lentitud neuronal y la degradación armoniosa (Tienson, 1987).

Es un hecho que la rapidez de las neuronas es bastante inferior si se compara con los computadores actuales. No obstante, “estamos capacitados para hacer un procesamiento muy sofisticado, tal como la recuperación de memoria, el procesamiento perceptual y el procesamiento lingüístico, en unos pocos cientos de milisegundos” (Tienson, 1987, p. 362). Esto significa que incluso con un sistema intrínsecamente lento, en comparación con los sistemas artificiales, somos bastante efectivos en la ejecución de tareas complejas. Esto ha llevado a suponer a algunos teóricos la llamada restricción de los 100 pasos, inicialmente planteada por Feldman (Tienson, 1987).

Según tal restricción, ninguna tarea cognitiva puede requerir, aproximadamente, más de 100 pasos secuenciales. Esta suposición es justificable si lo que se pretende es entender la velocidad del procesamiento humano. En otros términos, si una tarea implicase más de 100 pasos secuenciales, aproximadamente, con dificultad alcanzaría los niveles de velocidad típicamente humanos (Tienson, 1987).

Los sistemas artificiales difícilmente pueden cumplir con esta restricción; en tanto su procesamiento es sintáctico y secuencial, se ven reducidos a un tipo de cognición que demanda la ejecución de una gran cantidad de pasos. Tómese como ejemplo el diagrama de flujo planteado por Crane (2008) para dar cuenta del funcionamiento de los algoritmos, particularmente de un algoritmo para hervir un huevo:

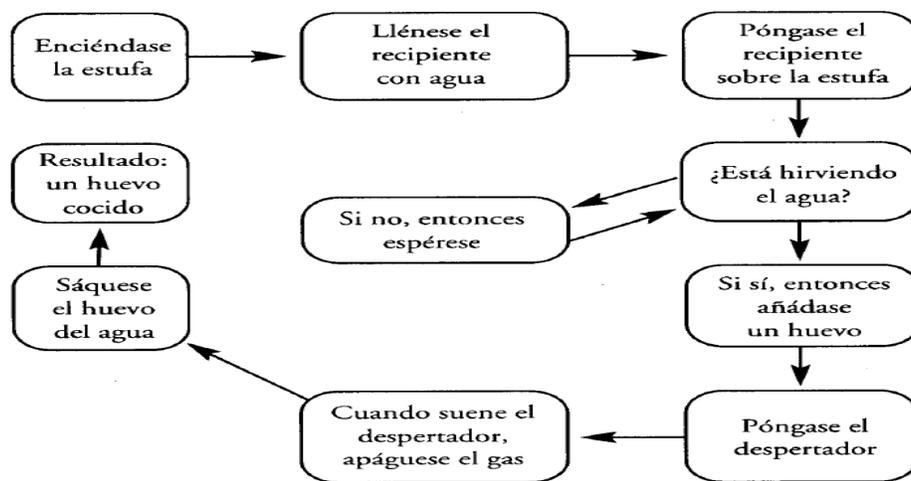


Figura 1: Algoritmo para hervir un huevo. Tomado de Crane, 2008, p. 154

Si este procedimiento demanda, por lo menos, siete pasos seriados, tareas cognitivamente complejas como la comprensión del lenguaje, por ejemplo, demandarían muchos más pasos que los 100 planteados por Feldman.

En suma, si la cognición humana fuese sintáctica y secuencial, teniendo en cuenta las limitaciones de velocidad del cerebro, sería imposible realizar en tan poco tiempo (como en efecto lo hacemos) tareas cognitivamente complejas. Lo anterior dado que dichas tareas demandarían una cantidad excesiva de pasos que el cerebro tardaría en ejecutar.

Por otro lado, la objeción de degradación armoniosa consiste en el hecho de que los sistemas naturales se deterioran de manera gradual tanto física como funcionalmente (Tienson, 1987). Así, “una lesión o una enfermedad cerebral puede afectar el funcionamiento,

pero no causar pérdida de recuerdos o de funciones específicas, antes de alcanzar un grado de severidad en el que se pierde una clase entera de capacidades” (Tienson, 1987, p. 362).

En el caso de los sistemas artificiales, “las lesiones causan la pérdida de funciones particulares y de todo lo que dependa de ellas” (Tienson, 1987, p. 362). Esto es así dado que en los sistemas artificiales una falla en el hardware compromete directamente la ejecución del software, imposibilitando su funcionamiento, incluso aunque se trate de una falla local. En otros términos, una falla estructural, por más mínima que sea, afecta el funcionamiento global del sistema. Además de esto, la pérdida de un fragmento del algoritmo del software también supondría una afectación general, de ahí que se planteen los sistemas artificiales como sistemas frágiles (Tienson, 1987).

Así pues, como señala Tienson, “la lentitud de las neuronas, la degradación armoniosa, y los dominios diferentes en los cuales cada cual está en sus mejores condiciones, sugieren que los sistemas naturales y los computadores convencionales hacen las cosas de maneras muy diferentes” (Tienson, 1987, p. 362). Es por esto que los teóricos del conexionismo optaron por una descripción de la cognición análoga al funcionamiento real del cerebro. A continuación serán presentadas las premisas básicas del modelo conexionista.

Se conocen ya algunas de las limitaciones del cognitivismo que funcionaron como motivación para el desarrollo del conexionismo. En general, puede decirse que el conexionismo “pretende alejarse del modelo del computador convencional y de lenguajes racionalista (...) para acercarse al modelo del cerebro animal y humano y a lenguajes como el de la neurofisiología” (Fernández, 1988, p. 25).

Para lograr tal transición, el primer paso consistió en redireccionar la atención puesta sobre los lenguajes formales y las representaciones hacia el funcionamiento de las neuronas. Así, la primera característica del conexionismo consiste en plantear que “el procesamiento

ocurre de modo simultáneo en una población de elementos simples” (Fernández, s.f., p. 30). Tales elementos simples, llamados nodos, poseen un funcionamiento semejante al de las neuronas. En suma, el primer rasgo del conexionismo es responsabilizar a una serie de estructuras simples el procesamiento cognitivo.

La segunda característica radica en plantear que el procesamiento cognitivo “consiste en que esas unidades elementales envíen sus señales de activación e inhibición a cada otra unidad” (Fernández, 1988, p. 30). Así pues, en términos de Tienson (1987), “cada nodo tiene conexiones dirigidas a varios otros nodos, de modo que obtiene señales de algunos nodos y envía señales a otros nodos, incluyendo, posiblemente, aquellos de los cuales obtiene señales” (p. 367).

En tercer lugar está el hecho de que el conocimiento es comprendido como la asociación y fuerza de conexión entre series de nodos (Fernández, 1988). Dicha fuerza de conexión es concebida como peso, así, un peso mayor significa menor resistencia en la activación de la red (Tienson, 1987). Esto implica que para el conexionismo no hay conocimientos ‘almacenados’; hay, sencillamente, patrones de conexión reforzados (Fernández, 1988).

Para ejemplificar la noción de peso, puede pensarse en la siguiente situación. Es probable que, luego de tomar una ducha, se siga un patrón definido para secarse aunque no haya conciencia de esto. El hecho de que, según este patrón, se inicie por el cabello (caso 1) y no por los pies (caso 2) obedece a que el patrón de conexión neuronal subyacente al caso 1 posee un mayor peso (probabilidad de ser emitido) que el patrón de conexión del caso 2. El que el comportamiento emitido sea el 1 y no el 2 es, en realidad, una arbitrariedad. Nadie nace con un patrón de conexión con mayor peso para un caso en particular (en el ejemplo descrito). Estos patrones se van reforzando con la experiencia.

La cuarta característica consiste en la concepción del aprendizaje como modificación del peso de las conexiones. Según esto, “aprender es modificar las conexiones de tal manera que se facilite el que un patrón de inputs antiguo evoque un patrón de outputs nuevo” (Fernández, 1988, p. 32). Una vez más, el aprendizaje deja de ser la adquisición de esquemas o representaciones y pasa a ser la reasignación de pesos a patrones de conexión entre redes de nodos.

En el caso de las redes artificiales, los pesos de conexión entre nodos pueden ser asignados por el programador. Usualmente son pesos de baja intensidad tomados al azar (Tienson, 1987). En el caso del cerebro humano, dichos ‘pesos’ corresponden a los umbrales de excitación necesarios para emitir el potencial de acción. Así, por ejemplo, neurofisiológicamente es posible dar una explicación de la memoria y el aprendizaje recurriendo a conceptos como la *potenciación sináptica a largo plazo*, el cual consiste en “un incremento sostenido de la eficacia de la transmisión sináptica tras estimular una vía aferente con estímulos de alta frecuencia” (Ortega & Franco, 2010). Dicho concepto es semejante a la noción de reasignación de pesos propia del conexionismo.

Ahora bien, una conexión simple puede ser representada de la siguiente manera, en donde N recibe una serie de inputs de otros nodos, emite un output hacia M, quien lo recibe y emite a su vez tantos outputs como conexiones posea. La W simboliza el peso (weight) de la conexión entre N y M.

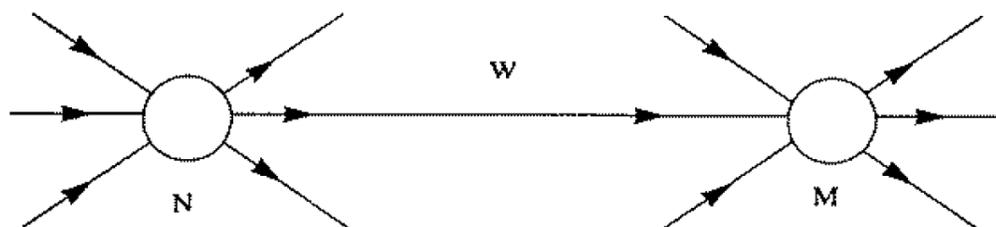
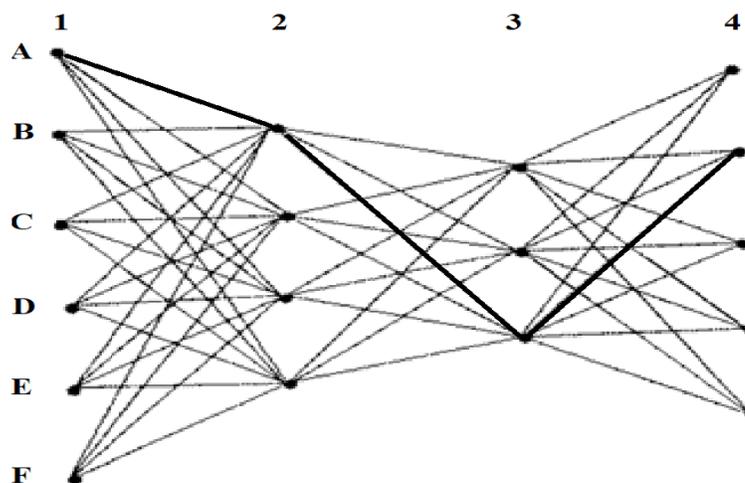


Figura 2: Conexión simple de nodos. Tomado de Tienson, 1987, p. 367

Una conexión compleja podría representarse de la siguiente manera, en donde los nodos de la capa 1 son de input y los de la capa 4 de output. Así, con la recepción de un input en la capa 1, la reacción en cadena se despliega por la capa 2 y 3 hasta llegar a la capa 4.



*Figura 3:* Red neuronal compleja. Tomado y adaptado de Tienson, 1987, p. 370

En este tipo de estructuras, el aprendizaje consistirá en la asignación de mayor peso a conexiones de nodos específicos, como la red de nodos 1A-2B-3D-4B. El mayor peso de esta red permitirá que ante una estimulación del nodo 1A sea más probable (menos resistente) la activación de los nodos 2B-3D-4B que otra combinación posible.

Ciertamente, la descripción de la cognición en estos términos supone una diferencia radical en relación a las premisas cognitivistas. Miremos ahora, brevemente, cuáles son las diferencias concretas y los puntos en que el conexionismo pretende subsanar las dificultades del cognitismo.

En primer lugar está el hecho de que la arquitectura conexionista es radicalmente distinta de la arquitectura cognitivista. Por arquitectura se entiende la conceptualización estructural del sujeto cognitivo. Así, tal como se planteó en el capítulo anterior, el cognitismo concebía al sujeto como un sistema que procesa información, constituido por

una memoria a largo plazo, un procesador central, efectores y receptores. A su vez, el conexionismo describe al sujeto como un sistema descentralizado, sin un procesador central en donde converjan todas las operaciones (Tienson, 1987). Para este modelo la manipulación de la información es local; cada nodo se ocupa de procesar unidades de información, siendo el resultado final producto del procesamiento en paralelo (Tienson, 1987).

Este tipo de arquitectura (la conexionista) soluciona los dos problemas planteados al inicio del apartado (la lentitud neuronal y la degradación armoniosa). En cuanto al primero, el hecho de que el procesamiento esté distribuido en paralelo, y no secuencialmente, posibilita al sistema ejecutar tareas complejas sin demandar una gran inversión de tiempo. Así, tal como señala Fernández (1988) un sistema intrínsecamente lento puede requerir poco tiempo en su funcionamiento “si dicho sistema consta de números masivos de unidades y conexiones en lugar de unas pocas macroestructuras que almacenan grandes cantidades de información simbólica y sobre las que opera un centro director del proceso” (p. 28).

Así mismo, en relación a la degradación armoniosa, el procesamiento descentralizado y autónomo de los nodos permite explicar por qué ante una lesión local no hay un compromiso general en el funcionamiento del sistema. En la arquitectura cognitivista, una lesión en alguna de las vías que transmite la información implica el deterioro inmediato de una función específica. Esto dado que el procesador central carecería de información relevante para la ejecución de una tarea. Peor sería el escenario en donde la lesión fuese directamente en el procesador central. En tal caso el sistema perdería su completa funcionalidad. La arquitectura conexionista está a salvo de tales problemas, esto dado que cada nodo tiene la capacidad de procesar por sí mismo fragmentos de información, sin necesidad de un procesador general (Tienson, 1987).

Ciertamente, el modelo cognitivista tuvo más dificultades de las ya señaladas. Algunas de ellas fueron el innatismo y el dualismo que parecía suponer (Tienson, 1987). De manera breve, puede decirse que el problema del innatismo deriva de la aparente independencia evolutiva de los algoritmos que se suponen fundamentan la cognición. En efecto, en autores como Fodor y Chomsky, existe la idea de un ‘lenguaje de máquina’, innato, a partir del cual el infante incluso se sirve para adquirir su lengua materna (Rodríguez, 1992). La promoción de un innatismo tal supone un gran problema, entre otras cosas, porque contraviene cualquier forma de psicología del desarrollo (tanto filogenético como ontogenético) (Rodríguez, 1992).

Por otro lado, si bien la idea de la superveniencia, detallada en el capítulo anterior, supone a los estados mentales una conexión necesaria con algún tipo de condición material (base de superveniencia), la idea de la autonomía funcional de los estados mentales hace sensato atribuirle al cognitivismo una condición emergentista más que monista (Ruiz, 2011). En otros términos, el hecho de que se conciba la mente como no reductible a interacciones puramente fisiológicas, y gobernada por leyes propias, le asigna una cierta independencia en relación a las condiciones materiales. Esta ‘cierta independencia’ ha sido nombrada emergentismo y para algunos puede incluirse dentro de los modelos monistas, pero para otros no es más que una suerte de dualismo residual o de propiedades (Ruiz, 2011).

Además de esto, ambos problemas se conectan entre sí (innatismo y dualismo). Esto es así dado que el hecho de que se suponga la existencia de un ‘lenguaje de máquina’ innato, que subyace a la cognición, abre una brecha entre el desarrollo de los procesos cognitivos y los procesos estrictamente biológicos, determinados por la evolución del organismo (Rodríguez, 1992).

En general, estas dificultades son mejor sobrellevadas por el modelo conexionista. Esto ya que dicho modelo posee mayor validez fisiológica (Fernández, 1988.). El conexionismo pone su atención en el funcionamiento del cerebro, de ahí que sus modelos sean más próximos al funcionamiento real de éste, en relación al cognitivismo, en algunos puntos excesivamente abstracto. En síntesis, para el conexionismo no hay porque suponer una autonomía funcional a la cognición. Ésta está directamente determinada por la interacción de los nodos o neuronas, de ahí que no sea algo que está ‘más allá’ de las condiciones fisiológicas y que, por lo tanto, obedece al desarrollo de las mismas.

Naturalmente, el conexionismo no está exento de críticas. Sea como sea, lo importante para este trabajo era señalar la forma como aborda lo que él mismo le reprocha al cognitivismo. Sus críticas están dirigidas, fundamentalmente, a la validez fisiológica de éste. No obstante, con esto no se agotan las críticas hechas al cognitivismo. Otro gran grupo de críticas están orientadas a un punto esencial de dicho modelo: su carácter sintáctico y formal. Veamos ahora el sentido amplio de dichas críticas.

## **2. Deducción y razonamiento semántico**

En términos generales, el enfoque cognitivista destaca ante todo las características lógicas, formales y sintácticas de la cognición humana. Esto se debe, según Rivière (1991), a la “propensión más racionalista” de los teóricos asociados a este enfoque (p. 140). No obstante, justamente dicho énfasis lógico-formal ha sido el principal blanco de objeción por parte de algunos críticos del Cognitivismo. A continuación, se presentará, en sentido amplio, la crítica hecha por Searle a este enfoque.

Searle (1994), con el argumento de la habitación china, que recrea una situación hipotética en la cual una persona, que desconoce el idioma chino, es capaz de responder a preguntas de hablantes nativos sirviéndose exclusivamente de un manual que contiene los

principios sintácticos para la manipulación de los símbolos de este lenguaje, lo que busca destacar es que “comprender un lenguaje, o ciertamente, tener estados mentales, incluye algo más que tener un puñado de símbolos formales”; implica, continúa, “tener una interpretación o un significado agregado a esos símbolos” (p. 39). En este sentido, la crítica de Searle (1994) puede resumirse, a mi juicio, con el siguiente pasaje:

Tener una mente es algo más que tener procesos formales o sintácticos. Nuestros estados mentales internos tienen, por definición, cierto tipo de contenidos. (...) Las mentes son semánticas, en el sentido de que tienen algo más que una estructura formal: tienen un contenido. (p. 37).

La justificación de esta crítica se halla en el hecho de que, como indica Searle (1994), aunque las respuestas dadas a los hablantes nativos sean acertadas, la manipulación sintáctica de símbolos no implica que haya una comprensión del idioma. De ahí que el autor afirme que tal comprensión demanda un componente puramente semántico, típico de la cognición humana, que los computadores están constitucionalmente imposibilitados de obtener, dado que “un computador tiene una sintaxis, pero no una semántica” (Searle, 1994, p. 39).

Si bien la crítica de Searle es ampliamente aceptada, el énfasis de este apartado no estará en ella por una sencilla razón: dicha crítica apunta hacia el problema de la conciencia, problema demasiado controversial y difícil como para querer evitar incluso su enunciación. Sea como sea, la idea de que la cognición es ante todo semántica, y no sintáctica como lo supone el cognitivismo, se expresó en varios campos de la psicología. Uno de ellos fue el campo del razonamiento, en donde halló gran validez experimental.

Históricamente, el estudio del razonamiento humano ha sido un campo de investigación básico en psicología. Entre otras cosas, esto se debe a la relevancia que posee el razonamiento dentro de la actividad cognitiva general, así como “al hecho de que su

estudio permite arrojar alguna luz en asuntos y polémicas tan importantes como el de la racionalidad o no del pensamiento humano” (García-Madruga, 1989, p. 324).

En relación a lo anterior, algunas investigaciones sobre el pensamiento han estado concentradas en “la búsqueda de correspondencias entre los modos en los que piensan los seres humanos y ciertos patrones normativos” (Aguirre & Moreira, 2016, p. 149). Específicamente, “en el caso del razonamiento deductivo interesó saber si los modelos provenientes de la lógica servían como modelo del razonamiento humano” (Aguirre & Moreira, 2016, p. 149).

El razonamiento deductivo puede ser definido como un tipo específico de operación en donde se “extraen conclusiones a partir de premisas o acontecimientos dados previamente” (García-Madruga, 1989, p. 324). A su vez, las conclusiones extraídas del razonamiento deductivo deben estar contenidas en las premisas, de ahí que sean *lógicamente* necesarias (García-Madruga, 1989).

Ahora bien, a propósito del problema de la primacía de la racionalidad en la cognición humana o no, las teorías explicativas del razonamiento tradicionalmente se han organizado en dos polos con énfasis distinto: en uno, se resalta el carácter lógico del razonamiento humano (Martín & Valiña, 2002) y en el otro, se destaca el papel determinante que tiene el contenido en los procesos de inferencia (Carretero & García-Madruga, 1995). Una de las teorías de naturaleza semántica con mayor aceptación es la teoría de los modelos mentales, expuesta por primera vez por Johnson-Laird (1983).

Como señalan Corral y Asencio (2014), la teoría de los modelos mentales “está basada en la idea de que la mente construye modelos internos del mundo externo, y asume como punto de partida que el sistema cognitivo no contiene reglas (...) sino que opera a partir del significado de las situaciones descriptas” (p. 106). Así pues, lo verdaderamente relevante

para el razonamiento, por ejemplo, es la construcción de modelos *análogos* del mundo, en donde el énfasis está en el contenido y no en la forma.

Un modelo mental consiste, como ya fue insinuado, en una representación análoga estructural del mundo (Corral & Asencio, 2014). En otros términos, los modelos mentales representan los objetos, sus propiedades y las relaciones de los mismos en situaciones concretas. Así, por ejemplo, ante la sentencia A es más grande que B y C, y B está a la derecha de A, a la izquierda de C, un modelo posible puede ser el siguiente:

A                      B                      C

*Figura 4:* Ejemplificación modelo mental. Elaboración propia.

Ahora bien, según esta teoría, la deducción está constituida por tres partes: 1) construcción de modelos mentales de las premisas, 2) formulación de conclusión tentativa y 3) búsqueda de modelos alternativos coherentes con las premisas y que falseen la conclusión (Corral & Asencio, 2014).

Específicamente, en el primer momento los sujetos construyen un modelo que dé cuenta de las condiciones relevantes de la situación descrita por las premisas. Dado que los modelos son conservados en la memoria de trabajo, idealmente deberán contener tanta información implícita como sea posible (Corral & Asencio, 2014). Así se evitará una ‘sobrecarga’, dadas las limitaciones de la memoria.

En segundo lugar, se derivará “una conclusión que describa de modo sencillo lo que está representado en el/los modelo/s, manteniendo la información semántica de las premisas y expresando algo que no se encuentre explícitamente en ellas” (Corral & Asencio, 2014, p. 108). Posteriormente se pasará al tercer momento, en donde el sujeto “buscará modelos

alternativos admitidos por las premisas, en los cuales la conclusión previamente obtenida es falsa” (Corral & Asencio, 2014, p. 108). Si es hallado un modelo que refute la conclusión obtenida, y que sea coherente con las premisas, entonces tal modelo se tomará como conclusión y será sometido al mismo escrutinio. Si no es hallado ningún modelo que falsee la conclusión, entonces ésta puede ser admitida como válida y necesaria (Corral & Asencio, 2014).

Así pues, la cuestión decisiva en relación al razonamiento deductivo, según la teoría de los modelos mentales, es la contrastación entre la conclusión y los modelos alternativos. Este procedimiento es puramente semántico, dado que lo que se corrobora es la coherencia de la conclusión con los elementos analógicamente representados por los modelos (Corral & Asencio, 2014).

Una de las principales inquietudes que ha motivado este tipo de modelos con énfasis semántico ha sido la sistemática tasa de errores en tareas de deducción, como en el caso del razonamiento silogístico (Santamaría, 1989). Esto generó una sospecha entre los teóricos del razonamiento. En concreto, la sospecha consistió en la no coherencia entre los resultados negativos de los sujetos en tareas deductivas y la supuesta primacía lógica de su razonamiento.

Todo esto condujo a la proliferación de teorías semánticas del razonamiento, como la teoría de los modelos mentales, en el razonamiento silogístico, o la teoría de razonamiento pragmático de Cheng y Holyoak en el razonamiento condicional (Valiña & Martín, 2014). Además de esto, el interés experimental por dichos fenómenos permitió hallar y conceptualizar variables determinantes del razonamiento, estrictamente semánticas, como los llamados *sesgos de creencias* (Corral & Asencio, 2014).

En efecto, se halló que las personas suelen dar respuestas que están determinadas por su credibilidad más que por la validez de las mismas (Corral & Asencio, 2014). Esto puede producir sesgos en el razonamiento, como el llamado *error de escrutinio selectivo*, en donde los sujetos aceptan una conclusión como lógicamente necesaria, sin someterla a corroboración, solo si es creíble, es decir, si no es incoherente (Corral & Asencio, 2014).

Desde la teoría de los modelos mentales, este error es explicado como un heurístico de prerrazonamiento, en donde los sujetos omiten el tercer momento de todo proceso de deducción, a saber, la búsqueda de modelos alternativos que refuten la conclusión (Corral & Asencio, 2014). Así, en síntesis, “cuando la primera conclusión es creíble, puede ser aceptada acríticamente, deteniéndose la revisión exhaustiva de modelos posibles” (Corral & Asencio, 2014, p. 115).

Todos estos avances, que demuestran el importante papel que juegan factores semánticos en la cognición, se han ido imponiendo como explicación aceptada de la cognición humana, haciendo a un lado las explicaciones cognitivistas, demasiado lógicas y con escaso apoyo empírico. Es así como, en general, las ciencias cognitivas, incluyendo la psicología, optaron por un viraje hacia los factores semánticos y contextuales del razonamiento, viraje que nos permitirá concluir este trabajo.

### **3. Significado y Psicología popular**

Para Bruner (1991), es paradójico que la psicología haya terminado ocupándose del procesamiento de la información, y que adoptara, además, la computabilidad como criterio fundamental para la explicación de la cognición. Lo paradójico consiste en el viraje que tomó la llamada revolución cognitiva, que pretendía rescatar los conceptos mentalistas, inadmisibles para el conductismo, hacia un nuevo tipo de reduccionismo (Bruner, 1991).

Concretamente, se pasó de un modelo concentrado en los estímulos y las respuestas, a un modelo enfocado en la relación *input-output*.

Con la aparición del paradigma cognitivista y su nueva forma de reduccionismo, “la computación se convirtió en el modelo de la mente, y en el lugar que ocupaba el concepto de significado se instaló el concepto de computabilidad” (Bruner, 1991, p. 23). Esto condujo a entender ‘la mente’, en sentido subjetivo, como un epifenómeno del sistema computacional, privándola de toda condición explicativa y causal en relación comportamiento (Bruner, 1991).

Todo esto condujo a la psicología a una suerte de *ensimismamiento*. Con esto lo que se intenta decir es que la psicología se concentró ante todo en el descubrimiento de los mecanismos computacionales de la cognición, sustrayendo al sujeto de su contexto de tal manera que se hicieron justificables, como se señaló hace poco, posiciones abiertamente innatistas. Para Bruner (1991), esta situación supuso un aislamiento del sujeto en relación a los factores culturales, también determinantes de la cognición.

Así pues, el impulso inicial de la revolución cognitiva, impulso que buscaba reivindicar los contenidos mentales, el significado, se desvió hacia los formalismos computacionales. De ahí que Bruner (1991) reclame una psicología cultural, en donde el significado sea el punto central, en donde se conciba la cultura no como una capa superpuesta a las condiciones biológicas, sino como la verdadera explicación del comportamiento humano.

La búsqueda de una psicología cultural y del significado tiene, por lo menos, tres justificaciones para Bruner (1991). La primera de ellas consiste en la admisión, difícil para la psicología pero bien conocida para otras humanidades como la antropología, de la cultura como contexto natural del hombre y a través del cual se desarrolla la cognición. La segunda

consiste en la concepción del significado, su construcción y utilización, como elemento conector entre el hombre y la cultura. Y la tercera consiste en reconocer la psicología popular como la condensación de la cultura y el significado. A continuación serán ampliados estos tres elementos.

Que Bruner abogue por una psicología cultural, en donde la cultura sea la variable determinante de la cognición humana, exige una conceptualización de lo que debe entenderse por cultura. Así, según Salazar (2009), la concepción de Bruner sobre la cultura puede ser reducida a dos elementos asociados entre sí: la concepción de la cultura como sistema simbólico compartido y como “consenso alrededor de determinadas creencias, valoraciones y prácticas” (p. 58).

El hecho de que la cultura, es decir, los sistemas simbólicos y las creencias y prácticas compartidas, determine la cognición humana, se debe a que justo en virtud de ella es posible la construcción de significados compartidos (Salazar, 2009).

La idea central de esta propuesta es que la cognición humana es ante todo significativa, o en otros términos, cargada de contenidos. Ahora, la forma como el individuo adquiere dichos contenidos es a través de su participación directa en la cultura, en donde los contenidos ya están dados previamente tanto en el lenguaje como en las tradiciones (Esteban, 2002). Todo esto, en palabras de Esteban (2009), puede ser planteado de la siguiente manera:

El auténtico aprendizaje, en este sentido, es ese que viene determinado por la significatividad, pero no aislada, sino compartida. Un mundo común de significados que no necesariamente puede ser objetivado y por tanto circunscrito o impuesto desde estrategias científicas reduccionistas. Un mundo en el que el sentido se negocia, se dialoga, se expone, se narra, se interpreta, se

explica, se comprende. Un mundo cultural de valores y significados que conforma, a través del relato, el desarrollo del individuo. (p. 256)

Así pues, la adquisición del significado, entendiendo por tal la carga semántica, representacional, de toda actividad humana, viene dada desde el exterior, desde la interacción del sujeto con todo el sistema simbólico que lo rodea. En efecto, “el niño entra en la vida de su grupo como participante de un proceso público de significados no mediante el ejercicio privado de procesos primarios, sino mediante un ejercicio vital de interpretación significativa compartida” (Esteban, 2009, p. 257).

Naturalmente, el lenguaje como sistema simbólico compartido facilita la ‘negociación’ de significados. No obstante, para Bruner existe una disposición intrínseca en el ser humano hacia la interpretación, es decir, hacia la atribución e identificación de significados (Bruner, 1991). Dicha disposición facilita a los infantes participar en toda la urdimbre de contenidos que es la cultura, incluso antes de adquirir un sistema lingüístico. En concreto, se trataría de una “predisposición cultural protolingüística para el significado” (Esteban, 2009, p. 258).

Dicha predisposición protolingüística está fundamentada en cuatro “rasgos gramaticales/léxicos/prosódicos” (Bruner, 1991, p. 85). Dichos rasgos son: “el interés agentivo por la interacción humana, la predisposición temprana por marcar lo inusual frente a lo usual, la secuenciación lineal de acontecimientos y la presencia de un narrador” (Esteban, 2009, p. 259).

Desarrollar estos rasgos tomaría un esfuerzo innecesario para este apartado. Esto en la medida en que lo que se pretende es presentar el giro propuesto por Bruner hacia el significado. No obstante, sí es importante señalar que la predisposición protolingüística hacia el significado puede ser comprendida, además, como una forma rudimentaria de psicología

popular, es decir, como “una capacidad natural para atribuir intencionalidad” (Salazar, 2009, p. 49).

En efecto, a juicio de Bruner (1991) la psicología popular debe ser el punto de organización de la psicología científica. Esto es así dado que es desde la psicología popular que se participa en la cultura (Bruner, 1991). Si no hubiese una atribución recíproca de metas, propósitos, creencias, etc., semejantes a las personales, difícilmente se podría participar en los significados compartidos de todo sistema cultural.

Ahora bien, tal atribución de metas, propósitos, creencias, y demás estados intencionales, parte de la sospecha, constitutiva de la psicología popular, de que los comportamientos tienen una causa mental (Rivière, 1989). No obstante, la atribución de estados mentales a partir de la percepción de comportamientos demanda un proceso interpretativo, de búsqueda de significado. Es esto “lo que hace que la interpretación y el significado sean fundamentales en la psicología cultural, o en cualquier psicología o ciencia de lo mental” (Bruner, 1991, p. 34).

Así pues, el significado adquiere un valor nuclear para la psicología en la medida en que posibilita lo que Rivière (1989) llamó el núcleo de la psicología popular, a saber: la atribución de estados mentales con valor causal en relación a la conducta. A su vez, la psicología popular adquiere un valor explicativo en tanto constituye el instrumento mediador entre el individuo y la cultura.

Este nuevo viraje, que apunta hacia el mismo sentido en que inicialmente apuntó la revolución cognitiva, pretende superar el ensimismamiento del cognitivismo. Así pues, tanto la demanda conexionista de un modelo más coherente con el funcionamiento del cerebro, el énfasis en la semántica como mecanismo básico de la cognición, y la reivindicación de la cultura como centro de análisis del comportamiento humano, permitieron superar,

progresivamente, las premisas cognitivistas, demasiado entusiasmadas con el sueño de la computación.

## Conclusiones

Durante las páginas precedentes fueron presentados algunos aspectos relevantes para la comprensión del llamado modelo cognitivista. En concreto, fueron señalados algunos hitos históricos determinantes para la consolidación del cognitivismo. Además de esto, se exploraron las premisas básicas de dicho modelo, haciendo énfasis en su visión computacional-representacional de la cognición y del sujeto cognitivo. Por último, se revisaron algunos de los señalamientos hechos a dicho modelo.

Dado que el objetivo de este trabajo consistió en presentar el modelo cognitivista de una manera puramente descriptiva, sin realizar valoraciones positivas o negativas sobre sus premisas o alcances, las conclusiones que se desarrollarán a continuación también serán estrictamente descriptivas. Con esto lo que se quiere advertir es que los hallazgos realizados son producto de los planteamientos hallados en la bibliografía consultada, y no de un proceso de experimentación con argumentos a favor o en contra del problema estudiado. Las conclusiones serán planteadas por capítulo, para luego dar una conclusión final de todo el trabajo.

Del primero capítulo se pueden concluir, esencialmente, dos cosas. La primera es relativa a la psicología fisiológica de Wundt, y consiste en destacar que si bien su modelo supone un primer esfuerzo para alcanzar la rigurosidad científica en psicología, en último término se ve enfrentado al problema del dualismo de sustancias. Esto fue así debido a la divergencia entre objeto y método de estudio. El hecho de no poder conciliar los constructos teóricos admitidos dentro de la competencia de la psicología (como pensamiento, lenguaje, etc.) con el método propuesto (la estimulación fisiológica), implicó la reapertura del viejo problema de la interacción mente-cuerpo.

No obstante, lo cierto es que con Wundt se avanzó en la solución de este problema. Lo anterior dado que el dualismo hallado en el modelo de Wundt no es un dualismo de sustancias, es decir, de realidades independientes. Se trata de un paralelismo psicofísico, en donde la mente y el cuerpo se conciben como dos elementos de una sola experiencia, realidad o sustancia.

La segunda conclusión extraíble del primer capítulo está asociada al conductismo. De éste puede decirse que supuso un cambio epistemológico al interior de la psicología. Así fue en la medida en que hizo objeto central de su interés un elemento que hasta entonces había sido periférico: la conducta. Este hecho se vio respaldado por una reconceptualización operante de conceptos mentalistas como *propósito* o *intención*.

Si bien con el conductismo se resolvió el problema dualista presente en Wundt, en tanto se eliminaron conceptos mentalistas con dudosa correspondencia empírica, se cayó en un reduccionismo, para algunos excesivo, que imposibilitó avanzar en la comprensión de elementos indudablemente psicológicos como la conciencia, irreductibles a conductas y contingencias de refuerzo. Este hecho fue el que precipitó la aparición de las ciencias cognitivas, y con ellas del cognitivismo como modelo imperante.

Del segundo capítulo son varias las conclusiones posibles. En primer lugar, puede decirse que el surgimiento del cognitivismo obedeció a los avances en computación e informática, además del reduccionismo conductista. Así, el que la computación estuviese en su apogeo facilitó a los investigadores ver en ella semejanzas tentadoras con la cognición humana, a tal punto de imponerle a ésta los principios de aquella, más claros e inteligibles.

Todo esto condujo tanto a avances como a dificultades. El principal avance fue conceder mayor claridad al problema del dualismo de sustancias. Esto fue así dado que no existía duda, en relación a los computadores, de que un sistema simbólico y abstracto podía interactuar con otro concreto y materialmente determinado. Así, no había por qué suponer que la relación mente-cuerpo fuese distinta a la relación software-hardware, de ahí que el problema del dualismo dejara de ser algo misterioso y desconcertante.

El precio de esta equivalencia fue imponer una visión logicista a la cognición. Dicha visión implicó un distanciamiento de la naturaleza esencialmente semántica y contextual del sujeto cognitivo, viéndose arrojado, paradójicamente, a un nuevo tipo de reduccionismo. Esto implicó una creciente insatisfacción con el modelo cognitivista, demasiado abstracto y solipsista, es decir, ajeno los determinantes ambientales.

Del capítulo tres, en donde fueron exploradas las críticas hechas al cognitivism, puede concluirse lo siguiente. Primero, que la solución dada por el cognitivism al problema dualista en realidad dejó una especie de dualismo residual. Esto dado que las premisas funcionalistas, nucleares en el cognitivism, permiten suponer una cierta independencia entre la cognición y las condiciones materiales sobre las que se realiza. Esta limitación fue subsanada por el conexionismo, modelo que ubica la cognición en los vínculos establecidos entre redes de nodos, semejantes al funcionamiento neuronal, y no en un plano funcionalmente autónomo. Esto confirió mayor validez fisiológica al conexionismo, por oposición a la abstracción e indeterminación cognitivista.

Además de lo anterior, también puede concluirse de este capítulo que existe en la actualidad una tendencia a resaltar las variables semánticas de la cognición en lugar de las formales (como lo hacía el cognitivism). Esta tendencia puede notarse tanto en la

investigación de procesos psicológicos como el razonamiento, así como en la comprensión general de las variables determinantes de la cognición, como en el caso de la psicología cultural propuesta por Bruner.

De todo lo anterior es posible extraer tres conclusiones generales. La primera es que el cognitivismo se enmarcó dentro de los numerosos esfuerzos de la psicología por hallar un lugar aceptable al interior de las disciplinas científicas. La prueba de esto fue su interés fundamental por resolver el problema del dualismo de sustancias, sirviéndose de la metáfora del ordenador como medio para lograr tal objetivo.

La segunda conclusión general no es más que una consecuencia de la primera. Se trata del carácter lógico-formalista que tuvo que aceptar el cognitivismo en virtud de la adopción de la metáfora del ordenador. Este hecho lo condujo a conceptualizaciones demasiado abstractas y difíciles de conciliar con las condiciones tanto fisiológicas como semánticas y contextuales de la cognición.

De ahí que la tercera conclusión consista en la progresiva renuncia por parte de los teóricos cognitivos a la dureza del cognitivismo, optando por modelos más sensibles a las variables ambientales y de contenido, así como más coherentes con las determinaciones físicas del sujeto cognitivo.

De todo esto puede plantearse una pregunta final: ¿es sensato adoptar una versión fuerte de la metáfora del ordenador o en su lugar debe aceptarse una versión débil, en donde la comparación sea hecha en los fines (el procesamiento de la información) y no en los medios (los algoritmos de cómputo)? Naturalmente, la respuesta a este problema queda en manos de las futuras investigaciones competentes a este campo específico.

### Referencias:

- Aguirre, R., & Moreira, K. (2016). Pensamiento y lenguaje. En A. Vásquez, Manual de introducción a la psicología cognitiva (págs. 147-178). Montevideo: Unidad de Comunicación de la Universidad de la República.
- Araujo, S. (2007). Actualidad de la posición de Wilhelm Wundt respecto al problema de mente - cuerpo. *Persona*, 99-108.
- Bechtel, W. (1988). *Filosofía de la mente*. Madrid: EDITORIAL TECNOS, S.A.
- Block, N. (2015). El funcionalismo. *Cuadernos Filosóficos. Segunda Época, XII*, 118-132.
- Bruner, J. (2009). *Actos de significado: Más allá de la revolución cognitiva*. Madrid: Alianza Editorial.
- Corral, N., & Asensio, M. (2014). Modelos mentales: una aproximación semántica al razonamiento. En M. Carretero, & M. Asensio, *Psicología del pensamiento. Teoría y prácticas* (págs. 103-132). Madrid: Alianza Editorial.
- Crane, T. (2008). *La mente mecánica. Introducción filosófica a mentes, máquinas y representación mental*. México D.F.: Fondo de cultura económica.
- Danziger, K. (1980). La historia de la introspección revisada. *Journal of the History of the Behavioral Sciences*, 241-262.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge: MIT Press.
- Esteban, J. (2002). La teoría narrativa de Bruner y sus implicaciones en una pedagogía hermenéutica. *Cultura y Educación*, 253-265.
- Fernández, J. (1988). El conexionismo. *Aldaba*, 25-40.

- Freixa, E. (2003). ¿Qué es conducta? *Revista Internacional de Psicología Clínica y de la Salud*, 595-613.
- García, J. (1982). Razonamiento silogístico e interpretación de premisas: un estudio evolutivo. *Infancia y Aprendizaje*, 175-204.
- González, R. (2011). Máquinas sin engranajes y cuerpos sin mentes. ¿Cuán dualista es el funcionalismo de Máquina de Turing? *Revista de Filosofía*, 183-200.
- Guerra, R. (1996). Ontología y Metafísica. *Theoría. Revista del Colegio de Filosofía*, 11-24.
- Johnson-Laird, P. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Harvard Univ. Press.
- Martínez- Freire, P. (2005). *La importancia del conocimiento*. La coruña: Netbiblo, S. L. .
- Martínez, A., López-Espinoza, A., Aguilera, V., Galindo, A., & De la torre-Ibarra, C. (2007). Observación y experimentación en psicología: una revisión histórica. *Diversitas: Perspectivas en Psicología*, 213-225.
- Martínez-Freire, P. (1995). *La nueva filosofía de la mente*. Barcelona: Editorial Gedisa.
- Moreno, K. (2011). Wittgenstein y la naturalización de la mente . *Tesis Psicológica*, 183-200.
- Moya, C. J. (2004). *Filosofía de la mente*. Valencia: Universitat de València.
- Neisser, U. (1967). *Psicología Cognoscitiva*. Bogotá: Trillas.

- Novoa, M. (2002). Algunas consideraciones sobre el dualismo en psicología. *Universitas Psychologica*, 71-80.
- Novoa, M. (2002). Algunas consideraciones sobre el dualismo en psicología. *Universitas Psychologica*, 71-80.
- Ortega, C., & Franco, J. (2010). Neurofisiología del aprendizaje y la memoria. Plasticidad Neuronal. *Archivos de medicina*, 22-29.
- Peña, T. (2010). ¿Es viable el conductismo en el siglo XXI? *LIBERABIT*, 125-130.
- Perner, J. (1994). *Comprender la mente representacional*. Barcelona: Ediciones Paidós.
- Putnam, H. (1967). Psychological predicates. En W. Capitan, & D. Merrill, *Art, Mind, and Religion* (págs. 37-48). Pittsburgh: University of Pittsburgh Press.
- Quiñones, E., & Ato, M. (1983). Wundt y la psicología cognitiva . *Anales de la Universidad de Murcia. Filosofía y Ciencias de la Educación*, 133-147.
- Ramos, J. (2001). Confusiones mentales acerca de lo mental. En J. Botero, *El pensamiento de L. Wittgenstein* (págs. 201-220). Bogotá: Unibiblos.
- Rivière, A. (1987). *El sujeto de la Psicología Cognitiva*. Madrid: Alianza.
- Rivière, A. (1989). Más a favor de la psicología popular. *Cognitiva*, 261-265.
- Rivière, A. (2003). *Obras Escogidas (Vol. I). Diálogos sobre Psicología. De los cómputos mentales a los significados de la conciencia*. Madrid: Médica Panamericana.
- Rodríguez, C. (1992). El lenguaje del pensamiento como lenguaje privado. Una crítica wittgensteiniana al innatismo de Fodor. *Psicothema*, 133-152.

- Ruiz, P. (2011). Filosofía de la mente; aportes teóricos y experimentales a la visión emergentista del vínculo mente-cuerpo. *Cuadernos de Neuropsicología*, 111-127.
- Salazar, J. (2009). *Jerome Bruner: mente, conocimiento y cultura* (Tesis de maestría). Universidad de Antioquia, Medellín, Colombia.
- Santamaria, C. (1989). Modelos mentales y razonamiento semántico: el silogismo. *Cognitiva*, 21-36.
- Skidelsky, L. (1999). Teoría semántica de la dependencia asimétrica: naturalización y reducción. *Epistemología e historia de la ciencia*, 440-445.
- Skidelsky, L. (2003). *Representaciones mentales: ¿Eslabón entre el individuo y el mundo?* Buenos Aires: FiloUba.
- Skinner, B. (1994). *Sobre el Conductismo*. Buenos Aires: Planeta Argentina.
- Tienson, J. (1987). Una introducción al conexionismo. En E. Rabossi, *Filosofía de la mente y ciencia cognitiva* (págs. 359-380). Barcelona: Ediciones Paidós Ibérica.
- Tomasini, A. (1990). Conductismo y filosofía. *Revista Mexicana de Análisis de la Conducta*, 53-63.
- Tortosa, F., & Civera, C. (2006). *Historia de la Psicología*. Madrid: McGRAW-HILL.
- Turing, A. (1997). Maquinaria computacional e inteligencia. En J. Haugeland, *Mind Desing II* (págs. 187-216). Cambridge: MIT Press.
- Villar, F. (2001). *Psicología Evolutiva y Psicología de la Educación*. Barcelona: Universitat de Barcelona.

Villena, D. (2017). ¿Qué es el funcionalismo? *Letras*, 129-155.

Wittgenstein, L. (2017). *Investigaciones filosóficas*. Madrid: CreateSpace.

Wundt, W. (1902). *Compendio de psicología*. Madrid: La España moderna.

Yela, M. (1996). La evolución del conductismo. *Psicothema*, 165-186.

Zumalabe, J. (2012). La transición del conductismo al cognitivismo. *eduPsykhé*, 89-111.