

**UNIVERSIDAD DE CHILE**  
Facultad de Filosofía y Humanidades  
Escuela de Postgrado  
Centro de Estudios Cognitivos (CEC)

# **SURGIMIENTO DE NUEVAS HIPÓTESIS POR RAZONAMIENTO BASADO EN MODELOS**

## **EXTERNALISMO ACTIVO EN ENFOQUES ACTUALES**

Tesis para optar al grado de Magíster en Estudios Cognitivos

AUTOR:

**Bernardo Antonio Pino Rojas**

PROFESOR GUÍA: Guido Vallejos Oportot

**Santiago, Chile 2011**



Epígrafe . .	4
AGRADECIMIENTOS . .	5
RESUMEN . .	6
INTRODUCCIÓN . .	7
CAPITULO I . .	13
1.1. RBM y el estudio de las prácticas científicas en ambiente real . .	13
1.1.1. Algunas nociones y distinciones preliminares . .	13
1.1.2. El quehacer científico como actividad cognitiva y el RBM . .	30
1.1.3. Esquema provisional de la externalización cognitiva en RBM . .	34
1.1.4. Principio del externalismo activo . .	45
CAPITULO II . .	48
2.1. Representación modélica y RBM . .	48
2.1.1. Modelos en la perspectiva de un razonamiento centrado en el agente . .	48
2.1.2. La persistencia metodológica de las representaciones externas . .	62
2.1.3. Modelos en la perspectiva de un razonamiento centrado en modelos . .	69
CAPITULO III . .	76
3.1. Formación de hipótesis por RBM . .	76
3.1.1. Hacia la propuesta de los “sistemas buscadores de oportunidades” . .	76
3.1.2. RBM como un método de inferencia y delegación cognitiva. . .	85
CAPITULO IV . .	92
4.1. Limitaciones y alcances de una teoría del surgimiento de hipótesis por medio de RBM . .	92
4.1.1. La tesis del RBM en el contexto de la interacción entre ciencias cognitivas y filosofía de la ciencia . .	92
4.1.2. Criterios para los estreñimientos de un RBM externalista . .	104
4.1.3. Condiciones de adecuación para una teoría del surgimiento de hipótesis por RBM . .	116
CONCLUSIONES . .	120
REFERENCIAS . .	125

## Epígrafe

***“...the cognitive sciences might come to play the sort of role that formal logic played for logical empiricism or that history of science played for the historical school within the philosophy of science. This developmet might permit the philosophy of science as a whole finally to move beyond the division between ‘logical’and ‘historical’ approaches that has characterized the field since the 1960s” – Giere, 1992, p. xv***

---

## AGRADECIMIENTOS

Este trabajo de tesis fue financiado por el proyecto FONDECYT número 1095020. Además, contó con el importante apoyo de la Beca de Vicerrectoría de Asuntos Académicos 2008 para pasantías cortas de investigación de la Universidad de Chile, obtenida como alumno tesista de postgrado y llevada a cabo en el *Computational Philosophy Laboratory* de la *Università degli Studi di Pavia*, Pavía, Italia.

El autor agradece el honesto, paciente y generoso apoyo que el supervisor de tesis, profesor Guido Vallejos, le ha brindado no sólo durante la realización de esta tesis, sino que de manera permanente a lo largo del programa de Magíster en Estudios Cognitivos. Es justo destacar que, como profesor y coordinador de este programa, don Guido nunca ha dejado de ofrecer oportunidades a quienes emprenden la temeraria tarea de querer entender más y mejor acerca de las materias y problemas a los que él mismo les ha dedicado gran parte de su vida académica. En este último ámbito, el filósofo ha sido una fuente de inspiración constante, lo que merece un sincero aprecio, a pesar de todas las rabias que pueda haber pasado debido a las vacilaciones y las ineptitudes del autor para avanzar más definitivamente en la educación que ha tenido el privilegio de recibir.

Del mismo modo, vaya un profundo agradecimiento al Dr. Lorenzo Magnani, quien junto a su familia y al equipo del Laboratorio de Filosofía Computacional en Pavía, hicieron todo lo posible para que la estadía del autor como alumno pasante se convirtiera en una preciosa experiencia de vida, dentro y fuera del Departamento de Filosofía. Gracias a Emanuele Bardone por compartir un ánimo común de debate productivo y apasionado que, aunque desde posturas intelectuales opuestas, ciertamente motivó aspectos centrales de esta tesis. A Tommaso Bertolotti y su familia se les estará siempre en deuda por su hospitalidad infinita. Nicola, Ambra y Serena completan el círculo de amigos de quienes el autor fue objeto de una espontánea y oportuna amabilidad. El autor espera con ansias la ocasión en que pueda volver a reunirse con todos ellos alguna vez.

La redacción de ninguno de los tres borradores que decantaron en esta tesis habría sido posible si el autor no hubiese contado con cierta flexibilidad en sus funciones laborales, pertinentemente coordinadas junto a su jefe, don Christopher Johnson. Muchas otras personas y amigos han contribuido con palabras de aliento permanentes para persistir en la tarea de llegar a puerto con esta esquivia investigación. Una de ellas es nuestra amiga y ex compañera de postgrado Soledad Vargas Ossa (“la jefa”).

La familia ha estado siempre al lado del autor, aún cuando el autor no estuvo con ellos. Es importante expresar gratitud por el esfuerzo de todos los familiares que de una u otra manera han contribuido a que el autor pueda estar alcanzando metas largamente anheladas. Varios ya no están, o están de alguna forma que el autor no logra comprender. De entre todos los que quedan, Claudia y Coni - esposa e hija, respectivamente - merecen el más especial de los reconocimientos. Constanza es la persona que el autor más ama en el mundo, pero Claudia la sigue muy de cerca, especialmente ahora que, sólo unos momentos antes de comenzar a redactar a estos agradecimientos, la mengana le dice al fulano que nuevamente será papá.

## RESUMEN

El tema general de esta tesis son los enfoques externalistas del razonamiento basado en modelos (RBM) y el rol que este tipo de razonamiento pueda cumplir en las investigaciones que lo sitúan al centro de la generación de hipótesis y el razonamiento creativo. Específicamente, en este trabajo se llevó a cabo un escrutinio crítico y clarificador de los planteamientos centrales de aquellas investigaciones conjuntas realizadas por filósofos y científicos cognitivos que conciben la actividad científica como un complejo proceso cognitivo susceptible de descripción y explicación. Una característica de los enfoques de RBM externalista es que asumen - con un grado variable de compromiso metodológico - aproximaciones experimentales que orientan la explicación hacia las prácticas científicas que ocurren en ambiente real. En este sentido, se ha intentado determinar criterios para constreñimientos de un RBM, probablemente complementarios a los de una inferencia de orden proposicional, que sirvan para establecer las condiciones de una posible teoría sistemática de la generación de hipótesis. En conformidad a lo anterior, se planteó una hipótesis que evalúa el posible carácter de sistema inferencial distribuido del RBM, y la viabilidad de la noción de agente epistémico autónomo, en tanto re/presentación artefactual, que actúe como una concepción modélica unificadora del RBM externalista. El supuesto que sirve al autor como referente organizador de la investigación es que muchas de las confusiones vinculadas a una probable teoría del RBM son el resultado de una inversión pragmatista de nociones metodológicas, entendida como un efecto nocivo de la interacción entre ciencia cognitiva y filosofía de la ciencia.

---

# INTRODUCCIÓN

El presente trabajo<sup>1</sup> cae dentro del ámbito de la filosofía de la ciencia, y constituye una exploración aclaratoria del rol que un Razonamiento Basado en Modelo (RBM, en adelante) pueda cumplir en aquellas investigaciones que sitúan este tipo de razonamiento en el centro de la generación de hipótesis y el descubrimiento científicos. Con el fin de delinear una lógica del descubrimiento en ciencia, estos dos últimos fenómenos han sido tradicionalmente tratados, sin éxito, sobre la base de un proceso inferencial abductivo<sup>2</sup>, formulado proposicionalmente y guiado por reglas explícitas. En esa línea, el estudio crítico de las propuestas conocidas<sup>3</sup> ha dado pie a la búsqueda de una teoría sistemática y alternativa de generación de hipótesis que, entre otros rasgos, pretende no limitarse solamente a la consideración de constreñimientos explícitos, sino también a otros de carácter implícito. Dicha tarea sería susceptible de ser abordada, particularmente, por medio de un RBM.

A partir de esa última tesis sobre RBM (que un RBM puede dar cuenta de los constreñimientos relevantes de la generación de hipótesis y el descubrimiento científicos), se examinarán críticamente algunos planteamientos y orientaciones centrales realizados por un grupo de filósofos experimentales y científicos cognitivos (e.g. Nersessian, Magnani, Knuutila, Gooding, entre otros). Estos planteamientos tienen que ver, por ejemplo, con el rol artefactual de los modelos, su estatus representacional, y el RBM como método de inferencia productivo, desde la perspectiva de la interacción entre Ciencia Cognitiva y la Filosofía de la Ciencia. Gracias a los aportes del RBM surgidos en el contexto de tales materias de discusión y experimentación – según se argumenta – sería posible una sistematización del rol de los constreñimientos implícitos supuestamente participantes en la formación de hipótesis por medio de un RBM.

Específicamente, se dará cuenta de aquellos enfoques del RBM que guardan mayores compromisos externalistas al proporcionar una explicación enfocada en *las prácticas científicas en ambiente real*. Como se verá, dichos compromisos externalistas hacen referencia a aquellas aproximaciones de base experimental según las cuales el ámbito de la agencia cognitiva desborda oportunistamente los límites establecidos por una postura individualista tradicional, desdibujando una demarcación precisa entre lo neuronal, lo corporal y lo ambiental, en tanto componentes del sustrato material determinante de procesos cognitivos específicos (ver Clark, 1997 y 2008; Hutchins, 1995; van Gelder, 1999, entre otros).

<sup>1</sup> Trabajo de tesis de postgrado financiada por proyecto FONDECYT número 1095020

<sup>2</sup> Dichos tratamientos se constituyen a partir de las formalizaciones inicialmente desarrolladas por Peirce (1960) y las consiguientes propuestas para una lógica del descubrimiento de Hanson (1958, 1960) y Harman (1966).

<sup>3</sup> Entiéndase por esto tanto aquellas que proponen la representación lógica de procesos inferenciales para justificar la hipótesis abducida (e.g. Gabbay et al. 2005, 2006); o la identificación de algún valor epistémico de dicha hipótesis agregando premisas adicionales al esquema original de Peirce para justificarla como “la mejor explicación” (e.g. Harman, 1966); o bien, la implementación computacional del razonamiento abductivo concebido como la satisfacción de ciertos constreñimientos de carácter formal (e.g. Thagard, 1992).

Una motivación central del presente trabajo es poder contribuir a la identificación de criterios sistemáticos e inequívocos de constreñimientos para un RBM externalista, y a la identificación de condiciones de adecuación mínimas para una teoría del surgimiento de hipótesis por RBM. Si se pudieran establecer estos requisitos de manera convincente, entonces lo que se propone como RBM podría constituir la teoría disponible más completa para explicar el surgimiento de nuevas hipótesis en el estudio de las prácticas científicas en ambiente real. Conviene aclarar que, al menos para este trabajo, el sostener la posibilidad de una *teoría disponible más completa* (y por lo tanto, probablemente verdadera) guarda cierto grado de compromiso con un realismo científico, cuestión que será relevante en la discusión sobre representación modélica, pero principalmente en cuanto a las confusiones tratadas en 4.1.2.3. acerca del estatus equívoco de los criterios de constreñimientos de un RBM. De todos modos, se intentará alcanzar una identificación de posibles constreñimientos y condiciones de adecuación, con el propósito de mostrar que el intento de una sistematización buscada por los investigadores del campo está sesgada por los efectos de una tensión filosófica subyacente. Dicha tensión se daría entre Pragmatismo y Racionalismo Cartesiano, toda vez que el primero constriñe las investigaciones actuales en ciencia cognitiva, mientras el segundo determina las visiones más tradicionales de aquel ámbito científico.

Dada esta última afirmación, se puede vislumbrar que, en el ámbito práctico de la investigación en ciencia cognitiva, los efectos de esta tensión tienden a manifestarse al menos de dos maneras relacionadas, a saber, o bien se invierten los supuestos centrales que subyacen a tal investigación (e.g. que el sustrato material de la cognición no está en la caja craneana sino que también en el ambiente, lo que invierte la dirección en la que se teoriza acerca de la mente y la cognición), o bien se sustituyen las preguntas de investigación básica (e.g. “¿qué es la mente y/o la cognición?” se sustituye por “¿qué tipo de sistemas y arreglos materiales, alternativos a un sistema simbólico formal, pueden realizar tareas típicamente consideradas cognitivas?”, lo que también ha dado pie a sustituir la noción de *agente psicológico* – o intencional - por la de *agente epistémico*). Ambos efectos caen dentro de lo que Bechtel (1998) a denominado la “crisis de identidad” de la ciencia cognitiva, caracterizada por el surgimiento de modelos que desafían la demarcación inicial del ámbito cognitivo en dos direcciones: “verticalmente hacia el cerebro, y horizontalmente hacia el ambiente” (p. 77). Aunque ambos efectos van de la mano, es la última dimensión de expansión horizontal la que presenta un interés particular para el presente trabajo, particularmente con respecto al generación de hipótesis por RBM.

Para entender mejor la hipótesis que presentaremos a continuación, será necesario tener en cuenta el *estado de la cuestión* que la presente investigación quiere evaluar clarificatoriamente. Este estado de cosas se puede resumir de la siguiente manera. En la investigación sobre RBM, dentro del contexto de la *Comunidad de RBM*, la noción de modelo no tiene una caracterización estable cuando se analiza desde una perspectiva general. El más claro indicio de esta situación es la noción de *continuo representacional* de doble lectura (a tratar en capítulo II): una relacionada a tipos de representaciones y otra relacionada a grados de representacionalidad, donde ambos aspectos se combinan de manera relativa a un mayor o menor compromiso instrumentalista (i.e con el grado de utilidad práctica de alguna teoría científica), constructivista (i.e. con la noción de conocimiento determinada por un proceso de construcción y adaptación mutua entre el agente cognoscente y el objeto cognoscido), y pragmatista (i.e. con el énfasis en las consecuencias prácticas de la representación) de los investigadores del campo. A lo largo de este trabajo denominaremos a todas estas concepciones como pragmatistas, en el entendido de que, de todas ellas, el uso que tienen en común por parte de los investigadores

---

aquí estudiados consiste en defender la tesis de que *lo esencial del pensamiento y la cognición no es su relación con las cosas del mundo que representan, sino las acciones que guían en un mundo materialmente complejo*. Así, por ejemplo, a un extremo de dicho continuo dual, se puede ubicar la noción de modelo según Giere (e.g. 2002, 2004), quien sostiene una visión instrumentalista (y por lo tanto realístamente agnóstica) de los modelos como representación de la realidad. En este caso, el instrumentalismo asociado a la noción de modelo se manifiesta por el hecho de éstos deben ser entendidos, no desde una perspectiva “traductora” de la representación, sino más bien en la actividad de representación como una actividad pragmática. Esto implica, según Giere (ver 2.1.1.), incluir a los agentes intencionales con sus metas y propósitos, ya que el científico (o el grupo de científicos) hace *uso* de un modelo determinado para representar ciertos aspectos del mundo para propósitos específicos. En un nivel intermedio del continuo representacional señalado anteriormente, habría que ubicar la noción de representación propuesta por Nersessian (e.g. 2008).

Nersessian considera que un estudio apropiado de las prácticas científicas exige combinar los hallazgos, métodos e interpretaciones provenientes tanto de la historia de las ciencias como de las ciencias cognitivas. A partir de esta postura metodológica, esta autora propone que los modelos deben ser entendidos como estructuras representacionales instanciadas en la memoria de trabajo, cuya factura incorpora mecanismos perceptuales de inferencia, sistemas con partes que interactúan entre sí, análoga y simulativamente respecto a lo que se representa, y la representación de tales interacciones. En esta concepción de modelos, se puede asumir la participación de modelos de distintos formatos que, aunque no claramente caracterizados, pueden ser agrupados como internos y externos. Si bien es cierto, las prácticas científicas observadas etnográficamente se centran en la importancia del uso de modelos físicos externos, se admite que las responsables de algún proceso de RBM debieran ser estructuras modélicas internas (Dogan y Nersessian, 2005). Esto último, sin embargo, no está libre de ambigüedades, especialmente cuando se defiende el rol de representaciones externas vinculadas (de alguna manera) a las internas como parte de sistemas inferenciales distribuidos. Por ahora, diremos que tal ambigüedad es efecto de un conjunto de concepciones laxas de modelo que permite suponer algún tipo de compatibilidad a lo largo del continuo representacional señalado (ver también sección e en apartado 1.1.1.), en cuyo extremo opuesto se puede situar la noción de modelos como “agentes cognitivos autónomos” (e.g. Magnani, 2009).

En este último caso, el formato de las representaciones modélicas adquiere mucha menor relevancia, ya que cualquier cosa que se considere una representación modélica a lo largo del continuo representacional es susceptible de jugar un papel productivo en lo que Magnani ha denominado “abducción manipulativa”. Este proceso de abducción manipulativa tiene que ver con los procedimientos de cognición distribuida (ver apartado ‘c’ en 1.1.) que supuestamente caracterizan a una *negociación epistémica* continua entre signos semióticos que son permanentemente modificables a través de una (supuesta) proyección causal bidireccional y productiva entre ambiente y cerebro. Tal interacción continua tendría sus bases en la noción de *affordance*, en virtud de lo cual el *agente activo* de Gibson (e.g. 1966, 1979) puede ser entendido como una versión de agente cognitivo en tanto “sistema buscador de oportunidades” (ver 3.1.1.). El componente pragmático o consecuencialista de esta última concepción es clara, pero también lo es la idea de una representación *constructivista reificada* defendida por Knuuttila et al. (2005), ya que según Magnani el proceso semiótico (o de atribución de significados) es totalmente dependiente del ser humano. Sin embargo, en este último caso, no es posible atribuirle una centralidad al agente intencional, porque el agente no dispone de algo que sea

representable independientemente de él mismo, y porque según la semiótica Peirciana que inspira la concepción de agente activo de Magnani, el hombre es un signo externo (con una capacidad intrínseca para atribuir significados). Más claramente, lo que parece asumirse cuando se defiende la tesis de que todo es dependiente del agente (o de lo mental si se quiere) es que o bien no habría una realidad material independiente del agente, o bien es inviable realizar afirmaciones con sentido acerca de una posible realidad externa. En ambos casos, sería irrelevante atribuirle una *centralidad* al agente cognitivo si tal concepto se considera vacío de contenido (i.e. cuando no hay algo independiente del agente con respecto a lo que se pueda atribuir dicha centralidad).

Con todo, el supuesto de un *continuo representacionalista* parece revelar al menos dos cosas: la incompatibilidad entre diversas nociones de representación modélica y la consiguiente inviabilidad de una teoría unificada del RBM como resultado de tal incompatibilidad. En otras palabras, no es del todo claro que las distintas nociones de representación modélicas puedan estar relacionadas entre sí por medio de un continuo de representaciones, especialmente porque dichas nociones participan en procesos que no ponen el énfasis productivo en los mismos aspectos (e.g. representación; acción de representar; manipulación; propósitos del agente intencional; *affordances* del ambiente en tanto “agentes cognitivos autónomos”, aprovechados por el “agente cognitivo buscador de oportunidades”, etc.). Luego, si la incompatibilidad en cuestión es correcta, las confusiones relativas a la liberación de concepciones modélicas sin formato definido, y, por consiguiente, de las interacciones causales relevantes, parecen socavar la posibilidad de unificar una teoría consistente de RBM sobre la base de las propuestas disponibles. No obstante lo anterior, es importante señalar que la propuesta de Magnani es la única que propone la posible unificación de un programa para una teoría epistemológica del RBM probablemente más completa, en el entendido de que sus planteamientos aportarían un aspecto extra-teórico central para que procesos de RBM más teóricos (y menos radicalmente externalistas) puedan disponer de una información tácita crucial en procesos creativos (ver apartados 2.1.1.; 2.1.3.1.; y 3.1.2)

En este sentido, se buscará ponderar la validez de la siguiente hipótesis: **si es posible identificar criterios sistemáticos e inequívocos de constreñimientos para el RBM en tanto *sistema inferencial distribuido*, y satisfacer condiciones de adecuación mínimas para la manipulación de representaciones externas sin formato definido, entonces los modelos como *agentes epistémicos autónomos* constituye la concepción unificadora disponible más adecuada para dar cuenta de una teoría del RBM externalista.**

Cabe señalar que, en este trabajo, la concepción de modelos como “agentes epistémicos autónomos” se considerará bajo esa denominación, o bien bajo la denominación de “artefactos cognitivos”, indistintamente. La utilización tal distinción terminológica estará simplemente determinada por el contexto temático en que se invoque. Así, por ejemplo, cuando se especifiquen las diversas nociones de modelos en RBM será relevante aludir a la concepción de modelo artefactualista no representacionalista (ver 2.1.3.1. y *sección e* en 1.1.1.), mientras que para especificar la noción de modelo asumida por una postura unificadora del RBM (según la hipótesis central de esta tesis) será relevante aludir a la supuesta autonomía epistémica de una concepción de modelos en tanto agentes (o prótesis) cognitivas.

Para informar de la tarea propuesta, este trabajo se organiza de la siguiente manera. En primer lugar, se explora un esquema preliminar para establecer el contexto de la presente investigación, así como el tipo de problemas relevantes al tema de la generación

de hipótesis por medio de un RBM. Luego, se explicitan algunos problemas centrales relacionados al rol y estatus representacional de la representación modélica del RBM en ciencia. A partir de esto, se abordan casos más particulares que vinculan la abducción al RBM, con el fin de ilustrar la incidencia de la teoría de los *affordances* y del denominado ‘principio de paridad’. Se intentará mostrar que estas dos últimas nociones – inicialmente inspiradas en el concepto de “agente activo” de Gibson (1966) - cumplen un rol central, con mayor o menor grado de explicitud, en la fundamentación teórico-experimental del RBM. Posteriormente, se articulan, críticamente, potenciales criterios de constreñimiento y posibles condiciones de adecuación para un RBM en el quehacer científico creativo o productivo. Las cuatro etapas señaladas hasta el momento son precedidas de un apartado aclaratorio inicial al comienzo del capítulo I, que busca facilitar y familiarizar al lector con la caracterización de términos y conceptos propios de la especificidad de los temas tratados a lo largo de esta investigación. Finalmente, se presentan las conclusiones e implicancias de la investigación realizada, sin dejar de proponer eventuales desarrollos futuros.

La razón por la que se ha decidido organizar este trabajo de la manera descrita pretende explicitar uno de los problemas centrales que subyacen al RBM aquí evaluado. Dado que no parece posible individuar ya sea un criterio de constreñimiento esencial, o al menos un conjunto de principios procedimentales, claramente compartidos entre las descripciones de un RBM, resulta lícito vincular dicho problema a la permanente discusión entre representacionistas y anti-representacionistas que ha matizado el estudio científico de la cognición durante las últimas décadas. Específicamente, se defenderá la idea de que distintas aproximaciones de un RBM asumen, no tan sólo alguna de las dos posturas señaladas, sino más bien la noción de un ‘continuo de *grados* y *tipos* de representacionalidad’ que oscurece el entendimiento del fenómeno que se pretende explicar. Aquí, al hablar de grados de representacionalidad se debe tener en cuenta dos nociones supuestamente relevantes en la construcción de modelos: por una parte, la hibridación de modelos que incorpora constreñimientos provenientes (pero no fielmente mapeados) de estructuras causales de un modelo fuente sobre un “modelo target”, y por otra parte, la noción constructivista de que un modelo no debiera entenderse como un sustituto representacional de parte de la realidad, sino más bien como una *re-presentación* literal de algún fenómeno que nunca es independiente del agente y sus procesos de adquisición de conocimiento. Adicionalmente, al hablar de tipos de representacionalidad se estará poniendo énfasis en concepciones más o menos externalistas, por ejemplo, en términos de una representación modélica de base corporalizada (ver sección c de 1.1.1.) con asiento en la memoria de trabajo, en comparación a un modelo artefactual físico con propiedades epistémicas atribuidas a su propia materialidad (tratados en 2.1.3.1.). La falta de claridad sugerida anteriormente vendría dada por la poca viabilidad de dicho *continuo dual* como una noción claramente justificada sobre la base de los antecedentes científicos como la denominada evidencia confiable proveniente de las ciencias cognitivas. Por lo tanto, trataremos de desafiar la postura que defiende lo contrario, basada en una posible lectura unificadora de la evidencia disponible en RBM realizada sobre bases pragmatistas.

La necesidad de contar con esta noción de continuo representacional tiene que ver, al menos tentativamente, con las dificultades conocidas para compatibilizar (o eliminar, según sea el caso) la visión clásica del sistema de representación mental con las interpretaciones e hipótesis provenientes desde las investigaciones prácticas en ciencia cognitiva que tratan de modelar y entender una “conducta adaptativa inteligente”. Según una versión general de la visión clásica, los estados mentales son relaciones entre los portadores de tales estados y representaciones mentales internas. Esto, en sí mismo, constituye una teoría metafísica de los estados mentales, pero no pareciera ayudar al quehacer práctico

de la ciencia cognitiva (y la psicología en tanto posible ámbito científico autónomo) en su cometido de explicar la forma que pueden tomar estas representaciones, cómo se pueden manipular, o cómo interactúan entre sí para mediar entre un input perceptual y un output conductual. Ciertamente, esta visión clásica tiene versiones actuales que intentan resolver estos problemas (e.g. Fodor, 1975; 2008), pero la línea filosófica-experimental que aborda la investigación en RBM corresponde, más bien, a una reacción más o menos radical a la visión tradicional. Como se dijo anteriormente, es posible asociar la noción de continuo representacional a una cierta motivación filosófica anti-Cartesiana, a partir de la cual se intenta subsumir la explicación de fenómenos cognitivos que los enfoques anti-representacionistas radicales (e.g. robótica situada, cognición como sistema dinámico, etc.) no logran proporcionar satisfactoriamente, aún cuando compartan el mismo tipo de motivación filosófica. La noción del continuo representacionalista, tal como argumentaremos, mina finalmente la posibilidad de confirmar la hipótesis planteada en este trabajo.

El trabajo concluye impugnando la hipótesis propuesta. Por un lado, según se discutirá en el último capítulo, los criterios de constreñimientos, aquí examinados, para un RBM con compromisos externalistas, no son del todo inequívocos. Por otro lado, las condiciones mínimas de adecuación que aquí se ensayan tentativamente no logran ser satisfechas para explicar el surgimiento de nuevas hipótesis por medio de un RBM. Finalmente, si los errores expositivos y la insuficiencia reflexiva del presente trabajo no son excesivos, el tratamiento que se propone sobre los asuntos anteriormente comentados debiera contribuir a dilucidar un poco más algunos de los desafíos que estos enfoques anti-individualistas enfrentan en la actualidad.

---

# CAPITULO I

## 1.1. RBM y el estudio de las prácticas científicas en ambiente real

### 1.1.1. Algunas nociones y distinciones preliminares

---

En este apartado abordaremos un conjunto de nociones y distinciones aclaratorias aludidas permanentemente a lo largo de la presente investigación. Será recomendable consultar las breves secciones que aquí se exponen cada vez que sea necesario, dado el grado de especificidad conceptual de los diversos temas relacionados al RBM en estudio. Específicamente, el apartado estará constituido por las siguientes subsecciones: individualismo y anti-individualismo; lógica del descubrimiento y lógica de la justificación; cognición clásica y enfoques alternativos de la cognición; abducción, razonamiento proposicional y RBM; y finalmente, la noción de continuo de representacional.

#### a) Individualismo y anti-individualismo

El anti-individualismo es la negación del individualismo. Hay versiones de cada una de estas posiciones (e.g. Burge, 1986, Clark, 2007, Rupert, 2004; y Putnam, 1975, Fodor, 1980, Stich, 1978, respectivamente) que para los presentes propósitos aclaratorios no tiene mayor relevancia especificar. Por lo tanto, haremos referencia a las características y supuestos centrales del individualismo metodológico en ciencia cognitiva, en el entendido de que la negación de aquellos proporcionará las bases de la postura contradictoria, también en términos generales.

El individualismo metodológico es una doctrina que ha regulado gran parte de la investigación en ciencia cognitiva clásica, tanto con respecto a la demarcación del ámbito de los estados psicológicos como a la individuación de los estados mentales por medio de sus posibles poderes causales intrínsecos. En el primer caso, la taxonomía de los estados psicológicos está constreñida por una demarcación metodológica cuyo supuesto central es el siguiente: la distinción entre los estados psicológicos de los individuos y sus ambientes físicos y sociales proporciona una base natural para demarcar científicamente el ámbito de las propiedades propiamente psicológicas. En el segundo caso, la posibilidad de que los estados mentales supervengan sobre procesos cerebrales es lo que motiva la prescripción epistemológica de restringir las premisas explicativas de la psicología a relaciones causales entre propiedades que estén dentro de la caja craneana.

En la actualidad, las posiciones anti-individualistas varían con respecto al rol de los factores ambientales en el procesamiento cognitivo. Un ejemplo de esto se da en el debate entre quienes defienden hipótesis más o menos externalistas de tal procesamiento (Clark, 2007; Rupert, 2004). A lo largo de este trabajo, veremos como la aceptación o rechazo de una perspectiva individualista ha repercutido en el surgimiento de concepciones alternativas de representación mental, procesamiento cognitivo, y causación mental, entre otros. Esto es reflejo de la complicada relación que existe entre individualismo y ciencia

cognitiva, marcada por distintas visiones acerca de cómo debieran entenderse las prácticas explicativas dentro de las actuales ciencias cognitivas.

Tradicionalmente en ciencia cognitiva y filosofía de la mente, el individualismo ha sido asociado al funcionalismo (ver sección *c* de este apartado), principalmente por la intuición de que los roles funcionales o causales de los estados psicológicos son intrínsecamente individualistas. Una “intuición cartesiana” asociada a esta posición dice relación con que los estados mentales de un organismo serían tal como son, aún si estuvieran situados en ambientes radicalmente distintos. Alternativamente, las investigaciones actuales en ciencia cognitiva y otras disciplinas relacionadas - específicamente en el tipo de RBM estudiado aquí - guardan un compromiso variable de corte anti-individualista. El tipo de base evidencial a la que normalmente se alude en este caso corresponde a las hipótesis empíricas que defienden una naturaleza corporalizada, situada y embebidas [*embedded*] de la cognición, en los términos que estos enfoques serán introducidos en el la sección *c* de este apartado.

### **b) Lógica del descubrimiento y lógica de la justificación**

Desde el siglo pasado los filósofos de la ciencia han hecho la distinción entre una lógica del descubrimiento y una lógica de la justificación. Como afirma Magnani (2009a), la mayoría de ellos ha concluido o bien que no existe una lógica del descubrimiento, o bien que un modelo racional del descubrimiento es imposible. Si esto es correcto, entonces el descubrimiento científico sería un fenómeno irracional, en el sentido de que no habría una manera posible de razonar hacia la producción de hipótesis [*there is no reasoning to hypotheses*]. La presente investigación trata, precisamente, de la evaluación de alguna de las propuestas que ciertos investigadores del RBM suponen como posibles de unificación, en el contexto de un modelo epistemológico unificado del descubrimiento científico donde la abducción cumple un rol significativo (ver sección *d* del presente apartado).

Para hablar de una lógica del descubrimiento, es necesario entender por qué tendría sentido tratar de buscar o delinear algo como aquello. Para este fin, podemos comenzar por hacer referencia a lo que se ha denominado una lógica de la justificación. Con ese propósito, caracterizaremos una versión general de la lógica de la justificación, para establecer el problema que enfrenta cuando se trata de abordar, por medio de las relaciones y elementos que la distinguen, procesos de razonamiento que no parecieran estar regulados racionalmente. Con esto, trataremos de establecer el contexto de lo que se busca cuando se intenta delinear lo que se ha denominado una posible lógica del descubrimiento.

Según Artemov (2008), una lógica de la justificación se basa en la lógica proposicional clásica regulada por un tipo de enunciados de justificación del tipo  $t:F$ , cuya lectura es “*t* es una justificación para *F*”, y sus orígenes se relacionan con el estudio del conocido problema (atribuido inicialmente a Platón) entre *creencia verdadera justificada* y *conocimiento*. Según este autor, esta afirmación ha tenido diversas interpretaciones, tales como: “*t* es aceptado por un agente como una justificación de *F*”; “*t* es un recurso suficiente para *F*”; “*F* satisface las condiciones *t*”, etc. Estas lecturas no son equivalentes, tal como se puede deducir de la última de ellas, donde *t* puede no ser todo lo que se requiera para que *la creencia de que F* cuente como *el conocimiento de que F*. Adicionalmente, aquello que un agente determinado considere como justificación de *F* puede incluso recurrir a un tipo de base evidencial que desafía el ámbito de la lógica tradicional, en la medida que relaciona la noción de justificación, por ejemplo, con algún tipo de *conocimiento empírico*. Este es, precisamente, el tipo de aspectos que abordaremos a continuación.

La distinción que nos interesa en esta sección debe ser entendida desde una perspectiva más amplia. Dicha perspectiva es la que en filosofía se conoce como la

distinción entre el contexto de la justificación y el contexto del descubrimiento. De acuerdo a Hoyningen-Huene (1987), la distinción puede ser históricamente rastreada hasta tiempos de Aristóteles y Euclides, pero en la forma que aquí nos interesa tiene que ver con el debate sostenido entre filósofos de orientación positivista (en un sentido amplio que incluye exponentes del empirismo lógico y del racionalismo crítico) e historicista. Asumiendo que esta última división hace justicia a la defensa o rechazo, respectivamente, de la distinción en cuestión, hay ciertas cuestiones centrales que se derivan a partir de la correspondiente controversia. Por un lado, para quienes han defendido la distinción de los contextos señalados, dicha distinción resulta fundamental para garantizar la autonomía de una filosofía de la ciencia, y con ello orientar la filosofía hacia los aspectos de la ciencia sobre los que se puede reflexionar filosóficamente, a saber, las ciencias entendidas como un cuerpo de conocimiento [*body of knowledge*]. Por otro lado, quienes la han rechazado sostienen que el aislamiento de la filosofía de otras ciencias de carácter empírico (e.g. historia, psicología, sociología de las ciencias, etc.) es extremadamente perjudicial para una provechosa colaboración interdisciplinaria que podría considerarse, alternativamente, como la correcta orientación y naturaleza de la filosofía de la ciencia.

Hoyningen-Huene presenta un interesante conjunto de distinciones relacionadas a la distinción de los contextos, en el entendido de que el surgimiento de tales distinciones ayudaría a explicar las confusiones que habrían marcado las controversias centrales entre defensores y opositores a la distinción entre contexto de justificación y de descubrimiento. La presentación de estas distinciones específicas que hace este autor parece apropiada, puesto que separa aquellas que surgen desde el punto de vista de los defensores de los contextos y desde los críticos, respectivamente. Por tal razón, resumiremos cada una de ellas en esos mismos términos, para, finalmente, enfocarnos en aquella distinción particular que permita caracterizar las motivaciones que actualmente dan lugar a la búsqueda de una lógica del descubrimiento científico *más allá* de la distinción en cuestión.

Desde el punto de vista de los defensores, la distinción de contextos ha tenido al menos cuatro versiones más o menos relacionadas entre sí. Una primera distinción puede entenderse en términos de procesos: el proceso de descubrimiento y el proceso de justificación de este descubrimiento. Aquí, es lógico pensar que el primero debe preceder al segundo en su caracterización, puesto que el descubrimiento es el objeto susceptible de justificación. Una segunda distinción se da entre proceso de descubrimiento y métodos de reconstrucción de la justificación (o de análisis o consideraciones relevantes a la justificación). En este caso, lo que se plantea es una distinción entre lo factual y lo normativo, puesto que ahora son los “métodos de justificación” y la “reconstrucción racional de la justificación” lo que se equipara al proceso histórico del descubrimiento, a expensas del proceso histórico de justificación mismo como ocurre en el primer caso. Hoyningen-Huene no deja de notar cierta ambigüedad en esta última distinción, en el sentido de no establecer claramente si los métodos de justificación de algún momento específico del pasado puedan ser apropiados en el presente o de manera permanente. Pero esta ambigüedad es la que se abordaría a través de una tercera distinción, a saber, aquella entre el análisis del descubrimiento como uno de carácter empírico y el análisis de la justificación como uno de carácter lógico. Aquí, de lo que se trata es de establecer una distinción entre métodos de justificación determinados por la lógica deductiva e inductiva del siglo XX, y el estudio de algún descubrimiento empírico particular, llevado a cabo, por ejemplo, por investigaciones históricas y psicológicas que no están determinados por la lógica, porque, según se supone en este caso, el proceso de descubrimiento no está sujeto a principios lógicos. Como se puede inferir, en esta distinción, el análisis lógico de la justificación hace irrelevante el problema de la temporalidad aludido en el caso anterior. La cuarta y última distinción se da

entre filosofía de la ciencia, por un lado, e historia, psicología y sociología de la ciencia, por otro. Esta distinción pone el énfasis no sólo en la diferencia de métodos correspondientes a cada disciplina académica, sino que principalmente a la relación de dependencia que las ciencias empíricas debieran tener con respecto a la filosofía de la ciencia. Se asume aquí que las disciplinas empíricas deben considerar en qué consiste la justificación filosófica para saber qué cosa requiere de una explicación empírica. La relación de dependencia se establece en la medida de que lo que cuente como justificación sea solamente una cuestión de razonamiento lógico, desde donde se desprende que la filosofía de la ciencia no tendría nada que aprender desde las disciplinas empíricas.

Desde el punto de vista de quienes critican la distinción, cada una de las distinciones parciales comentadas en el párrafo anterior reciben la siguiente contra-argumentación: (i) en tanto procesos, no es posible establecer una clara delimitación temporal entre los procesos de descubrimiento y de justificación; (ii) los procesos de descubrimientos no tienen aspectos lógicos, razón por la cual sólo pueden ser analizados en términos de las disciplinas empíricas. Hanson (1958, 1960) critica este supuesto, dando lugar a un tercer argumento en contra de la distinción; (iii) a partir de Hanson, se origina la visión de que la distinción entre descubrimiento y justificación es ambigua, ya que la noción de “descubrimiento” conlleva implícita alguna forma de justificación (*descubrir que p* implica haber adquirido algún tipo de conocimiento acerca de *p*). De aquí se sigue, de manera supuestamente plausible, la posible articulación de alguna noción que reconsidere algo que pudiera entenderse como “aspectos lógicos del descubrimiento”. Dicha articulación se daría en términos de, al menos, 3 etapas: la generación de una hipótesis, su indagación [*pursuit*], y su testeo crítico. Aquí las distinciones no son sólo temporales, sino que también metodológicas, aunque (como Hoyningen-Huene atribuye a Kuhn) no parece haber claridad con respecto a la caracterización metodológica de cada etapa; (iv) complementando el ataque anteriormente señalado en (ii), el cuarto contra-argumento se origina en los planteamientos de Kuhn según los cuales es necesario considerar factores sociológicos y psicológicos para describir adecuadamente las decisiones que tienen lugar en la elección de teorías. Dicho de otra manera, los factores que juegan un rol decisivo cuando una comunidad participa en la elección de alguna determinada teoría son sociológicos porque sólo pueden explicarse con referencia a los sistemas de valores cognitivos sostenidos por dicha comunidad (sean compartidos o no); y, a su turno, los factores psicológicos son relevantes cuando la decisión de un solo individuo puede únicamente ser explicada con referencia a la forma particular en la que este miembro interpreta los valores cognitivos de la comunidad.

Dado la situación general descrita acerca del denominado problema de la distinción de contextos de descubrimiento/justificación, es preciso considerar que la actual investigación acerca de una posible teoría epistemológica del descubrimiento en ciencia tiene relación con aquellas visiones críticas que sostienen que el descubrimiento empírico es relevante para el quehacer epistemológico. No pareciera haber claridad con respecto al grado de relevancia que tal tipo de conocimiento debiera tener, pero sí se sabe acerca del tipo de problemas que tal supuesto acarrea, tal como lo expresa Hoyningen-Huene:

***“Cuando la epistemología busca justificar la posibilidad del conocimiento empírico, entonces cualquier enfoque epistemológico que esté basado sobre la presuposición de un conocimiento empírico válido, concluye en la circularidad”.***  
***(p.510) (Todas las traducciones son responsabilidad del autor de este trabajo)***

El punto es que, hoy en día, tal circularidad no parece ser un impedimento para elaborar teorías que pretenden explicar la racionalidad del descubrimiento científico sobre una base

de conocimiento meramente empírico. Eso pasa, como se dijo, por re-evaluar lo que se considere como *aspectos lógicos del descubrimiento*. En el ámbito de los cambios en las teorías científicas, por ejemplo, el uso de aplicaciones computacionales ha permitido describir la (supuesta) operación de distintos tipos de procesos de descubrimiento. En estos casos, según Magnani (2009a):

**“Los métodos de descubrimientos son dependientes de los datos (generalizaciones a partir de la observación y los experimentos), de las explicaciones, y de la coherencia (establecida para superar contradicciones)” [Discovery methods are data-driven (generalizations from observation and from experiments), explanation-driven (abductive), and coherence-driven (formed to overwhelm contradictions).](p. 1)**

Para entender hacia dónde apunta el tipo de contribución señalada en la cita de Magnani, necesitamos entender el estado actual de las disciplinas empíricas de la cognición que reaccionan a la supremacía de una filosofía de la ciencia que defiende la distinción entre descubrimiento y justificación (o reconstrucción racional del descubrimiento). Eso es lo que resumiremos en la siguiente sección, para luego dar cuenta de la forma en que, como efecto de los desarrollos recientes de la ciencia cognitiva, el descubrimiento científico ha llegado a ser considerado por algunos como un fenómeno susceptible de explicación a través de la abducción y el RBM.

### c) Cognición clásica y enfoques alternativos de la cognición

Para entender mejor el estado de las ciencias cognitivas actuales que aquí nos importa, haremos una breve reseña de las visiones que han caracterizado el desarrollo de este campo hasta la actualidad. Si esta reseña resulta mínimamente aclaratoria, debiéramos terminar comprendiendo algunas motivaciones centrales que dan lugar a la noción de “ambiente real” y a la tendencia actual por priorizar explicaciones centradas en “prácticas cognitivas” involucradas en la resolución de problemas [*problem solving*]. Estas prácticas consisten, típicamente, en la generación, manipulación y propagación de representaciones salientes del “sistema cognitivo distribuido”, y su estudio implica examinar cómo fluyen los estados de representaciones a través de distintos medios (o dominios) con el fin de realizar un trabajo cognitivo determinado. A partir de esto, se haría ostensible la tendencia a enfocarse en la naturaleza del trabajo intelectual realizado en ciencia a través del *uso de modelos*, lo que para algunos científicos cognitivos (e.g. Knuutila et al., 2005) implica abandonar un interés tradicional de la filosofía de la ciencia caracterizado por el problema del realismo científico de los modelos (o el intento por dilucidar las confiabilidad relaciones existentes entre los modelos y los fenómenos que representan).

La denominada *revolución cognitiva*<sup>4</sup> estableció como idea básica la visión de que la conducta humana compleja está controlada, en gran medida, por estados representacionales internos. De acuerdo a dicha visión, estos estados son portadores de contenido, en el sentido de que son acerca de algo. Las operaciones (sintácticas) que se pueden realizar sobre estos estados internos constituyen un procesamiento que le permite al sistema realizar planes y resolver problemas. Esto implica dos cuestiones de alto interés para la filosofía de la ciencia cognitiva, a saber, la naturaleza de dichos procesos (i.e. la ‘arquitectura cognitiva’ postulada) y el contenido portado por los estados representacionales. Consideremos, por ahora, sólo el primero de estos aspectos.

<sup>4</sup> Ver Miller (2003)

La ciencia computacional inspiró una arquitectura cognitiva basada en la “metáfora del computador”, según la cual los estados del sistema cognitivo son estados representacionales con contenidos conceptuales que se corresponden con entidades como los nombres, los predicados y los cuantificadores de un lenguaje natural, entre otras. Según la teoría que se pretendía examinar, los procesos mentales son operaciones definidas sobre representaciones mentales estructuradas sintácticamente. El proyecto central detrás de esta estructuración es hacer uso de la hipótesis de que los estados mentales son lingüiformes, con el fin de proporcionar explicaciones para determinadas propiedades ubicuas de los estados mentales. Típicamente, estas propiedades hacen alusión a la ‘productividad’ (i.e. que uno siempre puede generar nuevos pensamientos) y la ‘sistematicidad’ (i.e. que el poseer ciertas capacidades cognitivas garantiza el poseer otras capacidades cognitivas relacionadas) de los estados mentales, y a que los procesos mentales tienden a ser “*truth-preserving*”, vale decir, que tienden preservar su valor de verdad ‘verdadero’. De esto se puede inferir el rol central que para esta arquitectura cognitiva tiene la apreciación del poder y el alcance de la lógica formal.

Clark (2001) resume las características principales de una lógica formal como un conjunto de símbolos y maneras de combinarlos que permiten expresar proposiciones y reglas complejas, las cuales especifican cómo derivar legalmente nuevos complejos simbólicos a partir de otros antiguos. Lo interesante de las lógicas formales – según remarca Clark –, es que la aplicación rigurosa de estas reglas garantiza que nunca se podrá inferir una conclusión falsa a partir de premisas verdaderas, incluso aunque se desconozca el significado particular de los símbolos involucrados en tales operaciones inferenciales. Además de las ventajas metodológicas que implica un entendimiento de la mente y la cognición en tanto sistema formal (e.g. superar cierto oscurantismo sobre la noción de mente y salvaguardar cierta autonomía de una ciencia de la mente), esta aproximación también constituye una propuesta seria con respecto a las herramientas posiblemente correctas para teorizar acerca la mente, tanto en psicología cognitiva como en inteligencia artificial.

El enfoque recién aludido tuvo una expresión en filosofía con lo que conocemos como *funcionalismo*. Según la tesis del funcionalismo, los estados mentales son estados funcionales que pueden interactuar entre sí causalmente, independientemente del tipo de materia en que éstos se realicen. Dicho de otra manera, lo importante es que tal organización funcional sea capaz de procesar información y controlar aquella clase de conducta que comúnmente nos parecería indicativa de una mente humana. Un arreglo funcional que es independiente de algún tipo de materia particular, pero que siempre es realizado en un *sustrato* material, apuntaría a superar los problemas propios de un dualismo de sustancia de base Cartesiano. Como se sabe, este tipo de dualismo hace referencia a la interacción causal que debiera existir entre la mente y el cuerpo, especialmente en el caso de asumir que la mente es alguna sustancia con propiedades tales como la intencionalidad o la consciencia (sea lo que fuere esta última) de naturaleza distinta a la material. El funcionalismo, por lo tanto, es una opción que tiende a satisfacer el requerimiento antidualista exigido por parte de quienes tienen el convencimiento de que alguna versión materialista (o fisicalista, si se quiere) de la mente debiera ser verdadera. La individuación de un arreglo material en términos funcionalistas equivale a sostener que lo importante no es de qué está hecho aquello que instancia un estado mental, sino, más bien, el arreglo causal (i.e. el rol funcional) en el que están tales instanciaciones. Luego, la identidad de un estado mental hace referencia a un rol causal/funcional determinado que, en sí mismo, no está determinado por el sustrato material particular en que se pueda realizar. Dicho arreglo funcional, además de proporcionar una teoría para especificar

*estados mentales con contenido* como las creencias, también sería capaz de dar cuenta de estados cualitativos (e.g. el dolor), siempre que fuesen especificados de la misma manera.

Típicamente, aunque no de manera exclusiva, la concepción funcionalista ha sido criticada sobre la base de problemas relacionados con la repercusión de que, sin excepciones, cualquier arreglo material apropiadamente organizado podría llegar a conformar una mente. Y dada la cercanía que el funcionalismo tiene a la visión que compara a los estados mentales con los estados funcionales o lógicos de un computador (Putnam, 1960), también han surgido críticas relacionadas que cuestionan la posibilidad de que las representaciones mentales (o su equivalente simbólico en el ámbito de la arquitectura computacional) puedan tener una intencionalidad no derivada<sup>5</sup>. Según Bechtel (2010), filósofos como Armstrong (1968) y Lewis (1980) han intentado combinar consideraciones funcionalistas con la denominada teoría de la identidad mente-cerebro. En tales casos, los estados mentales serían los “realizadores particulares” de un rol funcional, y no el rol funcional en sí mismo, descartando la posibilidad de una *realizabilidad múltiple* de los estados mentales, dada la identidad entre la realización de estados mentales y de los estados cerebrales. Otra posición funcionalista es aquella que defiende la idea de que el pensamiento y otros procesos mentales son operaciones computacionales. Esta visión corresponde a la propuesta que mejor representa el quehacer de la ciencia cognitiva clásica, porque en ella convergen productivamente los supuestos centrales de una tradición filosófica racionalista y los supuestos derivados de una noción sintáctica de procesos computacionales. Por un lado, la tradición racionalista sostiene que los procesos mentales tienen un carácter inferencial lógico, porque la vida mental está “amueblada”, para ocupar una expresión actualmente en boga, por estados mentales lógicos. Por otro lado, el descubrimiento de Turing establece un mecanismo según el cual una máquina puede detectar argumentos válidos siempre y cuando dicho mecanismo sea capaz de detectar relaciones y propiedades de carácter sintáctico. Una teoría que ha permitido dar cuenta de cómo los procesos mentales son sensibles a la forma lógica de los estados mentales (especialmente los pensamientos) es aquella que tomó la forma de lo que Fodor (2000) ha denominado una *psicología racionalista implementada por procesos sintácticos* (o Teoría Computacional de la Mente, TCM). Vale decir, una teoría donde convergen la idea de procesos mentales cuyo carácter inferencial típico estaría determinado por la forma lógica de los pensamientos, y la idea de que tales procesos son implementados como computaciones determinadas por la forma sintáctica de las representaciones mentales. A partir de esta convergencia, se asume que las operaciones sintácticamente especificadas son capaces de reconstruir relaciones de forma lógica, determinándose así el *rol computacional* de las representaciones sintácticas a partir del *rol causal* de dicho tipo de representaciones determinado por la forma lógica de los pensamientos.

Probablemente, las críticas más conocidas hacia el funcionalismo computacional en general no se dirigen certeramente a las propuestas centrales de lo que se propone como una teoría de lo mental, especialmente con respecto a la TCM. De hecho, mucho de las críticas suponen que tal visión clásica de la cognición debe ser entendida como una teoría exhaustiva o completa de la cognición. Sin embargo, en su defensa (Fodor, 2000) se ha argumentado que, aunque tal objetivo dista mucho de ser alcanzado, la teoría en cuestión puede tener parte de la verdad acerca de la cognición, especialmente si se puede mostrar que es la mejor teoría frente a cualquier otra alternativa disponible. No obstante la defensa aludida, resulta apropiado enunciar las críticas más típicas que recibe el funcionalismo

<sup>5</sup> Mas sobre esto último en Searle (1980).

computacional, con el fin de entender el tipo de (supuestas) limitaciones que motivarían el surgimiento de nuevas formas de caracterizar las representaciones y los procesos cognitivos.

Dos de los problemas más conocidos que se le atribuyen al denominado *computacionalismo* tienen que ver, por un lado, con las implicancias que tiene un posible carácter computacional de la actividad mental y, por otro, con la verosimilitud de la noción de representación mental involucrada en dicha visión cognitivista. En el primer caso, no sería del todo obvio que haya aspectos de la mente que computen ciertas representaciones mentales de una manera *serial* o “paso a paso”. Esto genera dudas relacionadas al ámbito y la factibilidad físico-biológicos de aquel procesamiento algorítmico, especialmente si se asume que una teoría computacional de la mente es, no sólo necesaria, sino que también suficiente para explicar cómo funciona una mente, y de manera particular el pensamiento humano. El segundo caso (i.e. el problema de un tipo de representación adecuado) tiene que ver con la pretensión experimental de vislumbrar un tipo de representación biológicamente más plausible que la de los modelos simbólicos, especialmente cuando se asume la existencia de procesos caracterizados por una manipulación sintáctica de símbolos lingüiformes explícitos, determinada por la propia estructura de tales símbolos. Así caracterizados, tal tipo de procesos sería incapaz de dar una explicación satisfactoria a fenómenos típicamente globales (i.e. no locales) como, por ejemplo, la abducción, en el entendido de que ésta se trate de un tipo de inferencia que opera sobre la base de información proveniente de diversas partes de algún sistema determinado (i.e. información no inmediatamente disponible en las representaciones sintácticas sobre las que opere algún proceso computacional). Asumir que (la totalidad de) los procesos mentales son computaciones, y especificables de manera exclusivamente local, es lo que ha llevado a sostener la inviabilidad del funcionalismo computacional, específicamente con respecto a la posibilidad de que los procesos cognitivos consistan en operaciones sintácticas locales realizadas sobre representaciones mentales locales.

En este contexto, surge la propuesta de un tipo de representación mental que proporcionaría modelos parecidos a la estructura cerebral subyacente. Dado el tipo de limitaciones recién aludidas del procesamiento computacional, la nueva arquitectura cognitiva tendría que adecuarse a un tipo de representaciones mentales menos locales y discretas (i.e. como pasos discretos de un algoritmo) que aquellas supuestamente requeridas para implementar subsistemas simbólicos asociados a áreas específicas de la cognición (e.g. lenguaje, intencionalidad, etc.). Esto explica que la interrogante de cómo debiera ser la arquitectura cognitiva alternativa que mejor faculta a un agente humano para *representar* la información que le permite comportarse de maneras sofisticadas haya sido uno de los problemas centrales en el campo de la Ciencia Cognitiva.

La alternativa más próxima vino de los métodos de resolución de problemas que se conoce como *conexionismo*, particularmente a partir de las publicaciones sobre “Procesamiento Distribuido Paralelo” de McClelland y Rumelhart (ver McClelland et al., 1986). Como resultante de una colaboración convergente entre la filosofía, las ciencias cognitivas y la neurociencia, la idea básica del conexionismo consiste en un procesamiento de la información realizado a través de *una representación numérica, en la forma de un vector que pasa por un conjunto de unidades interconectadas en paralelo*. Específicamente, el vector pasa por las conexiones establecidas entre dichas unidades denominadas micro-procesadores. En otras palabras, lo que se postula es un procesamiento donde un vector actúa como el *input* que atraviesa una matriz que representa los pesos de las conexiones en cuestión, para, finalmente, obtenerse otro vector como *output*. La variedad de funciones

que puede implementar una red conexionista dependerá de cómo y cuánto se puede hacer variar la eficacia de las conexiones, siempre en el formato de un mapeo que va de vector a vector. Así, se ha llegado a mostrar que cierto tipo de aprendizaje simple de resolución de problemas, por medio de ensayo y error, es modelable satisfactoriamente en formato conexionista. Se instaura, de esta manera, una posible alternativa (mínimamente viable) a la arquitectura clásica implementada por la inteligencia artificial tradicional, donde siempre se ha requerido contar con una programación explícita para modelar tales tipos de aprendizajes. Una de las implicancias más serias para la cognición que tiene la concepción de un *procesamiento distribuido paralelo* es que, al enfocarse en un nivel sub-simbólico de representación mental, se pierde la noción de un sistema representacional de sintáctica y semántica combinatorias, fundamental para explicar los fenómenos de productividad y sistematicidad del pensamiento. Luego, se abandona una concepción de procesamiento cognitivo (computacional) sin tener una clara alternativa capaz de dar cuenta de fenómenos como los recién mencionados.

Como se dijo, una de las ventajas más llamativas del conexionismo frente a la inteligencia artificial clásica tiene que ver con un cierto grado de plausibilidad biológica. En este caso, la idea es que los micro-procesadores conexionistas funcionarían de una manera parecida (o menos diferente que la opción clásica) a las neuronas biológicas. Esto exige considerar una analogía muy amplia o permisiva, donde los pesos de las conexiones serían análogos a la forma en que una neurona hace gatillar a otra. Sin embargo, son más las razones para no aceptar analogías como estas, especialmente si se alude a versiones de redes conexionistas donde cada una de los microprocesadores se considera portador de algún contenido proposicional, como se pensó inicialmente:

***“McCulloch y Pitts también hacen un vínculo con la lógica: las neuronas podrían ser asociadas con proposiciones, y dada la naturaleza binaria de estas unidades, sus estados de activación podrían ser asociados con valores de verdad”.***  
***(Bechtel et al., 1998, p. 31)***

El hecho concreto es que la plausibilidad biológica señalada parece sólo restringirse a los casos limitados en que, o se emplean algoritmos que describen fielmente el procesamiento de algún sistema neuronal muy simple, o bien donde las redes son utilizadas para *emular* algún sistema neuronal altamente conocido.

La arquitectura recién descrita es sólo la versión más conocida y aceptada de la arquitectura conexionista. Sin embargo, el éxito del conexionismo (en cualquiera de sus versiones) es aún una tarea inconclusa, especialmente cuando se contrapone a los logros de la inteligencia artificial clásica. La primera, por ejemplo, ha realizado aplicaciones típicamente aludidas como de “nivel inferior”, tales como la percepción y el control motor. La arquitectura clásica, por su parte, ha demostrado ser mucho más eficaz en modelos y aplicaciones asociadas a los denominados “procesos de nivel superior”, tales como la planificación y el procesamiento lingüístico.

La denominada contra-revolución cognitiva que surge a partir de la década de 1990, al alero de los planteamientos alternativos anidados durante décadas anteriores, no viene a desafiar la arquitectura conexionista que la precede inmediatamente, sino más bien a la arquitectura serial/representacional clásica. En este caso, la reacción se centra en contra de uno de los supuestos básicos subyacentes a dicha arquitectura, a saber, que la cognición se realiza en base a estados cognitivos lingüiformes. Según la visión de los grupos desafiantes, habría razones para pensar que dicho supuesto puede ser sustituido por otros, donde mecanismos más simples (más bien relacionados con nivel de procesamiento inferior al que

apuntaba el conexionismo<sup>6</sup>) podrían evitar el tener que suponer la existencia de sofisticados procesamientos de información y memoria en el sistema cognitivo. Una de las fuentes que inspiran estos enfoques es la investigación - poco atendida en la segunda mitad de 1960 – iniciada por J.J. Gibson (1966). Básicamente, la idea de Gibson era que los mecanismos relevantes para explicar la cognición son aquellos que, por un lado, explotan (directa y activamente) la *información* disponible en el ambiente, y por otro lado, se ajustan a ciertas *affordances* que serían revelantes para cada organismo. Algunas de las corrientes que mejor representan esta reacción son los denominados enfoques de la cognición embebida y corporalizada [*embedded/embodied cognition*], la robótica situada, la teoría de la cognición en tanto sistema dinámico, y la cognición distribuida.

Según el primero de estos enfoques alternativos, muchos problemas (que requieren una “conducta adaptiva inteligente”) pueden ser resueltos por mecanismos no representacionales simples que operan a través de la interacción corporalizada que tiene lugar en el ambiente estructurado en el cual la acción de un agente está alojada o “incrustada” [*embedded*]. Un ejemplo de esto sería la conducta característica de las abejas, donde se asume que, en vez de contar con un mapa del ambiente por el que navegan, bastaría con mecanismos que guían el vuelo con respecto a la posición relativa del sol. En el ámbito relacionado de la *robótica situada*, Brooks (1991) ha intentado mostrar la posibilidad de implementar sistemas robóticos que desempeñan determinadas tareas (que implican grados simples de aprendizaje) sin la necesidad de recurrir a un procesador central que “mastique” representaciones simbólicas de datos sensoriales con el fin de planificar rutas o ejecutar determinadas maniobras. La alternativa que explora el robotista es la de una *arquitectura de subsunción de capas* o sistemas independientes, vinculados directamente a sus propios sistemas sensoriales, con un mínimo de procesamiento simbólico.

En el enfoque de la teoría de sistemas dinámicos, por su parte, la cognición (o bien, la agencia cognitiva) es entendida como un sistema dinámico. El más destacado representante de esta visión es Tim van Gelder (e.g. van Gelder et al., 1995; van Gelder, 1995, 1998, 1999). La teoría de Sistemas Dinámicos constituye un aparato matemático elaborado para representar sistemas y la evolución de éstos en el tiempo. El sistema se representa como un conjunto de variables de estado. Este conjunto de variables de estado establece un espacio de estado, es decir, un espacio abstracto donde cada punto representa un estado posible del sistema, y el conjunto de todos los puntos representa todos los estados posibles del sistema. Las reglas de la evolución del sistema corresponden a su “dinámica”, y son las que especifican cómo el sistema ha de evolucionar en el tiempo. Dicho de otra manera, lo que esta dinámica especifica es hacia qué punto del espacio de estado se moverá el sistema, lo que se expresa como una función del punto en el que se encuentra actualmente. De esta manera, la dinámica establece un conjunto de rutas [*paths or trajectories*] a través del espacio que atravesará el sistema.

Lo relevante del enfoque de sistemas dinámicos es que los mecanismos de control de feedback funcionan porque están en una interacción causal continua con el ambiente, y que dicho sistema de control puede realizar complicados mecanismos internos que son innecesarios, a menos que se cuente con algún otro mecanismo de ajuste. Aquí es precisamente donde surge una alianza productiva entre el enfoque de sistemas dinámicos y el de la cognición embebida/corporalizada. Dado que los sistemas dinámicos se ocupan de explicar, principalmente, procesos de control de feedback continuo, el segundo enfoque - de base sensorio-motora – se apoyará en esta noción de interacción continua para fundamentar un “acople” entre el organismo y su ambiente, tal como sea

---

<sup>6</sup> Esto es una buena razón para no desafiar los postulados conexionistas.

requerido para la solución de alguna tarea determinada. Con todo, dicha combinación de enfoques representa una directa reacción a la visión representacional/computacional de la cognición. Mientras la última se caracteriza por atribuirle una centralidad a procesos de orden superior (e.g. razonamiento, lenguaje, etc.), los enfoques desafiantes privilegian las tareas de carácter sensorio-motor, tratando de prescindir de mecanismos y procesos representacionales clásicos. Esa dicotomía se refleja, obviamente, en lo que cada visión opte por considerar bajo el rótulo de cognitivo.

A su turno, la “cognición distribuida”, derivada de los planteamientos desarrollados por Hutchins (1995, 1999) en el campo denominado *antropología cognitiva*, cumple un rol particularmente ubicuo en el RBM. Tal como lo plantea Nersessian (2009), la incorporación de este enfoque está orientada hacia:

***“[...] un estudio integrado de las dimensiones socioculturales, artefactuales y corporalizadas de la cognición humana – un estudio en el cual el ambiente no sólo apoya la cognición, sino más bien, en el cual aspectos del ambiente juegan un rol integral en procesos cognitivos” (p. 731).***

Lo que se pone de manifiesto en esta cita es un compromiso poco claro con posturas anti-individualistas disímiles. Por un lado, se puede sostener, como se ha hecho (e.g. Burge, 1986), que el procesamiento cognitivo está determinado por factores ambientales. Esta postura no parece incompatible con lo afirmado en la cita. Por otro lado, también se puede sostener que lo anterior es sólo muestra de que los límites del procesamiento cognitivo se pueden ampliar de una manera que incorpore los factores que “juegan un rol en procesos cognitivos”, lo que forzosamente implica abandonar las nociones de procesamiento cognitivo tradicional y la dimensión intracraneal de lo mental. Esta postura tampoco parece incompatible con lo señalado en la cita. Trataremos de hacer ostensible las características y el alcance de este tipo de compromiso en la medida que abordemos los enfoques externalistas del RBM.

#### **d) Abducción, razonamiento proposicional y RBM**

La presente investigación no aborda detalladamente el aspecto sentencial de la abducción. Sin embargo, me parece apropiado aludir brevemente a lo que hace referencia por abducción sentencial, puesto que esta caracterización servirá de contraste para los desarrollos que se discutirán en el capítulo III, a partir de una relación entre abducción y RBM. Seguiré, para los propósitos de este apartado, lo resumido al respecto por Shelley (1996), no sin antes caracterizar la noción de abducción según Peirce.

Tanto en ciencia cognitiva como en inteligencia artificial, el razonamiento abductivo en ciencia se ha transformado en un importante foco de atención. Esta preocupación ha tenido una aproximación formal inicialmente planteada por el filósofo Charles Sanders Peirce, quien a la fecha ha actuado como punto de referencia principal en la materia. Para este filósofo, la abducción era, no solamente una forma de inferencia lógica complementaria a la deducción y la inducción, sino que también una suerte de etapa preliminar de la investigación científica en la cual, típicamente, se construyen explicaciones de anomalías detectadas en alguna teoría científica. En otras palabras, la abducción también constituiría, en la visión del filósofo, una posible “lógica del descubrimiento”.

Con el propósito de superar la división clásica de la lógica entre argumentos deductivos e inductivos (inferencias necesarias y probables, respectivamente), Pierce planteó la existencia de dos tipos de inferencias probabilísticas, a saber, la inducción y la abducción. Dicha conclusión fue alcanzada por medio del análisis de un silogismo deductivo básico

basado en el intercambio iterativo de sus premisas constituyentes (denominadas ‘Regla’, ‘Caso’ y ‘Resultado’) que, para efectos de síntesis y claridad, caracterizaremos sólo en términos de las conclusiones que tienen lugar al considerar la toma de muestras de objetos de una población determinada. Pensemos en una *novela de espionaje*, **M**, como miembro de una población de novelas de espionajes guardadas en un baúl. Asumamos también que el ser de *origen Británico*, **P**, corresponde a una propiedad que un miembro de la clase señalada puede tener. Finalmente, asumamos que un miembro **S** representa parte de una muestra sacada aleatoriamente de la población en cuestión. En los argumentos que se ilustran a continuación, (1) es un ejemplo de un silogismo válido (Regla > Caso > Resultado) representativo de una ‘deducción’, i.e. un ejemplo de inferencia necesaria en la forma de un argumento que va desde la población a la muestra aleatoria; (2) es un silogismo inválido (Resultado > Caso > Regla) representativo de la ‘inducción’, i.e. un ejemplo de inferencia probabilística en la forma de un argumento que va desde la muestra aleatoria a la población; y (3) sería un silogismo inválido (Regla > Resultado > Caso) representativo de la ‘abducción’, i.e. un ejemplo de inferencia probabilística que no va ni desde la población a la muestra aleatoria (deducción), ni viceversa (inducción):

- (1) **(REGLA)** **Todos los Ms son Ps, i.e.**  
 Todas las novelas de espionaje de este baúl son de origen Británico
- (CASO)** **Todos los Ss son Ms, i.e.**  
 Todas las novelas de espionaje de esta muestra particular fueron tomadas de este baúl
- (RESULTADO)** **Por lo tanto, todos los Ss son Ps, i.e.**  
 Por lo tanto, todas las novelas de espionaje de esta muestra particular son de origen Británico
- (2) **(RESULTADO)** **Todos los Ss son Ps, i.e.**  
 Todas las novelas de espionaje de esta muestra particular son de origen Británico
- (CASO)** **Todos los Ss son Ms, i.e.**  
 Todas las novelas de espionaje de esta muestra particular fueron tomadas de este baúl
- (REGLA)** **Por lo tanto, todos los Ms son Ps, i.e.**  
 Todas las novelas de espionaje de este baúl son de origen Británico
- (3) **(REGLA)** **Todos los Ms son Ps, i.e.**  
 Todas las novelas de espionaje de este baúl son de origen Británico
- (RESULTADO)** **Todos los Ss son Ps, i.e.**  
 Todas las novelas de espionaje de esta muestra particular son de origen Británico
- (CASO)** **Todos los Ss son Ms, i.e.**  
 Todas las novelas de espionaje de esta muestra particular fueron tomadas de este baúl

Como plantea Gooding (2004), el proceso abductivo de Peirce no dependía estrictamente de valores de verdad, sino que consistía en la generación no claramente explicada de una hipótesis utilizada junto a procesos deductivos para explicar hechos sorprendentes, lo cual era posteriormente validado por ‘inducción’, i.e. el tercer tipo de inferencia en el esquema de Pierce.

De todos modos, la posibilidad de comenzar a delinear una lógica de descubrimiento es lo que, posteriormente, da lugar a un interés por la teoría y el modelamiento de una lógica abductiva, concentrada en lo que se ha denominado la abducción sentencial, vale decir, un tipo de razonamiento abductivo basado en la representación lingüiforme de la información. La forma que tiene la abducción lógica desde esta perspectiva se puede ilustrar en la siguiente tabla, muy similar a la que Shelley utiliza para el mismo propósito:

Deducción	Inducción	Abducción
$\alpha \supset \beta$ [regla]	$\alpha$ [caso]	$\alpha \supset \beta$ [regla]
$\alpha$ [caso]	$\beta$ [resultado]	$\beta$ [resultado]
<u>Por lo tanto, <math>\beta</math></u> [resultado]	<u>Por lo tanto, <math>\alpha \supset \beta</math></u> [regla]	<u>Por lo tanto, <math>\alpha</math></u> [caso]

Tabla 1: Las formas silogística de la deducción, inducción y abducción

Como se aprecia en la tabla 1, la abducción se distingue por el hecho de que constituye una inferencia que va desde la regla ( $\square \supset \square$ ) y un resultado ( $\square$ ) hacia un caso inicial ( $\square$ ). Un ejemplo simple de abducción puede ser así:

Todos los vinos de esta caja son italianos	[regla]
<u>Estos vinos son italianos</u>	[resultado]
Por lo tanto, estos vinos son de esta caja.	[caso]

Se pueden destacar al menos tres observaciones a partir de este ejemplo: primero, que el caso abducido a partir de la evidencia tiene relación con la historia pasada o las condiciones iniciales de estos vinos; segundo, esta historia es una *explicación* potencial del por qué estos vinos (en contraste a otros, por ejemplo) son italianos; y tercero, que hay otras explicaciones posibles ya que los vinos italianos pueden venir de muchas lugares (o fuentes), de manera que la abducción no es una forma de inferencia estrictamente sólida [sound].

En los *Collected Papers* de Peirce (CP 2.755, 758<sup>7</sup>, según Shelley, 1996), se concibe también un rol del razonamiento abductivo, en tanto “lógica del descubrimiento”, que describe la formación de hipótesis y se ocupa de las potenciales causas de determinadas

<sup>7</sup> Esta nomenclatura hace referencia a volumen.párrafo-número en los *Collected Papers* que reúnen las conferencias de Peirce en Harvard (Hartshorne and Weiss, 1985), e.g., volumen 2, párrafos 755 y 758.

observaciones. Lo que se propone, entonces, es que un razonamiento abductivo apropiado licencia etapas posteriores de investigación inductiva y deductiva en ciencia, tal como se aprecia en las cuatro etapas de la investigación científica visualizadas por Peirce:

1. observación de un anomalía;
2. abducción de hipótesis con el propósito de explicar dicha anomalía;
3. testeo inductivo de hipótesis en la experimentación (y comparación con datos adicionales); y
4. confirmación deductiva de que las hipótesis seleccionadas realmente predicen la anomalía original (que ya deja de ser una anomalía).

En el ámbito de la inteligencia artificial, la caracterización de la abducción que propuso Peirce ha sido objeto de diversos intentos de generalización. Podemos ilustrar estos intentos a través de la teoría propuesta por Konolige (1990). Dicha teoría ha sido resumida por Thagard y Shelley (1997) en término de dos definiciones:

**Definición 1.1** Sea  $L$  un lenguaje de primer orden. Una teoría casual simple es una tupla  $\{C, E, \text{Sigma}\}$  donde:

**1.  $C$ , un conjunto de oraciones de  $L$ , son las causas; 2.  $E$ , un conjunto de oraciones de  $L$ , son los efectos; 3.  $\text{Sigma}$ , un conjunto de oraciones de  $L$ , es la teoría dominio.**

**Definición 1.2** Sea  $\{C, E, \text{Sigma}\}$  una teoría casual simple. Una explicación de un conjunto de observaciones  $O$  subconjunto de  $E$  es un conjunto finito  $A$  subconjunto de  $C$  tal que:

**1.  $A$  es consistente con  $\text{Sigma}$ ; 2.  $\text{Sigma} \cup A \neq O$ ; 3.  $A$  es un subconjunto minimal sobre conjuntos que satisfacen las primeras dos condiciones, i.e. no hay un subconjunto propio de  $A$  consistente con  $\text{Sigma}$  que implique  $O$ .**

Considerando la tabla 1, las generalizaciones en esta teoría se dan de la siguiente manera. El conjunto de causas  $C$  generaliza el resultado  $\square$ , el conjunto de efectos  $E$  generaliza el caso  $\square$ , y la teoría dominio  $\text{Sigma}$  constituye un conjunto de reglas que generaliza la regla  $\square \square$ . Los problemas de este tipo de teorías generalizadoras han sido discutidos por Thagard et al. (1997), en torno a lo que éste autor considera una concepción incorrecta de lo que se puede entender como una explicación científica (e.g. que la explicación no es deducción, que las hipótesis pueden tener múltiples capas, que la abducción puede ser creativa – y no sólo selectiva –, que las hipótesis pueden ser revolucionarias, etc.). En el ámbito de la investigación sobre RBM, este tipo de consideraciones críticas se asumen correctas, a partir de lo cual se propone la tesis de que el “razonamiento abductivo” (sea lo que fuere) no se puede limitar sólo al uso de representaciones sentenciales, o *proposicionales*. Suponiendo que el tipo de generalizaciones criticadas tienen una estructuración subyacente proposicional, investigadores como Thagard comienzan a explorar un tipo de evidencia que mostraría, por ejemplo, la relevancia de representaciones icónicas para entender mejor la inferencia abductiva. Esto daría pie para el estudio del rol de representaciones modélicas en el estudio de las prácticas científicas. Como el RBM es algo que se contrapone (o en el mejor de los casos complementa) al razonamiento proposicional, es necesario dar cuenta de éste último para entender el contraste señalado.

La noción de razonamiento proposicional sostiene que la mente está *equipada* (de alguna u otra manera) para llevar a cabo reglas formales de inferencias que le permiten realizar deducciones basadas en la hipótesis de que las representaciones mentales tienen una forma lógica, expresable a través de proposiciones. Esto, por sí sólo, corresponde a una metafísica de los estados y los procesos mentales, pero, desde la perspectiva práctica

del quehacer científico en ciencia cognitiva, eso no pareciera ser suficiente para dar cuenta de algún mecanismo físico que requiere explicitar poderes causales relacionados a tales procesos y estados de estructuración proposicional. A partir de esta teoría clásica del razonamiento humano, se han propuesto versiones de una TRM (teoría representacional de la mente) en la que convergen la lógica y la sintaxis. Cuando se asume que la forma lógica de las representaciones mentales puede ser instanciada a través de pensamientos con una sintaxis lógica, entonces también se asume que tales pensamientos pueden ser susceptibles de conectarse entre sí en virtud de la estructura sintáctica. La importancia de esto es que tal conectividad determinaría, por un lado, sus condiciones de verdad y su conducta inferencial, y por otro, sus roles causales en los procesos mentales. Un fuerte supuesto central, en este caso, es la hipótesis de que sólo algo que sea lingüiforme puede tener una forma lógica.

Ese es el supuesto central que pretende desafiar el RBM, si no completamente, al menos en lo que tiene que ver con la forma en que se llevaría a cabo un razonamiento científico donde lo que más importa es el rol de un tipo de representación modélica, alternativa a la lingüiforme o *sentencial*, la cual o bien se considera una ficción explicativa, o se confina a aspectos menos centrales del razonamiento científico (e.g. Nersessian 1999, Magnani 1999a). Para dar una cuenta preliminar de este tipo de razonamiento, tal como es el propósito de este apartado, seguiremos lo resumido por Li et al. (2006). Un primer elemento a constatar es que no existe un solo enfoque basado en modelos al que todos los investigadores del campo adhieran para dar cuenta de la práctica científica en tanto actividad cognitiva. Esto no es difícil de ilustrar, pero sí hace difícil la tarea de especificar aquello que se denomina razonar en base a modelos con miras a una posible teoría unificada. Una importante sugerencia metodológica proviene de Giere (2004), quien sostiene que un enfoque basado en modelos puede emplear hallazgos propios de la psicología cognitiva para tratar de entender el rol las teorías científicas (en tanto modelos utilizados como parte de una actividad pragmática de representación centrada en el agente) en la búsqueda de conocimiento científico. Li et al., resalta la seriedad de esta propuesta, porque aporta una mayor evidencia para lo que se ha dado en llamar la *hipótesis del continuo*. Este término puede ser un poco ambiguo si se considera que también suele utilizarse para hablar de un posible continuo entre una filosofía de la ciencia logicista y una historicista; o entre tipos de nociones de representación modélica más o menos constructivistas/instrumentalistas/pragmatistas (comentada anteriormente en la *Introducción*); o también como la que ahora se quiere aludir aquí, a saber, aquella entre la cognición científica y la cognición ordinaria. Lo que Li et al. señala es que la investigación en RBM debe incluir una perspectiva que incorpore tal continuo, toda vez que al soslayarla se corre el riesgo de crear la ilusión de que tanto la creatividad científica como el surgimiento de nuevas estructuras conceptuales sean consideradas un tipo peculiar de razonamiento científico, y por lo tanto caracterizadas por la existencia de mecanismos cognitivos ausentes en la cognición ordinaria. De aquí la idea defendida por estos autores de considerar (a pesar de no contar con una noción unificada de modelo) al RBM como una extensión del razonamiento ordinario, marcado por algún grado mayor de refinamiento y complejidad.

Otra consideración metodológica interesante es la distinción entre el “problema plural” y el “problema singular” de las representaciones que proponen Li et al. El primer caso tiene que ver con la naturaleza de los modelos en general, y no con la forma que éstos puedan tomar (ya se hable de estructuras o formatos). Un modelo en términos del problema singular tiene que ver con su estatus de modelo interno y a su funcionamiento en el razonamiento y la generación de nuevas estructuras internas. En su sentido generalizado, sostiene Li et al., un modelo mental (o modelo generalizado) cubre toda forma de estructura de conocimiento

en el contexto de un RBM en tanto proceso semántico. Alternativamente, en un sentido plural, los modelos mentales hacen referencia a la forma específica de una gran variedad de modelos mentales relacionadas a una variedad de clases de razonamiento. Esta distinción es relevante para estos autores porque ellos sostienen que es en la memoria de trabajo donde opera un RBM durante tareas cognitivas específicas, sobre la base de un ensamblaje específico de estructuras denominadas “modelos instanciados”, para diferenciarlos de otras posibles estructuras de conocimiento (modélicas o no) en la memoria de largo plazo. El rol crucial de la memoria de trabajo tiene que ver con la integración de información proveniente tanto de la memoria de largo plazo como de información perceptual, presentándose la oportunidad de generar nuevas estructuras modélicas instanciadas. Por esta razón, a la memoria de trabajo se le da el término de “aparato de ensamblaje” (p. 55). La idea de razonar en base a modelos, por lo tanto, tiene que ver con la hipótesis de que en la memoria de trabajo se instancia literalmente un modelo integrado (analógico o de simulación) a partir de distintos tipo de representaciones. Si se trata de información proveniente de la memoria de largo plazo, los autores en cuestión clasifican los modelos instanciados en la memoria de trabajo en al menos tres tipos: modelos teóricos (en la línea de los modelos de Gentner, entendidos como una clase de estructura de conocimiento de largo plazo), modelos esquemáticos (según Jonhson-Laird, i.e. estructuras de conocimiento construidas temporalmente en la memoria de trabajo), y modelos imaginísticos [*imagery models*] (en la línea de Kosslyn, vale decir, modelos en tanto representación de una imagen determinada neuronalmente).

La noción de RBM también ha estado restringida a aspectos creativos más acotados, como en el caso de Nersessian (2002), quien sostiene que el RBM causa el cambio conceptual en ciencia. Para esta autora el modelamiento visual y analógico pueden constreñir algún sistema conceptual actual en términos de los constreñimientos derivados del sistema *target*, generando nuevos constreñimientos que se integran en un sistema conceptual revisado. Esto equivale a sostener que para Nersessian los conceptos (sean lo que fueren) actúan como constreñimientos de modelamiento. Si comparamos esta visión con la noción de teorías científicas de Giere en tanto modelos incompletos, podemos entender que la caracterización de Nersessian no es claramente consistente con lo plantea Giere, donde el RBM parece tener un alcance más amplio (cosa que Nersessian (2008) no ignora). Consideremos que para Giere (1999) las leyes de la ciencia no son una afirmación acerca del mundo, sino más bien partes de una caracterización de modelos teóricos. Vale decir, dada la convicción de Giere con relación a la ambigüedad de las denominadas “leyes de la naturaleza”, las que más bien actuarían como generalizaciones empíricas de nivel inferior, las leyes debieran ser mejor entendidas como caracterizaciones parciales que se hacen verdaderas con respecto al modelo y no al mundo. De aquí que sólo al combinar estas partes con un contexto de problema específico se puede tener un sistema que puede ser comparado con un sistema real. En cualquier caso, el propósito de contrastar estas dos últimas visiones es que el RBM (como insiste Li et al.) es el de abordar posibles mecanismos de razonamientos de la cognición ordinaria como una forma de entender mejor las prácticas científicas significativamente productivas. Tal propósito es el que ha dado lugar a la idea de RBM como un tipo de proceso cognitivo guiado a metas derivadas de tareas cognitivas específicas. Tales metas determinarían qué tipo de constreñimientos de modelamiento son relevantes para un explanandum determinado y el nivel explicativo a alcanzar. Li et al. intentan ir más allá del ámbito de la memoria de trabajo en el que se circunscribe el RBM de Nersessian, y resumen un modelo de la explicación científica por RBM en las siguientes cuatro etapas: 1) abstracción genérica (i.e. la fijación de constreñimientos de modelamiento); 2) adscripción de constreñimientos [*feature constraint*] (búsqueda de una

relación apropiada de adscripción bajo alguna forma de representación de conocimiento en la memoria de largo plazo, y con respecto a alguna meta cognitiva específica – lo que contrasta con la forma de modelamiento mental de la generación de nuevas estructuras conceptuales según Nersessian); 3) la generación de un modelo instanciado en la memoria de trabajo; 4) la sensación de entendimiento [*feelings of understanding*] (*experiencia de entender el explanandum a través de un proceso de mapeo interno*).

### e) Noción de continuo representacional

Como se discutirá en 4.1.1.2., la noción de continuo representacional no hace referencia a una concepción necesariamente compartida por las diversas investigaciones convocadas en el estudio de un RBM en ciencia. Más bien, corresponde a una concepción aclaratoria que en este trabajo de tesis se propone como un supuesto necesariamente subyacente al modelo epistemológico unificado del RBM en ciencia que se estaría nutriendo prolíficamente del tipo de investigaciones mencionadas. Al asumir tal concepción, lo que se intenta es mostrar la inviabilidad de unificar, por un lado, nociones modélicas representacionistas alternativas a cualquier otra de estructuración posiblemente proposicional y, por otro lado, nociones modélicas artefactualistas que re-estipulan el estatus representacional de la representación en tanto re/representación (continua y productiva) de algún fenómeno determinado (ver capítulo II).

En términos generales, diremos que hay algo denominado RBM, y que eso está siendo estudiado en una diversidad de dominios y con una diversidad de herramientas metodológicas, en el entendido de que o bien se está estudiando un conjunto diverso de aspectos de un mismo fenómeno, o bien se están estudiando cosas muy distintas que sólo comparten en rótulo de RBM por razones más bien prácticas que programáticas. Con el fin de argumentar, por lo menos, que un proyecto programático caracterizado de manera tan heterogénea es poco viable (al menos por ahora), parece procedente tratar de identificar una noción de representación modélica que no soslaye aspectos que resultan centrales para la diversidad de aproximaciones que trataremos de hacer patentes. Dicha noción, me parece, puede ser descriptivamente caracterizada a lo largo de un continuo, cuyos extremos están marcados por el predominio de concepciones con características centrales significativamente disímiles. En el extremo más “conservador” de este continuo, se puede hacer referencia a una visión modélica cuyo carácter representacional guarda compromisos con una función representacional de los modelos determinada por algún tipo de estructura más o menos abstracta subyacente. En el extremo más “radical” del continuo, nos toparemos con una visión artefactual de los modelos cuyo carácter *re/representacional* (i.e. donde algo se vuelve a presentar literalmente) guarda compromisos con algún tipo de función *re/representacional* de los modelos determinada por la utilización de los modelos como instrumentos para interactuar con el mundo (también entendidos como *prótesis mentales* y/o artefactos en tanto *realidades paralelas artificiales*).

En la medida que logremos desenmarañar el denominado problema de la hibridación representacional de los modelos (capítulo IV), intentaremos hacer explícita la inviabilidad de unificar, por ejemplo, un tipo de concepciones de representación modélica fundadas en procesos de “abstracción genéricos” (o abstracción por modelamiento genérico), con otro tipo de concepciones de re/representación modélica fundadas en procesos (o más bien procedimientos) de “negociación epistémica” entre organismo y ambiente. La estrategia aclaratoria pasa por especificar el tipo de supuestos subyacentes que motivaría la aparente unificación programática. Con respecto a esto último, orientaremos el trabajo hacia los siguientes problemas: el surgimiento de la concepción artefactual de los modelos

(capítulo II), las concepciones de agencia cognitiva invocadas (como antecedentes de base experimental) en aspectos extramentales del RBM externalista (capítulo III), y las limitaciones y alcances de una posible teoría del surgimiento de hipótesis por RBM fundado en dichas concepciones, dentro del contexto de una poco saludable interacción entre filosofía y ciencia cognitiva (IV).

Como se puede apreciar, en este apartado hemos aprovechado de plantear algunas de las características recurrentes con respecto a la noción de RBM, pero, junto a ello, el desafío de ahondar en una mejor explicación de lo que se quiere entender por dicho tipo de razonamiento. Este es el objetivo que intentaremos alcanzar a partir de ahora, en el marco de la hipótesis contextualizada en la *Introducción*.

### 1.1.2.El quehacer científico como actividad cognitiva y el RBM

---

La actividad científica puede ser entendida no sólo como un objeto de justificación filosófica, sino también como un complejo proceso cognitivo sujeto a descripción y explicación. Esta última concepción del quehacer científico es la que ha guiado las investigaciones sobre Razonamiento Basado en Modelo (RBM) realizadas conjuntamente por una comunidad no despreciable de filósofos experimentales y científicos cognitivos durante las últimas décadas. Uno de los propósitos del presente trabajo de investigación es dar cuenta de las motivaciones que subyacen a esta última concepción de la actividad científica. Si es correcta, dicha concepción debiera constituir la mejor aproximación explicativa disponible de los procesos cognitivos que explican la productividad de las prácticas científicas que conducen al surgimiento de nuevas hipótesis.

Para los investigadores que adhieren las visiones más externalistas desarrolladas en ciencia cognitiva durante los últimos 30 años, este ámbito de investigación ha dejado de identificar la cognición humana con explicaciones “logicistas” o basadas en “reglas y representaciones” en ámbitos de resolución de problemas severamente restringidos, tal como lo suponía la Inteligencia Artificial clásica (GOFAI o *Good Old-Fashioned Artificial Intelligence*) de la Revolución Cognitiva. Alternativamente, el RBM debe ser entendido en el contexto de la Contra-Revolución Cognitiva (Grush, 2002), caracterizada en gran medida por la investigación empírica realizada en *ambientes del mundo real*.

Sobre la base de algunos de los diversos enfoques de esta Contra-Revolución que informan las aproximaciones teórico-descriptivas del RBM, los investigadores del caso - con grado variable de compromiso, como se dijo - argumentan que una concepción apropiada de la cognición científica debe ser *corporalizada, situada, distribuida, extendida*, y en gran medida *ecológicamente determinada*<sup>8</sup>, entre otras orientaciones metodológicas.

En una perspectiva preliminar, el RBM puede ser considerado como una innovadora aproximación descriptiva y explicativa al problema del descubrimiento en ciencia. Esta aproximación guarda un grado variable de compromiso con enfoques experimentales anti-individualistas de la cognición humana, surgidos como reacción a la visión computacional/representacional de la cognición abrazada por la ciencia cognitiva de “primera generación” (ver Grush, 2002). Dado este contexto, es importante considerar una particularidad relacionada al tipo de supuestos y principios metodológicos reguladores<sup>9</sup> que caracterizan las hipótesis alternativas de la cognición. Según lo que nos interesa evaluar

<sup>8</sup> Estas visiones se abordarán con distinto nivel de detalle a lo largo del presente trabajo.

<sup>9</sup> E.g. acción ecológica, cognición activa (basada en *affordances*), acción epistémica v/s acción pragmática, *principle of causal spread*, representación sensoriomotora, etc. (Ver Clark, 2008 ; Nersessian, 2008)

---

en este trabajo, veremos que muchos de los planteamientos utilizados para seleccionar la noción de agencia cognitiva que los investigadores en RBM externalistas consideren más apropiada para dar cuenta de ese tipo de razonamiento (en alguna comunidad dada) está determinado por supuestos metodológicos que también regulan las hipótesis surgidas de la investigación en ciencia cognitiva.

Como se sabe, esto último es reflejo de que los enfoques filosóficos que inspiran una metodología científica tienden a reflejar una concepción específica de la cognición humana. Lo que se pueda decir aquí acerca del RBM, nutrido de los actuales desarrollos experimentales, no es una excepción. Si uno piensa, como es el caso aquí, que el tipo de filosofía que inspira a los enfoques alternativos de la cognición influye en el tipo de supuestos metodológicos centrales que guían la investigación correspondiente, entonces sería posible evaluar dichos enfoques con respecto a las características que posea la concepción de agencia cognitiva que dicha simbiosis pueda reflejar. Luego, si tal concepción resulta, por cualquier razón, poco convincente, entonces es probable que la confiabilidad de los supuestos (o principios) metodológicos en cuestión pueda ser discriminados con relación a los problemas detectados a partir del soporte filosófico subyacente. Este es el problema que trataremos de especificar a lo largo del capítulo IV, junto con intentar establecer los términos en los que una probable teoría del RBM parece estar determinada por los efectos de una tensión filosófica (racionalista/pragmatista) subyacente a las propuestas epistemológicas de la comunidad de RBM.

No obstante aquello, es preciso señalar que al hablar anteriormente de un ciencia cognitiva de primera generación no se está reconociendo la existencia de alguna ciencia cognitiva que pueda ser claramente rotulada como de “segunda generación”. Lo que más bien sucede con respecto al RBM y a los investigadores del campo, es un compromiso explícito, aunque no unificado, con un “conjunto de visiones” que proponen mover los límites de la *representación* y el *procesamiento* fuera del individuo. Desde esta posición, el pensamiento científico se interpreta como un sistema cognitivo-cultural *distribuido*, y la explicación de las prácticas científicas concretas en ambiente real suelen ser objeto de interpretaciones fundadas sobre un conglomerado de tesis aparentemente sustantivas de la cognición. Un ejemplo de esto lo constituye un tipo de investigación donde se manifiesta explícitamente que uno de los principales propósitos de la *simulación basadas en modelo* en ciencia es “la función de modelos físicos como artefactos cognitivos en procesos cognitivos distribuidos” (Nersessian, 2009, p. 731).

En este respecto, el tratamiento más morigerado de Giere (2002, entre otros) sobre el rol de los modelos en ciencia actúa como un referente importante. Este filósofo se ha enfocado, precisamente, en el carácter distribuido de la cognición científica, sosteniendo que el conocimiento científico es un producto impersonal de toda la comunidad científica, lo que incluye tanto a los agentes humanos como a diversos tipos de artefactos. En esta perspectiva, los instrumentos y aparatos experimentales, así como los modelos materiales o los mismos computadores que se utilicen, son considerados como *implementos cognitivos* para la investigación científica. Este rol aparentemente fundamental que se le atribuye a los artefactos en la cognición distribuida los ha puesto, para propósitos de análisis cognitivo, al mismo nivel del científico que los manipula. El reconocimiento de un rol cognitivo de los artefactos utilizados en el trabajo científico (especialmente el conducente a descubrimientos revolucionarios) es lo que haría la diferencia entre un sistema cognitivo distribuido y la mera idea de la cognición social o colectiva. En este tipo de sistemas, por lo tanto, las actividades cognitivas como el cálculo matemático o la interpretación de datos estarían distribuidas, no sólo entre los científicos, sino que también literalmente implementados por los recursos

o instrumentos materiales disponibles - a los que incluso se les atribuye cierto grado de autonomía cognitiva<sup>10</sup>. La tesis que se defiende a través de estos planteamientos es que, *en tanto herramientas cognitivas, los artefactos afectan la manera en que los procedimientos científicos son implementados*. Podríamos decir que esta tesis tiene al menos dos lecturas relevantes para el análisis epistemológico:

1. Los artefactos son diseñados para ser operados con facilidad por humanos con cerebros comparadores de patrones [pattern matching], amplificando las habilidades cognitivas del científico.
2. En vez de “amplificar” [enhance] las habilidades cognitivas del científico, los artefactos transforman la tarea que éstos realizan, toda vez que la tarea o problema se representa en un dominio donde la solución o el camino a la solución se hace evidente, evitando parte importante del procesamiento interno.

Cuando se está dispuesto a aceptar alguna de estas lecturas, lo que también se hace, implícitamente, es comprar una de las posiciones adscritas al debate que en filosofía experimental se ha dado entre una visión débil o radicalmente externalista de la cognición, respectivamente (Clark, 2007; Rupert, 2004). Grosso modo, la primera visión defiende lo que se conoce como un carácter *embedded* del procesamiento cognitivo, según el cual se podría sostener una concepción anti-individualista de la agencia cognitiva, al tiempo que no se abandona la centralidad de procesamiento cognitivo del cerebro como locus biológico del razonamiento. La segunda visión, de carácter *extended* (como se verá en el capítulo IV), aboga por una suerte de paridad cognitiva entre componentes biológicos internos y artefactuales externos, asumiéndose que los procesos más interesantes no tienen mucho que ver con la centralidad de las nociones clásicas de razonamiento y representación, elementos que más bien estarían al servicio de procedimientos epistémicos y acciones ecológicamente exitosas.

En este contexto, Giere parece sostener la primera lectura de la tesis mencionada, es decir, un rol de los artefactos dentro de un sistema cognitivo distribuido, en el entendido de que éstos potencian el procesamiento interno del agente cognitivo, sin transformarlo sustancialmente. Alternativamente, Nersessian (con cierta ambigüedad), Gooding y Magnani entre otros - coincidiendo con Hutchins -, asumen la segunda lectura. Y en el RBM, el entendimiento del quehacer científico como actividad cognitiva, subyacente a las diversas investigaciones empíricas (ya sean etnográficas, histórico-cognitivas, o de modelamiento computacional), tienden a oscilar entre estas dos posturas. En este sentido, aquellas que se acercan a la postura de Giere pueden todavía reivindicar un papel más central de la agencia intencional en los posibles procesos generadores de hipótesis. Alternativamente, aquellos enfoques que adhieren a la segunda postura, vale decir, un rol cognitivo relativamente más autónomo de los artefactos<sup>11</sup>, tienden a defender la hipótesis de que parte importante del esfuerzo de procesamiento cognitivo es *literalmente descargado* en los artefactos (e.g. transformando problemas cognitivos complejos en problemas perceptuales, como cuando los científicos utilizan gráficos en vez de tablas estadísticas; o cuando incluso los buenos ajedrecista, según un ejemplo de Clark (Sterelny, 2004), prefieren practicar con tablero y piezas reales). Este es el sentido en que la cognición puede considerarse como un fenómeno *extendido*.

<sup>10</sup> Este carácter de autonomía cognitiva debe entenderse fuera de los cánones de una concepción individualista de cognición. Una caracterización más precisa será ponderada en 4.1.1.1. y 4.1.1.2.

<sup>11</sup> No sólo *en virtud* del supuesto rol facilitador, sino que también transformador por el que estos “implementos cognitivos” afectan las tareas realizadas por un agente.

Dadas estas nuevas aproximaciones de los estudios de la cognición en el dominio de las prácticas científicas, considero importante tratar de comprender tanto los criterios que justifican los constreñimientos específicos del RBM, así como las condiciones básicas de adecuación que tendría que satisfacer la manipulación de modelos en RBM para la generación de nuevas hipótesis. Con este doble objetivo, el presente trabajo se restringe básicamente a los enfoques de tendencia *externalista*. Esta tendencia resulta ubicua en gran parte de las investigaciones sobre el *razonamiento creativo* científico, al tiempo que dichos tratamientos resultan, en mayor o menor medida, tributarios de enfoques filosófico-experimentales actuales de mayor alcance que, sobre la base de una serie de heterogéneas generalizaciones empíricas<sup>12</sup>, buscan consolidar ciertas teorías empíricas de la cognición y la mente, desafiando el ámbito de la noción tradicional de la cognición de corte individualista. En otras palabras, y como será especificado en la sección siguiente (1.1.3.), los enfoques de RBM externalistas que aquí se analizarán críticamente persiguen una posible teoría unificada en la medida que establecen compromisos con aquello que se generaliza en términos de las teorías alternativas de la cognición (ver sección c en 1.1.1.), las que terminan actuando como antecedentes directos de la investigación en RBM.

Una de las características principales de esta nueva concepción de la actividad científica en tanto actividad cognitiva es un cambio de foco en el estudio de la “creatividad científica” en tanto actividad productiva (o ampliativa de conocimiento). A juicio de Nersessian (2008), esto implica pasar desde la centralidad que ha ocupado el cambio y la innovación conceptual en las denominadas “revoluciones conceptuales” hacia un énfasis en las prácticas de resolución de problemas que tienen lugar, precisamente, cuando se crean nuevas estructuras conceptuales. Al hablar de estructuras conceptuales, en un sentido tradicional o representacionista, se está hablando de alguna manera específica de representar la naturaleza. Sin embargo, ya no es éste el tipo de representación que interesa investigar en el RBM (al menos de manera exclusiva o central) para entender los procesos y mecanismos que puedan dar cuenta del descubrimiento científico. De hecho, para el caso de los enfoques más fuertemente externalistas de RBM, las representaciones que interesan son aquellas más directamente vinculadas a las prácticas científicas modeladoras y manipuladoras de modelos físicos. En este contexto, adquiere importancia el estudio de algún tipo de representación modélica que la comunidad de investigadores relevante considere que juega un rol facilitador de la manera en que los científicos razonan acerca del fenómeno determinado que se pretenda entender. Esta consideración no es una que se circunscriba al mero rol instrumental que los modelos puedan cumplir en dicha clase de episodios de razonamientos. Por el contrario, dentro de la investigación sobre el descubrimiento científico actual, hay quienes se apoyan en el supuesto de que una noción tradicional de razonamiento científico puede ser entendida sobre una concepción de liberalizada de RBM (Li et al., 2006). Vale decir, una concepción que no considera apropiado dar cuenta de alguna forma específica de representación modélica para definir tal razonamiento, “ya que los modelos mentales en cuestión son construidos con diferentes contenidos (tipos y niveles de información), estructuras, y formatos de representación interna, pudiendo tomar así muchas formas.” (p. 54). Algunos ejemplos de formato de representación interna a los que hace alusión son: imágenes, prototipos, marcos, esquemas, modelos mentales, símbolos preceptuales, y teorías del campo de la psicología.

Adicionalmente, cabe enfatizar que al enfocar la investigación en cuestión desde la perspectiva de un RBM, también se persigue “establecer una base cognitiva para las prácticas científicas en la cognición ordinaria, enriqueciendo a la vez un entendimiento de

<sup>12</sup> Esto ha sido explicitado en torno a visiones como las examinadas por M. Wilson (2002)

la cognición general” (Nersessian, 2008, p.2). Desde una perspectiva crítica, es posible que dicha *base cognitiva* esté fundada en un rol instrumental de los modelos en tanto “artefactos cognitivos” (usando la expresión de Hutchins, 1999). A eso, una posible respuesta de parte de los defensores del RBM externalista podría hacer alusión a la tesis de la cognición distribuida recién aludida, en sus dos lecturas. Como se ha mencionado, los artefactos cognitivos son tales porque están incrustados en sistemas socio-culturales más amplios que organizarían (o afectarían) las prácticas en las que son utilizados. En estos ambientes estructurados, es frecuente que los artefactos sean utilizados no sólo para lo que han sido específicamente diseñados, sino también para propósitos distintos y oportunistas (e.g. pilotos de aerolíneas que verifican el rango de desviación de velocidad por la manera en que miran un instrumental, y no por medio de los datos proporcionados por el instrumento para realizar los cálculos pertinentes (Hutchins, 1995)). Dado este sistema socio-cultural, el foco puede ponerse en el producto de la *actividad cognitiva*, o bien, en las prácticas que constituirían tal actividad cognitiva. En el primer caso, se constataría cómo los artefactos amplifican la actividad cognitiva interna (ej. resultados obtenidos a través calculadora). En el segundo, se haría ostensible la manera en que se modifica y reduce la actividad o procesamiento internos, dado el modo en que dicha actividad se ve afectada por las “oportunidades de acción” (o *affordance*) que presentaría el artefacto. Sin embargo, es importante tomar en cuenta lo siguiente. Por un lado, el rol de cualquier artefacto en la solución de algún problema (que requiera conductas complejas) depende siempre del conocimiento interno que posea el agente que lo manipula, sin importar cuánto se modifique dicho conocimiento en virtud de la materialidad o el uso del mismo artefacto. Por otro lado, lo que se pueda decir del rol cognitivo de un artefacto bien puede aplicar para otros aspectos del entorno material estructurado (en tanto “ambiente cognitivo estructurado”), desdibujando gravemente los límites de lo cognitivo. Estas últimas consideraciones hacen patente la dificultad llegar tempranamente a una conclusión decisiva con respecto al tipo o grado de compromiso que el RBM tenga con alguna de las dos tesis señaladas (facilitadora o transformadora de procesos cognitivos), asumiendo que tal distinción es metodológicamente plausible.

En el presente capítulo, así como en el siguiente, lo que se pretende mostrar es el tipo de problemas que se considera necesario atender y tratar en mayor detalle a lo largo de los capítulos posteriores. En este sentido, considero importante proponer un esquema provisional de la externalización cognitiva en RBM, para entender algunas de las motivaciones detrás de la consiguiente liberación de factores causales en el estudio del “razonamiento creativo” en ciencia.

### 1.1.3. Esquema provisional de la externalización cognitiva en RBM

---

La reflexión filosófica tradicional sobre la actividad científica se ha centrado en la reconstrucción racional de las teorías científicas basada en la lógica deductiva e inductiva (ver sección *b* de 1.1.1.). Al incorporar un RBM en una posible teoría sistemática del surgimiento de hipótesis en ciencia, se propone un proceso inferencial distinto y supuestamente complementario al de las inferencias proposicionales propias de la justificación tradicional. Dicha complementariedad, asumiendo que es posible, está lejos de ser un proceso que se pueda reconocer y descifrar fácilmente en los abundantes y variados informes de investigaciones realizadas por los especialistas miembros de la Comunidad de Razonamiento Basado en Modelos [*MBR Community*] <sup>13</sup>.

<sup>13</sup> [www.unipv.it/webphilos\\_lab/cpl2/](http://www.unipv.it/webphilos_lab/cpl2/)

La búsqueda de un modelo epistemológico unificador del descubrimiento en ciencia, como del que se intenta dar cuenta en el tipo de investigaciones producidas en el marco de la comunidad mencionada, tiene un punto de partida común: las aparentes limitaciones de una lógica del descubrimiento científico establecidas sobre la base de un proceso inferencial abductivo<sup>14</sup>, formulado proposicionalmente y guiado por reglas explícitas (e.g. propias de una lógica de primer y segundo orden). Alternativamente, investigadores como Nersessian (2009), tal como se ha señalado, adhieren aquella reconceptualización de los procesos cognitivos según la cual dichos procesos incluyen sistemas formados por humanos y artefactos:

***“Por supuesto, como Hutchins y otros han constatado, los cerebros forman parte de sistemas cognitivos y los procesos perceptuales y motores juegan un rol en la coordinación de representaciones artefactuales” (p. 745)***

Sin embargo, Nersessian también presta atención a uno de los problemas centrales de una posible teoría del RBM. Este problema tiene que ver con la naturaleza de las representaciones que tienen lugar dentro de los límites del componente humano. Nersessian parece asumir algo que podría considerarse un supuesto compartido entre sus pares, a saber, que la participación de este componente en un sistema cognitivo distribuido – según se sostiene - debe estar crucialmente facilitada por tales representaciones. Dada la centralidad de estas representaciones, y la generalizada aceptación del supuesto señalado, pareciera correcta la adscripción del carácter modélico de las mismas como base del razonamiento científico, a expensas de la formulación proposicional de los procesos inferenciales abductivos.

Desde Peirce, los estudios acerca de la generación de hipótesis se han centrado en la abducción. Ya sea desde la filosofía o desde la ciencia cognitiva, dichos estudios no han logrado dar cuenta adecuadamente de tal proceso de generación. El propósito en común que comparten los diversos estudios en cuestión<sup>15</sup> es explicar la generación de hipótesis mediante la satisfacción de constreñimientos o condiciones que se harían explícitos a través de reglas que prescriben la transformación de expresiones simbólicas en la modelación computacional. Dado que la abducción se ha articulado tradicionalmente en un formato de estructuración (supuestamente) proposicional, es entendible que aquel formato representacional exija el uso de constreñimientos explícitos (e.g. reglas de carácter lógico-sintácticas) cuando se trate de testear o dar cuenta de un proceso inferencial/computacional de generación de hipótesis.

El problema que surge, por lo tanto, tiene que ver con la búsqueda de otro tipo de representación que resulte más adecuada no sólo para abordar constreñimientos explícitos tradicionales sino también implícitos. Intuitivamente - para la comunidad de RBM -, es posible que ciertos procesos productivos (o de ampliación creativa del conocimiento) dependan no sólo de la abducción en términos de la tradición proposicional. Esto se puede interpretar, al menos, en dos sentidos: o se incluyen otras etapas al proceso de generación abductivo, o se restringe la abducción a determinadas etapas específicas. Según Magnani (1999a), el RBM propone una combinación de ambos. Vale decir, por un lado, se amplían las etapas del proceso de razonamiento creativo, y por otro, se confina la noción tradicional de abducción a un ámbito más periférico de tal proceso, bajo la etiqueta de “abducción sentencial” (ver capítulo III para un desarrollo más detallado). Las razones por las que

<sup>14</sup> Ver nota 2.

<sup>15</sup> Ver nota 3.

Magnani<sup>16</sup> le quita centralidad a este tipo de abducción tienen que ver con lo que considera un conjunto de aspectos del razonamiento creativo que no lograrían ser capturado por los “modelos sentenciales de la abducción”. A saber:

- El rol de las explicaciones estadísticas, donde lo que es explicado se sigue sólo probabilísticamente y no deductivamente a partir de las leyes y otros instrumentos que hacen la explicación.
- Las condiciones suficientes para la explicación.
- La idea de la existencia de clases de abducción *creativa* de nivel superior.
- La (supuesta) existencia de abducciones basadas en modelos (e.g. visuales y diagramáticas)
- El hecho de que usualmente las explicaciones no son completas sino que sólo dan cuenta parcialmente de la evidencia pertinente
- Que lo más importante de las nuevas hipótesis científicas (y teoría científica) es el poder que tienen para explicar *nuevos* hechos previamente *desconocidos*.<sup>17</sup>

La centralidad de la visión alternativa a la abducción sentencial que Magnani pretende desarrollar no es compartida por toda la comunidad de investigadores del RBM. Sin embargo, sí existe un consenso general con respecto a las supuestas limitaciones del razonamiento científico por medio algún proceso exclusivamente basado en representaciones de estructuración proposicional. Este es un consenso regulado por antecedentes experimentales de las ciencias de la cognición que nutren la reconceptualización de los procesos cognitivos aludida en la cita anterior de Nerssesian. Según esta autora (2008), el consenso señalado tiene que ver con la constatación de que aquellos procesos que guían el razonamiento científico hacia el surgimiento de nuevas representaciones conceptuales no se basan en sistemas axiomáticos o redes proposicionales, sino que más bien en la construcción, manipulación, evaluación y adaptación de modelos. Estas prácticas serían primarias en la formación de teorías y, por lo tanto, precursoras de la correspondiente abstracción que subyace a la expresión formal de las leyes y axiomas que requieren las teorías. Esto le asignaría una importancia central al estudio del rol que juegan los modelos como facilitadores del razonamiento acerca de los fenómenos del mundo que se pretenden entender y explicar [*target phenomena*].

Para entender la manera en que se conjugan nociones como la de distribución cognitiva (ver sección c en 1.1.1.) y el énfasis en prácticas vinculadas a los modelos en ciencia, se debe tener en cuenta lo siguiente. Tradicionalmente, los modelos en ciencia han servido o bien para representar aspectos del mundo, o bien para representar teorías en el sentido de que los modelos puedan interpretar las leyes y los axiomas de esa teoría. Sin embargo, para autores que adhieren aspectos de RBM tan diversos como los que estudian Nerssesian y Magnani (e.g. cambio conceptual y abducción manipulativas, respectivamente), el rol señalado de los modelos estaría incorrectamente constreñido por una concepción de razonamiento limitada por la lógica clásica. Lo correcto, para estos autores, es romper aquella relación de interdependencia, aceptar que no puede existir una “lógica del descubrimiento” en esos términos, y que tal relación entre razonamiento y lógica es sólo posible después que ocurren ciertos *procedimientos creativos* que incluyen la participación de representaciones modélicas, susceptibles de ser considerados como procesos de razonamiento, incluso aunque carezcan del carácter lógico de la visión tradicional. Esta visión tradicional del razonamiento creativo es la que Magnani ha rotulado

<sup>16</sup> Magnani, como se verá en capítulos siguientes, defiende la noción de *abducción manipulativa* como una dimensión extra-mental central de procesos inferenciales complementaria a la sentencial y a la basada en modelos.

como “abducción sentencial”, el que, resumidamente, no hace sino referencia a *la aplicación de algoritmos deductivos o inductivos sobre un conjunto finito de proposiciones*.

Luego, la ampliación de la noción de razonamiento y, por lo tanto, del proceso generador de hipótesis está marcada por la incorporación de un tipo distinto de representación, i.e. la modélica. Se sostiene que dicha representación podría capturar de manera más cabal los procesos centrales de la producción de nuevas hipótesis explicativas, toda vez que es capaz de adecuarse a las características de un *sistema inferencial distribuido* o bien al RBM como método de inferencia. Aunque esto último se tratará en detalle más adelante (ver capítulo III), es importante notar cómo comienzan a manifestarse al menos dos aspectos del enfoque del RBM que nos interesa evaluar críticamente en este trabajo. Por un lado, la posibilidad de una concepción ampliada (y de alguna manera no logicista) de razonamiento daría lugar a una inversión epistemológica aparentemente sustancial, a saber, aquella que coloca procedimientos de carácter epistémicos (i.e. ampliativos de conocimiento) como antecedendo a procesos de razonamiento lógico. Por otro lado, la liberación de factores constreñidores de la generación de hipótesis explicativas desafiaría una estrategia metodológica recomendada como un defensa débil del estudio científico clásico de la cognición, con respecto a la abucción más tradicional: concentrar los esfuerzos de investigación en áreas de procesamiento cognitivo donde los efectos de globalidad (u holísticos) se minimicen incluso hasta el punto de poder ignorarlos, sin que por ello se desatienda su potencial relevancia científica (Fodor, 2000).

Los dos aspectos señalados en el párrafo anterior pueden entenderse como regulados por una perspectiva pragmatista, tal como se caracterizó en la *Introducción* de este trabajo. De hecho, desde una posición racionalista de la ciencia cognitiva, uno bien podría esperar que, de ser cierta una teoría del RBM, ésta contribuya a establecer la manera en que determinados constreñimientos implícitos cumplan un rol relativo a procesos de razonamiento propocionales que se consideren más inherentemente cognitivos. En este sentido, no habría necesidad de desafiar ninguno de los dos aspectos señalados en el párrafo anterior, i.e. la noción de razonamiento logicista y la recomendación metodológica. Pero las consideraciones metodológicas que interesan al RBM como método de inferencia (dada la relación de dependencia que guarda con las hipótesis alternativas de la cognición) guardan compromisos laxos con una concepción alternativa y liberalizada de al menos tres nociones otrora centrales para la ciencia cognitiva clásica. Específicamente, nos referimos a la tríada “procesamiento cognitivo”, “información” y “representación”. Siguiendo a Gatti et al. (2005), la re-estipulación de estas nociones se resume de la siguiente manera:

***“Definimos un proceso cognitivo como cualquier proceso de recolección y procesamiento de información con el fin de alcanzar una respuesta o lograr un desempeño satisfactorio frente a una tarea específica [...] Definimos información como cualquier conjunto de datos [data] que incremente momentánea o permanentemente la base de conocimiento de un agente cognitivo. La información viene siempre dada de alguna forma y esta forma de la información puede ser considerada la representación de la información. Así, la definición general de representación que usamos es aquella de representación como la forma de la información.” (p. 233, énfasis original).***

Con respecto a esta cita, Gatti et al. (2005), explica que la “redefinición” de estos conceptos constituye una reacción al paradigma funcional/simbólico de la cognición, ya que pretende poner la atención precisamente en aquello que el paradigma clásico consideraba irrelevante: el “hardware” que realiza los procesos cognitivos. Tal como en el caso de la distribución cognitiva, diremos que esto expresa un claro compromiso con

la concepción corporalizada de la cognición (ver c en 1.1.1.), la que actúa como *punte* entre la consideración de representaciones internas y externas para dar cuenta de la noción de sistema cognitivo más apropiada para un RBM externalista. Uno de los antecedentes experimentales de este tipo de compromiso, lo constituye el programa de investigación denominado “constructivismo neuronal”, por el cual se pretende fundamentar la noción de representación modélica, independientemente de lo heterogéneo que resulte el formato de tal representación. Según el constructivismo neuronal (Clark, 2003), el crecimiento del cerebro, y especialmente el crecimiento sináptico y dendrítico, está determinado por el ambiente en el que crecimos. Gatti et al. reformula esta correlación afirmando que “la estructura del ambiente, la estructura de nuestra experiencia le da forma a la estructura de nuestro dispositivo de razonamiento biológico” (2005, p. 234). Y a partir de aquí se plantea una hipótesis más fuerte: *si el constructivismo neuronal es correcto, entonces se podría explicar porqué los seres humanos dependen tan profundamente del ambiente, poniendo atención a la naturaleza de las representaciones externas y las maneras en que éstas ayudan a darle forma a las tareas cognitivas.*

Mientras que la concepción de la cognición como sistema distribuido es uno de los enfoques que en el ámbito de las ciencias cognitivas proponen modelos que intentan mover los límites de la cognición “horizontalmente hacia el ambiente”, la concepción de corporalizada de la cognición actúan como uno de los enfoques que proponen mover el énfasis en una dirección “verticalmente hacia el cerebro”. Ambos enfoques, como veremos, actúan como antecedentes del RBM externalista. Esta es la razón por la que aquí nos interesa tratar de plasmar un esquema preliminar de la externalización en RBM, a partir de las limitaciones atribuidas a los enfoques abductivos tradicionales, y de los antecedentes filosófico-experimentales que, supuestamente, ayudarían a dilucidar a un sistema de representación e inferencia alternativo, que sobrepasa los límites del denominado “componente humano”.

Nersessian (2009) sostiene que la cognición distribuida actúa como un “marco” interpretativo para el RBM. El sentido que esta autora le atribuye a este marco es el de un intento por reconstruir un modelo del mundo que satisfaga criterios distintos a la mera *testeabilidad* como tal (i.e. cuan verdadero o falso es un modelo en tanto representación confiable de una parte de la realidad). Si bien el criterio de testeabilidad es requerido en el caso de las teorías, un marco adecuado debe ser consistente con los procesos en el mundo real [*in the wild*], en alguna forma que permita *anclar* dicho marco a los datos disponibles. Adicionalmente, Nersessian atribuye a este marco interpretativo el potencial de conducir a nuevas hipótesis que pueden ser exploradas y testeadas en la investigación experimental. El denominado enfoque corporalizado de la cognición también sería consistente con *procesos en el mundo real*, dado que, como se dijo anteriormente, es también un hecho claro para los investigadores del RBM que los procesos perceptuales y motores juegan un rol coordinador de las representaciones artefactuales (i.e. extracraneales y modélicas de formato variable, conjuntamente). Y un enfoque embebido [*embedded*] sería consistente con el anclaje señalado, toda vez es en el “estar ahí” – en terminología Clarkeana - donde se manipulan las representaciones externas (artefactos, modelos físicos, etc.) que se acoplan a las internas en virtud de alguna característica modélica intrínseca de estas últimas.

Sobre la base del marco interpretativo señalado, se desarrolla una amplia línea de investigación inspirada por la idea de que los procesos inferenciales involucrados en la generación de nuevas hipótesis son mejor caracterizados como RBM en tanto sistema inferencial distribuido. Sin embargo, una clara identificación de los constreñimientos de un RBM, así como de las condiciones de adecuación de una teoría basada en un RBM

---

para la generación de hipótesis, se ven mayormente obstaculizadas por la diversidad de procedimientos y principios metodológicos reguladores de las investigaciones relevantes. No obstante, parece apropiado comenzar a desenredar este problema a partir de un supuesto general que tiende a guiar los desarrollos teórico-experimentales en la ciencia cognitiva durante las últimas décadas. El supuesto en cuestión consiste en que la cognición puede caracterizarse como el producto causal entre los componentes de la tríada cerebro-cuerpo-mente, lo que queda bien resumido en la siguiente cita:

***“Los procesos cognitivos se extienden a lo largo del cerebro, el cuerpo y el ambiente: entender la cognición es entender la interacción entre estas tres dimensiones: los procesos de razonamiento internos no son más esencialmente cognitivos que la hábil ejecución de un movimiento coordinado o la naturaleza del ambiente en el cual la cognición tiene lugar” [Cognitive processes span the brain, the body, and the environment: to understand cognition is to understand the interplay of all three. Inner reasoning processes are no more essentially cognitive that the skillful execution of coordinated movement or the nature of the environment in which cognition takes place.] (van Gelder y Port, 1995, p. viii-ix)***

Es en este contexto que los factores causales relevantes en un RBM, así como la misma noción clave de modelo, tienden a ser objeto de una liberalización radical que hace difícil identificar el rol de los múltiples determinantes que suelen ser considerados en el surgimiento de hipótesis explicativas en ciencia. A lo largo de este trabajo de investigación, se intentará alcanzar cierta claridad en esta problemática, no sin antes delinear, sólo con un propósito didáctico e instrumental, un esquema sinóptico del contexto en que ocurre la liberalización de factores causales comentada. Una razón importante para este recurso inicial es la necesidad de ilustrar claramente la descentralización progresiva (aunque no total) del rol asignado a los agentes cognitivos, en tanto agentes intencionales, por parte del quehacer investigativo en la ciencia cognitiva actual. Conforme al siguiente esquema inicial, será importante reflexionar, finalmente, acerca del grado de plausibilidad general de las propuestas externalistas de la generación de hipótesis por medio de un RBM.

El esquema que se propone a continuación plantea que las explicaciones más o menos externalistas en RMB son el resultado de un conjunto de transiciones a lo largo de al menos tres ámbitos, cada uno asociado a distintas concepciones de la mente. Estos ámbitos guardan correspondencia con los tres factores casuales de la cognición anteriormente señalados (cerebro, cuerpo y mundo). Al mismo tiempo, se optará por sostener que las transiciones señaladas están motivadas por el surgimiento de distintas concepciones de agentes cognitivos, que actúan como el foco de procedimientos y principios de base pragmática subyacentes a la liberación de factores determinantes de un RBM. Estos factores pueden estar constituidos por la mayor o menor participación de estructuras extra-mentales, así como de cierta información sensorio-motora o bien propia de las emociones –subyugadas a un trasfondo de transacciones adaptativas –, lo que contrasta radicalmente con los procesos cognitivos “intracraneales” tradicionales de base representacional/simbólica. El siguiente diagrama ilustra lo que se quiere plantear como esquema de trasfondo del RBM:

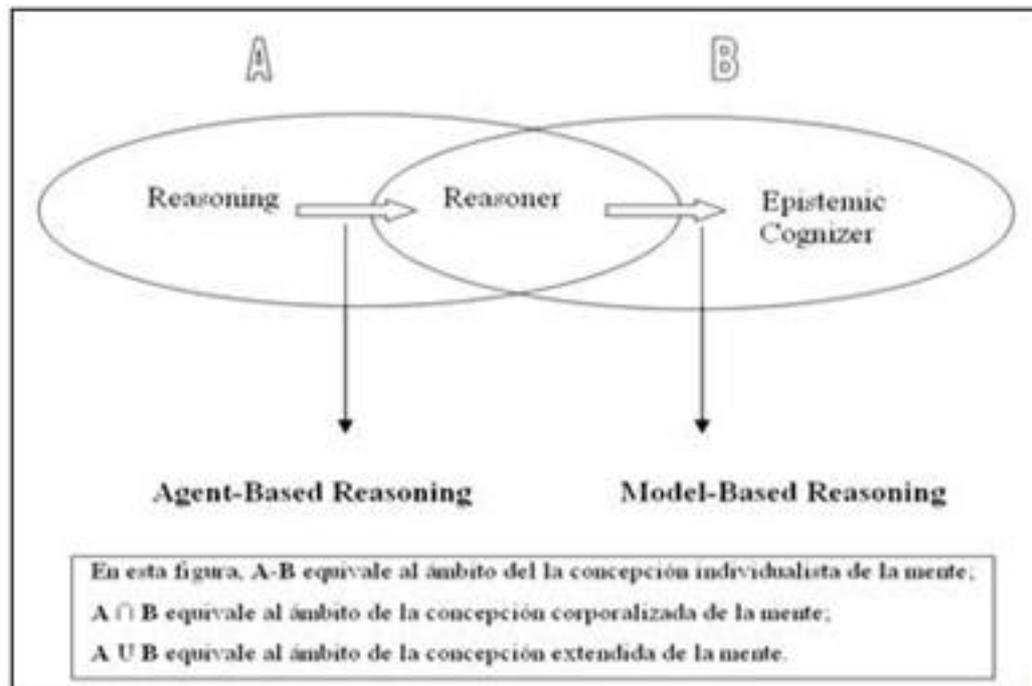


Figura 1: Concepciones de la mente subyacentes a la externalización en RBM.

Según el esquema propuesto, el cambio de foco desde una concepción del razonamiento a priori y descorporalizado [*Reasoning*] hacia una concepción pragmatista y corporalizada del agente cognitivo [*Reasoner*], tiene, como desarrollo ulterior, la noción de “agentes epistémicos” [*Epistemic Cognizer*]. Dicha noción es tributaria de al menos otras dos nociones conformadas a partir de la interpretación de ciertos hallazgos provenientes de la inteligencia artificial y la filosofía de la biología, respectivamente. En el primer caso se hace referencia a la noción de “acción epistémica” surgida a partir de los trabajos de Kirsh y Maglio (1994). En el segundo, a la noción de “ingenieros ecológicos”, según la caracterización de filósofos como Kim Sterelny (2004).

Según Kirsh y Maglio, las acciones de los agentes cognitivos al momento de resolver una tarea determinada pueden ser de dos tipos: pragmáticas o epistémicas. Mientras que las primeras se caracterizan por implementar alguna acción orientada a una meta específica, las últimas sólo tendrían por finalidad realizar acciones modificadoras del mundo, en el entendido que estas acciones son más dependientes del aparataje sensorio-motor que directamente del cerebro biológico como locus del procesamiento interno de nivel superior vinculado a la acción pragmática. El planteamiento original es que las acciones epistémicas ayudan a minimizar la carga de procesamiento interno, lo que se logra al reducir (i) la demanda de capacidad de memoria utilizada en la computación mental; (ii) el número de pasos involucrados en dicha computación, y (iii) la probabilidad de error de la misma. Sin embargo, tanto en los esquemas teóricos<sup>18</sup> como en las investigaciones subordinadas a determinados compromisos anti-individualistas, dicha noción ha servido para sostener la idea de acciones y procedimientos con propiedades cognitivas por derecho propio. El caso

<sup>18</sup> Término que ha sido propuesto para caracterizar las propuestas programáticas asociadas a la denominada Hipótesis de la Mente Extendida (Vallejos, 2008)

de los “mediadores epistémicos” y el de los “agentes ecológicamente ensamblados” son dos ejemplos relevantes<sup>19</sup>.

La noción de ingenieros ecológicos, por su parte, resalta más bien la idea de que la acción modificadora del ambiente en el que han habitado los seres humanos a lo largo de su evolución también ha provocado modificaciones en el propio agente modificador, sobre la base de nuevas presiones selectivas para determinadas conductas. Dicha acción modificadora, o *constructora de nichos*, también tiene lugar sobre el ambiente epistémico. De aquí se siguen los conceptos de “ingenieros ecológicos” y de “ingenieros epistémicos” atribuidos con especial énfasis a los miembros de esta especie, en tanto agentes capaces de transformar ambos tipos de nichos, y como consecuencia de aquello, al ingeniero mismo.

En términos más específicos, el esquema de trasfondo aquí explorado plantea el contraste entre dos tipos de razonamientos, alternativos a la idea tradicional de razonamiento. De esto no se debe asumir que haya una manera de caracterizar en detalle los tipos de razonamientos alternativos, sino solamente que hay ciertas motivaciones empíricas a las que se suele recurrir como justificación de su estatus alternativo. Los dos tipos de razonamientos aludidos son un ‘razonamiento basado en el agente’ (RBA) y un ‘razonamiento basado en modelos’ (RBM). Lo que propongo es entender el primero como la resultante del tránsito desde un énfasis en el *razonamiento* (descorporalizado) hacia un énfasis en el *razonador* (propio de  $A \cup B$ , donde el razonamiento está al servicio de los propósitos prácticos del agente). Consiguientemente, el segundo se puede entender como la resultante del paso desde un énfasis en el *razonador* hacia un énfasis en el *agente epistémico* (propio de  $B$ , donde, además de lo anterior, a ciertas prácticas y aspectos del mundo - relevantes en alguna tarea cognitiva- se les suele asignar un carácter cognitivo por derecho propio). Mientras que ambos tipos de razonamiento se inspiran en una postura pragmatista que reacciona frente una noción de razonamiento independiente del agente cognitivo, no es obvio que se pueda asumir una transición directa entre un RBA y un RBM. De hecho, es solamente en el último de éstos donde se reniega más claramente de una posible equivalencia entre razonamiento y lógica. Por esta razón, es necesario tener cautela con las posibles lecturas del esquema propuesto, y considerar ambos tipos de razonamientos como un efecto de las transiciones descritas, donde las nociones de agentes cognitivos subyacentes a cada una cumplen el rol de ilustrar el tipo de antecedentes tributarios de la ciencias cognitivas actuales, con respecto al posible desarrollo del programa de investigación unificado del RBM externalista.

En este sentido, hay ciertos antecedentes que pueden dar apoyo al esquema de transiciones propuesto para dar cuenta de los énfasis empírico-programáticos que parecen motivar el surgimiento de un RBM de tendencia externalista. Por el momento, se apuntará sucintamente a dos antecedentes distintos: por un lado, el estudio de *argumentos falaces*, asociado a un RBA, que pretende reivindicar un propósito en los “errores del pensamiento”; por otro lado, aquellos relacionados a la *cognición activa*, asociados a un RBM.

Con respecto a un RBA, resulta relevante referirse al tratamiento de la abducción como un aspecto del razonamiento independiente de su validez inferencial. Woods (2005, 2007), por ejemplo, ha intentado mostrar, a partir del estudio de la argumentación falaz, las razones por las que un razonamiento abductivo tiene un valor propio que trasciende el de una mera forma inválida o desviada de inferencia deductiva. En este caso, el proyecto que se plantea

<sup>19</sup> Para el caso de los mediadores epistémicos, ver Magnani, 2001a; Para el caso del rol de la noción de las acciones epistémicas en agentes ecológicamente ensamblados, ver Clark, 2008

es el de una lógica de la abducción construida sobre la base del siguiente tipo de directrices pragmatistas<sup>20</sup>:

- Los argumentos falaces que tiene la forma de “afirmar el consecuente” corresponden a lo que Gilbert Harman llama “inferencia a la mejor explicación”, o a lo que los científicos computacionales denominan “argumentos auto-epistémicos”, o a lo que Peirce formalizó como “Abducción”
- Una teoría del razonamiento debe estar al servicio de los razonadores (en tanto agentes cognitivos), y, por lo tanto, de sus objetivos cognitivos
- Los errores del razonamiento, y por lo tanto el razonamiento abductivo, depende de los estándares de consecución de los objetivos cognitivos del agente, y no de los estándares de validez lógico-deductiva
- Los argumentos abductivos son realizados por el agente cognitivo de manera intencionada, porque sirven como una base racional para la acción provisional
- Mientras que la deducción preserva la verdad y la inducción expande las probabilidades, la abducción preserva la ignorancia
- Los argumentos abductivos no remueven la ignorancia, porque los agentes no los usan para derivar una conclusión, sino más bien para licenciar su conjetura
- Una lógica para la abducción debe ser de carácter práctico

Con respecto al RBM, la siguiente cita de Wilson (2005) puede ayudar a resumir el tipo de contribución experimental asociada a los ámbitos corporalizados y extendidos<sup>21</sup> de la figura anterior (es decir,  $A \cap B$  y  $A \cup B$ , respectivamente), en base a la denominada ‘cognición activa’:

***“Los argumentos de la cognición activa se denominan así porque apelan al ejercicio activo de las capacidades cognitivas en el mundo real (ver Clark 1997, Haugeland 1995, Hurley 1998, 2001, Rowlands 1999: capítulos 5-8, Wilson 2004: capítulos 8-9). Los argumentos de la cognición activa han apelado especialmente a una gama de trabajos en psicología perceptual, que van desde los enfoques Gibsoneanos (Gibson 1966; ver Rowlands 1999) hasta el paradigma de la visión animada en psicología computacional de Dana Ballard (Ballard 1991, Ballard et al. 1997; ver Wilson 2004: capítulo 7) y el enfoque más reciente de la experiencia visual que involucra la exploración animada de nuestro ambiente y con ello la constatación de una sensibilidad hacia las contingencias sensorio-motoras en el ambiente del organismo (O’Regan y Nöe 2001). Como uno podría esperar, dada la diversidad de estos trabajos, los argumentos de cognición activa a favor de la mente extendida también han recurrido a otras áreas de la psicología: a la resolución de problemas en casos donde esto incluye explotar las características táctiles o visuales de nuestro ambiente para completar la tarea de resolución de problema (Clark 1997, 2003; Wilson 2004: capítulo 8); a casos reales o posibles de inversión sensorial (Hurley 1998, 2001); y al enfoque mediacional del desarrollo cognitivo iniciado por Vygotsky y Luria, liderado en la literatura contemporánea por Michael Cole (1996) y James Wertsch (1998, 2002)”.* (p. 6)**

<sup>20</sup> Por esto se asume la adhesión a una interpretación liberalizada de la *máxima pragmatista* Peirciana, entendida como una regla para la clarificación del contenido de las hipótesis determinado por sus “consecuencias prácticas”

<sup>21</sup> Para conocer el origen de los planteamientos de la Hipótesis de la Mente Extendida, ver Clark and Chalmer, 1998

Para ilustrar mejor la incidencia de este último tipo de contribuciones, basta colocar atención al hábito establecido de darle una liberalizada interpretación unificadora a una amplia y diversa gama de evidencia empírica, tal como se constata en la defensa que Clark (2008) realiza de la Hipótesis de la Mente Extendida<sup>22</sup> (HEM). En esta hipótesis, la idea básica que el filósofo experimental quiere asentar es la de que un determinado cuerpo, cuando está equipado con el tipo apropiado de dinámica pasiva<sup>23</sup>, puede co-evolucionar con su entorno de una manera profundamente transformadora, en el entendido que la distribución de la carga de trabajo entre un agente y su ambiente - al momento de la solución de problemas determinados - debiera ser mejor interpretada como la incorporación literal a la agencia cognitiva de los recursos relevantes. Para fundamentar esta visión, se vinculan hallazgos de distinta naturaleza experimental, sin especificar convincentemente algún tipo de procesamiento alternativo al representacional/simbólico de la ciencia cognitiva clásica, que constituye la visión desafiada en este caso.

Un ejemplo de lo anterior ocurre cuando se alude a investigaciones que asimilan el movimiento ocular (sacádico) a la manera en que un computador von Newman hace cambios de referencia entre algún conjunto determinado memorias. El rol computacional de los movimientos sacádicos espontáneos se haría manifiesto en la medida que, frente a una tarea experimental determinada, estos movimientos actúan como desplegando “estrategias de memoria mínima” que *complementan* programas cerebrales ya creados para minimizar la cantidad de memoria de trabajo requerida. Esta complementación requiere aceptar la estrategia explicativa denominada Descomposición Funcional Distribuida, según la cual una referencia implícita, como la realizada por el movimiento ocular o el uso de los dedos para sumar, constituye una codificación de tipo deíctica (o indexical) que permite explotar partes del mundo como dispositivos de almacenamiento externo. Aquí, tal como en la descomposición funcional de la IA clásica, no importa (el material o) la forma específica que toman los elementos que permiten la referencia implícita, sino más bien el rol (o *función*) de los elementos mecánicos constituyentes de algún todo que desborda el límite de los elementos relevantes en un enfoque estrictamente internalista. Para reforzar el argumento, dicha interpretación se suele combinar con la investigación que en robótica situada se realiza en torno a cierto tipo de robots cuyo funcionamiento y *aprendizaje* estaría regulado, precisamente, por la dinámica pasiva anteriormente señalada. El punto es que a partir de este tipo de interpretaciones laxamente combinadas<sup>24</sup>, alcanzaría cierta plausibilidad aquello a lo que se le ha llamado el “Principio de Ensamblaje Ecológico” (Clark, 2008), según el cual un agente cognitivo tiende a reclutar, en el acto, cualquier combinación de recursos solucionadores de problema que permitan un resultado aceptable con un mínimo de esfuerzo. En versiones similares, este principio también está presente en los enfoques más externalistas del surgimiento de hipótesis por RBM, confirmándose el compromiso pragmatista de acuerdo al cual lo central de los pensamientos no es cómo éstos representan alguna parte del mundo de forma relativa a su valor de verdad, sino más bien cómo éstos ayudan a guiar *acciones en el mundo* de

<sup>22</sup> En muchos casos, se trata de un tipo de interpretación claramente más liberalizada e innovadora de lo que se está dispuesto a aceptar. Para una discusión crítica de la HME, ver Adam y Aizawa, 2001, 2008 y 2009.

<sup>23</sup> Una dinámica pasiva consiste en el tipo de acción realizada por el agente que es causada, según esta visión, por la cinemática inherente al dispositivo físico, cuando interactúa en (o más bien dicho *junto con*) el entorno físico determinado.

<sup>24</sup> Pero crecientemente aceptadas por una gran parte de investigadores en ciencia cognitiva como se desprende de lo planteado por Wilson (2005), y también autores como Magnani (2006) con respecto a la abducción multimodal.

manera adaptativamente exitosa<sup>25</sup>. Una postura conservadora de este principio se da con respecto al modelamiento mental y las prácticas científicas de RBM que Nersessian (2008) considera “minimalistamente” dependiente de alguno de los argumentos de cognición activa señalados en la cita anterior. Específicamente, con respecto al cambio conceptual en ciencia, esta autora admite que:

***“El razonamiento se realiza por medio de la construcción y manipulación de modelos. En el proceso de construir, manipular y modificar modelos mentales, la información en variados formatos, incluyendo lingüísticos, visuales, auditorios, y kinestésicos, puede ser utilizada para construir y animar el modelo. La interacción con representaciones externas durante el razonamiento, que yo formulo como acople representacional, lleva a la noción de que los modelos mentales en la memoria de trabajo tienen aspectos significativamente modales (‘modelos mentales perceptuales’), aunque un argumento conclusivo a favor o en contra de esto no puede obtenerse ya sea desde literatura modal o desde la amodal”. (p. 128)***

Para lo que interesa en este trabajo, la relevancia del formato de las estructuras modélicas que eventualmente se consideren participantes de algún procedimiento inferencial de base abductiva parece relativa al lugar en que se les considere a lo largo de las transiciones ilustradas en el diagrama de la figura 1. Un ejemplo de esto es, precisamente, el compromiso que se tenga con formatos más o menos radicalmente modal de los modelos (e.g. Magnani, 2006a). No obstante, el carácter y la naturaleza misma de lo que pueda considerarse una estructura modélica, especialmente no lingüiforme, tiende hacia un tratamiento más o menos (exclusivamente) metodológico en la medida que se aleja del ámbito individualista. Esto podría entenderse así, como se constata en la cita de Wilson, dada la difícil tarea que representa la unificación de hallazgos experimentales muy diversos y de disciplinas con heterogéneos procedimientos de recolección de evidencia. Por esta razón, nos limitaremos a considerar principalmente el carácter metodológico de los modelos extra-mentales, lo que, ciertamente, tampoco será garantía de concepciones claramente unificadas. Básicamente, la problemática que nos lleva a circunscribir la investigación al ámbito de lo metodológico tiene que ver con la dificultad de establecer procedimientos esencialmente compartidos entre las diversas aproximaciones al RBM. Dicho de otra manera, si los procedimientos que podrían dar cuenta de procesos de ampliación de conocimiento son significativamente diferentes, resulta inviable tratar de especificar cuál o cuáles de ellos son los que, estando ausentes en cualquier ejemplar de RBM, harían imposible llevar a cabo el propósito ampliador (o epistémico) en cuestión.

En el siguiente capítulo, se intentará dar cuenta de algunos problemas centrales relacionados al rol de la representación modélica y al RBM en ciencia. La importancia de aclarar este tipo de cuestiones tiene que ver con ciertas preguntas básicas que nos importa tratar de responder. Estas preguntas dicen relación con cómo los modelos permiten solucionar problemas y cuál es el rol esencial que juegan en la generación de hipótesis explicativas. De particular importancia para el presente trabajo es poder alcanzar cierta claridad respecto de los criterios subyacentes a los constreñimientos que determinan el rol de las denominadas “representaciones (modélicas) externas”, toda vez que a partir de dicho

---

<sup>25</sup> Este es un argumento que Fodor (2008) desarrolla sobre la base una clase de inferencia pragmatista inválida, por la que se pretende llegar a conclusiones metafísicas (e.g. que la verdad de los pensamientos no está determinada por la relación entre mente y mudo) a partir de premisas epistemológicas, típicamente relacionada a alguna variedad de escepticismo epistémico acerca de las creencias empíricas.

rol se podrían esbozar planteamientos acerca de lo que realmente formaría parte de un RBM externalista en tanto sistema inferencial distribuido. Resulta importante, por lo tanto, una breve caracterización previa de un principio metodológico que supuestamente regularía uno de los supuestos generales de la externalización cognitiva. Como se señaló anteriormente, este supuesto sostiene que la cognición puede caracterizarse como el producto causal entre los componentes de la tríada cerebro-cuerpo-mente. A falta de una caracterización precisa de aquello que pueda determinar el “acople causal” de aspectos neuronales, corporales y ambientales (tema central de la sección 4.1.1.1.), la externalización cognitiva se fundamenta en el denominado ‘principio del externalismo activo’.

#### 1.1.4. Principio del externalismo activo

De manera muy amplia, el concepto de *externalismo activo* se puede entender en oposición al de *individualismo*. En una perspectiva metodológica, haremos referencia a esta oposición como principios que se niegan mutuamente. De este modo, el principio del individualismo metodológico (en cualquiera de sus formulaciones más conocidas) niega el del externalismo activo, en el entendido de que el primero regula una clara demarcación científica del ámbito de estudio de los estados psicológicos de las personas, haciendo una clara distinción entre el locus internalista de estos estados y el ámbito físico-social de los individuos (ver 1.1.1.). Por su parte, el principio metodológico del externalismo activo desafía, precisamente, la demarcación de tales ámbitos, defendiendo un continuo causal entre los mismos. Dado el hecho de que este último principio rechaza la noción de un límite intracraneal de los estados psicológicos, propia del individualismo, las dos posiciones en cuestión también han sido caracterizadas como internalismo y externalismo, respectivamente. No ahondaremos aquí en las simplificaciones o imprecisiones que tal terminología pueda implicar. Lo que sí resulta relevante es el hecho claro de que defender una u otra posición conlleva importantes implicancias para dar cuenta de cuestiones centrales como la representación y la causación mentales.

La tradición dominante en ciencia cognitiva ha sido, de manera explícita o implícita, principalmente individualista, tendencia que no ha estado libre de importantes controversias. Aparentes incompatibilidades entre un enfoque científico individualista y la psicología de sentido común han sido fuente de constantes argumentaciones y contra-argumentaciones poco concluyentes. Lo mismo se puede decir con respecto a la ‘intencionalidad’ de las propiedades mentales, especialmente a partir de las reivindicaciones articuladas en términos de una defensa realista de algunos conceptos propios de la psicología popular como las creencias y deseos (e.g. Fodor, 1987). El principio del externalismo activo es una de las reacciones a la visión individualista hegemónica que reflejan la difícil relación que ésta ha tenido últimamente con las ciencias cognitivas, dada la variedad de visiones sobre cómo entender las prácticas explicativas en este ámbito científico.

Según Clark y Chalmers (1998, 2008, entre otros), es arbitrario sostener que la mente está contenida dentro de los límites de caja craneana, ya que el ambiente en el que se encuentra un individuo jugaría un rol activo en el funcionamiento y la constitución de los procesos cognitivos<sup>26</sup>:

**“[...] el organismo humano está vinculado con entidades externas en una doble interacción, creando un sistema acoplado que puede ser considerado un sistema cognitivo por derecho propio. Todos los componentes en el sistema**

<sup>26</sup> En el RBM externalista, dicho rol activo de aspectos extramentales actúa como un supuesto explicativo central.

***juegan un rol causal activo, y juntos gobiernan la conducta en la misma manera que la cognición usualmente lo hace. Si quitamos el componente externo, la competencia conductual del sistema decrecerá, tal como sucedería en el caso de extraer parte de su cerebro. Nuestra tesis es que este tipo de proceso acoplado cuenta igualmente como un proceso cognitivo, ya sea que tenga lugar, o no, completamente dentro de la cabeza***<sup>27</sup>. (2008, p. 222)

En el capítulo 4, discutiré ciertas observaciones interesantes vinculadas a aquella parte de la cita anterior que sostiene que “Todos los componentes en el sistema juegan un rol causal activo, y juntos gobiernan la conducta en la misma manera que la cognición usualmente lo hace”. Trataremos ahí de mostrar que no resulta convincente el criterio paritario subyacente de que los objetos externos deban *funcionar* con el mismo propósito de los componentes internos de la cognición. Por ahora, basta considerar que el principio externalista en cuestión es el que sustenta la visión de un organismo capaz de “utilizar” partes del mundo de una manera tal que, en ciertas ocasiones, pueden ser consideradas como una extensión cognitiva extracraneal de los procesos cognitivos.

A modo de resumen, en este primer capítulo han sido presentadas alguna de las nociones y distinciones metodológicas preliminares para llevar a cabo un escrutinio crítico de enfoques de RBM externalista. Junto con esto, se ha dado cuenta del tipo de inversión conceptual y metodológica que subyace al estudio del quehacer científico en tanto actividad cognitiva. Asumiendo que tal inversión está regulada por principios filosófico-experimentales de corte pragmatista, hemos delineado un esquema de transcurso de la externalización del RBM. Con esto se ha intentado establecer una posible relación simbiótica entre distintas nociones de agencia cognitiva, vinculadas a los enfoques de la cognición alternativos a la visión abrazada por la ciencia cognitiva clásica, y el surgimiento de concepciones de razonamiento alternativas a la concepción racionalista tradicional. Hemos planteado que las nociones de agencia cognitiva en cuestión actúan como antecedentes de las concepciones de razonamiento alternativas, sobre la base de transiciones relativas a la supuesta interacción causal entre los componentes de la triada cerebro-cuerpo-ambiente que regula los argumentos de cognición activa. Estos argumentos serían consistentes con una reconceptualización laxa de nociones como *procesamiento cognitivo, información y representación*. Dicha reconceptualización es relevante tanto para la descripción de supuestos sistemas cognitivos distribuidos como para la visión de la cognición en términos de la Hipótesis de Mente Extendida (fundada en el principio metodológico de *externalismo activo*), según los cuales todas esas nociones pueden concebirse como parte de una concepción cognitiva que desborda los límites de un procesamiento intracraneal. De aquí la importancia de explicitar el lugar que ocupa la noción de “agente epistémico autónomo” con respecto a un RBM, con miras a la caracterización de este tipo de razonamiento en tanto sistema inferencial distribuido que permita evaluar la relación de ambos en el contexto de hipótesis propuesta en esta investigación.

Tanto la noción de modelo en términos de un agente epistémico autónomo como la noción de un RBM como método inferencial distribuido han quedado vagamente especificadas en este capítulo. La especificación de ambos será el objetivo último de los

<sup>27</sup> “[...] the human organism is linked with [...] external entit[ies] in a two-way interaction, creating a coupled system that can be seen as a cognitive system in its own right. All the components in the system play an active causal role, and they jointly govern behavior in the same sort of way that cognition usually does. If we remove the external component the system’s behavioral competence will drop, just as it would if we removed part of its brain. Our thesis is that this sort of coupled process counts equally well as a cognitive process, whether or not it is wholly in the head.” (Clark, 2008, p. 222)

dos capítulos siguientes, respectivamente. Los temas centrales abordan la caracterización artefactual de los modelos en un RBM externalista, y el RBM como método de inferencia. Ambos temas serán tratados por separado, en términos de un desarrollo que va desde las perspectivas más conservadoras hacia las más radicalmente externalistas. Vale decir, desde un RBA hacia un RBM en el primer caso, y desde el tratamiento del surgimiento de hipótesis por abducción sentencial hacia el rol de la denominada delegación cognitiva en el segundo caso. Será en el capítulo cuarto donde se evaluarán integradamente las limitaciones y el alcance de una teoría del surgimiento de hipótesis posiblemente subsumida por un RBM externalista que incorpora la noción de *modelo como agente epistémico o artefactual autónomo*, con miras a una probable teoría epistemológica unificada del descubrimiento en ciencia.

## CAPITULO II

### 2.1. Representación modélica y RBM

Tal como se señaló en el capítulo anterior, una doble lectura de la tesis central de la cognición distribuida parece atravesar las diversas concepciones de representación modélica y el rol que éstas juegan en el RBM. En lo que sigue, trataremos de abordar las visiones más características de este tipo de representación, con el fin de aclarar - si es posible - el surgimiento de la concepción más externalista, a saber, la de carácter artefactualista. De especial importancia será ponderar el hecho de que una versión artefactualista de la representación modélica plausible es metodológicamente equivalente a la noción de una agencia epistémica autónoma, aún cuando esta última termine siendo impugnada por las razones que se discutirán en cuarto y último capítulo de esta tesis.

Lo que interesa en este capítulo no es proporcionar un tratamiento exhaustivo de visiones distintas, sino más bien resaltar aquellos puntos de contraste que reflejan la influencia - a ratos demasiado ecléctica - de las concepciones alternativas de la cognición ya comentadas con anterioridad (ver sección c de 1.1.1.). Cabe hacer notar que en algunos casos, como en lo relacionado con las propuestas de Johnson-Laird, no se debe suponer que se trata de una etapa más de algún tránsito incremental hacia la externalización. Esto sería tan erróneo como pensar que el esquema didáctico descrito en 1.1.3. expresa una transición similarmente fluida entre los tipos de razonamiento aludidos ahí (i.e. desde RBA a RBM). La razón por la que aquí se incluye la perspectiva de Johnson-Laird tiene el propósito de servir como punto de contraste a la noción de representaciones extra-modales de orden modal. Si este objetivo se cumple, entonces debiéramos poder explicitar inicialmente las características y las complicaciones de las nociones de representación modélica más aceptadas dentro de la comunidad del RBM.

#### 2.1.1. Modelos en la perspectiva de un razonamiento centrado en el agente

---

Según Giere (2004), una relación *traductora* entre entidades semejantes a las del lenguaje humano y el mundo resulta inadecuada para describir las prácticas de representación científica. Por esta razón, Giere propone enfocarse en lo que denomina la representación como una actividad pragmática. La visión traductora mencionada, junto al desarrollo de la lógica matemática moderna, habría dado origen al consenso epistemológico de que el “lenguaje de la ciencia” puede ser descrito y explicado en términos de una sintaxis y una semántica. Estas categorías metodológicas guardan consistencia con un enfoque lógico-proposicional del lenguaje humano, cuya semántica derivaría de la semántica de las representaciones mentales (estructuradas) del pensamiento que el lenguaje refleja al momento de ser enunciado. A su turno, la semanticidad o intencionalidad de las representaciones mentales, constitutivas de cualquier acto de pensamiento, tendría un estatus ‘original’ o ‘no derivado’, supuesto central del enfoque computacional de la mente que, como señala Wilson (2005), ha estado fuertemente constreñido por una visión naturalista de la mente. Según este constreñimiento naturalista, la intencionalidad

original de la mente debiera ser explicada de manera no circular, i.e. sin que dicha explicación suponga la intencionalidad de otra cosa. A partir de aquí, tendría lugar el surgimiento de una especie de programa que intenta mantenerse dentro de los márgenes del realismo intencional, posición filosófica según la cual la intencionalidad original de las representaciones mentales sería parte del orden natural de las cosas. En este sentido, los esfuerzos por explicar el problema de la semántica de las representaciones mentales – según Wilson- han desplazado la noción de intencionalidad en favor de otras como ‘causación’, ‘dependencia contrafactual’, ‘función biológica’, ‘historia filogenética’, y, al mismo tiempo, han dado origen a propuestas como la Semántica informacional de Dretske (1981, 1988), la Biosemántica de Millikan (1984, 1993), y, en términos más generales, la Teoría Causal del contenido de las representaciones mentales de Fodor (1984, 1987).

Lo descrito en el párrafo anterior puede entenderse como un planteamiento subordinado al principio del *individualismo metodológico*. Recordemos que, según este principio, los procesos cognitivos causalmente relevantes para la explicación de la conducta inteligente son aquellos que están en una cierta relación de dependencia causal respecto del cerebro, y que, a la vez, estarían absolutamente desvinculados del ambiente físico y social en el que tienen lugar. Alternativamente, Giere adhiere a los planteamientos que confinan la sintaxis y la semántica a un estatus de categorías metodológicas emergentes, determinadas por el uso del lenguaje, y, por lo tanto, por una actividad pragmática. Con estas ideas en mente, se abre la posibilidad de estudiar el quehacer científico en una perspectiva que invierte los niveles metodológicos tradicionales. Consiguientemente, el énfasis no se coloca en el lenguaje en sí mismo, sino en las prácticas científicas en las que el lenguaje es utilizado. Este cambio de foco implica el tránsito desde un enfoque centrado en las representaciones propiamente tal a otro centrado en la actividad misma de representación, lo que, según Giere, debiera incluir a los agentes *intencionales* con sus metas y propósitos. La expresión conocida es “S utiliza X para representar a W para propósitos P”. En otras palabras, el científico (o el grupo de científicos) hace uso de un modelo determinado para representar ciertos aspectos del mundo para propósitos específicos. Los valores que puede tomar X, continúa Giere, pueden ir desde las palabras (en su estatus de símbolos materiales) hasta las imágenes generadas por computadores, siendo de mayor interés para este autor aquel medio de representación científica tradicional conocido como las ‘teorías científicas’. De aquí, el rol de *modelos* que Giere le asigna a las teorías científicas, en contraste a nociones más tradicionales donde el uso de modelos se confina a etapas posteriores del desarrollo de teorías, como algo que está al servicio de la explicitación de axiomas, leyes o, finalmente, teorías científicas.

Desde su concepción de modelos, Giere sostiene algo que contrasta con lo defendido por los exponentes de los enfoques más externalistas, como, por ejemplo, aquellos liderados por Nersessian (de manera no tan radical), Gooding, Magnani y Knuuttila. El planteamiento de Giere es que dichos modelos son representaciones abstractas análogas “previamente interpretadas” de algún aspecto del mundo. Vale decir, no es en el modelo donde que instancia la representación, sino en el agente que utiliza la estructura modélica en cuestión. En el presente trabajo, se entiende que los planteamientos de Giere tienen un alcance pragmatista que no transgrede los supuestos teóricos que sitúan al agente intencional como en el núcleo de cualquier razonamiento basado en modelos (internos y externos). En la figura 1 presentada en el capítulo anterior, podemos situar dicha postura dentro de los límites de **A**, no porque la visión de representación modélica de Giere sea compatible con la visión “traductora” inicialmente comentada, sino porque se centra en el agente y su acción pragmática. Alternativamente, en el caso de las aproximaciones situadas y distribuidas, dichos compromisos teóricos se disipan, razón por la cual parece

más apropiado sostener, en contraste a lo sostenido para la visión de Giere, que el alcance pragmatista de estos enfoques más radicales está mejor contenido en torno al *principio de paridad* propio del ámbito extendido de AUB. Mientras que el principio de externalismo activo apunta a un posible rol causal por parte de ciertos aspectos físicos del mundo - externos al agente intencional -, este nuevo principio pretende enfatizar una supuesta irrelevancia de la delimitación intra/extra craneal. Según el principio de paridad (Clark, 2008)<sup>28</sup>:

**“si, confrontados a una tarea, una parte del mundo funciona como un proceso que, de ocurrir en la cabeza, no dudaríamos en aceptar como parte del proceso cognitivo, entonces esa parte del mundo es (para dicha ocasión) parte del proceso cognitivo” (p. 77, y originalmente en Clark y Chalmers 1998, p. 8”. El énfasis de esta cita es responsabilidad del autor de este trabajo).**

Como se detallará más adelante, nociones como las de ‘construals’ (Gooding, 1990) y ‘mediadores epistémicos’ (Magnani, 2001a-b) son ejemplos que apuntan, con no pocas ambigüedades, en la dirección defendida a través del principio citado. En el caso de los *construals*, la idea es reconocer la relevancia de ciertas regularidades extra-rationales que de alguna manera estarían corporalizadas en ciertas plantillas de conducta. En el caso de los *mediadores*, se pretende dar cuenta de la manera en que surge lo que se concibe como nuevas oportunidades de razonamiento inferencial, en la medida que el agente cognitivo trasforma diferentes aspectos del ambiente físico en mediadores epistémicos. El hallazgo de *información* no disponible, imposible de obtener por medio de procedimientos explícitos (i.e. reglas de transformación inferencial tradicionales), se realiza por medio de la interacción entre procedimientos tácitos y procesos inferenciales. En este caso, cuando se habla de procedimientos tácitos se hace referencia, por ejemplo, a plantillas de conducta no razonadas que constituirían una suerte de interacción epistémica extracraneal, en el sentido de que a tales interacciones se les pueda atribuir algún estatus informativo susceptible de *causar* inferencias intracraneales<sup>29</sup>. Magnani denomina *abducción manipulativa* a este proceso de interacción continua, caracterizado por un efecto de externalización semiótica, determinado por la habilidad del agente cognitivo para redistribuir recursos cognitivos externos y externos (ver 3.1.1.). Si bien es cierto que, como se dijo, las representaciones externas adquieren un cierto estatus cognitivo (no intencional), hecha ostensible en la relevancia las oportunidades de acción que proporcionaría su propia materialidad (en tanto *affordance*<sup>30</sup>), las oportunidades que surgen de la abducción manipulativa siguen ubicando al agente intencional como núcleo del razonamiento inferencial. Dicho de otra manera, aunque se estipule un carácter cognitivo asociado a la manipulación de objetos externos o la interacción en tanto oportunidades de acción, ambos procesos requieren de un “manipulador” de un agente intencional para el que tenga sentido el aprovechamiento de tales oportunidades de acción. Lo que resulta adicional en estos planteamientos es el rol constitutivo que se le asigna a las representaciones externas, en su calidad de signos materiales. Dicha relevancia se manifiesta en la noción de *pensamiento a través de la*

<sup>28</sup> En el capítulo 4 se discutirán los alcances de este principio y la relación implícita de subordinación que impone sobre ciertos aspectos del RBM.

<sup>29</sup> Esta es sólo una posible interpretación de lo que Magnani denomina un proceso de diálogo o negociación epistémica entre organismo y aspectos extra-mentales del mundo. Desde su perspectiva semiótica, este filósofo experimental no admitiría la delimitación literal entre ámbitos semióticos internos y externos, ya que según ésta toda atribución de significado (y por lo tanto toda posible información disponible en el mundo) es dependiente del organismo y su capacidad intrínseca de atribución de significados.

<sup>30</sup> *Tanto las fundadas biológica/ambientalmente (Gibson, 1979) como las fundadas perceptualmente (Norman, 1988)*

*acción*, en contraste a la idea de *pensamiento acerca de la acción*. Consistentemente con la cita anterior, hasta donde alcanzamos a entender, “pensar a través de la acción” hace referencia al hecho de que dicha acción se considera imprescindible al momento de hacer explícita alguna información, no disponible internamente, que resulta crucial para la instanciación de modelos mentales.

Adicionalmente, la dicotomía formalidad/funcionalidad de las estructuras modélicas (mentales o físicas) en particular, y de las representaciones externas en general, es otro elemento que, aparentemente, puede caracterizarse en base al esquema que ilustra el diagrama de la figura 1. Así, para comprender la práctica científica, entendida como una actividad cognitiva en sí misma, Giere sostiene que los modelos no deben entenderse como una representación abstracta en la que se instancian los axiomas de una teoría determinada – coincidiendo en eso con Neressian (ver 1.1.3.) -, sino como una representación de la realidad. Para esto, concibe (i) los aspectos formales de los modelos como una condición de adecuación de su poder representacional de la realidad, y (ii) la funcionalidad de los modelos como condición de adecuación a los fines perseguidos por los científicos. Ambos aspectos de dicha dicotomía tienen distintos énfasis o relevancia a lo largo de las transiciones propuestas, siendo la funcionalidad (relativa a los propósitos de la actividad pragmática) el aspecto que adquiere mayor prominencia hacia extremo más externalista.

Lo anterior se puede ilustrar si se contrasta la noción de representación como actividad pragmática de Giere, recién aludida, y la noción de mediación epistémica de Magnani (2008). Mientras que para Giere las condiciones de adecuación (i) y (ii) son relevantes para explicar cómo se utilizan los modelos para representar la realidad, Magnani enfatiza el hecho de que el valor pragmático del uso de modelos tiene su mayor importancia en una manipulación epistémica extramental de objetos materiales, relativa a las condiciones de (ii). Por esto último debe entenderse que la representación es dependiente de las acciones epistémicas no razonadas que regulan la productividad de acciones pragmáticas razonadas con asiento en el agente intencional. Como la visión del agente cognitivo de Magnani está supuestamente inspirada en la semiótica de Peirce, la dependencia recién mencionada se podría entender como una interpretación metodológica revisada de la visión que el filósofo norteamericano tenía con respecto a las creencias en tanto reglas de acción (Peirce, 1878). Según este filósofo, el pensamiento es esencialmente acción, y su propósito último es el de fijar alguna creencia. No deja de ser interesante la aparente simetría entre *pensamiento como acción* planteada por Peirce y la noción de *pensamiento a través de la acción* que defiende Magnani, así como entre la fijación de *creencias en tanto reglas de acción* y la fijación de configuraciones neuronales a partir de la creación de signos semióticos externos en tanto *agentes cognitivos autónomos portadores de información relevante*, propuestas por Peirce y Magnani, respectivamente. Las siguientes citas de Magnani (2007a) parecen sugerir lo señalado:

**“Si una acción manipulativa realizada sobre el ambiente está dirigida a crear una configuración de signos que portan información relevante, esa acción bien podrá ser considerada como un proceso semiótico cognitivo y la configuración de elementos que se crea bien podrá ser considerada una representación externa”.** (p. 22, énfasis agregado) **“Este modelo externo es un tipo de agente cognitivo autónomo que admite nuevos interpretantes del problema/objeto en cuestión. El modelo puede ser seleccionado posteriormente y aprehendido por el agente humano a través de la estabilización [fixation] de una nueva configuración neural [...]. La presencia de este “poder autónomo” [autonomous power] explica por qué yo le atribuyo al sistema de representación un estatus de agencia cognitiva**

***similar al de una persona humana, aunque por supuesto sin aspectos como intencionalidad directa y responsabilidad”.* (p. 33, énfasis agregado)**

El contraste entre posturas tan distintas como las de Giere y Magnani refleja al menos dos tipos de problemas poco especificados en el contexto de una probable teoría unificada de RBM externalista. El primer problema tiene que ver con el carácter representacional de la noción de representación modélica. El segundo dice relación con aquello que se enfatiza como central en los procedimientos donde tiene lugar el uso supuestamente productivo de tales representaciones. En este sentido, se puede decir que claramente no se está hablando de un mismo tipo de representación modélica, y tampoco de un mismo tipo de procedimientos epistémicos participantes en la explicación de probables procesos productivos de RBM.

El primer problema se puede explicitar de la siguiente manera. Para Giere, los modelos son utilizados por el agente intencional para llevar a cabo la representación de (algún aspecto de) la realidad en función de propósitos particulares del agente. De aquí se sigue que este tipo de representación depende del uso instrumental que el agente intencional haga de modelos interpretados (i.e. no formales o abstractos). Planteado de esta manera, el proceso que se lleva a cabo no es necesariamente incompatible con la posibilidad de que, aunque el modelo interpretado en sí mismo no tenga un carácter directamente *traductor* o sustituto de la realidad (ver 2.1.2.), el proceso inferencial genuinamente productivo se lleve a cabo en base a un formato de representación no modélico. Vale decir, el uso instrumental de los modelos es al fin y al cabo un instrumento para algo más (que no se agota en su mero rol instrumental por sí solo), y ese algo más puede ser un proceso inferencial interno genuinamente productivo (e.g. de estructuración proposicional). Ciertamente, esto último no es algo que Giere pareciera estar dispuesto a aceptar. Sin embargo, me parece saludable señalar la pertinencia de posibles procesos productivos de formatos más tradicionales mientras no haya razones concluyentes para descartarlos.

Algo distinto ocurre en el caso de Magnani. Aunque este investigador compartiría el rechazo a la visión traductora de los modelos que señala Giere, su visión del carácter representacional de los modelos en tanto representaciones externas está más cercana a lo que se concibe como una *re-presentación* literal (o genuina) de algún fenómeno de la realidad (ver secciones 2.1.1.2. y subsiguientes en este capítulo). Aquí, la idea de una representación interpretada de la realidad se sustituye por la de una especie de *prótesis artificial* de la realidad (o nuevo interpretante, para ocupar otra expresión equivalente de Magnani) cuya manipulación debiera ser entendida como una legítima presentación del fenómeno sujeto a explicación. Y la razón por la que esta visión tiene sentido para Magnani es porque se está aceptando la visión de procedimientos epistémicos “gatilladores” de procesos de razonamiento de orden superior, y por lo tanto, de una inversión conceptual anti-racionalista según la cual la *acción es anterior al pensamiento* que se pueda instanciar acerca de tal acción. Esto es algo que trae a primer plano el segundo problema anteriormente señalado, a saber, el de los procedimientos epistémicos relevantes en la explicación de un RBM claramente productivo o ampliativo de conocimiento.

Con respecto a esto último, es importante también considerar las diferencias de niveles explicativos a los que apunta cada una de las visiones en contraste. Mientras que la visión de Giere, como se dijo, podría llegar a compatibilizar procedimientos de uso instrumentalista de los modelos con probables procesos inferenciales internos de carácter genuinamente productivo, Magnani apuntaría a procedimientos de orden *subpersonal* o de *cognición online*<sup>31</sup>. Con estos términos sólo se alude a procedimientos de manipulación epistémica

---

<sup>31</sup> Esta expresión hace referencia a procesos típicamente constreñidos por el aparataje sensorio-motor.

---

de naturaleza corporalizada y situada, susceptibles de ser integrados – según estipula Magnani, sin especificar claramente esta posible integración – en sistemas de orden distribuido que permitan explicar el origen de procesos de razonamiento de orden superior.

La conclusión provisional es que en la medida que se enfaticen procesos de orden inferior, como en el caso de la denominada manipulación de representaciones externas, el valor epistémico y productivo del RBM se estaría desplazando, aparentemente, desde los aspectos formales (propios de un nivel superior) hacia aspectos relacionados con el uso y la funcionalidad de los modelos (propios de nivel inferior). Sin embargo, habría que hacer hincapié en el carácter aparente de este desplazamiento, ya que la identificación de concepciones discrepantes en el carácter representacional de los modelos genera una crisis frente a la posibilidad de unificar el tipo de visiones como las contrastadas. Por esta razón, es necesario ahondar en otro tipo de distinciones que permitan determinar de mejor manera la posible viabilidad del enfoque de RBM externalista que nos interesa especificar. Ese será el objetivo del resto de las secciones que componen el presente capítulo, para lo cual se considera importante especificar el rol y el carácter modal y amodal de los modelos. Como se asumirá en las secciones que vienen a continuación, un formato modal de los modelos hace referencia a lo que Nersessian (2008) denomina *representaciones mentales preceptuales*, dado el rol determinante de la representación analógica que jugaría el estado perceptual del cual se extraen. A su turno, por formato amodal se hará referencia a una transducción arbitraria de los estados preceptuales, como la relación que hay entre la palabra “sillón” y los aspectos preceptuales de un sillón cualquiera. La distinción entre estos formatos de representación es importante, porque hace la diferencia, por ejemplo, entre el rol más o menos determinante que se le atribuya a ciertos procesos motores y preceptuales en la representación y manipulación de cada cual.

### **2.1.1.1. Modelos como representaciones mentales amodales**

Según la ‘teoría de los modelos mentales’ de Johnson-Laird (2005), la mente construye modelos del mundo y los utiliza para razonar. Estos modelos pueden ser construidos a partir de la percepción, la imaginación, el conocimiento, y la comprensión del discurso. En tanto representaciones mentales, los modelos mentales tienen una estructura que se corresponde con la estructura de lo que representan, lo que constituye una característica distintiva de este tipo de representaciones. Sobre la base de esa condición de adecuación de la representación modélica, se establecería una correspondencia entre propiedades del mundo y la representación como *tokens* de tales propiedades. Del mismo modo, los rasgos o atributos de tales de las propiedades del mundo, así como las relaciones que pueden existir entre ellas, pueden ser representados en el modelo como *tokens* de tales rasgos y relaciones, respectivamente. Una razón de fondo para establecer estas correspondencias es que Johnson-Laird considera que los modelos mentales son representaciones de relaciones espaciales, así como también de eventos y procesos sustentados según la hipótesis de que estos modelos pueden haber evolucionado como el *output* final de procesos perceptuales (2001).

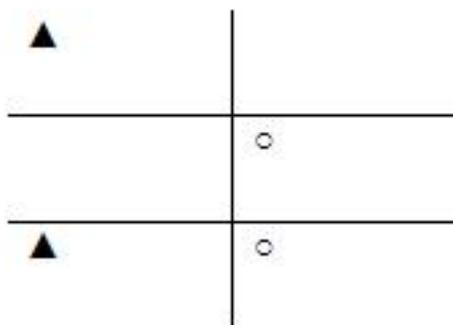
Según Johnson-Laird, dado que un aspecto crucial del razonamiento es establecer una conclusión, la fortaleza de esta conclusión depende de si existe un modelo de las premisas que pueda refutarla. Esto implica una búsqueda de tal modelo en términos de un proceso típicamente inferencial, en el entendido de que, si no se encuentra dicho modelo, la inferencia se juzga válida, mientras que su hallazgo implica una de dos posibilidades: o bien la revisión de una conclusión obtenida previamente para que se haga verdadera en todos los modelos construidos hasta el momento, o bien la especificación probabilística de alguna

conclusión provisional. De este modo, la teoría proporcionaría una explicación unificada de lo que *necesario*, *probable*, y *posible*. Consiguientemente, su teoría de modelos plantea que una conclusión es necesaria si se puede sostener para cada uno de los modelos de las premisas; que una conclusión es probable si se puede sostener para la mayoría de los modelos de las premisas; que una conclusión es posible si se puede sostener para al menos uno de los modelos de las premisas. Esto hace alusión a otro aspecto importante de la teoría de modelos de este autor, y que contrasta con aquellas visiones que resultan más centrales para nuestra investigación. Si bien es cierto que la base *corporalizada* de la visión de representación modélica propuesta podría concebirse como un aspecto compartido por los enfoques estudiados por la *Comunidad de RBM*, Johnson-Laird vincula su teoría al razonamiento asumiendo que los modelos son manipulados por algún sistema capaz de derivar inferencias deductivas e inductivas a partir de ellos. De esto se sigue que la diferencia entre un posible tipo de razonamiento llevado a cabo a través de la manipulación de representaciones de estructuración proposicional y un posible tipo de razonamiento basado en la manipulación de modelos no es la presencia o ausencia de relaciones lógicas. Más bien, la diferencia central tiene que ver con el tipo de representaciones sobre el cual operan dichas relaciones lógicas.

Desde una perspectiva formal, la lógica oracional sostiene que una inferencia es válida cuando su conclusión es verdadera en virtud del hecho de que sus premisas también son verdaderas. Pero también es el caso que no todas las inferencias válidas desde la perspectiva de un sistema semántico-oracional pueden ser probadas utilizando un sistema formal consistente (i.e. uno que excluye la validez de proposiciones contradictorias). En este último caso, Johnson-Laird alude a los descubrimientos de Gödel, según los cuales para todo conjunto de axiomas de la aritmética construible por el hombre existe una fórmula que se obtiene de la aritmética pero es indemostrable en ese sistema. Otra manera de caracterizar los teoremas de este lógico matemático es que, dado un sistema formal, es siempre posible construir una proposición verdadera que no se puede ni demostrar ni refutar. Sin embargo, Esto es algo que no ocurriría en el caso de la lógica de conectivos oracionales, y esta es una de las razones por las que Johnson-Laird restringe su estudio de modelos mentales al rol que estos juegan en el razonamiento y las denominadas *inferencias ilusorias* (i.e. un tipo de inferencia que, en la interpretación de conectivos, da lugar a falacias sistemáticas que los sujetos investigados no pueden eludir, aparentemente debido a un efecto ilusorio insalvable).

El supuesto sobre el que descansa la teoría de Johnson-Laird es que los seres humanos tienen una disposición natural a pensar en posibilidades. Aplicada a los conectivos lógicos, esta teoría considera la negación, la conjunción y la disyunción inclusiva como fundamentales. Así, las posibilidades alternativas se representan como una disyunción de posibilidades; y cada modelo de una posibilidad representa una conjunción de proposiciones afirmativas o negativas.

El siguiente ejemplo ilustra un set de modelos asociado a las posibilidades alternativas “hay un triángulo en la pizarra o hay un círculo en la pizarra o ambos”:



En esta línea de investigación, Johnson-Laird aborda el estudio de la interpretación de los conectivos temporales en el lenguaje natural. Según este autor, un sistema de interpretación humano adecuado de dichos conectivos no puede fundarse en interpretaciones donde la verdad de un conector es funcional a la verdad de los constituyentes de la oración en que ocurren (o de la proposición subyacente), ya que al comprender una oración del lenguaje natural las personas no llegan a su valor de verdad. De hecho, sostiene Johnson-Laird, puede que las personas jamás sepan el valor de verdad de los constituyentes de una oración, el que depende de la relación entre lo que sus constituyentes significan y el estado del mundo.

Alternativamente, lo que propone Johnson-Laird es que la comprensión comienza con la construcción del significado de una oración, y dicha construcción debe recuperar sus referentes, sus propiedades y las relaciones existentes entre ellos. Este proceso, que puede depender del conocimiento, termina con la representación de situaciones posibles a las que hace referencia la oración. La conclusión que defiende Johnson-Laird es que ninguno de los conectivos del lenguaje natural (ej. “y”, “y entonces” [*and then*], “antes”) son interpretados en un modo de ‘verdad funcional’, ya que muchos conectivos (“si”, “o”, “y”) tiene usos que no expresan un significado lógico. Por ejemplo, en inglés “and” puede ser interpretado con el mismo significado de “and then”. O bien, mientras que la disyunción “They played soccer or they played some game” puede ser formalmente válida, aún en el caso de que la segunda oración sea falsa, es poco probable (sostiene Johnson-Laird) que en la vida real una persona angloparlante infiera como verdadera la primera, sabiendo que la segunda es falsa.

Para Johnson-Laird, un modelo mental representa una *posibilidad* o, más precisamente, **la estructura y el contenido del modelo mental captura lo que es común a las distintas maneras en las que puede ocurrir una posibilidad**. Experimentos relacionados que exponen a diversos grupos de sujetos a tareas que impliquen utilizar un set de modelos de posibilidades permitirían concluir que por limitaciones de memoria las personas razonan sobre la base de un modelo a la vez. Así entendidos, los modelos mentales cumplen con el principio de ‘parsimonia’, vale decir, con la propiedad de representar lo que es posible, pero no lo que es imposible. Es este principio el que le daría plausibilidad a los modelos mentales, dado que frente a la capacidad de memoria observada, permitirían minimizar la sobrecarga de la memoria de trabajo y problemas de “cuello de botella”.

Es importante señalar que los modelos no representan solamente las cláusulas mencionadas en las premisas, ya que aquello daría pie a sostener que a cada premisa le corresponde un modelo, independientemente de los conectivos que ocurran en ellas. Johnson-Laird elabora este punto al plantear que el principio de parsimonia funciona en dos

niveles: por un lado, los modelos representan sólo lo que es posible (c.f. una tabla de verdad representa además lo que es imposible); por otro, una posibilidad en tanto modelo mental representa una cláusula en las premisas, ya sea afirmativa o negativa, sólo si la cláusula permite tal posibilidad. En este último caso, se asume que la información irrelevante no se hace explícita en el modelo.

También es importante señalar qué ocurre con la información no explícita en la representación mental. Según Johnson-Laird, dicha información implícita estaría disponible pero de manera periférica, en el sentido de que sin estar presente en la representación modélica, sí estaría eventualmente disponible como una “*footnote*”, la que es olvidada si la tarea demanda mucho trabajo o si la oración contiene muchos conectivos.

Especialmente en el caso del conectivo condicional, la diferencia entre valores de verdad y posibilidades, sostiene Johnson-Laird, es importante porque los sujetos responden de manera distinta en cada caso. Mientras que los juicios de verdad o falsedad exigen que los individuos relacionen el modelo mental con posibilidades externas, los juicios de posibilidad o imposibilidad exigen que los individuos entiendan las oraciones y recurran a modelos mentales para explicitar modelos de un condicional<sup>32</sup>.

Johnson-Laird propone ciertos mecanismos de construcción de modelos, todos los cuales han sido implementados computacionalmente, con el propósito de mostrar que un razonador en base a modelos mentales debe hacer suposiciones “por el bien” del argumento. En el análisis de los programas computacionales, las posibilidades basadas en suposiciones que generan una contradicción constituyen un “modelo nulo”. En el caso correspondiente de un razonador prudente, la conclusión final está constreñida por la consideración de que cualquier conclusión depende de una suposición. De acuerdo a esto, la generación de un modelo nulo lo haría apreciar el hecho de que la suposición es imposible, dada la verdad de las premisas.

Lo anterior es ilustrado por medio del contraste entre las conclusiones generadas por la teoría (i.e. el programa computacional) y las generadas por un grupo de sujetos humanos, sobre la base del tipo de inferencias que estos últimos *pueden* hacer. Según Johnson-Laird, la teoría muestra una incidencia de lo que se ha denominado un “inferencia ilusoria”, un tipo de inferencia falaz cuya alta incidencia en el caso de los sujetos estudiados – según delata el programa que procesa la misma inferencia - la hace parecer una ilusión cognitiva.

Ejemplo con disyunción excluyente:

*O bien Juana está arrodillada junto a la chimenea y está mirando la televisión, o Marco está parado en la ventana y observa el jardín*

*Juana está arrodillada junto a la chimenea*

*¿Se sigue de todo lo anterior que Juana está mirando la televisión?*

La mayoría de los sujetos responden que “sí”, aunque el programa arroja que la sexta conclusión también es válida<sup>33</sup>:

---

<sup>32</sup> Recuérdese el supuesto de que, según Johnson-Laird, los seres humanos tenemos una disposición natural a pensar en posibilidades.

<sup>33</sup> *O bien Juana está arrodillada junto a la chimenea y no está mirando la televisión, o Marco está parado en la ventana y observa el jardín*

Juana:	arrodillada	mirando TV	Marco: $\neg$ parado	$\neg$ observa
Juana:	arrodillada	mirando TV	Marco: $\neg$ parado	observa
Juana:	arrodillada	mirando TV	Marco: parado	$\neg$ observa
Juana:	$\neg$ arrodillada	$\neg$ mirando TV	Marco: parado	observa
Juana:	$\neg$ arrodillada	mirando TV	Marco: parado	observa
Juana:	arrodillada	$\neg$ mirando TV	Marco: parado	observa

El mismo tipo de experimentos se ha realizado con respecto a la interpretación del conectivo de condicionalidad, proporcionándose mayor evidencia de la robustez del fenómeno de las inferencias ilusorias, en el contexto de la construcción de modelos basados en la interpretación lógica de conectivos. La teoría de modelos de Johnson-Laird permitiría predecir dichas inferencias ilusorias, proporcionando así cierta plausibilidad a la tesis de que razonar solamente acerca de lo que es posible constituye un supuesto más adecuado para la explicación psicológica que sobre la base de reglas formales de inferencia.

La investigación de Johnson-Laird en modelos mentales puede considerarse una aproximación que antecede al estudio del RBM en el contexto de la cognición cotidiana. Sin embargo, el RBM defiende la tesis basada en los desarrollos actuales de la ciencia cognitiva según la cual la instanciación de modelos mentales, así como su manipulación, son una actividad epistémica que debe considerar el actual entendimiento del agente cognoscente al menos en su carácter corporalizado, situado y distribuido. De hecho, esa es la línea de investigación que promueve e intenta profundizar el trabajo de Nersessian (2009):

***“La naturaleza de los mecanismos que conectan los componentes humanos y artefactuales es aún una pregunta abierta para la ciencia cognitiva. La investigación debe orientarse tanto a los casos co-ocurrentes de manipulaciones interactivas de los modelos del investigador y los modelos artefactuales, como a los casos en que la simulación se lleva a cabo por el modelo artefactual, y cuyos resultados conducen a cambios en los modelos de largo plazo del investigador”.***  
(p. 749)

Los mecanismos señalados en la cita son mecanismos que deben ser compatibles con la idea de un ensamblaje cognitivo literal de componentes internos y externos. Aquí convergen los intereses teórico-prácticos del RBM y los de la Hipótesis de la tesis de Mente Extendida (Clark, 2008), al menos en lo que tiene que ver con el denominado problema de la *complementariedad*. Esta convergencia requiere poner el énfasis en un tipo de representación distinta a la teorizada por Johnson-Laird, a saber, una representación modal de carácter modal.

Como veremos en las siguientes secciones de este capítulo, y a diferencia de lo comentado acerca de la representación amodal, el rol que adquieren las estructuras modales en el RBM se aparta de la centralidad que Johnson-Laird atribuye al razonamiento. Consecuentemente, esto supone rechazar la visión de que el razonamiento está determinado por modelos mentales compatibles con las mismas relaciones lógicas que prescriben la manipulación de representaciones proposicionales. La razón detrás del rechazo de tales concepciones tiene que ver con la reconceptualización de lo cognitivo

según lo que proponen ciertas aproximaciones empíricas de las ciencias cognitivas que actúan como antecedente del RBM (ver apartado c en 1.1.1.). Para estar en condiciones de abordar el carácter alternativo de los enfoques de RBM más externalistas como método de inferencia (Capítulo III), necesitamos primero dar cuenta de las características centrales de la concepción de modelo relevante para un RBM, con el fin de poder especificar la medida en que dichos enfoques pueden concebirse como el resultado de una inversión empírico-conceptual de inspiración pragmatista (Capítulo IV).

### 2.1.1.2. Modelos como representaciones modales extra-mentales

Si hay algo que puede considerarse como un mínimo acuerdo metodológico compartido, entre visiones externalistas interrelacionadas del RBM, es la hipótesis general de que algunas representaciones mentales de formato modal están organizadas en unidades que contienen conocimiento de estructura espacio-temporal, conexiones causales, y otras estructuras relacionales. (Nersessian 2008).

Lo anterior puede parecer demasiado general, pero la diversidad de investigaciones dentro del “marco de investigación” del RBM - como Nersessian prefiere llamarlo - hace difícil encontrar generalizaciones de este tipo. Nersessian asume una representación modélica ensamblada a partir de la contribución, no sólo de otros modelos, sino que también de la acción sensorio-motora asociada a los diversos modelos (e.g. el modelo como producto híbrido de la interacción entre manipulaciones perceptuales internas y motoras externas). Asumir esto genera dudas respecto no sólo de cómo literalmente se podrían ensamblar estructuras representacionales internas y externas (especialmente en virtud de las particularidades materiales de cada cual), sino que también respecto al rol cognitivo que se le atribuye a las partes del ensamblaje extendido en el cual el científico (o cierta comunidad de científicos) es sólo uno de los componentes. Específicamente, un posible cuestionamiento podría ser el siguiente: si es el caso que un ensamblaje extendido [humano(s) + artefacto(s) + dinámicas de coordinación distribuida + todo lo que se quiera agregar] realmente causa alguna instanciación interna de representación modélica en el componente humano, ¿por qué aceptar que tal representación es también de naturaleza híbrida? O bien, si fuera realmente el caso de que un ensamblaje extendido tuviera el poder causal necesario para generar “modelos internos híbridos”, ¿por qué aceptar que un posible razonamiento inferencial basado en estos modelos tenga de regularse por los aspectos externos que ayudaron a generar tales modelos?

En este punto, parece relevante hacer una distinción entre el “carácter extendidamente híbrido” de las representaciones modélicas y el “carácter genéricamente híbrido” de las mismas. El primer caso hace referencia a un ensamblaje literal entre algún componente externo del modelo y otro interno, donde ninguno de estos componentes es constituyente de estructuras representacionales de orden abstracto<sup>34</sup>. El segundo caso se hace referencia a alguna estructura modélica interna susceptible de ser construida como parte de un proceso de abstracción genérica que incluye la revisión modificación distintos ejemplares que median entre *modelos fuentes* y *modelos target*. Para Nersessian, por ejemplo, estos modelos se instancian en la memoria de trabajo y tienen la cualidad de integrar constreñimientos en la forma de estructuras causales en tanto abstracciones de distintos modelos fuentes (externos o imaginados). La principal diferencia entre modelos híbridos extendidos y genéricos es que sólo los genéricos son modelos a partir de los cuales se pueden suscitar procesos de inferencia, probablemente tradicionales. Los extendidos, por

<sup>34</sup> Este ensamblaje es equivalente al tipo de extensión cognitiva regulada por el principio del externalismo activo, cuya versión paradigmática cae dentro de la denominada Hipótesis de la Mente Extendida (ver secciones 1.1.4. y 4.1.1.1.)

su parte, admiten un tipo de interacción causal fundada en procesos que diluyen una clara distinción entre acople y causalidad, postergando aún más el rol del razonamiento lógico como etapas menos interesantes del procesamiento cognitivo. No obstante, en cualquiera de los dos casos de hibridación, no se tiene una especificación con respecto al modo en que se haría ostensible el vínculo con procesos inferenciales clásicos.

La cuestión anterior parece no tener respuestas por el momento, pero uno podría cambiar el foco del cuestionamiento y sustituir el “por qué” por un “para qué”. De esta manera uno puede también considerar matices de orden más práctico detrás de las generalizaciones empíricas. Quizás, sólo se trate de motivaciones metodológicas orientadas a delinear algún programa de investigación productivo que permita teorizar sobre la base de los hallazgos experimentales disponibles. Dichas necesidades parecen priorizar excesivamente la dimensión científica descriptiva por sobre la normativa, propia de la justificación filosófica más tradicional. Algo de esto puede que esté reflejado en las dificultades para elaborar una teoría unificada del RBM sobre la significativa variedad de aproximaciones investigativas y objetos de estudio. En este contexto, Nersessian opta por desplegar lo que denomina el “Método Histórico Cognitivo” para dar cuenta de un modelamiento mental apropiado en el RBM. A través del prisma de este método, trataremos de identificar las bases de la noción de modelo marcada por medio de lo que ha llamado el “acople representacional”. Sostendremos que de este acople depende la concepción de un modelo extra-mental modal.

La noción de representación modal que aquí interesa asume un compromiso con la tesis de la cognición corporalizada. En términos generales, los enfoques corporalizados pueden tener aproximaciones más o menos radicales con respecto a la participación del cuerpo biológico en lo que sea que “amueble” o constituya la vida mental de agentes como nosotros. Uno puede tomar, por ejemplo, una posición moderada y sostener que el rol de nuestros cuerpos en la conformación de un sistema cognitivo extendido es importante pero no siempre necesario (Clark, 2008). O bien, sostener que sin cuerpos como los que tenemos, no podríamos percibir y cognocer las cosas como lo hacemos, lo que es equivalente a sostener que agentes con cuerpos distintos no pueden tener el mismo tipo de mentes. En el RBM, ambas visiones oscilan, nuevamente, a lo largo de gran parte del espectro subyacente graficado en la figura 1. Donde haya un énfasis en la cognición “on-line”<sup>35</sup>, habrá una tendencia hacia la visión más radical, y ocurrirá lo opuesto en el caso de enfatizar la denominada cognición “off-line”<sup>36</sup>. Tanto Nersessian como Magnani han intentado desarrollar aproximaciones y métodos que sirvan para caracterizar un alcance más global del RBM. Gran parte de la literatura producida por Magnani aborda posibles maneras en que la ‘abducción manipulativa’ (ver 3.1.1), propia de un nivel *online*, pueda resultar decisivamente compatible con la ‘abducción teórica’, propia del nivel *offline*. Por su parte, Nersessian ha llevado una prolífica investigación etnográfica para dar cuenta de cómo las prácticas científicas centradas en modelos externos ayudan a generar “modelos funcionales híbridos” como el producto de la contribución (física e históricamente distribuida) del quehacer científico de una comunidad determinada. Esto incluye la evidencia recopilada sobre la base de entrevistas, registros históricos, usos alternativos del instrumental y espacio de trabajo, etc.

Concretamente, podríamos decir que una representación externa es modal cuando se asume que su estructura no es de carácter arbitrario, sino que más bien está constreñida

<sup>35</sup> Aquella que ocurre anclada al ambiente en el que cierto agente intenta resolver una “tarea cognitiva”

<sup>36</sup> Aquella que donde el procesamiento cognitivo puede ocurrir con (relativa) independencia del ambiente físico y social que

lo motiva.

por el aparataje sensorio-motor (especialmente en un supuesto nivel *online*), y también por aspectos del mundo con iguales derechos causales (si se acepta la noción de procesos epistémicos extendidos). Pero este aspecto modal sería objeto de una exploración demasiado parcial sino se considera con respecto a la problemática relacionada con el realismo representacional (o representacionalismo). Consideremos, por lo tanto, lo central acerca de un sistema representacionalista clásico, y luego un par de argumentos críticos en su contra.

El representacionalismo (en una versión estándar) asume que la mente de un individuo no puede tener acceso directo a los objetos del mundo externo. La única forma de aproximarse a este mundo externo es a través de algún sistema representacional (probablemente constituido por estructuras representacionales básicas que pueden componer estructuras complejas) que reproducen realísticamente (en un sentido no ficticio) lo que hay fuera de la mente, si es que lo hay. Si esto es así, uno de los problemas que se le imputan a este representacionalismo tiene que ver con que la mente sólo puede “suponer” tal realismo representacional, precisamente, debido a la imposibilidad de acceder directamente a los objetos del mundo. Una segunda imputación relacionada es que, dada la posible disyuntiva de que, o bien una mente no pueda “capturar” nunca los objetos externos, o bien sólo pueda proporcionar un sustituto imperfecto, aquello sería prueba de que simplemente no habría nada susceptible de ser capturado “allá afuera”<sup>37</sup>.

Esta última imputación es la que da pie a un entendimiento de las representaciones tributario de tradiciones no sólo provenientes de la filosofía o la ciencia cognitiva. Como daremos cuenta en la sección 2.1.2., las investigaciones agrupadas dentro del campo de la denominada “percepción activa” han hecho converger visiones claramente opuestas en la filosofía y la ciencia cognitiva, a saber, representacionalismo y constructivismo. Una representación modal, por lo tanto, podría corresponder a cualquier objeto externo que no necesita ser reproducido por un sustituto representacional intrínsecamente interno. Basta que se considere el resultado (aún no especificado claramente) de un proceso que se construye a través de una suerte de negociación epistémica llevada a cabo por un sistema distribuido como el defendido por Nersessian. Asumiendo esta cualidad de la naturaleza de las representaciones modélicas externas, es posible estar preparado para aceptar el hecho de que la complementariedad que “acopla” un ensamble modélico híbrido [*interlocked models*] es dependiente de procesos simuladores internos y externos que sólo difieren entre sí por tratarse de distintos grados representacionales de un continuo construccionista. Sin embargo, incluso asumiendo todo esto, parece difícil de aceptar que algún “proceso modélico” de razonamiento inferencial pueda tener lugar sin apoyarse en otro (proposicional, quizás) que respete algún tipo de procesos de transformación regimentados formalmente, que dé cuenta de procesos ubicuos del pensamiento como, por ejemplo, la propiedad de ser composicional. En la figura 2, para dar un ejemplo simple, se podría asumir que los diagramas dados causan simulaciones inferidas a partir un problema planteado. Incluso sin atender a los enunciados lingüísticos, se podría interpretar lo que pregunta el (digamos) diseñador del ejercicio. Parafraseando a Peirce, la duda que inicia determinados trenes de pensamiento sólo tiene sentido si se considera aquello que *simboliza* el signo de interrogación “?”. Difícilmente, se puede aceptar que las supuestas inferencias basadas en simulaciones modélicas dependen, precisamente, de las características icónicas del “?” que nuestro sensorio puede captar. Eso sería equivalente a

---

<sup>37</sup> Cabe señalar que en esta segunda imputación no se está proporcionando un argumento concluyente acerca del supuesto realista representacional. Nótese que esta segunda postura anti-realista se basa en lo que opté por denominar como una “posible disyuntiva de que X”.

sostener que la materialidad de los signos lingüísticos impresos en el texto de la solución del problema (como la tinta en el papel o la imagen digital en la pantalla de mi computador) - en la parte baja de la figura 2 - es la que causa las simulaciones modélicas necesarias para entender la solución del problema. Dicho de otra manera, el proceso más interesante para resolver la tarea propuesta no parece inicialmente “gatillado” por los modelos que uno podría imaginar, o por analogía o por simulación, a partir de los modelos fuentes propuestos. La razón principal detrás de esta afirmación es doble. Por un lado, los modelos fuentes, por sí solos, no motivan un propósito iniciador de algún tipo de manipulación modélica. Por otro, el valor epistémico que se le atribuye, ya sea a los modelos fuentes o los híbridos que resulten a partir de una abstracción genérica, no ofrece respuesta a la pregunta normativa básica de cómo uno *debería* proceder cada vez que se ve confrontado con un problema que requiere un RBM en tanto sistema inferencial distribuido (problema tratado en Capítulo III). Ambas cuestiones parecen requerir tanto la noción de agente intencional como la existencia de relaciones lógicas subyacentes y/o reguladoras de un RBM.

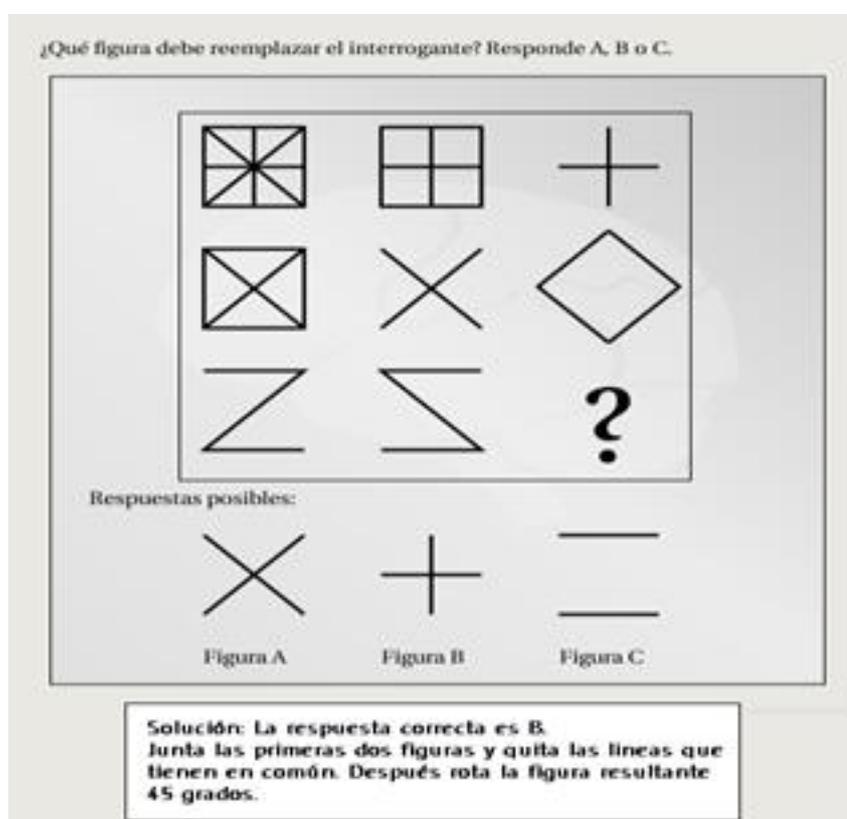


Figura 2 (Gentileza de [live.gnome.org](http://live.gnome.org) y Software libre [gbrainy](http://gbrainy.org)).

Pareciera más razonable sostener cierto escepticismo frente al rol autónomo de los modelos, especialmente si se admite que una comunidad científica determinada requiere compartir algún conocimiento previo respecto al uso de modelos, típicamente subordinado a los propósitos del o los científicos en cuestión. Este conocimiento compartido no sólo sería necesario, sino que también constreñiría el diseño y la materialidad que se estimen pertinentes para la construcción de tal diseño. Luego, una cuestión importante es especificar cómo las simulaciones, incluso las manipulativas, pueden prescindir de procesos inferenciales más tradicionales para jugar un rol autónomamente epistémico.

Dicho problema puede incurrir en el mismo tipo de imputaciones que se le hace al representacionalismo. Vale decir, cómo se podría probar que las simulaciones son confiables si sólo están basadas en modelos, considerando que no se dispone de una explicación acerca de la manera en que los modelos se “anclan” realísimamente al mundo externo. Frente a esto, la respuesta es meramente metodológica: se opta por no suponer que se requieren estructuras representacionales que “sustituyan” el mundo externo y se propone un programa en el que se asume un “continuo experiencialista”. Dado este nuevo supuesto, las investigaciones en RBM aquí estudiadas se han concentrado en heurísticas particulares que recurren a *formas extra-teóricas de pensamiento*<sup>38</sup>. Una manera de dar cuenta de los constreñimientos de este proceso es a través de la descripción de plantillas [*templates*] de conductas manipulativas relacionados a las “acciones cognitivas y epistémicas” más comunes, por ejemplo, cuando ocurren casos de descubrimiento serendípico. Estas plantillas también expresarían un tipo de simulación modélica, y como tales, según lo planteado anteriormente, constituirían un tipo de representación, pero en el sentido de una “re-presentación”, y no de una “sustitución” representacionalista de dudoso anclaje con una realidad independiente del agente cognitivo. ¿Cuál es la diferencia entre simulaciones epistémicas externas e internas? La respuesta que nos urge examinar es aquella que defiende una diferencia de *grados de re-presentación*.

En lo que viene a continuación, trataremos de caracterizar aquello que podría asumirse como las características compartidas de la noción de representación externa en el campo del RBM. Es cierto que existen diversas aproximaciones (ej. más o menos externalistas y/o anti-representacionalistas), pero no deja de ser equívoco el hecho de que esta noción sea invocada persistentemente como si estuviera fundada en un entendimiento cabalmente compartido.

### 2.1.2. La persistencia metodológica de las representaciones externas

---

Desde los primeros desarrollos interdisciplinarios de la ciencia cognitiva clásica, la noción de representación ha jugado un rol central en la explicación de la cognición humana. La centralidad de esta noción, tal como la entendían los fundadores de la *coalición de disciplinas* que conformaban la ciencia cognitiva de primera generación, sigue en pie, a pesar de los continuos ataques que ha recibido en las últimas décadas. De manera particularmente frontal e interesante, estas embestidas han surgido principalmente desde algunos enfoques inspirados en la teoría de sistemas dinámicos - como es el caso de los enfoques sub-simbólicos de la cognición - (ver sección c de 1.1.1.), así como también desde alguna de las áreas de investigación más innovadoras en el campo de la inteligencia artificial - como es el caso de la robótica situada (Ibíd.). En esta sección, se intentará esbozar la idea de que si bien la noción de “representación interna” tradicional ha sufrido una aparente osificación metodológica, aquello que se plantea como alternativa no logra desplazarla definitivamente.

No obstante lo anterior, y como algo más pertinente al presente trabajo, se intentará delinear una caracterización general de aquello a lo que se hace alusión como “representación externa” en los círculos de investigadores del RBM. A partir de esto, interesará defender la idea de que, en ciertos pasajes importantes de los planteamientos externalistas, ambos tipos de representaciones (internas y externas) parecen correr el

<sup>38</sup> Ratificándose, de pasada, la afirmación de que “la mayoría de las veces, las teorías basadas experimentalmente no se ocupan de proporcionar los fundamentos adecuados a los supuestos preliminares sustanciales con los que se comprometen cuando formulan sus teorías” (Vallejos 2008, p. 30)

riesgo de una suerte de trivialización metodológica. La razón principal de lo anterior no es sólo la imposibilidad de determinar un carácter claramente compartido de las “representaciones externas” en las investigaciones más representativas del RBM en ciencia, sino que también una falta de claridad con respecto a qué es aquello que, en cualquier caso relevante, no podría llegar a cumplir el rol de alguno de estos tipos de representaciones modélicas (internas o externas). Dicho de otra manera, como resorte de la liberación de factores (tentativamente) determinantes de un RBM, pareciera no haber procedimientos que permitan descartar como potencial representación interna a cualesquiera propiedades física en el ámbito del procesamiento intracraneal y, a su vez, descartar como representación externa a todo y cualquier cosa que no ocurra en dicho ámbito, en la medida que pueda cumplir el rol de “andamio cognitivo”.

Aún sin haber consenso claro sobre de la naturaleza misma de las representaciones internas (representaciones de estructura lingüiforme, conexionistas o modélicas, por ejemplo) dentro de los márgenes del individualismo metodológico, sí suele haber acuerdo respecto a lo imprescindible que resulta ser esta noción. Knuuttila y Honkela (2005) resume esto de la siguiente manera:

***“Tan fuerte es la sujeción a esta visión representacional de nuestros cerebros que, como resultado de las diversas maneras de conceptualizar y modelar nuestras mentes, cosas formidablemente diferentes han llegado a ser consideradas como representaciones”.* (p. 209)**

Para la ciencia cognitiva clásica - siguiendo a estos autores -, las representaciones internas corresponden a estados simbólicos estructurados, a partir de lo cual se asume que a cada estado del mundo le correspondería una unidad discreta de representación. Esta idea básica inspiró lo que Miller (2003) ha expresado como “el sueño de una ciencia unificada que descubra las capacidades de la mente humana y su realizaciones estructural y funcional” (p. 144). Esta ciencia inmadura e interdisciplinaria se materializó en dos aspectos principales: por un lado, el desarrollo de una teoría de la conducta humana en base a tareas específicas (Ej. resolución de problemas, comprensión del lenguaje, tareas de memoria) y, por otro, el desarrollo de un dispositivo computacional que desarrolle dichas tareas. Esto es lo que normalmente se entiende como el doble producto del modelamiento simbólico de la cognición (Lewis, 1999). Lo que la mente haría, al menos en una medida importante, es realizar determinadas tareas computacionalmente. Aquí, el carácter simbólico viene dado principalmente por el atributo de poseer las propiedades de un sistema simbólico tal como técnicamente lo especificó Newell y Simon (1976) en la Hipótesis de los Sistemas de Símbolos Físicos (HSSF). La HSSF proporciona una hipótesis acerca de las condiciones necesarias y suficientes para que un dispositivo físico pueda realizar tareas inteligentes. El requerimiento es que un sistema conformado simbólicamente debe ser capaz de manipular y componer estructuras simbólicas, con procesos asociados que le dan a este tipo de estructuraciones el poder de denotar otras estructuras simbólicas similares o símbolos que son instrucciones. El rasgo distintivo de este procesamiento es que es capaz de producir nuevas estructuras, lo que para sus defensores constituye una característica distintiva plausible para explicar la cognición. Este rasgo se caracteriza en términos del principio conocido como el de composicionalidad. Dicho principio hace referencia a los procesos de productividad y sistematicidad inherentes al procesamiento en cuestión, y evidentes en la conducta lingüística de los seres humanos. Dado que la composicionalidad de las representaciones mentales dependería de su estructura sintáctica constitutiva, es razonable la tendencia a suponer un carácter representacional lingüiforme del procesamiento cognitivo interno.

Por su parte, la alternativa conexionista propone, a través del uso de redes simuladas de microprocesadores simples, entender la representación como algo que está distribuido en una red de ese tipo. Específicamente, el conexionismo consiste en un procesamiento de la información realizado a través de una representación numérica en la forma de un vector que pasa por un conjunto de micro-procesadores interconectados en paralelo, de modo que dicho vector actúa como el *input* que atraviesa una matriz que representa los pesos de dichas interconexiones, con el fin obtenerse otro vector como *output* (ver sección c de 1.1.1.). En este caso, la representación es dependiente de la activación y la conectividad de la red completa. Aunque esta opción también tiene un carácter individualista (o internalista), es importante enfatizar la dificultad que ya se hace presente aquí para discriminar entre aquello que puede o no contar como una representación propia o genuina de lo que se quiera entender por cognición. Conocer la naturaleza y constitución de las estructuras representacionales internas y saber cómo éstas hacen la representación (o los procesos que determina su poder representacional) son dos problemas íntimamente vinculados. Si este problema resulta poco resolutivo en el caso de dichas representaciones mentalistas, la situación no es mejor cuando se incorpora la participación de representaciones externas en la explicación cognitiva. En este caso, la variedad de elementos y procedimientos subyacentes a la explicación tiende a ser tan amplia que estas nuevas estructuras sólo pueden intentar dar cuenta de aspectos menos sustanciales. Uno de ellos se puede plantear de la siguiente manera. Si de algo hay claridad, eso tiene que ver con lo difícil y complejo que puede resultar la individuación de una propiedad intrínseca que permita discriminar el carácter cognitivo de los objetos o partes del mundo en tanto representaciones externas. Por lo tanto, pareciera más adecuado, por su factibilidad, aspirar a dar una cuenta preliminar de algún rasgo (o conjunto de rasgos) común a ambos tipos de representaciones (internas y externas), sobre el cual puedan considerarse su estatus cognitivo, de manera inevitablemente arbitraria. En otras palabras, la tarea sería intentar establecer no el por qué alguna representación externa determinada posee un carácter cognitivo - dada la falta de claridad de lo que se sostiene por proceso cognitivo en cada caso relevante -, sino más bien por qué se tiende a hablar de *representaciones*, independientemente del dominio intra y extracraneal al que se haga alusión (para usar expresiones ilustrativas del límite individualista soslayado por la ciencia cognitiva actual).

En los enfoques externalistas de RBM, la noción de representación pareciera estar motivada por una interpretación liberalizada de la “máxima pragmatista” de Peirce (1878), a partir de la cual los modelos cumplen roles cognitivos susceptibles de ser discriminados por sus consecuencias prácticas, independientemente del ámbito en el que se le ubique. Es importante señalar que lo que aquí se opone al principio del individualismo metodológico es el denominado *principio de paridad* que Clark (2008) ha presentado como fundamentación del “externalismo activo” que defiende como principio regulador de la extensión mental (ver 1.1.4.). Este último principio se ha desarrollado sobre la heterogénea base experimental de los “argumentos de la cognición activa” aludidos en la cita de R. Wilson presentada en 1.1.3. Tal como la caracteriza Clark (1998), la perspectiva de un externalismo activo se basa en la idea de que el organismo humano está vinculado con el entorno en una interacción de doble sentido, constituyéndose así un “sistema acoplado”. Este sistema acoplado se debe entender como un sistema cognitivo por derecho propio, dado que todos sus componentes juegan un rol causal activo, gobernando la conducta tal como suele hacerlo la cognición según el entendimiento tradicional. Aquí la semejanza de gobernabilidad se da con relación a una aparente dependencia causal, no a características finamente detalladas del tipo de procesamiento involucrado en cada caso (intra o extra-craneal). De hecho, lo que licencia el principio de paridad es un argumento contrafactual: frente a una tarea determinada,

procesos y aspectos del mundo funcionan como procesos cognitivos intracraneales, porque si ocurrieran en la cabeza, no dudaríamos en considerarlos cognitivos.

Aquí, habría que hacer una salvedad con respecto al vínculo entre representaciones mentales y extra-mentales, ya sean modélicas o no. Si bien la existencia de ambos tipos de estructuras ha sido invocada para explicar algún tipo de procesamiento cognitivo alternativo y/o complementario al computacional simbólico, la “paridad cognitiva” comentada anteriormente puede ser igualmente relevante para concepciones de con motivaciones conceptuales muy distintas. Específicamente, tal paridad se puede concebir siendo parte de procesos cognitivos externalistas mediados o no mediados, sin que aquella distinción parezca ser muy importante. Por ejemplo, para Clark la conformación de un sistema cognitivo extendido no implica proceso de mediación alguno, dado que todos los aspectos físicos relevantes en (la resolución de) alguna tarea determinada pasan a formar parte del sistema. De ahí que cualquier aspecto del mundo que según el criterio de paridad pueda tener un legítimo estatus cognitivo al momento de resolver la tarea es algo que siempre actúa *con* el agente, y no, por ejemplo, al servicio del agente, o si se quiere, entre el agente y el mundo. Desde una perspectiva individualista de mente, se podría decir incluso que la noción de agencia cognitiva se torna una categoría irrelevante en ese caso. Otro ejemplo es la noción de “prótesis mental” o mediador epistémico de Magnani, la que sería dependiente de la atribución de significado del agente, y actuaría como agente epistémico autónomo en la medida que admite el rol de *interpretante semiótico-cognitivo* de algún problema u objeto susceptible de manipulación extra-mental. Para este autor, dicha mediación es un caso de externalización cognitiva, aún cuando no le asigna importancia a la demarcación intra/extra craneal de los procesos cognitivos.

Uno puede plantearse algunas preguntas relativamente obvias con respecto a los ejemplos del párrafo anterior. ¿Por qué la mediación es de carácter cognitivo en el caso de Magnani? ¿Por qué la extensión cognitiva propuesta por Clark no es considerada simplemente un caso de uso instrumental de aspectos físicos del mundo? ¿Por qué en ambos casos se habla de una *dimensión externa* si no hay algo claro con respecto a lo que dicha externalidad se pueda contrastar? Esto último parece más controversial si se considera que para Clark el mundo ajeno al sistema cognitivo extendido parece siempre prescindible (o siempre imprescindible, porque el límite de un sistema extendido no parece relevante), y que para Magnani parece no haber más que una única dimensión, siempre dependiente del agente (o de lo mental) e indiferente a lo intra/extra craneal.

Con fines clarificatorios, Knuuttila y Honkela (2005) - siguiendo a Prendergast (2000) - hace alusión a la distinción semántica entre el significado de representación como “representación”<sup>39</sup> y como “sustitución”. En el primer caso, representar significa la reaparición literal de una persona o cosa ausente, lo que también incluye la idea de hacer presente algo nuevamente por medio de un simulacro. En el segundo caso, representar significa sustituir, reemplazar o simbolizar [*stand for*], como cuando un término “b” representa o sustituye un término ausente “a”. Interesantemente, la sustitución puede tomar la forma de una simulación, pero en este caso sólo será una forma de re-presentar (según el primer sentido) si dicha sustitución “logra producir una ilusión de presencia en virtud de ser una copia del objeto real” (p. 211).

El segundo significado descrito – prosiguen estos autores - es el que se asocia al representacionalismo (o realismo representacional), visión filosófica según la cual una

<sup>39</sup> Aunque la simbiosis filosófica-experimental entre la filosofía fenomenológica y los enfoques corporalizados/embedidos de la cognición no se abordará en este trabajo, es pertinente hacer notar que la concepción de *re-presentación* es consistente con la noción de *interacción habitada* de inspiración fenomenológica desarrollada por Dourish (2001).

mente capaz de conocer no puede tener acceso directo a los objetos del mundo, sino sólo indirectamente a través de ideas que los representen. De aquí se sigue la visión de que el conocimiento es una suerte de ensamblaje de representaciones que reproducen o sustituyen fielmente lo que yace fuera de la mente. Consecuentemente, quienes adhieren a esta concepción del conocimiento suelen sostener que un conocimiento es verdadero o genuino en la medida que nuestras representaciones mapean correctamente el mundo, es decir, representándolo tal como sería, independientemente de dichas representaciones. Esto es lo que se ha caracterizado como una metáfora pictórica del conocimiento, y las críticas que recibe normalmente tienen que ver con la clara asociación que mantiene con un funcionamiento intuitivo del sistema visual.

Considerando las dos concepciones descritas, es preciso señalar un aspecto curioso de los desarrollos experimentales del estudio de la visión en ciencia cognitiva, a saber, la asociación entre representacionalismo y constructivismo, a pesar de la diametral oposición existente entre estas visiones en ámbitos como las ciencias sociales y la filosofía. Como se sabe, el constructivismo niega que la representación y lo representado pertenezcan a ámbitos ontológicos diferentes, sosteniendo en cambio que ambos se construyen en el acto de producción de conocimiento. Así es como debe entenderse la idea proveniente desde la ciencia cognitiva actual según la cual los movimientos sacádicos de los ojos, en tanto procedimiento epistémico (de carácter indexical y como estrategia de optimización de memoria biológica), tendrían un rol importante en la construcción de las distintas partes de alguna imagen perceptual determinada. Esto sería, por un lado, evidencia de que nuestros cerebros biológicos nos re-presentan el mundo, y, por otro, cierto argumento para sostener que si la percepción visual “directa”<sup>40</sup> es algún tipo de re-presentación, también podría serlo el resto del “amueblado” mental. El punto es que, según esta visión, la *ilusión de presencia en virtud de ser una copia del objeto real* que alude Knuuttila, parece adquirir un estatus re-presentacional que ya no sería considerado como una ilusión.

Esta concepción alternativa, fundada en la simbiosis entre representacionalismo y constructivismo recién comentada, tiene una interesante manifestación en el contexto de los modelos y la representación científica. Esta manifestación tiene que ver con la posibilidad de atribuir un estatus re-presentacional a un objeto de factura constructivista (o constructivístamente reificada), que es, al mismo tiempo, susceptible de ser manipulado sin importar el formato que pueda tomar. Esto último tiene como antecedente la noción laxa de representación que Gatti et al. (2005) estipuló como dependiente de la noción de ‘información’ en tanto *cualquier conjunto de datos que incrementalmente momentánea o permanente la base de conocimiento de un agente* (ver 1.1.3.). Así concebida, esta noción liberaliza la concepción misma de ‘proceso cognitivo’ en tanto procesamiento información regulado por un criterio de desempeño exitoso frente a cualquier tarea específica. De aquí, la diversidad de concepciones asociadas a las representaciones modélicas cuyo único rasgo compartido aparente es el de ser parte de una reacción metodológica al énfasis que la filosofía de la ciencia convencional ha puesto en las teorías y en los experimentos, probablemente constreñido por una aproximación científica comprometida con un realismo representacional. La reacción en cuestión, según lo que alcanzo a comprender, es de carácter pragmática porque se ajusta a la caracterización amplia especificada en la *Introducción* de este trabajo<sup>41</sup>, y por lo que tendría su expresión a partir de la primera

<sup>40</sup> Esto debe entenderse como una manera de capturar el contenido de la experiencia distinta a la de una copia empirista de las impresiones sensoriales, dado que en este caso la experiencia sería parte (literalmente) constituyente del contenido mental.

<sup>41</sup> Según la tesis pragmatista de que *lo esencial del pensamiento y la cognición no es su relación con las cosas del mundo que representan, sino las acciones que guían en un mundo materialmente complejo.*

transición ilustrada en la figura 1, donde, como resultado de dicha transición, el rol del agente cognitivo y sus propósitos prácticos comienzan a ser reivindicados con miras a una explicación alternativa de la cognición. Esta transición, se podría caracterizar como se expone a continuación.

La aproximación canónica de la ciencia se funda en la idea de que nuestras mejores teorías son descripciones más o menos verdaderas de los objetos y procesos que se pretenden explicar. Esto se manifiesta en una visión diádica entre las representaciones científicas y los fenómenos naturales, donde no participan el agente (o sus interpretaciones y propósitos) como determinantes en la expansión del conocimiento. La reacción pragmatista postula que la aproximación científica debe ser mejor entendida como una relación triádica que incluya, además de lo anterior, al agente en un rol que se puede evaluar conforme al “éxito/fracaso de logro” que su participación (o conducta) tenga con respecto a las prácticas productivas o de expansión de conocimiento. Esto implica asumir al menos dos cosas: por un lado, el supuesto pragmatista según el cual no habría cosas que puedan representar otras cosas en sí mismas, independientemente de la participación del agente; y por otro lado, que, dada la relevancia que tiene el científico en la construcción del nuevo conocimiento, las prácticas de este agente importan tanto o más que la tarea tradicional de testear hipótesis derivadas hipotético-deductivamente a partir de algún marco teórico formal. Luego, resulta entendible que una nueva aproximación al descubrimiento en ciencia ponga atención a la más importante y frecuente de las prácticas científicas: el modelamiento.

Es importante hacer notar que este nuevo énfasis en el modelamiento como práctica científica conlleva un problema significativo. La discusión sobre la que se basa esta transición tiene una correspondencia con aquellas acerca de la representación mental anteriormente aludida. Un supuesto fuerte del realismo representacional es que basa la adquisición de conocimiento genuino en un carácter representacional intrínseco de las representaciones. Esto es, sobre la base de un realismo intencional posiblemente determinado por leyes naturales peculiares de la intencionalidad mental en tanto fenómeno relacional, el cual permitiría anclar el contenido de la representación a algún aspecto u objeto del mundo. Luego, un énfasis en las prácticas de modelamiento debiera proporcionar evidencia para una explicación que subsuma esta probable ventaja, en el entendido que una explicación de dicho fenómeno relacional no viene normalmente de la mano con las explicaciones del RBM que no adhieren al realismo intencional. Para salvaguardar esta desventaja, se incurre en la atribución de algún grado representacional *ad hoc* a los modelos utilizados (reificados, para usar la expresión de Magnani) por los agentes científicos. La visión pragmatista acepta que para obtener conocimiento genuino es necesario preservar la noción de representación, pero como también rechaza el compromiso realista de las representaciones (relativo a la naturaleza misma de las representaciones), entonces parece inevitable abandonar la idea de que los modelos deban ser entendidos exclusivamente como entidades mentales diádicas<sup>42</sup> (ver Suárez, 2002). Al menos, esto es así con respecto al entendimiento de modelos metales de tipo semántico-estructuralistas donde el modelo en tanto figura abstracta guarda un isomorfismo con lo representado. En la concepción semántica, el poder representativo de una estructura modélica depende de dicho isomorfismo. No obstante, cuando se atiende, precisamente, al tipo de estructuras consideradas como determinantes para el descubrimiento en ciencia (diagramas, ecuaciones matemáticas, programas computacionales, organismos, población

<sup>42</sup> Por *diádico* debe entenderse una relación entre dominio fuente y dominio *target*, por ejemplo, en términos de similitud, causación o denotación.

de científicos de un laboratorio, etc.), el problema que surge es que no parece haber consenso respecto de qué depende la proyección que hace posible hablar de isomorfismo, aun teniendo la convicción de que la cognición y el conocimiento deben estar basados en algún tipo de representación.

El problema planteado, según se ha sugerido por Giere (2004) y otros (e.g. Knuuttila y Honkela, 2005), puede tener un tratamiento alternativo si se abandona la centralidad que se le ha dado al carácter representacional intrínseco de la representación, enfocándose más bien en los usos prácticos que el agente le da a esta variedad de modelos, toda vez que la relación entre la representación y lo representado es algo que debe ser construido. Esta construcción es lo que se considera como el logro específico resultante de la actividad científica, y es también lo que motiva el estudio de las denominadas “prácticas representacionales” en contextos científicos. Es preciso señalar que dicho programa de investigación resulta tributario de otros de mayor alcance, donde se pretende dar forma a planteamientos alternativos relacionados a la cognición humana en general, tal como fueron delineados en el primer capítulo de este trabajo (apartado c en 1.1.1). Por ahora, será pertinente comenzar a examinar la manera en que la concepción pragmatista delineada hasta el momento se torna un argumento para fundamentar el carácter artefactual de los modelos científicos. Asumiendo que este tipo de modelos es susceptible a determinadas formas de manipulaciones, esta concepción también juega un rol central a la hora de dar cuenta del estatus de mediadores que se le atribuye a los modelos, algo que atraviesa gran parte de la investigación en RBM externalista.

En resumen, en esta sección hemos comentado y especulado acerca la centralidad que ha tenido la noción de representación en el contexto de las ciencias de la cognición desde sus inicios. En este sucinto escrutinio se abordó el tipo de antecedentes que nos parecen relevantes para una concepción inicial de representación externa que parece jugar un rol afín a al RBM dentro de los enfoques externalistas. Se comenzó con especificar las características de la noción en cuestión en el marco de la ciencia cognitiva de corte mentalista. Según dijimos, la ciencia cognitiva clásica concibió la representación como estados simbólicos estructurados dentro de un sistema simbólico formal que tiene su más clara especificación en términos de la HSSF. Al mismo tiempo, caracterizamos la representación conexionista como una representación numérico-vectorial distribuida en una red de microprocesadores, dependiente de la activación y la conectividad de la red completa. Dado el carácter descriptivo-experimental que subyace a la investigación en RBM, planteamos que el problema de la naturaleza de las representaciones, así como el problema del poder representacional de las representaciones, han sufrido un cambio de énfasis marcado por la propuesta de una noción de representación distinta a las anteriores. Esta noción es modélica pero caracterizada por no poseer un formato definido que aplique para todos los casos de RBM. Dicha diversidad hace controversial el hecho de que se pueda estar hablando de un mismo fenómeno bajo el rótulo de RBM, pero refleja la centralidad que sigue teniendo la noción de representación.

Para delinear inicialmente la noción de representación modélica de corte pragmatista con la que queremos caracterizar al RBM externalista, hemos resumido cuatro tipos de antecedentes. Primero, la idea de una representación modélica como una categoría regulada por una interpretación liberalizada de la Máxima Pragmatista de Peirce. Segundo, el denominado Principio de Paridad como antecedente controvertido de las representaciones externas en tanto “andamio cognitivo” de formatos (e.g. físicos, semióticos) y como parte de *procesos cognitivos* (e.g. mediados, no mediados) variados. Tercero, la distinción idiosincrática dentro de la *comunidad de RBM* (externalista) entre dos

acepciones del término ‘representación’: como “re-presentación” y como “sustitución”. Y cuarto, la asociación conceptual entre representacionalismo y constructivismo, y su efecto en la concepción pragmatista de re-presentación como copia (o emulación) genuina de un objeto o evento del mundo que requiere explicación por RBM.

La concepción modélica pragmatista correspondiente se puede delinear como a una re-presentacionalista constructivista reificada, cuyo formato oportunista está determinado por el formato que posea aquello que se considere como el tipo de información participante del procesamiento cognitivo, regulado por el propósito de guiar la acción exitosa de un agente cognitivo en el contexto situacional de una tarea determinada. Como esta información se define laxamente como cualquier tipo de *datos* que incrementa la base de conocimiento del agente, la forma de la información es la representación. Esta concepción, más bien un esbozo general, de representación modélica de un RBM externalista está en la base de aproximaciones más radicales que abogan por una noción libre del problema representacional más tradicional. Ahora abordaremos una noción anti-representacionalista de modelo como agente epistémico.

### 2.1.3. Modelos en la perspectiva de un razonamiento centrado en modelos

Para comenzar, insistiremos en la siguiente afirmación. Una de las razones por la cual se ha hecho tanto hincapié en el tema de la representación y el representacionalismo es la dependencia tradicional que la ciencia ha tenido del carácter representacionalista de la representación científica o “traductora”, para emplear la expresión de Giere (2002, 2004). La visión más común de las teorías científicas es que son algún tipo de representación estructurada más o menos precisa de la realidad, ya sea que se considere a dicha representación como algo totalmente independiente de la realidad o no. Como las representaciones parecen jugar un rol crucial en el estudio de la ampliación del conocimiento, no es un misterio que la relación entre representación, conocimiento y cognición esté en el centro de las investigaciones en RBM. Aceptándose la centralidad de esta interrelación, abordaremos una concepción radicalmente revisada de representación y su rol en la producción de conocimiento, motivados por distintos entendimientos de la cognición.

Como veremos, los enfoques anti-representacionalistas de la representación modélica tienden a resistir la tentación de fundar el poder representacional de los modelos en la noción de representación inmaterial o ideal, como sería el caso de estructuras formales abstractas, ya sean de estructuración proposicional o como modelos mentales según los entiende Johnson-Laird. Lo que normalmente se propone a cambio es analizar el poder representacional de los modelos desde un punto de vista de las prácticas epistémicas implementadas como parte de un *sistema cognitivo acoplado*:

***“En general, hay un amplio reconocimiento de que las prácticas científicas y ordinarias de solución de problemas indican la necesidad de re-considerar la relación entre los mundos internos y externos, para ser entendidos como formando un sistema cognitivo acoplado. En esta visión, las inferencias se realizan por el sistema cognitivo que incluye personas y una gama de artefactos.” (Nersessian, 2008, p. 116. Énfasis original)***

Lo anterior ha dado pie a un programa amplio y heterogéneo de investigación que se centra en la producción y el uso de representaciones físicas, así como también en supuestos

rasgos epistémico asociados a su particular materialidad. De ahí la noción de modelos como artefactos epistémicos (en tanto agentes epistémicos autónomos) que veremos en el resto de este capítulo. Para este fin, será útil tener en cuenta que el carácter mediador y artefactual de las *representaciones* modélicas representa una reacción a ciertos supuestos básicos del representacionalismo que se consideran errados o incorrectos para guiar una investigación adecuada de un RBM elemento central de la actividad científica en tanto proceso cognitivo. Knuuttila (2005) resume esto en tres rasgos asociados al representacionalismo: el primero es que el conocimiento consiste en un conjunto de representaciones internas; el segundo es que estas representaciones se corresponden con partes de la realidad de manera precisa; y el tercero es que la realidad consiste en objetos cuya existencia es independiente de las representaciones internas.

Como alternativa a los rasgos mencionados, se propone una serie de tesis de relevancia metodológica. Estas tesis actúan, a su turno, como nuevos supuestos desde un punto de vista anti-representacionalista. Primero: la función representacional de los modelos no debe ser entendida en términos “representacionalistas realistas”, sino más bien como artefactos externos utilizados por los humanos para interactuar con el mundo. Segundo: los modelos no son entidades abstractas, sino entidades que se materializan en algún medio físico. Tercero: el valor epistémico de los modelos se funda significativamente en su dimensión material, lo que explicaría las funciones adicionales que puede tener el modelo además de representar el mundo. En lo que sigue, veremos cómo estos supuestos se hacen manifiesto las concepciones mediadoras y artefactuales de los modelos.

### 2.1.3.1. Los modelos como mediadores epistémicos y artefactos cognitivos

La persistencia de los modelos como representaciones materiales de carácter artefactual, y susceptibles de ser manipulados, es lo que da lugar a la concepción de modelos como “mediadores epistémicos”. En este caso, se puede hacer alusión a la última transición de la figura 1. Según esta transición, lo que adquiere preponderancia en un RBM es el énfasis en lo que se denomina la agencia epistémica.

Un modelo con carácter artefactual – habría que asumir - resulta más apropiado para describir el rol práctico que puede tener este tipo de representaciones externas en las prácticas científicas. Según Morgan (1999), los modelos pueden ser considerados como agentes autónomos, en el sentido de que a través del proceso de construcción ganarían una independencia del marco científico de justificación tradicionalmente asumido por la filosofía de la ciencia<sup>43</sup>. Una generalización asociada a la visión tradicional los modelos es que éstos tienden a ser considerados abstracciones o réplicas teóricas de algo que es externo a ellos mismos, y que de alguna manera logran instanciar. Estos autores reaccionan a esta visión destacando un rol mediador instrumental de los modelos. Este rol instrumental no sería trivial, sino más bien uno que les permite ser categorizados como “instrumentos investigativos”, lo que representa una versión de modelos más consistente con el carácter “extendidamente híbrido” que con el “genéricamente híbrido” de las estructuras representacionales que mencionáramos en la sección 2.1.1.2. Sin embargo, lo interesante esta categoría extensionalista de instrumentos investigativos se funda en el mismo tipo de supuesto externalista de inspiración empírica que regula la hibridez genérica. Nersessian (2008) resume este supuesto claramente:

---

<sup>43</sup> Probablemente, esto también asume las distinciones planteadas por Magnani entre una concepción de “agente activo”, propia de la cognición activa, y una de “agente estático” [*stagnant agent*], propia de un cognitivismo-representacionista clásico. (Según comunicación personal con el autor)

**“Una razón por la que es importante tomar en cuenta los recursos externos es que la posibilidad de razonamiento simulativo, en casos que no sean los más simples, a menudo soslaya representaciones y procesos demasiado detallados y complejos para ocurrir ‘en la cabeza’”. (p. 115)**

De manera congruente con la transición general anteriormente señalada, este cambio en la concepción de los modelos se puede describir también como una transición entre la visión formalista de los modelos y una visión radicalmente funcional de modelos manufacturados a partir de fuentes diversas. Mientras que en la primera visión lo que importa es una estructura abstracta subyacente a los modelos, en la última visión los modelos suelen construirse con una multiplicidad de ingredientes: analogías, metáforas, nociones teóricas, conceptos matemáticos, técnicas matemáticas, datos empíricos, etc. Dicho de otra manera, los que aquí se quiere enfatizar es que una visión permisiva de los modelos como mediadores (donde, además, el modelamiento computacional ha jugado un rol central) es la consecuencia radical de una transición entre los intentos por establecer qué son los modelos en tanto formalizaciones aplicables de manera general y universal (e.g. Suppes, 1961; van Fraassen, 1980), y los intentos por establecer el rol pragmático, o la función, de los modelos dentro del quehacer científico. El resultado es lo que caracterizaremos como mediadores epistémicos en tanto artefactos cognitivos.

En esta concepción alternativa, los modelos computacionales no son entendidos como algo que representan algún sistema objetivo [*target system*], al menos como una réplica fiel e independiente de lo representado. Más bien lo que importa es que son utilizados para obtener algún output determinado o algún resultado interesante. Esto explicaría, en cierta medida, el cambio de énfasis desde un interés en la estructura de un fenómeno determinado hacia la “simulación” de su comportamiento, tal como se expresa en la siguiente cita:

**“Así, el interés en el funcionamiento interno de los modelos es canalizado a través de su desempeño [*performance*] – lo que uno típicamente quiere es saber qué tipo de comportamiento es causado por cierto tipos de factores [*features*] y soluciones. Esto nos obliga a tener una nueva mirada de los desafíos epistémicos del modelamiento” (Knuuttila y Honkela, 2005, p. 216).**

Las razones para fundamentar esta nueva concepción pragmatista de los modelos se puede resumir como sigue. Para conocer más acerca de un fenómeno determinado (u obtener nueva información), no basta construir modelos idealizados del mismo con el fin obtener una descripción de su funcionamiento, precisamente porque se sabe poco del fenómeno. Más bien, los modelos deben tener una función ilustrativa o simuladora del fenómeno en cuestión, que no actúa como una representación del fenómeno, sino como un sistema artificial auto-contenido a partir de cuyo funcionamiento podemos *inferir* cosas nuevas acerca de cómo podría ser dicho fenómeno. En consecuencia, los modelos como mediadores no actúan como una imagen [*picture*] de alguna parte de la realidad, sino más bien como una suerte de *realidades artificiales paralelas*. Cuando esta última noción se toma en un sentido real o literal (y no sólo metafórico, por ejemplo), se puede entender por qué suele ser considerado un tipo de *agente epistémico autónomo*. Pero entender por qué se hace esa atribución radical no es garantía de que se esté proponiendo una alternativa claramente viable para un RBM de la noción en cuestión.

La concepción de mediadores epistémicos se funda en la idea de que ciertos modelos externos son intencionalmente contruidos sobre la base de recursos materiales que permiten ser utilizados (cf. Giere 2004, en 2.1.) en una variedad de formas. Por ejemplo, un mismo modelo - o incluso algún método de modelamiento - puede ser utilizado en

distintas maneras tanto como herramienta o como objeto de estudio. Como los modelos son *interpretados* de distinta forma en la medida que son utilizados en distintas disciplinas, se dice que estos modelos sólo alcanzan su valor epistémico definitivo cuando incorporan conocimiento y datos propios de cada disciplina (Ej. cierto aspecto físico de un mismo modelo puede cumplir la función de campo electromagnético en el ámbito de la física, y el de vórtices producidos por un objeto en el agua en el ámbito de la vida artificial). Para alcanzar un formato contingente consolidado, dichos modelos son construidos en base a la explotación de medios materiales específicos, los que obviamente constriñen la manera en que los problemas son diseñados y la medida en que pueden facilitar ciertos razonamientos y soluciones. Esto es lo que permite correlacionar el grado de productividad epistémica de los modelos con el grado de apertura [*open-endedness*] - o su carácter multifacético. Los modelos son productivos epistémicamente si el rol epistémico de su materialidad cruza una variedad de usos científicos.

El carácter múltiple de los usos y contextos a los que se aplica un modelo, en virtud de su propia materialidad, es lo que determina su carácter autónomo en tanto mediador epistémico. El producto de los modelos así entendidos son los “problemas inesperados” o las líneas de investigación o las nuevas aplicaciones a las que puede dar lugar. Como se ha señalado, una aproximación radical y anti-representacionista de los modelos externos requiere poner el foco en los constreñimientos materiales de dichos instrumentos de investigación, los que normalmente se conciben como un tipo de *affordance* (ver Capítulo III) capaz de portar conocimiento. A su vez, este planteamiento resulta central en el énfasis sobre los agentes epistémicos de la última transición aludida en la figura 1, por cuanto se funda en una base evidencial típicamente aludida por la visión extendida de la cognición. Según esta hipótesis de la cognición, nuestro conocimiento estaría íntimamente determinado por nuestra habilidad para construir herramientas y manipular nuestro entorno material inmediato (Clark, 1997, 2008 y otros). Estos planteamientos darían fuerza al supuesto de que los modelos, en tanto mediadores materiales externos, constituyen un objeto operativo [*working object*] cuya manipulación determina procesos de aprendizaje científico. Obviamente, este aprendizaje está centrado en el agente psicológico, pero, para entender la visión aquí comentada, habría que asumir que hay algo sustantivo (i.e. el permanente debate de una supuesta complementariedad entre procesos intra/extra craneales) en el uso de modelos interpretados en tanto modelos físicos.

En esta visión alternativa de los modelos como mediadores epistémicos, el vínculo entre conocimiento y representación se debe entender desde la perspectiva constructivista. Según esta perspectiva, ninguna de estas cosas está dada previamente, sino que son realidades construidas por los agentes epistémicos participantes de la actividad mediadora y manipuladora (e.g. humanos o artefactos), que de alguna manera debiera ser concebida como un proceso que regula la transacción causal entre el cerebro, el cuerpo y el ambiente. Si bien es cierto que esta postura intenta impugnar cualquier visión representacionista, parece hacerlo a un precio que puede ser muy alto, por cuanto asume una concepción oscura de conocimiento. Específicamente, sería una concepción oscura si se compara con una concepción de conocimiento de naturaleza representacionista, donde se asume una relación realista entre lo representado y la representación (i.e. la explicitación de la descripción científica). Para que esto quede más claro, es preciso decir algo más sobre la supuesta dimensión artefactual y cognitiva de los modelos, porque aún parece posible explotar con fines clarificatorios ciertas implicancias que surgen a partir de esta noción de modelos.

A diferencia de la visión pragmatista morigerada de los modelos defendida por Giere (i.e. los modelos deben ser similares a su *target*, en ciertas maneras que deben ser especificadas; ver 1.1.1.), la concepción de los modelos como artefactos cognitivos epistémicos tiene sentido para un defensor del anti-representacionalismo científico sólo cuando se abandona completamente la idea de que los modelos cumplen un rol mediador, exclusivamente, entre teorías y datos (o entre teoría y mundo). Ahora bien, el abandono de la relación diádica señalada no es algo motivado por una concepción artefactual inequívoca de modelos en ciencia, sino que más bien debe entenderse como una etapa ulterior de un *giro* constante que ha estado ocurriendo en el estudio de los modelos en ciencia, especialmente informado por hallazgos empíricos provenientes de disciplinas como la psicología cognitiva. Este giro comienza con una reacción a lo que Bailer-Jones (1999) denomina el desprecio del empirismo lógico hacia situaciones donde ciertos aspectos cognitivos del uso de modelos podrían conformar un campo de investigación bien establecido, seguido por el tipo de reivindicaciones que filósofos como Giere hacen no sólo del uso de los modelos sino que también de una epistemología naturalizada de que se centra en el agente cognoscente y sus propósitos. Adicionalmente, habría que adoptar la perspectiva de que los modelos son instrumentos cognitivos debido a su *relativa independencia* mediadora. Esto último apunta al supuesto externalista de que es la “materialidad” misma de los artefactos la que proporciona cierta robustez necesaria que les permite conservar una identidad a través de distintos campos científicos, donde diversas actividades y actores convergen en torno a tales objetos. Un ejemplo de esto se podría constatar en el modelamiento computacional, donde los modelos tenderían a actuar como “objetos delimitadores” entre grupos de científicos de disciplinas distintas (e.g. modelamiento interdisciplinario de la transmisión de alguna enfermedad infecciosa).

Lo que se asume detrás de la noción de independencia mediadora, y los roles adicionales de los modelos autorizados por esta independencia (i.e. delimitador de prácticas y sub-comunidades de investigadores), es que uno *aprende con* los modelos no sólo al construirlos y manipularlos, sino también debido a su propia dimensión material. Gracias a su materialidad, uno puede interactuar y trabajar con artefactos concretos. Aquí, la noción de materialidad se reduce a una cosa física auto-contenida que ocupa tiempo y espacio, y es susceptible de interactuar con otras entidades similares y con las personas. Luego, si los modelos se restringen a cosas materialmente (y artificialmente) construidas, no es difícil entender que aquello que constriñe su carácter de artefacto epistémico puede ser reducido a las denominadas “*affordances* materiales para propósitos epistémicos”<sup>44</sup>. La razón por la que estos aspectos materiales adquieren centralidad se debe al giro metodológico anti-representacionista del RBM externalista, el cual relaciona un cierto valor epistémico de los modelos a la *productividad epistémica*, y no a la representación de sistemas sociales o naturales preexistentes.

Por supuesto, uno podría sostener que el uso meramente instrumental de los modelos puede propiciar en el agente intencional el diseño de experimentos que bien pueden ser entendidos como un caso de productividad epistémica, sin tener que asumir la centralidad o el rol cognitivo que se le atribuye a las *affordances* materiales del modelo. Eso es algo que estamos dispuestos a aceptar desde la perspectiva crítica de este trabajo de investigación, pero, eso no impide tratar resaltar el hecho de que un entendimiento de la visión criticada pasa por reconocer la disyuntiva que se manifiesta en posturas dicotómicas como la siguiente. Si uno piensa que el asiento natural de la cognición está en la cabeza, dada una relación de determinación causal legaliforme entre lo mental y ciertos procesos biológicos

<sup>44</sup> Esta noción se atenderá con más detalle en el capítulo siguiente.

cerebrales, entonces lo obvio es que es el individuo el que está *aprendiendo de* o a partir de los modelos. Sin embargo, el rol cognitivo que se le atribuye a la *affordances* de un modelo artefactual (como instrumentos investigativos) parece entenderse por parte sus defensores como parte de un sistema distribuido, donde tienen lugar procesos capaces de transformar las tareas cognitivas, en el entendido de que tanto los científicos como los instrumentos de investigación son responsables causales del aprendizaje que pueda ocurrir. A la expansión de conocimiento que ocurre por medio de este tipo aprendizaje es lo que hemos denominado productividad epistémica, asumiéndose que en este caso el individuo (o un conjunto de ellos) *aprenden con* los artefactos (modelos físicos, instrumentos, etc.).

Según la perspectiva alternativa en cuestión, la concepción representacionista no sólo corresponde a un prejuicio adquirido de la tradición realista de los modelos científicos, sino que además constituye una concepción estrecha del rol de las representaciones en ciencia. Como veremos en el Capítulo III, lo que se propone a cambio es un carácter epistémico de los modelos determinado por la manera en que se intersectan sus constreñimientos materiales y sus usos intencionales. Dicho de otra manera, el valor epistémico de los modelos derivaría del hecho de que son objetos materiales intencionalmente construidos y con propósitos constreñidos por su materialidad.

Algunos autores como Magnani (e.g. 2008 y otros) y Knuuttila (e.g. 2005 y otros), defienden la tesis de que la intencionalidad y materialidad de los artefactos epistémicos se *acoplan* o convergen en la noción de *affordance* propuesta por Gibson (1979) en su teoría ecológica de la percepción. Una ventaja de esta noción es que iría más allá de la dicotomía entre lo objetivo y lo subjetivo, toda vez que, al acentuar la *complementariedad* entre el ambiente y el organismo, se funda simultáneamente en ciertas propiedades materiales del ambiente y en sus consecuencias para el organismo. De aquí que la aparente obviedad de que la materialidad de un modelo constriñe los usos a los que pueda ser destinado por parte de los científicos sea considerada una forma de constreñimiento determinante del razonamiento científico.

En el siguiente capítulo abordaremos esta concepción epistémica de los modelos desde el punto de vista de las prácticas manipuladoras extra-mentales sobre las que se hipotetizan supuestas etapas cruciales en el surgimiento de nuevas hipótesis. Veremos que, a través de un proceso denominado *abducción manipulativa*, sería posible construir “prótesis mentales” o *mediadores epistémicos*, específicamente por medio de la interacción constructiva con objetos externos. De esta forma se crearía *conocimiento tácito* o información implícita a través de la acción, y con eso se producirían, por ejemplo, oportunidades para detectar anomalías. Esta información implícita y corporalizada cumpliría un rol crucial en el proceso subsiguiente de comprensión y descubrimiento científicos.

Siguiendo la investigación epistemológica de Polanyi (1996), autores como Magnani destacan la idea de que gran parte del conocimiento no es explícito, sino que tácito. Es decir, sabemos más de lo que podemos decir, y no podemos saber nada sin apoyarnos o confiar en aquellas cosas que no estamos en condiciones de decir o reportar. Se sostiene, además, que este conocimiento tácito es un conocimiento de orden práctico requerido para llevar a cabo alguna tarea. De aquí la importancia de apoyarse en los desarrollos provenientes de la “percepción activa”, relacionada a la tradición del denominado “realismo ecológico” propuesto por Gibson hacia la segunda mitad de la década 1960. Según este enfoque, los sistemas cognitivos deben entenderse como en términos de su *situacionalidad* ambiental [*environmental situatedness*]. Para Gibson (1966), el organismo que percibe debe ser entendido como un **agente activo** capaz de coordinar sus propias capacidades de acción o de movimiento con la información disponible en el entorno inmediato. En este

contexto, las capacidades perceptuales del agente son utilizadas para extraer y obtener “cualquier” información específica necesaria para su comportamiento en el mundo.

Este proceso implica un ajuste continuo de recolección de información tácita, ecológicamente controlado. La responsabilidad de este control es compartida, ya que los agentes no sólo habitan su entorno, sino que también lo modifican activamente. Esta acción modificadora, o *constructora de nichos*, también tiene lugar sobre el ambiente epistémico, y es lo que se haría manifiesto en las prácticas científicas “buscadoras de oportunidades epistémicas”, en el caso del pensamiento científico abductivo. Aceptándose estos aspectos extra-mentales de un posible razonamiento abductivo, se estaría preparado para entender el tipo de mediación epistémica material al que normalmente se recurre en RBM (e.g. el uso de diagramas matemáticos para enseñar conceptos matemáticos difíciles de entender). Si bien es cierto que aquellos modelos externos no proporcionan un conocimiento explícito completo, sí orientarían al agente cognitivo para entrar en un *diálogo epistémico continuo* entre dichos diagramas y el conocimiento interno, con el fin de entender algún tipo de información ya existente, o “crear” algún nuevo conocimiento.

---

## CAPITULO III

### 3.1. Formación de hipótesis por RBM

En este capítulo, nos detendremos a considerar la denominada *abducción basada en modelos*, y luego los planteamientos anti-representacionistas que supuestamente completarían una mejor teoría del surgimiento de hipótesis, alternativa a la meramente sentencial. Como veremos, los planteamientos anti-representacionistas de formación de hipótesis tienden a asumir que la manipulación de estructuras materiales (internas o externas) conforman un sistema donde dichas estructuras cumplen roles simuladores susceptible de perfeccionamiento sobre la base de una interacción productiva (o generadora de hipótesis). En esta interacción, se asume la (re)combinación *creativa*<sup>45</sup> de modelos híbridos que no tendrían asiento exclusivo en un ámbito interno o externo, sino más bien en un continuo de re-presentaciones materiales simuladoras de un fenómeno que también sería parte del mismo ámbito metafísico asumido por tal continuo re-presentacional (i.e. el de la experiencia, o el del *aquí y el ahora*).

Finalmente, y en un sentido más amplio, abordaremos algunos aspectos y problemas centrales relacionados a un posible RBM en tanto sistema inferencial distribuido, ponderando críticamente la noción de *delegación cognitiva* como fundamentación del carácter epistémico autónomo de las denominadas representaciones artefactuales externas.

#### 3.1.1. Hacia la propuesta de los “sistemas buscadores de oportunidades”

---

Asumiendo que el tratamiento de base proposicional descrito en la sección *d* del apartado 1.11. es insatisfactorio para dar cuenta de procesos abductivos, lo que autores como Thagard proponen es una alternativa asociada al reconocimiento de procesos inferenciales determinados, de alguna manera, por un tipo no lingüiforme de representación. Específicamente, dicho reconocimiento pasa por aceptar un tipo de estructuración modélica cuyas bases representacionales estarían informadas por una amplia gama de evidencia experimental. Para aclarar este punto, veremos algunos aspectos de una posible teoría del descubrimiento alternativa, donde se propone combinar diversas contribuciones teórico-experimentales.

Según Magnani (2007a), el RBM puede fundarse en la visión Peirciana de que todo pensamiento se realiza en base a signos, y de que toda inferencia es una forma de actividad semiótica. Peirce (1931-1958, CP 5.283, citado en Magnani, 1999a) sostenía que los signos pueden incluir “sentimiento, imagen, concepción y otras representaciones”. Adicionalmente, la *síntesis* Kantiana también es para Magnani una forma de inferencia, en

---

<sup>45</sup> Esta interacción creativa no asume que se esté hablando de un proceso creativo original. Imagínese el caso en que el cálculo infinitesimal haya pasado por razonamientos modélicos no idénticos para llegar al mismo descubrimiento, pero a través de procesos de interacción modélicos independientes entre sí. Aquí, habría que hablar de un RBM creativo productivo, aunque lo que se produce (i.e. las hipótesis) ocurra en más de una ocasión.

la medida que apunta a la unificación de todo el material proporcionado por los sentidos. Esta vinculación surge no a partir de un entendimiento equivalente en lo que Peirce y Kant consideran como razonamiento sintético<sup>46</sup>, sino al hecho de que Kant hace referencia a *reglas y procedimientos* cuando describe la manera en la que el intelecto unifica y constituye un fenómeno a través de la imaginación. Según Magnani, se sabe que las reglas y los procedimientos representan características centrales en la concepción moderna de inferencia, algo que podría ratificarse en la tendencia que las personas tienen a resolver problemas geométricos proporcionando pruebas por medio de diagramas e ilustraciones<sup>47</sup>. Otro antecedente aludido por Magnani es que Peirce consideraba el problema de *cómo es posible el razonamiento sintético* como una cuestión central en la filosofía, así como también el hecho singular de que la mente presenta una tendencia a unificar aspectos exhibidos por los fenómenos<sup>48</sup>. Con estos antecedentes, Magnani sostiene que la cognición como actividad semiótica podría subsumir posibles procesos de inferencia no tradicionales dependientes de la experiencia sensorial.

No obstante lo anterior, el RBM adquiere su relevancia creativa cuando se considera en el contexto de la abducción. La abducción basada en modelos está constituida por una variedad de tipos. Aquí, es necesario aclarar que se tendrá en cuenta aquella noción de modelo en tanto mediador epistémico, discutido en el capítulo anterior. Recordemos que en tal caso, los modelos son considerados epistémicamente productivos sobre la base del rol epistémico que juega su propia materialidad. A continuación, caracterizaremos brevemente algunos tipos de procesos abductivos - agregando un breve comentario, cita o ejemplo relevante a cada uno -, donde se asume una hibridez modélica entre representaciones de algún conocimiento ya almacenado internamente, y la materialidad de algunas estructuras externas<sup>49</sup>. Dicho a la manera de Magnani, la abducción sería una actividad semiótico-cognitiva interpretable como una interacción productiva continua entre símbolos semióticos materiales, donde el ser humano es entendido como un signo externo más, y donde la productividad se manifiesta como nuevos interpretantes semióticos híbridos (e.g. modelos físicos con características miméticas). La hibridez de estos interpretantes es el resultado de una construcción semiótica determinado por alguna modalidad o multi-modalidad sensorial.

**Abducción visual:** ocurre cuando las hipótesis se derivan instantáneamente de una serie de experiencias similares previamente almacenadas. Basándose en la caracterización que Peirce hace de la percepción como un “procedimiento descontrolado de producción de conocimiento”, Magnani sostiene que la percepción activa es un vehículo para la recuperación [*retrieval*] instantánea de conocimiento, previamente estructurado en nuestras mentes a través de procesos inferenciales. Por medio de la percepción, las construcciones de conocimiento se organizan de manera tan instantánea que se tornan habituales y difusas y no requieren mayor evaluación [*testing*]. La expresión relevante de Pierce sería la siguiente: “[...] una inferencia interesante, simple, y completamente aceptada tiende a remover cualquier reconocimiento de las premisas complejas e irrelevantes de las que ésta deriva” (CP 7.37). Un ejemplo de este tipo de abducción sea halla en el trabajo de Shelley

<sup>46</sup> “By *synthesis*, in its most general sense, I understand the act of putting different representations together, and of grasping what is manifold in them in one knowledge”—Kant, [A77=B103], citado en Magnani, 2007.

<sup>47</sup> Magnani aclara que el atributo de creatividad de la inferencia en este caso no quiere decir que los procedimientos de abducción por RBM se estén realizando por primera vez, o de que sea un caso de creación original en el desarrollo de algún conocimiento.

<sup>48</sup> “the function of conception is to reduce the manifolds of sensuous impressions to unity” (CP 1.545)

<sup>49</sup> Esto es consistente con lo que Magnani considera un “interplay productivo” entre tales representaciones.

(1996). Hay ejemplos más triviales con respecto al diseño de páginas Web y el rol de ciertas convenciones de diseño informadas por estudios acerca del rol la percepción activa y el funcionamiento relacionado a ella de la denominada abducción visual. Uno de estos ejemplos puede ser el uso de características de formato especial que ayudarían a capturar la atención del agente humano a través del uso de caracteres o signos en negrita (el signo de interrogación “?” en la figura 2). Otro ejemplo, sería el cambio de color de los *vínculos visitados* cuando se navega en los resultados que arrojan los motores de búsqueda como Google®Search, lo que gatillaría una estrategia de memoria indexical del “historial de la interacción” entre usuario e interfase, para ayudar a inferir fácilmente algún sitio Web que ya ha sido visitado.

**Abducción táctil:** “Un hombre puede distinguir diferentes texturas de un tejido por la sensación: pero no inmediatamente, porque requiere mover los dedos a lo largo del tejido, lo que muestra que está obligado a comparar sensaciones de un instante con aquellas de otro instante” (Peirce, 1931-1958, CP 5.221, citado en Magnani 2006a y b). Claramente, esta cita seleccionada por Magnani hace alusión a la *reacción* que Peirce sitúa en la segunda categoría de su Faneroscopia, i.e. a la “secundidad”.

**Abducción auditiva:** la percepción de un tono surge a partir de una actividad mental sólo después de haber notado la rapidez de las vibraciones de las ondas sonoras, pero la posibilidad de individuar un tono ocurre sólo después de haber escuchado varios de los impulsos sonoros y después de haber juzgado su frecuencia. En consecuencia, la sensación de la nota clave de una melodía [*pitch*] es posible por las experiencias previas y cogniciones almacenadas en la memoria, de modo que una mera oscilación en el aire no podría producir un tono.

La visión que sustenta los tipos de abducción mencionados es que todo conocimiento es *inferencial*, y el acto de inferir no es instantáneo, sino que sucede después de un proceso que requiere una actividad de comparación que involucra **muchos tipos de modelos** en un lapso de tiempo. En consecuencia, dado el supuesto de que todas las sensaciones y percepciones participan en la naturaleza de una hipótesis unificadora, la abducción basada en modelos pareciera consistente con el carácter intuitivo que Peirce le atribuye a la creatividad y, a su vez, con la idea de que la creatividad no posee un carácter irracional. Pero eso no es todo, porque uno también podría impugnar una falta de economía de parte de esta noción de abducción en posibles procesos de razonamiento abductivo modélicos. Considérese, nuevamente, el caso de la figura 2. La tarea propuesta ahí bien podría resolverse a través de la manipulación literal de las figuras, hechas con fósforos, alfileres, o lo que sea. Pero la correspondiente “manipulación epistémica” requiere de conocimiento adicional que no podría estar disponible solamente en formato modélico, aunque fuese de carácter interno. De lo contrario, habría que asumir procesos de manipulación internos y externos, muchos de los cuales debieran ser redundantes para permitir “comparar” posibles soluciones o vías de solución. De hecho, las instrucciones de la figura 1 (aún si sólo consideramos el símbolo “?”), suponiendo la manipulación literal de los mismos diagramas dibujados, podría ilustrar el tipo de combinación representacional requerida para manipular modelos internamente. Es cierto que esta manera de fundamentar la impugnación en cuestión es meramente especulativa. Sin embargo, es pertinente tener en cuenta si acaso tiene sentido asumir componentes adicionales (internos o externos) que pueden redundar en problemas ya conocidos (cuellos de botella informativos, etc.). Estas cuestiones son relevantes para aclarar cómo el RBM puede llegar a constituirse en un método de inferencia por derecho propio, tal como veremos en la sección 3.1.2.

Por ahora, digamos que el razonamiento modélico asumido anteriormente corresponde a un tipo de actividad hipotética que es inferencial pero no verbal, basada en modelos de tipo sensorial: visuales, auditivos, etc. Conviene aclarar que cuando se habla de modelos externos sensoriales, se hace referencia a “signos” materiales externos que son principalmente *capturados* por alguna modalidad sensorial particular del agente cognitivo intencional. Del mismo modo, cuando se habla de modelos internos sensoriales, se hace referencia a posibles representaciones materiales directamente instanciadas por las anteriores. Las representaciones determinadas por alguna modalidad sensorial, así como multi-modales, son persistentemente defendidas como signos materiales, como una manera de desligarse de cualquier atisbo de representacionalismo tradicional. Sin embargo, esto implica reducir todo procesamiento interno a la materialidad de los denominados patrones de activación neuronal. Obviamente, dichos patrones no podrían dar cuenta, por sí mismos, de todos los procesos de la cognición superior que parecen reflejar una estructuración representacional composicional. Esta es la razón por la cual no se descarta del todo un tipo de razonamiento proposicional, sino que más bien se confina a algún tipo de procesos menos centrales del razonamiento creativo científico.

Aún así, el problema de las representaciones modales no está libre de cuestionamientos importantes, como los que plantea Pylyshyn (2002), de acuerdo a lo cual se puede distinguir (sensatamente) que una cosa es la existencia de imágenes internas (o representaciones modales, para el caso) y otra es la existencia de pensamientos acerca de “cómo algo podría lucir en formato de imágenes”. Otro problema, más de fondo, sucede cuando la semántica de los modelos modales se confunde con la ontología del formato de las representaciones que se defienden (Fodor, 2008). Esta confusión no es ingenua, sino que de inspiración anti-representacionista, y responde a lo que hemos aludido en la *Introducción* como el tipo de inversión pragmatista de nociones metodológicas, propias de un efecto nocivo de la interacción entre filosofía de la ciencia y ciencia cognitiva (ver 4.1.). En este caso, lo que se defiende es la “interpretación” del valor epistémico del formato en tanto representación interpretada, asumiéndose que el pensamiento es la internalización de la acción corporalizada (y no la acción como externalización del pensamiento, que es como lo entendería un filósofo racionalista). Para los defensores del RBM, esas cuestiones son relativamente soslayadas, confinándose a lo que se informe empíricamente desde los pertinentes desarrollos experimentales futuros, siempre en la línea de las posturas metodológicas alternativas de la cognición.

A partir de lo planteado anteriormente, se puede presentar la idea de un RBM dependiente de la construcción y manipulación de variados tipos de representaciones, no necesariamente sentenciales o formales. Esto incluye la construcción y manipulación de representaciones visuales, experimentos de pensamientos, razonamiento analógico, etc. No obstante lo anterior, una manera alternativa de entender esto es que, como ciertos procesos inferenciales son requisitos para la Abducción Basada en Modelos, quizás los mismos tipos de RBM que son relevantes para la explicación del razonamiento creativo en ciencia, pueden depender de procesos inferenciales básicos de tipo sentencial, en el entendido de que estos procesos primarios debieran ser considerados como la clase de *procesos razonados* (i.e. procesos que forman parte de un razonamiento inferencial por derecho propio). Esto último, a diferencia de lo que sostienen los defensores de un RBM, sería consistente con la posibilidad de que cualquier RBM pueda depender originalmente de representaciones estructuradas más tradicionales, que, a diferencia de un RBM por sí sólo (si es que existe), es susceptible de prescribir normativamente cómo se debiera proceder al enfrentarse a un proceso inferencial determinado. Sin embargo, dicha interpretación no

parece lo suficientemente persuasiva si se acepta una base cognitiva continua entre la abducción humana y animal.

Dada la continuidad sentencial-modélica señalada, lo que se requiere, entonces, es determinar las “estrategias” que permitirían la interacción entre un nivel proposicional-verbal a uno basado en modelos. Por un lado, esto se aborda con alusión al denominado “concimiento implícito”, y por otro, a partir de cierta sensibilidad a las denominadas “inconsistencias” o “anomalías empíricas” del razonamiento científico (Magnani, 2007a). Siguiendo la investigación epistemológica de Polanyi (1996), Magnani destaca la idea de que gran parte del conocimiento no es explícito, sino que tácito. Es decir, sabemos más de lo que podemos decir, y no podemos saber nada sin apoyarnos o confiar en aquellas cosas que no estamos en condiciones de decir o reportar. Esto no debiera ser inconsistente con un tipo proposicional de conocimiento tácito, si no fuera porque, además, Polanyi sostiene que este conocimiento tácito es un conocimiento de orden práctico requerido para llevar a cabo alguna tarea. Esto implica la necesidad de poner atención en heurísticas particulares que recurren a *formas extra-teóricas de pensamiento*. En este sentido, algunas características comunes de las plantillas [*templates*] de abducción manipulativa se relacionan con:

1. 1. sensibilidad hacia aspectos de los fenómenos que pueden ser curiosos o anómalos; se espera que las manipulaciones permitan introducir inconsistencias potenciales en el conocimiento recibido, y así abrir nuevas oportunidades de razonamiento;
2. 2. Sensibilidad preliminar al carácter dinámico de los fenómenos, y no a las entidades o sus propiedades. Un objetivo común de las manipulaciones es reordenar de manera práctica la secuencia dinámica de eventos en alguna secuencia espacial estática que promueva una visión panorámica [*bird's eye view*] fructífera para otros resultados;
3. 3. Redireccionamiento [*referral*] hacia las manipulaciones experimentales que explotan aparatos artificiales con el propósito de liberar nuevas fuentes de información repetibles y estables acerca de constreñimientos y conocimiento tácito (e.g. el arreglo artefactual de una torre alfileres le habría mostrado a Davy, uno de los fundadores de la electroquímica, que la magnetización estaba relacionada a la orientación, y que no requería contacto físico);
4. 4. Una variedad de formas de acción epistémica, tales como mirar desde diferentes perspectivas, corroborar cierta información disponible, comparar eventos consecutivos, elección/descarte/imaginación de otras manipulaciones, reordenamiento y cambio de relaciones en el mundo por medio de una evaluación implícita de la utilidad de algún nuevo ordenamiento. En resumen, toda la actividad manipulativa está orientada a la creación de mediadores epistémicos.

Lo anterior, por cierto, requiere una teoría de rol de las anomalías en la formación de hipótesis. Dicha teoría, según los defensores del RBM, no está disponible aún, principalmente porque la explicación de formación hipótesis ha estado confinada a un nivel teórico y formal, donde se habrían descuidado importantes aspectos relacionados a la complejidad de la agencia humana. Esta complejidad debe ser entendida sobre la base experimental que actualmente informa un nuevo entendimiento del agente cognitivo. Según Gooding et al. (1999), el tipo de situación científica que requiere ser apropiadamente modelada debiera dar cuenta de un proceso de descubrimiento donde algo inicialmente creído por alguien, o que forma parte de su experiencia subjetiva, se traslada al ámbito del dominio público o los hechos objetivos. Las experiencias iniciales están constreñidas por los métodos y preconcepciones del mundo que adopte cada observador, razón por la

cual son subjetivas y diferentes en cada individuo. Asimismo, el lenguaje descriptivo de un campo científico emergente es comparable con el estado inicial de un lenguaje natural que aún no alcanza cierta estandarización, por lo que también es esperable una alta variedad descriptiva. Sin embargo, el tránsito desde el dominio personal al público requiere de un lenguaje descriptivo claro y estable, lo que requiere operacionalizar los descriptores en términos de los procedimientos que hacen posible observarlos y comunicarlos:

***“En otras palabras, la formulación, la clarificación, la comunicación, y la validación de hipótesis están inextricablemente vinculadas a las actividades en las que los observadores están involucrados”. (p. 106)***

Magnani (1999a y b) ilustra lo anterior haciendo referencia a los procedimientos que debieron permitir el surgimiento de la geometría no Euclideana – también referida como la revolución no Euclideana. El cambio desde una geometría a otra representa un caso significativo de cambio de sistemas conceptuales, donde un tipo de “abducción visual” cumpliría el rol de una estrategia para resolución de anomalía relacionada a una forma de “pensamiento visual productivo y explicativo”. El problema conceptual que presentaba la geometría Euclideana tenía relación con el quinto postulado de la misma. Dicho postulado era considerado poco evidente, lo que en sí mismo constituía una anomalía, despertando sospechas acerca de la confiabilidad de la teoría de las paralelas en su totalidad, por cuanto dicha teoría consistía en teoremas que solamente se derivaban del quinto postulado. Durante dos milenios, los intentos por resolver la anomalía dieron lugar a demostraciones más o menos falaces del quinto postulado, llegando incluso a proponer demostraciones de este postulado a partir de los otros restantes. El descubrimiento de la geometría no Euclideana, tal como lo analiza Magnani, pasa por considerar las estrategias de manipulación de símbolos que Lobachevsky utilizó para resolver la anomalía, lo que le habría permitido reconstruir los principios, derivar nuevas pruebas y proporcionar un nuevo sistema [*apparatus*] matemático. En este caso, la conclusión de que las hipótesis creadas por el matemático Ruso se basaban en imágenes, y por lo tanto fueron realizadas por una abducción visual relevante a la generación de hipótesis y al descubrimiento científico, se sustenta en el estudio y análisis de los procedimientos epistémicos que la literatura reporta sobre este descubrimiento (ver Magnani 1999b, para un análisis detallado).

En términos más clarificatorios, la manera figurada que se utiliza para caracterizar la inferencia abductiva es considerarla como las respuestas a las preguntas, explícitas o implícitas, que el investigador en tanto interrogador le hace a la naturaleza, cuando éste se haya frente una fuente definida de respuestas. Magnani (1999a) propone enfatizar el carácter interrogador del razonamiento abductivo, para entender mejor el hecho de que, si la abducción consiste en la construcción [*making of*] de un conjunto de respuestas posibles, entonces la elección que se haga de las posibles preguntas también es decisiva. Lamentablemente, las ciencias cognitivas no contarían aún con una buena teoría de cómo producir preguntas fructíferas. Para este autor, dicha teoría debiera ser capaz de involucrar aspectos emocionales, éticos y pragmáticos de la cognición.

Comenzar a delinear dicha teoría implicaría atender debidamente al rol de las *inconsistencias* en el razonamiento científico. Las diversas maneras en que un RBM parece participar en la generación de hipótesis explicativas de algún fenómeno (o problema conceptual)<sup>50</sup> que produce una interrogante se vincularían a la actividad misma de encontrar dicho fenómeno (o problema conceptual). Estas anomalías empíricas no pueden ser apropiadamente subsumidas por un tratamiento logicista de la abducción, por cuanto dicho

<sup>50</sup> Los casos son vastos, y gran parte de ellos han sido reportados en Magnani et al. (eds.) (1999, 2002, 2007, entre otros), y en un tratamiento más in extenso se recapitula en Nersessian (2008).

tratamiento sólo puede resolver anomalías a nivel teórico. Lo apropiado, según autores como Magnani, Gooding, o Nersessian, es atender el nivel de los mecanismos cognitivos, donde tanto las anomalías empíricas como los problemas conceptuales pueden ser mejor caracterizados en su rol de “gatillos para generación de hipótesis”. Como se ha dicho, esto pasa por aceptar que la generación de una hipótesis imaginística [*imagery*] pueda ser considerada un tipo de abducción inferencial creativa, lo que prometería describir cómo las representaciones modélicas pueden conducir a descubrimientos científicos.

A partir de lo anterior, Magnani sostiene que la noción de abducción puede extenderse para describir, de manera fructífera, todas aquellas inferencias hipotéticas realizadas a través de acciones que consisten en “astutas manipulaciones”, tanto para detectar nuevas oportunidades de acción [*affordances*] como para crear objetos externos que ofrezcan nuevas oportunidades de acción e inferencia. Esta es, precisamente, la idea que regula su concepción de la cognición humana en tanto sistema buscador de oportunidades [*humans as chance seekers*], tal como intentaré caracterizarla a continuación.

Asumiendo la tesis recurrente en los enfoques cognitivos alternativos de la cognición de que los humanos tienden a distribuir funciones cognitivas en el ambiente con el efecto de diluir sus límites adaptativamente (ver 1.1.1.), Magnani (2008) trata de dar cuenta del diseño de modelos, representaciones y estructuras mediadoras sobre la base de la noción de *nichos cognitivos*<sup>51</sup>. Dicha noción sería el resultado de una acción modificadora del ambiente intencionada por parte de este tipo de agentes cognitivos.

El carácter intencionado de la creación de nichos cognitivos estaría fundado en la necesidad de crear *affordances artificiales*, las que constituyen una concepción elaborada a partir de la noción de base biológica/ambiental propuesta originalmente por Gibson (1979) para ilustrar el carácter híbrido que tendría la percepción visual. Estos apoyos artificiales son el resultado de un entendimiento del agente cognitivo según el cual éste actúa y modifica activamente el ambiente en el sentido de un “ingeniero ecológico” (ver caracterización en 1.1.3.). Estos procedimientos activos de construcción de andamios o apoyos de la cognición, según Magnani, serían indicios interesantes de una capacidad animal para detectar, crear mayores “oportunidades de acción”. De manera especialmente interesante, la abducción podría describirse en base a todas aquellas inferencias hipotéticas que los humanos realizan, precisamente, a través de acciones y manipulaciones astutas, cuya finalidad no es otra sino la de fabricarse objetos externos que proporcionen nuevas *affordances* y potenciales nichos cognitivos.

Tal como se expuso en el capítulo II, las representaciones externas manipuladas para constituirse en el tipo de apoyos en cuestión suelen tener un carácter simulador, emulador, o mimético. Estos elementos son representacionales no en el sentido de substituir (o traducir) alguna parte del mundo por algún tipo de estructura mediadora interna, sino que lo hacen en el sentido de una *re-presentación* simuladora de dicha parte del mundo<sup>52</sup>. Esto equivale a decir que la representación modélica externa es en sí misma una instancia genuina de alguna parte del proceso inferencial, en el caso que dicha representación proporcione información o conocimiento adicional. De esto se desprenden dos constreñimientos importantes: (a) la vinculación entre conocimiento tácito y representación externa que resulta ubicua a toda idea de representación modélica; y (b) la motivación pragmatista que licencia dicho vínculo. Magnani (2008) caracteriza esto de la siguiente manera:

<sup>51</sup> Concepto que se atribuye originalmente a Tooby et al. (1987), y posteriormente reutilizado por Pinker (1997).

<sup>52</sup> Ver notas 37 y 38.

**“Por ‘recursos externos’ quiero decir todo aquello que no está dentro del cerebro humano, y que podría ser de ayuda en el proceso de decidir, pensar acerca de algo, o utilizar alguna cosa” (p. 920)**

En un sentido real, Magnani asimila estos procedimientos epistémicos a la idea de conjetura, dado el rol que juegan en la tarea usual de toma de decisiones confiando en la “información incompleta” de que se dispone. Considerando (a) y (b), la conjetura en cuestión debe ser entendida como un “diálogo” o negociación epistémica continua entre *agente cognitivo intencional* y *artefacto cognitivo externo*, toda vez que dicha negociación permite transformar información implícita en explícita. Tal transformación ocurriría como resultado del *interplay* entre las estructuras representacionales de conocimiento internas y externas. Estos procesos completarían la información requerida para “gatillar” procesos de razonamiento inferencial productivos. De hecho, para este autor, poseer información incompleta significa que:

1. las deliberaciones y decisiones que normalmente toman las personas nunca son la mejor respuesta posible, pero por lo menos son satisfactorias.
2. las conclusiones son siempre retractables [withdrawable] o revisables.
3. gran parte de la actividad cognitiva de las personas corresponde a la elaboración de conjeturas o hipótesis con el propósito de obtener información más adecuada.

Las conjeturas pueden ser fruto de una selección abductiva de un conjunto de hipótesis previamente almacenadas o de la creación de nuevas hipótesis, como en el descubrimiento científico. En el caso de las *affordances*, estas también pueden ser seleccionadas de manera “diagnóstica”, vale decir, en consideración a las pistas o claves que uno puede obtener de un objeto, como, por ejemplo, una silla que presenta, en tanto “oportunidad de acción”, la posibilidad de sentarse en ella (e.g. por la forma, rigidez, robustez, superficie plana, etc.).

La abducción es caracterizada como el proceso de inferir ciertos hechos y/o leyes e hipótesis que hace plausible algunas sentencias, que explica o descubre algún fenómeno u observación (eventualmente nuevos). Sobre la noción semiótica de signo señalada anteriormente en esta misma sección, la concepción de abducción es extendida más allá de la dimensión sentencial. De manera particular, el caso de la abducción visual presenta una fuente ilustradora del rol que jugarían las *affordances* del entorno, tanto las ambientales como las perceptuales. Así, asumiendo ciertos supuestos de la cognición corporalizada según los cuales la cognición es dependiente de los mecanismos de procesamiento sensorial y control motor, se asume que lo que uno ve es lo que nuestro aparato visual puede “explicar”. Lo que se sigue de esto es que la “consulta activa” que el organismo realiza hacia su entorno configura un proceso propio de lo que se ha dado en llamar la cognición *on-line*. De aquí que Magnani (2008) caracterice a las *affordances* participantes de un proceso abductivo como sujetas a estabilización, en el entendido que esta estabilización se constituye, a través de la experiencia, por medio de “*affordances* altamente exitosas”. La liberalización que se hace a partir de los planteamientos de Peirce (i.e. que la abducción depende de un continuo de signos materiales externos e internos que sólo difieren entre sí por grados de re/presentación) permite considerar esta estabilización como un tipo de hábito, en tanto regla de acción. Sin embargo, habría que señalar que estos hábitos están sujetos a revisión en la medida que las *affordances* también lo están. Así como el entorno es susceptible de continua modificación, también las capacidades perceptivas lo son cuando se enriquecen a través de las habilidades cognitivas de orden superior, propias de una cognición de tipo *off-line*.

La conclusión de lo expuesto en esta sección puede entenderse de la siguiente manera. Asumiendo que el tratamiento de la abducción sobre la base de representaciones de estructuración supuestamente proposicional es insatisfactorio, los defensores de un RBM proponen considerar dos aspectos de un razonamiento creativo distintos al tradicional. Estos aspectos se han denominado, respectivamente, abducción basada en modelos y abducción manipulativa, y tendrían un rol central en la formación de hipótesis. De lo anterior se sigue que los tratamientos proposicionales o bien no tienen la centralidad asumida originalmente a partir de las propuestas de Pierce, o bien podrían ser complementarios (de una manera poco interesante para los investigadores en RBM) a los aspectos alternativos del razonamiento creativo señalado.

La abducción basada en modelo es parte de un enfoque no externalista de RBM y consiste en un proceso principalmente centrado en el agente intencional, caracterizado por la comparación continua de modelos multimodales, es decir, modelos determinados por una combinación de modalidades sensoriales. La abducción manipulativa, por su parte, reconoce los aportes del RBM, específicamente como éste es caracterizado por Nersessian (2008 y otros), pero incorpora un aspecto supuestamente soslayado por la abducción basada en modelos, a saber, aquel donde lo central del proceso creativo se explicaría en términos de una explotación crucial del ambiente. Específicamente, esto último implica aceptar la idea de un proceso inferencial dependiente de un proceso de interacción entre representaciones materiales internas y externas. Asumiendo la posibilidad de este proceso de interacción, autores como Magnani han defendido la idea de que cierta información implícita podría estar codificada en la manipulación literal de objetos externos, los que actuarían como mediadores epistémicos entre el agente intencional y el surgimiento de potenciales interpretantes semióticos que, eventualmente, concederían accesibilidad a tal información para procesos inferenciales de generación y selección de hipótesis más tradicionales. En consecuencia, con esto se estaría dando cuenta de un tránsito de un aspecto internalista del razonamiento creativo hacia lo que se ha denominado un proceso de pensamiento y descubrimiento a través de la acción.

Finalmente, habría que resaltar una par de observaciones críticas que mantendremos por lo que queda de este trabajo de investigación. En primer lugar, no es del todo claro por qué se tendría que asumir un compromiso ontológico con algún tipo de formato modélico de representaciones internas a partir de la adscripción de cierta semánticidad de lo que se ha denominado representaciones modales. Especialmente si dicha semánticidad se puede interpretar como algo dependiente del hecho de que las personas también pueden pensar en *cómo algo podría lucir en tales o cuales formatos*. En este último caso, uno incluso podría reivindicar una distinción entre una intencionalidad original o genuina y una derivada, toda vez que la semánticidad de alguna posible representación modélica dependa de la semánticidad de otro tipo de representación más tradicional. En segundo lugar, se mantiene la permanente oscuridad con respecto a una posible *complementariedad* entre aspectos exclusivamente internos del razonamiento creativo (e.g. modélico y proposicional) y aspectos radicalmente externalistas del RBM. Mientras que las representaciones internas están dotadas de un carácter más o menos representacionalista (con matices que pueden hacer crisis al incorporar las representaciones corporalizadas), las externas tienden a abandonar un compromiso representacionalista, a favor de lo que se ha caracterizado como una re-presentación epistémica artificial. En este último caso, incluso se habla de la modificación del ambiente externo con la finalidad de crear un nicho cognitivo de *affordances* como la base de la interacción entre representaciones externas e internas. Nos parece difícil concebir una interacción literal *ad hoc* entre concepciones representacionalistas y anti-representacionalistas de los vehículos de la

---

cognición, especialmente si se evade el problema de especificar claramente cómo tiene lugar la complementariedad en cuestión.

Diremos que esta última dificultad tiene que ver con dos problemas relacionados. El primero de estos problemas es que la complementariedad no cuenta con una lógica sobre la cual dependan prescripciones normativas que responda, por ejemplo, a “cómo se debiera proceder al enfrentarse a un proceso inferencial de tales características”. El segundo problema tiene que ver con la inversión pragmatista que hemos mencionado con anterioridad. Esta inversión consiste en abandonar la intuición de que la acción del agente cognitivo corresponde a la externalización (en tanto manifestación conductual) del pensamiento, y, por el contrario, aceptar que el pensamiento (en tanto producto de procesos de corporalización y distribución cognitivas) se deben entender como la internalización de la acción de un agente.

### 3.1.2. RBM como un método de inferencia y delegación cognitiva.

---

En esta sección, elaboraremos ciertos aspectos y problemas centrales relacionados a un posible RBM en tanto sistema inferencial distribuido. Un elemento controversial, aunque crucial si se quisiera aceptar el rol epistémico de objetos y procedimientos extra-mentales, es la noción de *delegación cognitiva*. Dado que dicha noción es subyacente a la concepción de modelos en tanto agentes epistémicos autónomos (o artefactuales) - propia de una posible teoría unificadora del RBM -, consideramos importante evaluar el alcance y los desafíos que ésta pueda tener en el contexto de un método inferencial externalista y alternativo a un razonamiento proposicional.

Según Nersessian (1999), el término (y la noción) de “razonamiento basado en modelos” se usa para indicar la construcción y manipulación de variados tipos de representaciones, las que no son sólo sentenciales y formales, sino que principalmente modélicas y relacionadas a mediadores externos. Para entender cómo se llega a esta visión, es necesario entender lo que esta autora propone como el “método histórico-cognitivo”, el cual surge a partir de la necesidad de explicar el cambio conceptual en ciencia.

Nersessian (2008) ha planteado que un estudio apropiado de las maneras en que nuevos conceptos son creados en ciencia exige combinar los hallazgos, métodos e interpretaciones provenientes tanto de la historia de las ciencias como de las ciencias cognitivas. En este contexto, la dimensión histórica contribuye no sólo con el registro de las prácticas investigativas de los científicos, sino que también de la manera en que éstas se desarrollan en el tiempo. Dicho desarrollo debe ser examinado de una manera que coloca a los individuos creativos dentro de las situaciones de problemas de sus comunidades locales y también dentro de su contexto cultural más amplio. De este modo, quienes practican el denominado método histórico-cognitivo buscan interpretar las prácticas de investigación a la luz del entendimiento científico que se tiene de las capacidades cognitivas que pueden subyacer, facilitar y constreñir dichas prácticas. Desde lo que se propone como una perspectiva naturalista, esto es lo que Nersessian afirma que tiene lugar en los procedimientos de *bootstrapping* de la cognición ordinaria y, de manera paradigmática y significativa, también en las prácticas científicas más creativas.

La concepción de la cognición como *bootstrapping* hace referencia a la manera en que los agentes cognitivos, dado un contexto de problemas y procedimientos o métodos disponibles, adopta la solución más adecuada, sin que aquello constituya necesariamente la mejor de las soluciones posibles. Específicamente, Nersessian (2008) caracteriza el *bootstrapping* como “ciclos de construcción, simulación, evaluación, y adaptación de

modelos que sirven como interpretaciones temporales del fenómeno a explicar” (p. 184). Una diversidad de elementos puede contar como fuente de datos empíricos acerca de las prácticas científicas: registros históricos, cuadernos, bitácoras, correspondencia, bosquejos, publicaciones, artefactos (tanto instrumentos como modelos físicos). Lo que se tiene, entonces, es un examen de las prácticas científicas informadas a partir de la mejor comprensión científica vigente del sujeto cognoscente, tal como ésta tiene lugar en el ámbito de las ciencias cognitivas. En este último ámbito es precisamente donde se pueden encontrar interpretaciones acerca de los factores causales que parecen cumplir el rol regulador apropiado para fundamentar el método de Nersessian. Uno de estas interpretaciones tiene que ver con el tipo de estructuración conceptual que puede subyacer a procesos de cambios conceptuales dinámicos y flexibles, como los que a esta autora le interesa defender. Específicamente, dicha defensa tiene relación con sustentar la idea básica de que la construcción, manipulación, evaluación y adaptación de modelos son los medios primarios a través de los cuales los científicos crean nuevas representaciones conceptuales. De aquí que la innovación conceptual dependa, a la luz del método histórico-cognitivo, del cambio y creación de nuevos modelos.

Una explicación general del cambio conceptual comentado hasta este punto se puede resumir de la siguiente manera. Un concepto, como una forma de representación, puede especificar un conjunto determinado de constreñimientos para el surgimiento de una clase de modelos. En este sentido, la formación de conceptos, así como el cambio conceptual, es un proceso que implica la generación de nuevos constreñimientos, y la modificación de otros antiguos, siempre relativos a la construcción de modelos. El rol que juega un RBM en este proceso tiene que ser entendido como un tipo efectivo de recursos que hacen evidente los constreñimientos de los sistemas representacionales existentes, dirigiendo la atención hacia ellos y permitiendo con esto la discriminación selectiva de ciertos constreñimientos. Una parte significativa de este proceso tiene que ver con la facilitación que el RBM permite de la integración de múltiples dominios. En este sentido, el RBM cumple el rol de andamio para un *bootstrapping* orientado a representaciones satisfactorias.

La noción de modelo que plantea Nersessian, inspirada en la noción presentada inicialmente por Craik<sup>53</sup> (1943), puede concebirse como la representación (interna o externa) de un sistema con partes que interactúan y con representaciones de dichas interacciones. A partir de esta caracterización, razonar sobre la base de modelos incluye la manipulación (interna o externa) de variados tipos de modelos: conceptuales, físicos, matemáticos y computacionales, por nombrar los más conocidos. La manipulación interna puede tomar la forma paradigmática de la simulación mental de estados futuros de aquello que modela. La manipulación externa, por su parte, estaría fundada en una relación de acoplamiento entre las representaciones internas y externas, siempre que estas últimas jueguen un rol esencial en la simulación.

Resulta interesante la asociación que Nersessian hace con los planteamientos de Barsalou (1999). Según este autor, patrones de activación (de manera *bottom-up*) en áreas sensorio-motora del cerebro pueden ser parcialmente reactivadas (de manera *top-down*) para formar símbolos perceptuales. Al mismo tiempo, y a través de la denominada “atención selectiva”, representaciones esquemáticas de los componentes perceptuales serían extraídos de la experiencia y almacenados en la memoria. Según sostiene Barsalou, es en este punto que la memoria implementa un *simulador* versátil capaz de generar un sinnúmero de simulaciones de los componentes perceptuales señalados, lo que

---

<sup>53</sup> Craik sostenía la idea de que un modelo era un constructo funcional, conductual y estructuralmente análogo a algún fenómeno del mundo real.

proporcionaría la base para diferentes aspectos de la cognición avanzada (incluyendo la categorización, las inferencias categoriales, productividad, proposiciones, conceptos abstractos, etc.). Lo que se intenta con esta teoría es la posibilidad de un sistema perceptual totalmente funcional que pueda constituir una teoría del conocimiento perceptual que evite suponer que las representaciones cognitivas son inherentemente no perceptuales (o amodales). Según esto cualquier representación modélica tendría que estar en última instancia ligada al ámbito sensoriomotor. Nersessian parece compartir esto en una importante medida.

Asumir lo anterior es comprometerse con ciertos supuestos del denominado enfoque corporalizado de la cognición (ver 1.1.1.). Si bien es cierto que el planteamiento de una estructura conceptual modal de base sensoriomotora que defiende Nersessian tiene una naturaleza principalmente visual, aceptarlo autoriza considerar aspectos sensoriomotores adicionales. De esta manera se entiende la alusión a la abducción multimodal comentada en la primera sección de este capítulo, cuestión que se puede elaborar de la siguiente manera.

Si es posible asumir que el formato de las representaciones modélicas internas tienen una base sensorio motora, entonces es posible asumir que las prácticas manipuladoras y simuladoras de dichas representaciones están significativamente constreñidas por lo que Koch y Crick (2001) ha ilustrado como el “inconsciente cognitivo” (o el “zombie” sensoriomotor que guía gran parte de nuestra vida diaria). Según estos autores, la ciencia ha proporcionado evidencia creíble a favor de la existencia de sistemas sensoriomotores en el cerebro de primates como nosotros que funcionan en ausencia de toda consciencia. En otras palabras, estos sistemas suelen desempeñar tareas rutinarias complejas de manera automática y sin la necesidad de un input consciente. La “acción adaptativa” a la que hace alusión en este caso corresponde al nivel de la cognición denominada *online*, y se funda en la distinción que los investigadores han planteado entre “visión para la percepción” y “visión para la acción”<sup>54</sup>. Aunque aún permanece abierta la controversia de si ambos tipos sistemas operan sobre la base de sistemas neuronales claramente distinguibles, dicha distinción metodológica es útil cuando se correlaciona con aquella otra distinción entre “acciones pragmáticas” y “acciones epistémicas” comentada en la sección 1.1.3.

Asumir la existencia de *programas sensoriomotores autónomos* (i.e. rutinas activadas sin input consciente) asociados, por un lado, a un tipo de actividad visual y, por otro, a ciertas rutinas de acción es lo que permite dar cierta plausibilidad a los planteamientos de Nersessian (2008), Gooding (1990) y Magnani (2001a, 2007a-b y 2008, entre otros) con respecto al reconocimiento de la relevancia de ciertas regularidades extra-rationales codificadas o “encarnadas” [*embodied*] en ciertas plantillas de conducta [*templates of behavior*]. Luego, al defender la posibilidad de “nuevas formas de inferir” hipótesis explicativas cuyo surgimiento no está regido por reglas lógico-proposicionales explícitas, lo que se asume es una “negociación epistémica” (ver “sistemas buscadores de oportunidades” en 3.1.1.) entre el agente solucionador de problemas y ciertos objetos externos (representaciones modélicas de variado tipo), regulada por procedimientos implícitos básicos como la acción epistémica realizada a través de programas sensoriomotores autónomos. Magnani (2006a) expresa lo anterior como sigue:

**“La abducción manipulativa se refiere a una conducta extra-teórica que se orienta a la creación de explicaciones comunicables de nuevas experiencias**

<sup>54</sup> Básicamente, esta distinción surge de los estudios que contrastan el desempeño en la realización de tareas determinadas por parte de sujetos experimentales sanos y enfermos (o en un mismo sujeto que sufre episodios de sonambulismo o epilepsia) y el posterior reporte retrospectivo de dichos sujetos.

### ***para integrarlas en sistemas de prácticas experimentales y lingüísticas (teóricas) previamente existentes” (pp. 194-195)***

Las conductas extra-teóricas aludidas en la cita no tienen un propósito pragmático (o dirigido a metas, si se quiere) en sí mismas. Investigadores como los recién mencionados sostienen que estos sistemas cumplen un “rol conjetural autónomo”, en el sentido de que cumplen la función de apoyos (vía *bootstrapping*) para algún tipo de manipulación provisional, a partir de la cual se facilitarían nuevas y productivas simulaciones internas. Lo que parece más radical en el planteamiento de estos investigadores es que tanto los patrones de acción epistémica como los modelos externos que pueden ser objeto de manipulaciones serían susceptibles de una suerte “delegación cognitiva”. Lo que pareciera licenciar dicha delegación de propiedades cognitivas es simplemente su rol autónomo (i.e. no dependiente de una *acción pragmática* sino más bien de una *acción epistémica*) en algún proceso creativo cualquiera de RBM. Dicho de otra manera, la participación de aquellos factores de regulación implícita determina críticamente nuevas oportunidades de inferencia y eventual surgimiento de hipótesis explicativas, razón por la cual se les atribuye un rol cognitivo delegado (i.e. no intencional). En este caso, la argumentación contrafactual es que, de no poder contar con tales manipulaciones y con el resultado de tales manipulaciones, sería imposible acceder a cierta información implícita únicamente susceptible de ser codificada en la acción epistémica y en los artefactos creados por ella.

Claramente, lo que se está asumiendo en esta manera de dar cuenta de factores reguladores implícitos de un RBM no es sólo una concepción corporalizada de la cognición, sino también una concepción situada y distribuida<sup>55</sup>. Tal como Nersessian (1999) ha planteado que las representaciones visuales externas suelen tomar la forma de diagramas susceptibles de manipulación, Magnani (2007a) ha planteado que esa es sólo una de las tantas formas de “externalización” semiótica (o descorporalización cognitiva) que caracteriza nuestra capacidad de redistribución del esfuerzo cognitivo y epistémico para acceder a objetos e información que no están inmediatamente representados o disponibles internamente. Junto con aclarar que toda la actividad de manipulación está orientada a construir variados *mediadores epistémicos* externos, Magnani admite haber desarrollado esta noción de mediadores a partir de la noción de “estructuras mediadoras” propuesta por Hutchins (1995, 1999) para referirse a las diversas herramientas externas que pueden construirse para apoyar cognitivamente la actividad de navegación tanto en contextos modernos como primitivos.

Ahora bien, las posibles regularidades que constreñirían el surgimiento de nuevas hipótesis explicativas en ciencia son equivalentes a las que se asumen en las concepciones de la cognición alternativas al marco clásico, y que suelen subyacer o apoyar a la misma investigación sobre RBM. El punto se puede aclarar mejor de la siguiente manera. Uno de los requisitos del naturalismo epistemológico que comparten los investigadores en RBM ya mencionados es que las teorías filosóficas del conocimiento científico necesitan estar informadas por la mejor comprensión científica disponible del sujeto cognoscente. Sin embargo, esa misma “mejor comprensión” suele estar fundada en mecanismos biológicos, cognitivos y sociales que aún son deficientes o incompletamente entendidos. ¿Cómo, por ejemplo, se podrían compatibilizar claramente los hallazgos de una visión

<sup>55</sup> Estas nociones fueron contextualizadas en 1.1.1. Resumidamente, una cognición corporalizada es aquella que supuestamente está constreñida y determinada por el cuerpo de los agentes; la situacionalidad de la cognición viene dada porque, además de la corporalidad, los contextos socio-situacionales concretos en que ocurre la actividad cognitiva también determinarían la conducta humana inteligente; la cognición distribuida sostiene que las funciones cognitivas son llevadas a cabo, no sólo por el agente humano intencional, sino que principalmente en conjunto con una multiplicidad de individuos y artefactos compartidos.

---

corporalizada, socialmente distribuida y bio-ambientalmente incrustada [*embedded*] de la cognición con la arquitectura conexionista? Este parece ser un requisito metodológico para darle plausibilidad a la extensión literal que se propone de los sistemas cognitivos solucionadores de problemas.

En la primera parte de este capítulo dijimos que un RBM hace referencia a la construcción y manipulación de un variado tipo de representaciones internas y externas. Una posibilidad importante es la idea de una extensión o delegación cognitiva no literal sino que derivada, en el sentido de que esta última no pierde de vista la participación claramente demarcada de un “solucionador de problemas”. Si bien es cierto, hemos visto que se suele convenir en la extensión extracraneal del sistema cognitivo, también hemos especificado que se suele aludir a las nociones de acción y negociación epistémicas. Dicha negociación epistémica - asumiéndose una interpretación correcta de nuestra parte - estaría constituida por intercambios causales distribuidos entre modelos internos, externos y acoplados. Al mismo tiempo, variadas prácticas corporalizadas y situadas, en la medida que presentan regularidades extra-racionalmente estabilizadas como *affordances* del contexto físico/social/cultural, constriñen la solución de problemas. Aquí, lo importante, a mi parecer, es que la externalización (y construcción) de mediadores epistémicos tiene su origen en el del agente solucionador de problemas. Los mediadores epistémicos, como se dijo, pueden distribuirse en una comunidad científica física, social y culturalmente demarcada, *en la medida que un conjunto de agentes decodifique un valor epistémico compartido de dichos mediadores*. La manipulación de modelos construidos como mediadores epistémicos presenta oportunidades para inferir hipótesis, pero dicha inferencia debiera dar cuenta, en algún momento, de la contribución que puedan tener estructuras de conocimiento previamente almacenadas en la memoria de largo plazo. Magnani (2001a, 2006a-b, 2007a, 2008, entre otros) propone un *interplay* entre procesamientos online y offline que pretende dar cuenta de dicha contribución, asumiendo la existencia de un procesamiento sintáctico combinatorio interno. En este punto, se asume una externalización y reproyección continua de estructuras modélicas externas sobre patrones neuronales, no porque haya razones para defender una identidad entre estructuras representacionales internas y dichos patrones, sino más bien porque se tiende a evadir la posibilidad de cualquier representación interna abstracta o simbólica. Esto último puede parecer demasiado especulativo, pero, por ahora, lo que interesa es sostener que la necesidad de un negociador epistémico, cuyos límites no se diluyan con los de sus propios mediadores epistémicos, hace plausible la defensa débil de una extensión derivada o mediada, por sobre una extensión literal del sistema cognitivo. En otras palabras, la “delegación cognitiva” en RBM no implicaría claramente un compromiso fuerte con la extensión cognitiva literal.

En consecuencia, la atribución de funciones cognitivas a objetos y prácticas determinadas (i.e. la delegación cognitiva), cuando se realizan alguna tarea cognitiva, no se podría explicar sólo con referencia a enfoques cognitivos que estudian principalmente la manera en que se coordina la generación y manipulación de representaciones artefactuales externas. Para explicar (y no sólo asumir) la existencia de sistemas inferenciales distribuidos es más importante especificar las características distintivas del supuesto acople representacional entre lo externo e interno. El marco investigativo de la cognición distribuida, por ejemplo, puede servir de base para dar cuenta de algún micro-sistema socio-cultural donde haya tenido lugar el surgimiento de soluciones candidatas para un problema determinado. Se podrá constatar, además, que tal sistema reúne a seres humanos y artefactos modélicos que congregan a un grupo de científicos. Pero esto sólo tiene que ver con la comprensión de que en tales circunstancias ha tenido lugar un sistema inferencial

exitoso. De aquí no se sigue un entendimiento sustantivo de cómo funciona y qué constituye exactamente tal sistema inferencial. Para redondear este punto, consideremos lo siguiente:

**“Un experimento con un artefacto modelo proporciona un extensión de la práctica científica de experimentación mental [...] más completa y potencialmente menos sujeta al error y sesgo personal.” (Nersessian, 2009, p. 749)**

Aunque aquí se alude a una extensión de la experimentación mental, se admite cierta independencia del funcionamiento del modelo con respecto a tal experimentación mental. De esto, uno puede suponer que es positivo para la búsqueda de posibles soluciones (y posibles descubrimientos, eventualmente) que los científicos cuenten con una *descorporalización* de sus propios modelos mentales con el fin de poder tener un mejor control de una amplia gama de manipulaciones posibles. Sin embargo, hay al menos dos líneas posibles de mayor cuestionamiento hacia una posible externalización cognitiva. Por un lado, surge la pregunta de si tal externalización de modelos mentales nos dice algo respecto de la naturaleza y características reales de la supuesta representación mental que lo “motivó”. Por otro lado, y suponiendo que se trate de la primera versión del artefacto, es fácil notar que tal artefacto no fue necesario para la instanciación del modelo mental que lo originó. Si es así, entonces son, al menos, dos las implicancias que se siguen de aquello:

1. Si los modelos internos existen, entonces no es siempre necesario la participación de un modelo externo para generarlos.
2. Los modelos mentales que participan de un experimento de pensamiento preceden a los artefactos que (supuestamente) “extienden” la práctica científica de la experimentación mental

El argumento que se puede esgrimir en este momento es que, aunque los artefactos presenten oportunidades para instanciar nuevas fases de un experimento mental, de 1 y 2 no se sigue que tales representaciones externas se acoplen literalmente con posibles modelos internos, precisamente, porque se desconoce la *naturaleza* del modelo interno. Nersessian defiende lo opuesto. Según ella, los modelos artefactuales sirven de interconexión [*interlocking*] entre distintas disciplinas, entre artefactos e investigadores, y entre las diversas configuraciones que pueda tomar el modelo en experimentos particulares. Esas serían buenas razones para proponer que el “modelo experimental que hace las predicciones” es, precisamente, aquel sistema modélico distribuido e interconectado que acopla al investigador y al artefacto. Obviamente, tal argumento se basa más en la descripción de prácticas científicas observables que lo que se conoce acerca del posible funcionamiento de supuestos modelos internos. Lo que surge – según me interesa plantear aquí - es el riesgo de una trivialización, no sólo de la noción de representación modélica, sino de lo que se pretende proponer por RBM en cuanto tal.

La trivialización en cuestión tiene que ver con atribuirle un carácter cognitivo a determinados objetos típicamente presentes en las prácticas científicas, tal como si se le atribuyera, por ejemplo, un *carácter solidario* al dinero manipulado en las prácticas recolectoras de las campañas de caridad callejeras. Mientras que en el primer caso la agencia cognitiva está distribuida entre humanos y objetos, en el segundo caso la *agencia moral* también estaría literalmente distribuida entre los donantes y el dinero donado. Obviamente, las situaciones aludidas no son completamente equivalentes. Sin embargo, la comparación sirve para considerar lo poco confiable que puede resultar suponer, sobre la base de procedimientos manipulativos externos, que el funcionamiento de ciertos procesos internos, no sólo tendría una naturaleza similar al externo, sino que además estaría determinado por aquellos.

A continuación, se intentará dar cuenta del tipo de limitaciones que hacen aún más controversial a un RBM, elaborando críticamente los aspectos (y principios metodológicos) que subyacen a tales limitaciones. El análisis de estas limitaciones tiene como propósito tratar de establecer un conjunto de criterios tentativos de constreñimientos y condiciones de adecuación que permitan evaluar la hipótesis central de este trabajo de investigación.

---

## CAPITULO IV

### 4.1. Limitaciones y alcances de una teoría del surgimiento de hipótesis por medio de RBM

Este último capítulo está orientado a explicitar dos condiciones de adecuación para una posible teoría del surgimiento de hipótesis por medio de un RBM externalista. La primera de estas condiciones dice relación con una concepción apropiada de representación modélica, y la segunda tiene que ver con la concepción de sistema cognitivo subyacente al surgimiento de hipótesis por RBM. Lo que se intentará argumentar es que ambas concepciones se basan en criterios de constreñimientos equívocos, derivada de una individuación mecanicista de procedimientos epistémicos.

Para dar cuenta de lo anterior, será necesario considerar, por un lado, lo que entenderemos por individuación de mecanicista de procedimientos epistémicos, y por otro, cuáles son las razones que hacen equívocos los criterios de constreñimientos del surgimiento de hipótesis por RBM. Por individuación mecanicista de procedimientos epistémicos se debe entender *el desafío de establecer qué o cuáles (tipos de) procedimientos epistémicos podrían ser aquellos sin los cuales sería imposible la expansión del conocimiento científico caracterizado como una instancia de descubrimiento por RBM en tanto sistema inferencial distribuido, determinada por aspectos del contexto temporoespacial y cultural en que se sitúa una (o un tipo de) tarea cognitiva dada*. Sostendré que tal noción mecanicista es un producto inicuo de la interacción entre ciencias cognitivas y filosofía de la ciencia.

Al mismo tiempo, trataremos de elaborar dos argumentos críticos de los criterios de constreñimientos frecuentemente aludidos en el RBM externalista. En este sentido, se propone que tales criterios están basados en dos duplas conceptuales implícitamente contradictorias. Primero, la dupla externalismo activo y *principio de paridad*. Segundo, la dupla representaciones híbridas y *continuo representacional*. Cabe aclarar que no se defenderá ni descartará, al menos en este trabajo, la veracidad de los dos primeros componentes de cada dupla. Más bien, se atacará la viabilidad de los segundos componentes como justificación apropiada de los primeros. Lo que de ahí se seguirá es que tanto el externalismo activo como la posible hibridación de modelos internos/externos carecen de confiabilidad para fundamentar y/o regular la manipulación de modelos en RBM como la defendido por la comunidad de investigadores pertinentes, donde exponentes como Knuuttila y Honkela expresan las visiones más radicalmente anti-representacionalistas.

#### 4.1.1. La tesis del RBM en el contexto de la interacción entre ciencias cognitivas y filosofía de la ciencia

---

La tesis del RBM que aquí nos ha interesado sostiene que la generación de hipótesis es un fenómeno que involucra constreñimientos explícitos e implícitos que pueden explicarse y describirse de mejor manera en función de un razonamiento basado en modelos. A lo largo del capítulo II, tratamos de abordar distintas concepciones modélicas con el fin de

llegar a precisar aquella noción artefactualista que motiva la denominada externalización (o delegación) cognitiva. En este sentido, es importante contar con una mejor especificación de qué es aquello que realmente se entiende por RBM externalista. En otras palabras, dada una noción artefactualista de los modelos, la pregunta es cómo es que se razona en base a ellos. Esta pregunta hace necesario evaluar posibles tipos de razonamiento que subyacen al problema de cómo la mente podría operar sobre vehículos cognitivos internos y externos a la vez. Una posibilidad, relativamente conservadora, es que un razonamiento formal (e.g. sintáctico o semántico) sea “causado” a partir de modelos físicos externos. En tal caso, es pertinente exigir también una especificación de los mecanismos que harían posible tal realización formal, asumiendo que el *input* sensorial se vincula causalmente con tales modelos sintácticos o semánticos internos. Otra posibilidad, menos conservadora, es que los modelos (en tanto artefactos) externos “motiven” la instanciación de modelos internos analógicamente. En este caso, se puede asumir una relación de isomorfismo mediada por el carácter realista del sistema representacional interno, lo que a su vez tiende a presuponer (por parte del agente cognitivo) un conocimiento previo de aquello que se requiere representar, y (por parte del investigador) una teoría causal de cómo es que se representa un objeto externo<sup>56</sup>. Una tercera posibilidad es la que hemos asociado a una extensión cognitiva, donde los modelos externos estarían en una continuidad material con los internos (sea lo que sean), para lo cual habría que aceptar una postura anti-representacionalista de las estructuras modélicas así acopladas.

Limitándonos a las tres posibilidades en cuestión, es apropiado hacer la distinción entre representaciones modélicas (y sus posibles mecanismos de representación) y la noción de un proceso de razonamiento basado en tales estructuras modélicas. Incluso si aceptamos el hecho de que la mente humana pueda representar (o re-presentar<sup>57</sup>) situaciones *modélicamente*, de eso no se sigue con claridad la manera en que una mente pueda operar con tales estructuras internas y sus supuestas instanciaciones materiales externas. Por esta razón resulta importante no despreciar por completo la distinción entre visiones formales y artefactuales de modelos, aun cuando las conocidas versiones formales resulten ser insatisfactorias para dar cuenta del fenómeno productivo generador de hipótesis que se quiere explicar.

Las visiones formalistas aludidas, en el contexto de la filosofía de la ciencia, tienen dos versiones contendoras entre sí: la sintáctica y la semántica. Según Knuuttila et al. (2005), ambas visiones tienen como motivación común la de establecer, dentro de un marco formal, qué son los modelos científicos. La primera concibe a los modelos como la interpretación de un formalismo (o cálculo) no interpretado. Este formalismo no interpretado (o parcialmente interpretado) corresponde a la teoría científica, la que también puede ser entendida como una estructura puramente sintáctica constituida por un conjunto de axiomas. Siguiendo a Nagel (1960), Knuuttila (2005) caracteriza la especificación de un modelo según la visión sintáctica de la siguiente manera:

***“Interpretar una teoría era especificar un modelo para ella, que hace verdaderos (o falsos) a todos los axiomas de la teoría. En consecuencia, un modelo para una teoría T podría ser definido como un conjunto de proposiciones verdaderas que posee la misma estructura formal o cálculo que T.” (p. 37)***

Alternativamente, la concepción semántica de los modelos sustituye la formulación sintáctica de la teoría por los modelos de la teoría. En otras palabras, lo que sucede es que

<sup>56</sup> Proyecto que, en filosofía, se enmarcaría dentro de la búsqueda de una teoría naturalizada (y convincente) del contenido mental.

<sup>57</sup> Ver nota 37.

se sustituyen las entidades lingüísticas por entidades extra-lingüísticas. Esta sustitución se nutre de cierta contribución proveniente de las ciencias cognitivas actuales que la primera visión soslayaba, razón por la cual no la podríamos entender a cabalidad sin primero atender a dicha contribución.

Los autores anteriormente señalados abordan dos versiones de la concepción semántica de los modelos. Dichas versiones tendrían distintos compromisos constructivistas, a saber, uno de inspiración empirista y otro realista, y estarían representadas por exponentes como van Fraassen (e.g. 1980) y Giere (e.g. 1988) respectivamente. El primero ha defendido una concepción de la teoría científica como una familia de modelos constituida de subestructuras empíricas capaces de representar directamente un fenómeno observable. En este sentido, la adecuación empírica de la teoría depende de que haya algún modelo cuya descripción experimental tenga una relación de isomorfismo con las subestructuras empíricas mencionadas. Giere, por su parte, niega la necesidad de una relación de isomorfismo entre un modelo y un sistema real. La alternativa propuesta se puede resumir de la siguiente manera: los modelos no son verdaderos o falsos con respecto al mundo, ni se corresponden con el mundo por medio de un isomorfismo. Más bien, los modelos guardan una relación de *similitud* con el mundo, donde dicha noción se debe interpretar como un “buen calce” [*good fit*] entre los modelos y los sistemas reales del mundo. De aquí que, para Giere, el rol de las teorías en tanto familia de modelos sea el de intentar establecer un buen calce entre modelo y sistema real. Para este fin, por lo tanto, las teorías deben estar constituidas por dos elementos: una población de modelos, y una serie de hipótesis que vinculan tales modelos a los sistemas del mundo real. Estos dos elementos harían de la teoría científica el tipo de modelo más interesante, si se considera que ya no habría nada en la estructura de un modelo que pueda determinar si pertenece a una familia de modelos en particular.

Lo que hace interesante y controversial la concepción de modelo de Giere es que el grado de similitud de un modelo con respecto a un sistema real depende de lo que juzgue alguna comunidad científica determinada. Esto conlleva dos implicancias interrelacionadas. Por un lado, el modelo pasa de ser un “modelo de” a un “modelo para”. Por otro lado, esta última adscripción depende directamente de la noción de similitud o calce anteriormente señalada, ya que esta *condición de semánticidad* es la que haría relevante el uso de modelos determinado por los propósitos del agente cognitivo. Y es aquí, precisamente, donde tiene lugar la contribución de las ciencias cognitivas y su evidencia confiable disponible. Según el propio Giere (1988), la evidencia acumulada de las ciencias cognitivas permitiría afirmar, con cierta certeza, que la cognición humana y la percepción operan sobre la base de algún tipo de “métrica de similitud” [*similarity metric*]. En otras palabras, es sobre esta evidencia que Giere da por supuesta la robustez de aquel mecanismo que vincula modelo y mundo.

Gran parte de las ciencias cognitivas actuales se caracteriza principalmente por establecer una concepción del agente cognitivo sobre la base del estudio de procedimientos epistémicos en contextos de resolución de problemas prácticos. Es entendible que la concepción de “modelos para” determinados propósitos *representacionales* del agente tenga un valor epistémico asociado a una visión de modelos enfocada en la práctica. En consecuencia, lo que se persigue al estudiar el rol de los modelos en el razonamiento científico es la descripción de prácticas que puedan estar constreñidas por una individuación mecanicista de procedimientos epistémicos. Tal individuación se ve reflejada no en el poder representacional de los modelos (aunque se insista en el uso del término), sino, precisamente, en una concepción de modelo que evite, entre otras cosas:

- suponer que el poder representacional (abstracto y/o idealizado) de un modelo está en la base de su rol epistémico
- asumir que el modelo es una representación fiel y claramente demarcada de algún sistema real del mundo, en vez de un artefacto material que le sugiere al científico mecanismos, procesos y soluciones tentativos que pueden funcionar como la base de variadas inferencias posteriores
- aceptar que un agente cognitivo pueda modelar algo erróneamente, como si se contara con un conocimiento previo de lo que se quiere modelar, en vez, más bien, de concebir la estructuración y el funcionamiento de modelos como un proceso de conjetura que se apoya en otros modelos artefactuales más familiares (lenguajes formales, diagramas, métodos de modelamiento, etc).
- soslayar el hecho de que los modelos tienen usos distintivos adicionales a la mera representación (e.g. estudiar el comportamiento de la estructura de un modelo variando ciertos aspectos del mismo, “reciclar” el uso de modelos exitosos en dominios distintos, etc.)
- desvincular el valor epistémico (o productivo) de un modelo a los constreñimientos materiales que se intersectan con sus usos intencionales.

A partir de esto es posible afirmar que razonar en base a modelos no hace alusión a la manera en que los científicos puedan implementar “trenes de pensamientos” relacionados entre sí por algún tipo de “proceso inferencial modélico” interno (sea lo que fuere que aquello). Más bien, al hablar de RBM se hace alusión un posible rol epistémico de la manipulación de modelos materiales supuestamente autorizado por una argumentación contrafactual como la siguiente: *de entre los agentes intencionales que razonan (sin especificarse cómo), aquellos que participan en el quehacer regular de alguna comunidad científica tienden a usar modelos físicos cuando producen conocimiento. Dado que tal conocimiento es productivamente más significativo que el de los agentes intencionales ordinarios, es posible concluir : (a) que sin el uso de tales objetos no sería factible producir (ni entender) tales grados expansión del conocimiento, y (b) que la aparente centralidad de los modelos externos, dada su ubicuidad en las prácticas conducentes a la expansión de conocimiento, es una buena razón para asumir el carácter modélico de las estructuras representacionales internas.* Esta última tesis es la que he tratado de exponer como un producto poco saludable de la interacción entre filosofía de la ciencia y ciencias cognitivas.

Sobre la base de la evidencia proveniente de las ciencias cognitivas actuales, se ha intentado establecer una concepción empírica de sistema cognitivo íntimamente vinculada a las tecnologías, artefactos y ambientes externos al organismo humano (e.g. Clark, 2003). Lo que ahora importa de esta concepción es que, de aceptarse, debiera entenderse como una concepción basada en ciertos principios de *extensión cognitiva literal*. Aunque no suele explicitarse claramente, parece evidente que los criterios de constreñimientos para un RBM, en cualquiera de sus versiones más o menos externalistas, suelen estar regulados por el mismo tipo de principios. Además, tales principios tienen un grave problema de justificación, toda vez que para su fundamentación se recurre a otros principios que resultan compatibles con los primeros sólo en apariencia. Veremos, sucintamente, el caso de las dos duplas de principios reguladores de la extensión cognitiva, señalados en la introducción de este capítulo, a partir de lo cual procederemos a evaluar críticamente ciertos criterios de constreñimientos implícitos en RBM.

#### 4.1.1.1. Externalismo activo en RBM y la falacia del principio de paridad

En la sección 1.1.4., al dar cuenta del principio del externalismo activo, destacamos la tesis defendida por Clark y Chalmers (2008) de que en un sistema cognitivo extendido:

**“Todos los componentes en el sistema juegan un rol causal activo, y juntos gobiernan la conducta en la misma manera que la cognición usualmente lo hace.” (p. 222, énfasis agregado)**

En esta sección, se intentará mostrar que tal principio no sólo subyace ubicuamente como justificación implícitamente oportunista del RBM, sino que además se basa en una falacia de paridad entre componentes internos y externos a la caja craneana, con implicancias posiblemente fatales. En otras palabras, mostraremos que lo que menoscaba la plausibilidad del externalismo activo también menoscaba la plausibilidad del RBM como sistema inferencial distribuido. Para empezar, veremos un caso determinado dónde se puede constatar el rol implícito del principio externalista. La falacia que conlleva tal principio, y su efecto en el RBM, serán discutidos inmediatamente a continuación de este caso.

En Nersessian (2009), se intenta proporcionar mayor evidencia acerca de las actividades de investigación que podríamos traducir como *interconexión de modelos* [*interlocking models*]. Para Nersessian, esta noción actúa como una unidad de análisis capaz de capturar las prácticas en las que tienen lugar una interacción interdisciplinaria de “conceptos, métodos, discursos, artefactos, etc.” (p. 734). El estudio en cuestión se restringe, inicialmente, a tres preguntas de investigación centrales que reflejan el carácter heterogéneo del nivel explicativo asumido por la interconexión de modelos:

1. ¿Dónde, a lo largo del quehacer investigativo interdisciplinario seleccionado, tienen lugar las prácticas cognitivas bajo estudio (e.g. formulación de problemas y razonamiento)?
2. ¿Cuáles son las prácticas de investigación a través de las cuales se realiza una aproximación a los problemas de investigación?
3. ¿Cuáles son las prácticas de interacción en la cultura de laboratorio que facilitan y sostienen el aprendizaje y la investigación?

Esta aproximación *filosófico-experimental* acerca del quehacer de una comunidad científica determinada asume una relación fundamental entre facetas culturales y cognitivas, centrando su análisis en tres tipos de prácticas: cognitivas, de investigación, y de interacción. El primer tipo agrupa prácticas como la formulación y solución de problemas. El segundo dice relación con “la naturaleza del trabajo diario de los miembros del laboratorio en la medida que intentan desarrollar entendimiento, técnicas y tecnologías” (p. 736). Las prácticas de interacción, a su turno, tienen que ver con las configuraciones socio-culturales que dan cuenta del tipo de relaciones que se forman en el laboratorio y la forma en que éstas se sustentan en el tiempo. Según Nersessian, este último tipo de prácticas es el que soporta y mantiene a los dos primeros.

Para Nersessian, el hecho de que las prácticas interaccionales sean las que soportan tanto a las prácticas investigativas como a las cognitivas no es trivial. De hecho, la noción de interacción sería ubicua con respecto a las dimensiones investigativas y cognitivas del quehacer científico estudiado. Por un lado, esta ubicuidad se hace patente en torno a prácticas de resolución de problemas y, también, a aspectos modélicamente emulados de algún fenómeno transversal a la división *in vivo/in vitro*. A su vez, tal ubicuidad también se manifestaría con respecto a lo que se denomina experimentación a través de “modelamiento simulativo”. Es importante extenderse un poco más sobre las supuestas manifestaciones de esta ubicuidad, porque es en ellas donde mejor se puede mostrar aquello que menoscaba las tesis centrales del enfoque epistemológico en cuestión.

En la ingeniería biomédica se presenta una dificultad ética frente a la experimentación realizada directamente sobre seres humanos. Por esta razón, los investigadores del campo se ven obligados a diseñar y construir modelos físicos que permitan simular fenómenos *in vitro*. Según Nersessian, cuando estos modelos tienen características (sustancialmente<sup>58</sup>) compartidas con aquello que emulan, se puede hablar de un *proceso paralelo de simulación*. Dicho de otra manera, lo que se está asumiendo es un grado de identidad entre ciertos aspectos o rasgos seleccionados del modelo físico y ciertos aspectos o rasgos propios del fenómeno supuestamente emulado. Claramente, si entendemos bien, el paralelismo aludido se haría manifiesto porque cierto conjunto R de rasgos (e.g. funcionales, estructurales, materiales) de algún modelo físico emulador A, también estarían presentes en el fenómeno B que es emulado por A, pero no como una simulación arbitraria, sino como una instanciación genuina de R. Luego, si algún R determinado (e.g. la fuerza del flujo sanguíneo en una arteria humana) es efectivamente instanciable en A y B, se estaría cumpliendo con una supuesta condición de adecuación imprescindible. A nuestro modo de entender, tal condición podría explicitarse de la siguiente manera: para que un modelo físico sea un modelo simulativo que forma parte de un “sistema solucionador de problemas distribuido”, tal modelo debe “ser (algún tipo de) X” y “referirse a X” al mismo tiempo. De aquí, la defensa recurrente que hace Nersessian de la noción de un sistema híbrido donde los modelos físicos [*devices*] cumplirían el rol de “socios cognitivo” con respecto a la participación del organismo biológico en tal sistema:

**“La experimentación con modelos físicos requiere configuraciones en las que los componentes biológicos e ingenieriles convergen en la investigación, y a las que los investigadores llaman un ‘sistema modelo’.” (2009, p.738)**

Esta “sociedad cognitiva” - según me interesa argumentar - no se sería otra cosa que una versión más del principio de paridad. Y la transversalidad que se le atribuye tanto a las prácticas de resolución de problemas como a los rasgos de emulación modélica, no sería más que otra versión de lo que se ha comentado anteriormente como una extensión cognitiva literal. Ambas cosas se pueden corroborar sobre la base de la evidencia presentada por Nersessian y su equipo de investigadores. De hecho, la metodología de investigación escogida por ellos incluye la codificación de prácticas cuya categorización se establece sobre la base de dos procedimientos: el análisis de reportes lingüísticos seleccionados de las entrevistas realizadas a los miembros del laboratorio, y la caracterización descriptiva de observaciones etnográficas en terreno. Resulta interesante enfocarse en la categoría superordinada que estos investigadores denominaron “Cognición basada en modelo”. La tabla 2 muestra esta categoría,

incluyendo sus correspondientes subcategorías y ejemplos:

**Tabla 2: Ejemplos de cognición basada en modelo, según Nersessian (2009, p.737)**

<b>CATEGORÍA SUPERORDINADA Cognición basada en modelo</b>	
<b>CATEGORÍAS</b>	
Entendimiento basado en modelo Hacer referencia, indicar, explicar, dar evidencia	“Entonces ideamos a este flujo helicoidal [ <i>flow loop</i> ] como una manera de imponer una presión de corte bien definida a lo

<sup>58</sup> Por esto habría que entender algún tipo de naturaleza o rasgo esencial compartidos entre las instanciaciones de los modelos.

de comprensión en términos de una representación organizada	largo de una gran población de células de modo que la respuesta global se deba a dicha presión de corte bien definida.”
Razonamiento basado en modelo Evidencia de razonamiento a través de la construcción y manipulación de modelos, marcadas por inferencias (“si...entonces”, “así”, “parece como si”, “quizás”, “sospecho”, “por lo tanto”, etc.)	“[las neuronas] Están sintiendo y tocando y probando alrededor de su ambiente hasta que se topan con las compañeras correctas con las cuales formar una red. Y sospecho que gran parte de la decisión tiene que ver con la actividad que tiene lugar en la placa [de Petri] también. Por lo que, quizás, si reciben una cierta señal, una señal eléctrica en el mismo momento que hacen contacto, pueden decidir formar una conexión permanente. Y la razón por la que sospecho esto es porque existe muy buena evidencia en los animales...”
Explicación basada en modelo Uso de modelos para comunicar, transmitir entendimiento	“Si miras un vaso sanguíneo, a esta escala parece un tubo. Si tú estatura fuera de un milímetro, entonces se verían muy planas. Así es que utilizamos ese [constructo plano] como una aproximación a la superficie tubular.”

Es interesante notar que Nersessian (2009) no sólo correlaciona el uso de un “lenguaje publico figurado” con el posible rol constitutivo de la cognición que puedan tener las representaciones modélicas. Una segunda correlación que se presenta, como evidencia de lo mismo, es aquella que se da entre el grado de *expertise* (y por lo tanto de experiencia y familiaridad) de los científicos del laboratorio en cuestión y la tendencia a atribuirle una agencia a los modelos físicos que utilizan, especialmente en contraste a los novatos o recién llegados. En lo que viene, se argumentará que tal evidencia no cumple el propósito señalado, sino que más bien tiende a expresar un grado de comprensión mayor o menor por parte de los sujetos estudiados con respecto a los procesos, eventos y sistemas que tienen lugar en el contexto estudiado.

La noción de *interlocking models* comentada al inicio de esta sección es la unidad de análisis que Nersessian utiliza para defender la tesis de que ciertas relaciones entre diversos componentes del laboratorio (biológicos y no biológicos) pueden conformar “sistemas inferenciales distribuidos”. Al expresar estos sistemas como un sistema cognitivo socio-cultural, lo que se hace es tratar de aproximarse a una unidad multidimensional que permita **especificar las maneras en que una simulación determinada cumple el rol de pivote** o *hub*. Este rol es ilustrado por Nersessian haciendo la analogía con un sistema de transporte donde diversas líneas de servicio de intersectan en estaciones centrales. En este sentido, la función crucial que tienen los modelos (en tanto *hubs*) es la de articular el razonamiento basado en modelos que estaría distribuido en al menos tres dimensiones: la interdisciplinaria; la de los componentes del sistema conformado por artefacto e investigador; y aquella de las configuraciones de modelos en los experimentos.

Es interesante notar que las dimensiones mencionadas se cohesionan en torno a dos hipótesis, supuestamente complementarias. Por un lado, se asume que la evidencia confiable disponible sería indicativa de que los investigadores son capaces de instanciar

modelos mentales o conceptuales internos (acerca) de procesos y entidades cuando se requiere llevar a cabo tareas cognitivas. Las transcripciones de las investigaciones lideradas por Nersessian proporcionarían evidencia del tipo ilustrado en la tabla 2. Por otro lado, la observación etnográfica de las prácticas científicas mostraría que gran parte de esta actividad tendrían como componente central (por su ubicuidad) el uso, construcción y manipulación de modelos físicos artefactuales. No obstante lo anterior, uno podría seguir invocando afirmaciones en torno de estas dos hipótesis sin hallar una argumentación sustancial relacionada a la naturaleza de la complementariedad aludida. De hecho, el problema del “ensamblaje” entre modelos internos y externos no es distinto al problema más general involucrado en la polémica acerca de la extensión mental. Para aceptar dicho ensamblaje se deben aceptar “supuestos no negociables” que, precisamente, parecen esconder la falacia que explicitaremos a continuación.

La plausibilidad de un sistema inferencial distribuido (o bien, de un RBM externalista) depende de la plausibilidad de la Hipótesis de la Cognición Extendida (o bien, de la posibilidad de que el sustrato físico de los procesos cognitivos estén constituidos por algún ensamblaje entre humanos y artefactos). Lo que hace equivalente ambas relaciones de dependencia es, en primer lugar, el supuesto de que la cognición estaría determinada por componentes materiales extra-craneales, y, en segundo lugar, el supuesto de que algún sistema constituido de tal manera pueda conformar una unidad de análisis metodológico que transgrede de la noción individualista tradicional de la cognición. De estos último, sin embargo, no se sigue que haya una manera clara de establecer qué es aquello que le otorga unidad (o completitud, si se quiere) a dicha supuesta unidad de análisis alternativa. La argumentación, como sabemos, es contrafactual, pero fatalmente paradójal. Específicamente, la paradoja consiste en tratar de darle plausibilidad a un principio de paridad cognitiva que pretende disolver un límite conceptual entre lo interno y lo externo, al tiempo que retiene el mismo límite metodológico para justificar la existencia de un posible continuo representacional.

Tal paradoja ha sido explicitada, para fines distintos a los nuestros, por Di Paolo (2009). En nuestro caso, diremos que el problema se resume de la siguiente manera: mientras que, por un lado, los defensores de la extensión cognitiva están comprometidos con la idea de que los límites de la caja craneana son irrelevantes para demarcar lo cognitivo, por otro lado, proponen un test que considera, precisamente, la equivalencia funcional de lo externo con procesos que ocurren dentro de los límites de la caja craneana. Si, como se ha dicho acerca del RBM en ciencia, la continuidad representacional de las estructuras modélicas analógicas (internas y externas) se fundamenta en la relevancia de las “prácticas inferenciales distribuidas”, esto equivale a sostener dos cosas. Primero, que lo que está fuera de la caja craneana *puede actuar* como lo que está adentro. Segundo, y dado lo anterior, que hay procesos inferenciales idiosincrásicos de la actividad científica que no pueden ser sólo internos o sólo externos. Ambas afirmaciones incurren en la misma paradoja metodológica-conceptual anteriormente señalada. Es por este motivo que, a nuestro parecer, resulta razonable sostener que la falta de plausibilidad del RBM externalista (o distribuido) está determinada por la falta de plausibilidad de la HCE (Hipótesis de la Cognición Extendida), toda vez que la plausibilidad de la HCE depende de la plausibilidad de un principio de paridad que resulta ser falaz.

Creo que hasta el momento hemos aclarado una parte del problema, vale decir, lo referente a la dupla externalismo activo y *principio de paridad*, siempre y cuando, como se dijo en la introducción de este capítulo, quede en evidencia la pérdida justificación del primer componente de la dupla sobre la base del segundo componente. Debiera

resultar más convincente ahora el hecho de que términos como “externalismo activo en RBM” y “razonamiento modélico distribuido” hacen referencia al mismo tipo de distinciones metodológicas. Sin embargo, aún corresponde abordar la segunda parte del problema en cuestión, a saber, aquella relacionada a la dupla representaciones híbridas y *continuo representacional*. De eso se trata la siguiente sección.

### **4.1.1.2. El problema de las representaciones híbridas y el continuo representacional**

En lo que sigue trataré de desarrollar un argumento en contra de la noción de continuo representacional, tal como ésta suele asumirse metodológicamente para sostener la relevancia de representaciones híbridas determinantes de la cognición. Como hemos visto a lo largo de este capítulo, la supuesta interacción productiva entre estructuras representacionales internas y externas estaría determinada por el carácter modélico de tales representaciones. Dado este supuesto del RBM, trataremos de mostrar que asumir una organización representacional híbrida, basada en la noción de continuo representacional, no es distinto a sostener que tal continuo implica, de manera no especificada, una etapa transitoria de componentes “internos” y “externos” donde:

1. bien a la interacción (productiva) entre vehículos representacionales internos y los modelos simuladores externos se le atribuye un estatus arbitrario de constitución;
2. bien los vehículos representacionales (internos y externos) interactúan en una caja negra (o interfaz entre lo interno y lo externo) donde las funciones de “vehículo” y “representación” se hibridan en algo que las hace indistintas.

Ciertamente, uno puede ensayar más de alguna hipótesis interpretativa con respecto a estas dos posibilidades. Del mismo modo, también podríamos sostener resolutivamente que el problema central tiene el mismo vicio discutido en la sección anterior, pero con eso no ahondaríamos en el propósito aclaratorio del presente trabajo. Por esta razón, considero apropiado contextualizar primero el problema de las denominadas representaciones híbridas, para luego explicitar los argumentos que debieran aclarar un poco más la inviabilidad de un continuo representacional como justificación de tal tipo de estructuras modélicas. Esto debiera sentar las bases para el carácter de los criterios de constreñimientos de en un RBM externalista que pretendo finalmente desafiar.

Mucha de la investigación relacionada a la contra-revolución (ver sección 1.1.1.), estaría en lo correcto cuando denuncia que las representaciones entendidas solamente como símbolos estructurados en la línea de lo que se conoce como una lógica de primer orden, y manipuladas por reglas inferenciales, no parecen dar una cuenta cabal de diversos aspectos de la cognición. Sin embargo, esto no licencia la hipótesis radicalmente anti-representacionalista de la contra-revolución, razón por la cual parece plausible replantear al menos algo de lo que se propone en cada una de las trincheras opuestas. Este replanteamiento pasa por tomar partido entre aspectos conceptuales y metodológicos, y adaptar lo primero a lo segundo parece tener una viabilidad claramente más práctica. Esto podría explicar la tendencia por privilegiar los hallazgos provenientes de las diversas ciencias cognitivas actuales abocadas al estudio de lo que cada una considera como mecanismos cognitivos, buscando luego una fundamentación teórica apropiada, y suficientemente flexible, para acomodar los hallazgos correspondientes. Dicha tendencia metodológica no puede ser indiferente a la noción de representación, dada su rol explicativo necesario en procesos de orden superior. De ahí que un resultado, siempre provisional,

tome la forma de un compromiso programático basado en la interpretación de hallazgos experimentales para la elaboración de cierto entendimiento empírico de la cognición.

Ahora bien, el costo de mantener un esquema explicativo que incorpore la noción de representación es el de tener que modificar la concepción misma de representación, en los términos que impone aquello que parecieran informar las disciplinas experimentales. La diversidad de aproximaciones, finalmente, produce diversas concepciones de la noción en cuestión, elaboradas incompletamente con la agravante de que también son incompletos aquellos programas particulares de investigación desde los que han surgido. Los compromisos “teóricos” y metodológicos así adquiridos asumen nuevas nociones de representación sobre una justificación del tipo “y qué otra cosa puede ser”. Sin embargo, la representación asumida puede perder un rol explícito en la explicación de los mecanismos cognitivos. En otras palabras, las explicaciones así producidas en RBM mantienen compromisos con la noción de representación presupuesta, sin explicitarse el formato y el rol que juega dicha noción. Así, en la práctica parece no ser un problema realizar investigación bajo un mismo paraguas programático general, aún cuando se realice sobre distintas nociones de representaciones modélicas que resultan conceptualmente confusas.

Lo expresado en el párrafo anterior puede resultar, y con razón, altamente especulativo. Sin embargo, la discrepancia entre distintas nociones de representaciones se puede constatar sobre la base de la distinción entre “modelos *sobre* los que razona” o “modelos *con* los que razona”. Ambas concepciones suelen, incluso, estar presentes en un mismo caso. Nersessian (ver 1999, 2008, como casos paradigmático del RBM), ilustra el uso de su método histórico-cognitivo en el estudio del desarrollo de la teoría electromagnética de Maxwell. En general, este estudio describe procesos de RBM donde, supuestamente, razonamientos modélicos analógicos y visuales pueden expresar los constreñimientos de un sistema conceptual contingente en términos de constreñimientos derivados del “target” (i.e. del fenómeno simulado por analogía), y luego generar nuevos constreñimientos que son integrados en un sistema conceptual revisado. Lo que completa la tarea de expresar e integrar constreñimientos de distintas fuentes es un proceso de “abstracción genérico” (o abstracción por modelamiento genérico). Dicho proceso de abstracción, de base modélica en sí mismo, contrasta con la modelación sensorial que se aprecia en otra etapa del mismo estudio:

**“Como se usa en el modelamiento en física, las representaciones visuales externas tienden a tomar una forma diagramática. [...] Por ejemplo, en la construcción inicial del concepto de campo de Faraday, el modelo visual [...] proporcionó una representación idealizada de la líneas de fuerza adyacente a la barra magnética.” (1999, p. 18)**

Algo interesante es que la misma “representación visual” del modelo de líneas de fuerza funcionó después, de manera más general, como la representación de un modelo teórico dinámico de fuerzas, sin que ese fuese el propósito inicial (caracterizándose, por lo demás, la supuesta autonomía de los modelos). Si a esto se suman aquellos otros modelos multimodales de carácter *directamente* sensorial, propios de la abducción manipulativa online, el asunto de la variedad y confusión de formatos modélicos se hace aún más evidente.

Dado el contexto descrito, corresponde aclarar qué es lo que se sigue de cada una de las dos posibilidades i) e ii) señaladas anteriormente. En primer lugar diremos que mientras la primera tiene una clara orientación metodológica (o práctica, si se quiere), la segunda pretende asumir una conceptualización empíricamente motivada. Es decir, si la relación interacción/constitución tiene fines explicativos de orden pragmático (i.e. discriminados por el efecto o consecuencia de algún “valor epistémico” asociado a algún tipo de mecanismo

cognitivo), la hibridación en cuestión se inspira en la necesidad de contar con un continuo representacional de carácter sustantivo, asumiendo que tal noción estuviese en condiciones de desafiar el supuesto clásico de los vehículos cognitivos representacionales de la Ciencia Cognitiva. Naturalmente, no es el propósito de este trabajo el dilucidar esta *dicotomía de sustantividad*, pero es importante considerar que la adhesión a cualquiera de estas nociones tiene implicancias significativas, en el sentido de que a partir de ellas se pueden ofrecer hipótesis acerca de la naturaleza de los estados mentales y, por consiguiente, de la autonomía que los mismos puedan tener en alguna ciencia especial. Por ahora, bastará poner atención a la viabilidad de una posible representación híbrida fundada en una noción continua de representaciones mentales y extramentales, con supuestos poderes causales a lo largo de todo el continuo.

El problema de la relación interacción/constitución no es distinto al de la polémica desarrollada en torno a la *falacia acoplamiento/constitución* sostenida en Adams, F. y Aizawa, K. (2001, y otros). Ciertamente, afirmar que la interacción - por íntima que sea - entre elementos biológicos y artefactuales de algún sistema determinado constituye la integración de un *sistema cognitivo* es sólo una interpretación de base empírica. Esta interpretación está sesgada por el argumento (cada vez más habitual) de que dicho sistema participa en la solución de problemas adaptativos, en último caso motivados por el aprovechamiento de oportunidades de sobrevivencia. Luego, la falacia consistiría en atribuirle arbitrariamente un carácter cognitivo a la mera interacción de componentes, entendiéndolos como “socios cognitivos” por el sólo hecho de interactuar contingentemente al momento en que al solución de alguna tarea o problema es observada. Esta generalización del aparente estatus cognitivo señalado es la misma que ocurre cuando investigadores como Nersessian (ver sección anterior) justifican la distribución de un sistema solucionador de problemas que gira en torno a la “interconexión de modelos” (o al rol de los modelos físicos como *hubs*). En ambos casos, se asume el supuesto de que no habría nada controversial entre la vinculación (o complementariedad) literal entre vehículos causales (modélicos) internos y externos. Y eso es precisamente lo que aquí se intenta contradecir categóricamente.

Para dicho propósito, me concentraré en impugnar la posibilidad (ii), asumiendo que a partir de eso se hará innecesario (o menos plausible) sostener la posibilidad (i). En primer lugar, debemos tener en cuenta la posición artefactualista presentada en la sección 2.1.3. Se podrá recordar que esta visión anti-representacionista de los modelos esgrimía un carácter de mediador epistémico de los mismos como una reacción en contra de tres supuestos básicos del representacionismo: que el conocimiento consiste en un conjunto de representaciones internas; que estas representaciones se corresponden con partes de la realidad de manera precisa; y que la realidad consiste en objetos cuya existencia es independiente de las representaciones internas. Todos estos supuestos giran en torno a la tesis de que la función principal de los modelos es la de representar. Ahora bien, el considerar los modelos como estructuras meramente materiales es una estrategia que permite resaltar cierta “virtud epistémica” vinculada a los procesos de manipulación y construcción de modelos. Normalmente, dicha virtud se expresa en función de la independencia del *objeto* (o modelo físico), toda vez que dicho objeto u artefacto presenta oportunidades de inferencias sobre la base de constreñimientos físicos que pueden actuar como *affordances* para propósitos epistémicos. De dónde surgen tales modelos artefactuales con virtudes epistémicas es algo que interesa examinar ahora, porque no es del todo compartido que un modelo en tanto “artefacto epistémico” sea la única versión de modelo relevante en un RBM externalista. Nersessian, por ejemplo, acepta la (necesaria) posibilidad de que la memoria de trabajo albergue estructuras modélicas que

---

importan más por sus características analógicas y simuladoras de un fenómeno que por su propia materialidad.

Sin embargo, es preciso sostener que en ambas visiones se asume un canal representacional continuo entre procesos intracraneales y extracraneales. Y aunque, ciertamente, los vehículos de la interacción de ambas dimensiones debieran variar según cada una de esas visiones, hay coincidencia en atribuirle una importancia significativa a las estructuras modélicas híbridas. Tales estructuras, supuestamente resultantes de algún proceso de interacción causal constructivo o de continuidad modélico-representacional, debieran tener un locus externo al organismo. No podría ser de otra forma, si se quiere considerar relevante la existencia de socios cognitivos distribuidos o re-organizados extracranealmente. Y según lo que nos interesa impugnar, diremos que, aún aceptando la relación parcial/totalmente isomórfica y/o analógica entre modelos externos e internos, cualquier hibridación constructorista de los modelos queda confinada a una “caja negra” que deja fuera todo lo que no es híbrido en cualquiera de las direcciones del *canal representacional continuo* en cuestión. Si estamos cercanamente a un razonamiento correcto, podríamos hablar ahora de una dicotomía de orden metodológica, en los términos que se exponen a continuación. Por un lado, ambas visiones (más/menos externalista) aceptan que los modelos híbridos tienen su mayor importancia en virtud de la manipulación epistémica de la que pueden ser objeto. Por otro lado, ambas visiones admiten que, en su versión de objetos materiales epistémicos, los modelos pueden también tener un poder menor o mayormente representacional, en virtud de su carácter analógico y/o emulador. En este caso, la dicotomía es metodológica porque tiene como interés la delimitación externalista de la “unidad de análisis” de lo cognitivo. Luego, si es el caso que tanto la manipulación epistémica como la representación (funcional y estructuralmente) analógica son cruciales en la producción de conocimiento, entonces no es poco controversial lo que hemos señalado como el funcionamiento de una caja negra como la aludida hasta el momento.

A partir de lo anterior, se vislumbra una discrepancia - oportuna para nuestros propósitos - entre teorías de RBM que suscriben una posible unificación programática. Desde la perspectiva más conservadora, los modelos físicos externos tienen una importancia epistémica central porque su manipulación determina inferencias internas que también operan por medio de estructuras internas modélicas (al menos hasta un nivel de memoria de trabajo). Desde una perspectiva más radical, los modelos físicos externos tienen una importancia epistémica central porque se alinean con una concepción pragmatista de la representación, a saber, aquella que confina las estructuras modélicas a su realidad material continua con el resto del mundo, donde la cuestión principal es el propósito práctico de determinar *cómo se usan* los modelos para re/presentar, en vez de *cómo representan*. La discrepancia principal entre estas aproximaciones es que mientras ambas defienden la noción de modelos híbridos (distribuidos y corporalizados), sólo una visión defiende la pertinencia del poder representacional de los modelos internos, ya que la otra visión defiende un anti-representacionalismo que abandona el rol del componente representacional interno (a menos que ciertas estructuras internas, probablemente neuronales, sirvan para re/presentar aspectos de un fenómeno tal como lo harían las re/presentaciones externas).

La noción de continuo representacional, utilizada para sostener el surgimiento de representaciones híbridas determinantes de la cognición científica, resulta menos consistente en su versión más radical, o más artefactualista. La razón de esta inconsistencia es que se estaría defendiendo una eliminación de la noción de representación interna

sobre la base de una caja negra, como se dijo, sin especificar los mecanismos por los que esta caja lleva a cabo la hibridación de las representaciones externas/internas. En el caso de propuestas como las de Nessesrian y Magnani, entre otros, efectivamente tiene lugar un mecanismo de interacción que puede defenderse metodológicamente. Es decir, es posible que tal interacción (eventualmente productiva) permita la descripción de sistemas distribuidos donde se complementa el organismo (o un conjunto de organismos) y alguna serie de estructuras físicas modélicas. Sin embargo, de esto no se sigue necesariamente cuál sea el locus de los procesos de expansión de conocimiento, y si esto no está claro, lo mismo se puede decir del carácter cognitivo de la interacción en cuestión. Además, tampoco se puede decir mucho del carácter predictivo de tales sistemas, hasta el momento. En consecuencia, es una cuestión más o menos obvia que estas relaciones de interacción (no cognitivas) puedan especificar relaciones *mecanicistas* que permitan describir el funcionamiento de un sistema de contribuciones distribuidas, pero no es obvio que asumiendo una caja negra – como la aludida – se especifiquen (paradojalmente) sus propios límites y supuestos mecanismos cognitivos idiosincrásicos (distintos a los tradicionales).

### 4.1.2. Criterios para los constreñimientos de un RBM externalista

---

En esta sección, trataremos de explicitar ciertos criterios para formular constreñimientos de un RBM que podrían desprenderse de los temas tratados hasta este punto. En lo que sigue, se hará referencia a la centralidad que tienen las aproximaciones pragmatistas en el sentido discutido en la sección 2.1.2<sup>59</sup>. Dada esta aproximación, haremos patente la contribución de la noción “artefactualista” de los modelos, con el propósito de evaluar aquello que parece estar dándose por sentado en los enfoques de RBM examinados en este trabajo, a saber, que la nociones pragmatistas de “representación modélica extra-mentalista” (donde prima una supuesta autonomía de componentes extra-craneales que negocian epistémicamente con el agente intencional) y de “representación modélica híbrida” (donde prima una supuesta complementariedad intra/extra craneal no aclarada) pueden formar parte de un programa de investigación unificado de RBM.

Específicamente, los criterios que a continuación se discuten se relacionan entre sí sobre la base de una cuestionable línea argumental. Esta argumentación sostiene que la teoría de RBM que se pretende desarrollar, para la explicación de procesos de descubrimientos científicos, puede ser justificada por las virtudes epistémicas que se logren atribuir (y fundamentar experimentalmente) a las prácticas analizadas en alguna área de investigación determinada. En este sentido, se asume una conexión entre ciertas “capacidades de manipulación epistémica” - supuestamente presentes en todos los mamíferos – y una suerte de “amplitud de oportunidades cognitivas” relativas al acrecentamiento del mencionado valor epistémico de diversas prácticas corporalizadas y ambientalmente situadas, por medio de la creación de “nichos cognitivos”. Intentaremos contra-argumentar que dicha conexión es al menos controversial cuando se evalúa en torno a los denominados “experimentos de pensamiento visual” como evidencia empírica confiable.

A partir de la discusión de estos criterios, desarrollaremos el tema de la última sección, donde se impugnará la posibilidad de encontrar regularidades confiables en la descripción

---

<sup>59</sup> Concretamente donde se especificó que «... la aproximación científica debe ser mejor entendida como una relación triádica que incluya, [...], al agente en un rol que se puede evaluar conforme al “éxito de logro” que su participación tenga con respecto a las prácticas productivas o de expansión de conocimiento».

de capacidades epistémicas que puedan cumplir el rol de condición de adecuación mínima para la manipulación de representaciones modélicas sin formato unificado en un RBM externalista. Dado el hecho de que la investigación empírica asociada a esta posible teoría del descubrimiento soslaya el trasfondo filosófico normativo que puede guiar las respectivas descripciones de un RBM, diremos que no es viable aún plantear una clara continuidad entre aproximaciones científicas sustentadas en posiciones filosóficas incompatibles. Sin ahondar en el trasfondo racionalista que pueda existir en las visiones más tradicionales de la actividad cognitiva, sostendremos que un énfasis pragmatista de las visiones alternativas presupone una discontinuidad sustancial, expresada por discrepancias conceptuales y metodológicas irreconciliables a la hora de plantear un programa de investigación unificador del razonamiento creativo en ciencia.

#### 4.1.2.1. Funcionalidad y contexto de uso en RBM

Pohjola (2007), ha proporcionado una caracterización de los artefactos técnicos como interfaces entre dos ambientes: uno interno que es propio de la sustancia y la organización del artefacto, y otro externo, que es el ámbito en que se utiliza el artefacto en virtud de tal interfaz (a la que parece atribuírsele un sentido indexical real). Esta distinción, según un aspecto que me importa enfatizar, tiene una pertinencia particular para las interpretaciones pragmatistas del razonamiento generador de hipótesis explicativas. Esta pertinencia se hace explícita a través del rol que Giere (2004) destaca de los modelos y de los propósitos que los científicos, en tanto agentes intencionales, tengan al momento de utilizar dichos modelos para representar algún fenómeno natural. En la medida que los artefactos son entendidos como interfaces al servicio de los propósitos prácticos del agente, lo que se está defendiendo (o asumiendo) es que la acción de dichos agentes es relativa al uso del artefacto a través de dicho tipo de interfaces, y que sólo en función de una acción así especificada adquieren significancia los procesos modeladores. Dicho de otro modo, las propiedades interesantes que se les puedan adscribir a los objetos físicos en la construcción de modelos, en general, no vienen dadas por lo que son o por la sustancia de la que están hechos solamente, sino principalmente porque esas propiedades adquieren significancia de acuerdo al uso en virtud de una interfaz de interacción. Estos usos tienen su lugar en la externalidad de los objetos artefactuales, fundándose con esto cierto grado de plausibilidad para la tesis de que el ambiente externo a la interfaz es tanto o más importante que el interno. Tal como se concluyó en el capítulo anterior, aquí no basta aceptar una “inversión” radical de los aspectos que puedan ser más centrales a los procesos cognitivos, sino sólo asumir que ambos ambientes, el externo y el interno, no pueden prescindir el uno del otro.

Es interesante hacer notar que para Clark (2008, entre otros), al cuerpo biológico de agentes como nosotros también se le puede atribuir un carácter de “interfaz”. La importancia de esta alusión es que tal carácter sirve para fundamentar que el cuerpo, en tanto interfaz, no establece los límites de todo aquello que debiera considerarse como cognitivamente relevante en cualquier aproximación correcta al estudio de nuestra conducta inteligente. Del mismo modo, para un investigador de la comunidad de RBM que comparta dicha visión, la acción intencional humana se considera tal no porque posea propiedades cognitivas intrínsecas, sino porque las propiedades cognitivas más interesantes para la construcción de modelos serían aquellas que adquieren relevancia (o son significativamente funcionales) con respecto a los usos que se materializan en algún contexto externo. De aquí, la importancia de los “contextos de uso” en los que se podría verificar la distinción entre propiedades cognitivas, propiamente tales, y propiedades meramente físicas o estructurales. El rol cognitivo de los modelos en tanto artefactos, sólo susceptible de ser discriminado en contextos o ambientes externos a la interfaz

de interacción, es el mismo que se suele atribuir a los objetos físicos cuando se les denomina “objetos intencionales”, entendiéndose que dicho rol es discriminado por los mismos procedimientos contextuales que harían patente alguna virtud epistémica productiva. A partir de aquí, es posible vislumbrar, por lo tanto, un a su vez de criterio general para los constreñimientos de un RBM inferencial en ciencia. Dicho criterio se podría explicitar de la siguiente manera:

**Criterio de funcionalidad epistémica contextual:** *Para que algo (e.g. objetos físicos, prácticas modeladoras, investigadores, etc.) sea cognitivamente constituyente de un proceso cognitivo (i.e. un socio cognitivo), tanto en la manipulación de modelos como en la formación de nuevas hipótesis determinadas por dicha manipulación, debe ser primero reconocido como algo cuyo valor epistémico está en relación a sus usos contextuales.*

El criterio aludido suele ser reportado explícitamente en las investigaciones de RBM, normalmente con alusión a una fundamentación filosófica construida en torno a ciertos planteamientos fundamentales de Peirce. Según señalan autores como Magnani (2002) y Pohjola (2007), entre otros, el RBM tendría directa relación con la visión de Peirce acerca de cómo los pensamientos llegan a ser representaciones y sistemas de representaciones. Por lo que logramos comprender de lo planteado por Pohjola, el reconocimiento de las *condiciones preliminares del pensamiento* que Peirce especifica en su Faneroscopia (i.e. la primera categoría de disciplinas filosóficas para Peirce) para la producción de un pensamiento con significado [*meaningful thought*] tendría alguna similitud con el reconocimiento de la funcionalidad contextual de un modelo a la cual los científicos le atribuyen un valor epistémico (en tanto *affordance* bio-ambiental y/o perceptualmente basada<sup>60</sup>). La razón de esta similitud, a partir de la cual se defienden los derechos causales del ambiente en la cognición, puede expresarse de la siguiente manera. Para Peirce, la producción de un pensamiento con significado requiere, por un lado, que dicho pensamiento sea representado conceptualmente, y por otro, que posea un contenido cualitativo - que es lo que haría posible la identificación y el reconocimiento. Para Pohjola, este contenido (propio de la primera de tres categorías preliminares del pensamiento que Peirce llama “primeridad”<sup>61</sup>) también está a la base de un RBM, toda vez que es consistente con la evidencia disponible que apoyaría cierta estructuración del razonamiento en términos analógicos. Luego, uno podría concluir, provisionalmente, que se está frente a una postura donde la adscripción de funciones epistémicas que una comunidad científica determinada hace con respecto a determinados objetos, o dispositivos físicos, se entiende como parte de lo que constituye el “Fanerón” o (como se caracteriza en alusión a Peirce) “*la totalidad de todo lo que de algún sentido o manera se le presenta a la mente, sin importar si aquello corresponde a algo real o no.*” (Pohjola, 2007, p. 317)

Ciertamente, la Faneroscopia en la teoría Semiótica de Peirce cumple un rol distinto a lo recién planteado, ya que Peirce intenta dar cuenta de cómo las ideas (específicamente de la terceridad) se reflejan y representan en la lógica y en las formas gramaticales de las

---

<sup>60</sup> La aproximación Gibsoniana a la noción de *affordance* ha sido denominada como de carácter bio-ambiental, en contraste a la aproximación perceptual propuesta por Donald Norman (1998). Según lo que plantea este último psicólogo cognitivo, los « recursos externos » perceptuales, a pesar del estatus convencional que puedan tener, son tanto o más importantes que los otros.

<sup>61</sup> La segunda categoría fundamental de “segundidad” se refiere a la *reacción*. Esta categoría se preocupa de las distinciones, y hace posible la diferenciación de las experiencias, tal como tener o no tener dolor. La “terceridad”, por su parte, es la categoría de la representación.

proposiciones<sup>62</sup>. Pero para dejar bien en claro aquello que se plantea desde la perspectiva de RBM en cuestión, con relación al criterio anteriormente propuesto, daremos alguna ejemplificación.

Comencemos con un ejemplo proveniente desde del ámbito de la teorización acerca del diseño y creación de tecnologías, que sirve para ilustrar aspectos del rol que los modelos desempeñan en un ámbito científico. Vermass y Houkes (2006) han propuesto que las funciones de un artefacto técnico dependen de una “adscripción de funciones” que conecta ciertas propiedades estructurales de un artefacto técnico con los usos intencionales de un agente o diseñador. La intencionalidad, en este caso, debe entenderse sobre la base de una “acción dirigida a una meta”. De aquí se sigue la noción de “plan de uso” como un factor que regula el razonamiento. Según esta noción, el éxito que tenga la creación y el diseño de un artefacto técnico dependerá de la manera en que el agente diseñador logra comunicar a los potenciales usuarios la funcionalidad que dicho diseñador le ha adscrito a ciertas partes estructurales del artefacto, y con ello garantizar, hasta cierto punto, la consecución del plan de uso. A modo de ejemplo, la medición de la actividad cardíaca y el medidor de pulso (a veces llamado “computador de muñeca” por su complejidad), utilizados para el análisis de rendimiento del entrenamiento deportivo, pueden ser descritos en términos de adscripción de funciones. En este caso, la medición de cierta tasa de actividad cardíaca cuenta como efecto del entrenamiento en el contexto de un deporte de exigencia. Claramente, esta descripción depende de hechos sociales tales como las funciones del artefacto y el propósito de sus usos<sup>63</sup>.

De acuerdo al desarrollo expositivo que se ha realizado sobre las aproximaciones al RBM hasta este punto, no resulta difícil asimilar lo recién planteado al rol del modelamiento científico. Los planteamientos instrumentalistas de Giere sobre teorías científicas en tanto modelos, así como las concepciones de modelos en tanto objetos dinámicos de simulaciones corporalizadas y situadas socio-ambientalmente que defienden Nersessian y Magnani, entre otros, parecen tener una expresión común en la noción general de estructuras modélicas con funciones socialmente determinadas y orientadas a metas de uso. Sin embargo, hay un punto adicional que conecta esta concepción pragmatista y contextualista de los modelos como artefactos o interfaces, y este dice relación con la noción de “razonamiento práctico”. Como se verá un poco más adelante, este tipo de razonamiento hace alusión a un proceso reflexivo con poder deliberativo, vale decir, un proceso que, teniendo alguna acción como contenido intencional, pueda al mismo tiempo dar lugar directamente a dicha acción. Desde esta perspectiva, lo que caracteriza al razonamiento práctico es que el sistema normativo subyacente está orientado a la regulación de la acción intencional. Por cierto, hay mayores consideraciones relacionadas con la controversia de si este tipo de razonamiento, tal como se ha caracterizado brevemente aquí, corresponde o no a un tipo de razonamiento genuino, pero ese no es el propósito que nos interesa por ahora. Más bien, mostrando cómo esta noción sirve de puente entre el rol de la acción abductiva extra-teórica y el descubrimiento en ciencia, lo que intentaremos es llegar a un segundo criterio general para los constreñimientos de un RBM en la generación de hipótesis explicativas. Al igual que en el caso del criterio ya propuesto, el nuevo criterio supondría

<sup>62</sup> En este caso, las últimas son dependientes de las primeras. Así, las formas generales del pensamiento se representan en formas simbólicas y el significado es dependiente tanto de la estructura como del contenido del pensamiento representado en los símbolos

<sup>63</sup> De hecho, incluso se puede plantear, con relación a este tipo de aproximaciones, un vínculo con la concepción de una regla constitutiva propuesta por Searle (1995), en el contexto de una ontología social: “X cuenta como Y en contexto C”. Según esta regla, la creación de una institución involucra necesariamente un acto performativo que la hace accesible y aceptable por otros.

un tipo de supuesto metodológico y conceptual subyacente a las aproximaciones al RBM en cuestión.

### 4.1.2.2. Desde la manipulación epistémica a la cognición eco-cognitiva

Para explicitar lo recién señalado, es conveniente reforzar la idea de que, si bien las propiedades estructurales de un objeto físico entendido como interfaz requieren la atribución y comunicación de funciones epistémicas, la intencionalidad (descrita como una acción orientada a la consecución de ciertas metas de uso) no es estrictamente un fenómeno socialmente distribuido. Tal como en el caso del diseñador de un artefacto tecnológico, no es necesariamente a partir del quehacer colectivo y distribuido entre un grupo de científicos que fabrica o hace uso de dispositivos modélicos desde donde *emergen* serendípicamente las propiedades funcionales más interesantes del modelo. Más bien, se requiere que los investigadores, individualmente, aporten las innovaciones preliminares sobre las que se han de interpretar virtudes epistémicas de aspectos salientes de lo modelado. Dado este requisito, lo que parece darse por sentado es que sólo después que dichos atributos son comunicados (o “socializados”) en la comunidad, y aceptados por la misma, podría tener lugar la explotación de modelos como apoyos u oportunidades para la generación de inferencias y el eventual surgimiento de hipótesis explicativas. De hecho, así es como Nersessian (2008) hace concesión de este punto, en el contexto del análisis de los aspectos creativos del cambio conceptual - proceso de cambio que ocurriría por medio de un continuo *bootstrapping* apoyado y constreñido por un RBM:

***“Aunque el cambio conceptual en ciencia es un fenómeno que tiene lugar a través de comunidades, son principalmente los científicos como individuos los que crean las innovaciones conceptuales que son investigadas, mayormente desarrolladas, y adoptadas por la comunidad” (p. 183)***

Dichas innovaciones conceptuales, como resultado de un proceso creativo, estarían determinadas por un RBM que hace explícitos los constreñimientos que orientan selectivamente los usos de ciertas características o funciones salientes de representaciones preexistentes. Estos usos, como se dijo, estarían guiados por una acción intencional del tipo “plan de uso”. Aquí, habría que aclarar que por ‘función’ se hace referencia a la capacidad relacional de uso que pueda observarse entre los agentes humanos y los modelos (o ciertas partes de ellos). Mientras que por ‘uso’ se hace referencia a aquellas regularidades que, a partir de ciertas características funcionales y estructurales del modelo físico, puedan ser objeto de atribución de virtudes epistémicas en contextos de uso habituales. Los constreñimientos que se hacen explícitos suelen ser, entre otros, de tipo espacial, temporal, causal, funcional, categorial, y matemáticos. Tanto a los modelos como a la construcción de modelos se les caracteriza como dinámicos, en el sentido de que apoyan o facilitan [*afford*] la realización de inferencias a través de procesos simulativos continuos. De aquí la convicción compartida por los investigadores del caso de que un RBM corresponde a un tipo genuino de razonamiento.

Las manipulaciones susceptibles de proporcionar simulaciones apropiadas para procesos inferenciales creativos se constituyen a partir de una regularidad en el uso práctico de estructuras modélicas dinámicas. Estos usos han dado lugar a la noción de “razonamiento práctico”. Dicha noción, derivada inicialmente de la contribución realizada por von Wright (1963 y 1971) al ámbito de la filosofía de las ciencias sociales, hace alusión a la acción intencional reconstruida como un razonamiento práctico, haciendo referencia a “medios” y “fines”, lo que se ha descrito en términos de un silogismo práctico:

**“Si una persona *p* se orienta a alcanzar cierta meta *G*, y posee la creencia *B* de que puede alcanzarlo haciendo *A*, entonces *p* comienza a hacer *A*. Este razonamiento práctico puede ser descrito como un silogismo de la siguiente manera: *p* intenta causar *G*; *p* considera que no puede causar *G* a menos que haga *A*. Luego, *p* comienza a hacer *A*” (citado en Pohjola, 2007, p. 325).**

El propósito de von Wright era describir las condiciones necesarias y suficientes para la acción intencional. Luego, para que una persona *p* realice intencionalmente *A*, ella debe considerar *A* como los medios necesarios y suficientes para causar *G*. El silogismo práctico expresa una exigencia lógica para el razonamiento práctico. Claramente, las personas suelen fracasar en la consecución de metas, y no siempre hacemos un razonamiento práctico para todas nuestras acciones, lo que dejaría como condición mínima *que una persona *p* necesite estar convencida de que hacer *A* causará *G**.

Aún cuando restringiésemos el silogismo práctico al contexto de un plan de uso, lo importante es que las concepciones de razonamiento práctico y la acción intencional (analizada en términos de medios y metas) permitirían defender una postura donde la el contenido de la acción intencional (y en particular de las hipótesis generadas para resolver alguna tarea determinada o explicar cierta anomalía específica) está determinada por sus consecuencias prácticas. De esto no se sigue inmediatamente que las manipulaciones epistémicas - supuestamente constitutivas de procesos cognitivos - sean ecológicamente determinadas, pero el compromiso conceptual con tal puente o conexión puede reconocerse (no muy convincentemente) si se examina lo que autores como Magnani (2009b) proponen como la dimensión epistemológica y eco-cognitiva de la abducción. Brevemente, la historia sería como sigue.

La abducción sería parte del un proceso explicativo en el que un agente cognitivo participa cuando se enfrenta a un fenómeno que no puede explicarse sólo a base de inferencias deductivas e inductivas. Específicamente, la *cognición abductiva* se trataría de un proceso de razonamiento invocado para explicar una observación confusa, oscura y/o anómala. Dada la conclusión aparente de que los procesos inferenciales activos en una lógica del descubrimiento y una lógica de la justificación tradicionales no han sido capaces de proporcionar un modelo del descubrimiento científico, aproximaciones multidisciplinarias acometen un enfoque alternativo. Dicho enfoque guarda compromisos experimentales que darían luces acerca de las denominadas “inferencias expertas y creativas”. En el ámbito del RBM, ya es un lugar común que las inferencias en general, y la abducción en particular, se realizan por medio de la creación, manipulación, adaptación y evaluación de modelos. Para Nersessian, por ejemplo, el foco está en los modelos conceptuales, entendidos como sistemas que la imaginación produce como análogos estructurales, funcionales o conductuales de los fenómenos que se pretenden representar explicativamente. En este contexto, las prácticas modeladoras están subyugadas a la solución de problemas que pueden conducir a la innovación conceptual. La base corporalizada y socio-culturalmente situada de dichas prácticas modeladoras se completan con un tipo de acción epistémica de carácter extra-teórico, que no sería incompatible sino que complementaria con los procesos inferenciales tradicionales. De hecho, para abordar dicha complementariedad, sería pertinente considerar lo ya parece un conjunto laxo de compromisos compartidos en el campo del RBM, que Nersessian (2008) resume en los siguientes puntos:

- El RBM es ampliativo y puede ser creativo.
- El RBM involucra *bootstrapping* <sup>64</sup> .
- Por medio de *bootstrapping*, el RBM es selectivo.

- El RBM está íntimamente ligado a la analogía.
- En el RBM, el *formato* de la representación posibilita distintos tipos de manipulaciones, posiblemente por medio de reclutamiento de procesos motores y perceptuales.

La acción epistémica de carácter extra-teórico mencionada estaría determinada por la naturaleza ecológica de la cognición. Inspirándose en este supuesto de la cognición, el estudio de la denominada cognición abductiva se ha caracterizado por al menos dos tendencias metodológicas relacionadas: por un lado, se ha puesto énfasis en el cierre de la brecha tradicional entre el estudio de los humanos y el resto de los animales. El cierre de dicha brecha, se fundamentaría en los hallazgos de la investigación comparativa en la que convergen áreas como la ciencia cognitiva, la etología animal, la psicología evolutiva, y la arqueología cognitiva. Por otro lado, aunque relacionado a lo anterior, se ha puesto el énfasis en el estudio de procesos motores y perceptuales, en el entendido de que este tipo de procesos estaría en la base de un RBM de tipo creativo y ampliativo de conocimiento. En consideración a lo anterior, se ha propuesto dar cuenta de una suerte de tránsito incremental que llegaría a procesos de generación de hipótesis de orden superior, partiendo de supuestos procesos relacionados de interacción ecológica de orden inferior. Eso nos da pie para desarrollar el segundo criterio que anunciaremos a continuación.

**Criterio de afinación eco-cognitiva:** La abducción está fundada en un proceso de afinación [attunement] eco-cognitiva, entre un organismo y un estado de cosas socio-bio-ambiental contemporáneo. Este proceso le proporciona al organismo grados variables de “potenciales poderes abductivos”, facilitados por la manipulación de mediadores epistémicos que flexibilizan la selección o creación de nichos cognitivos artificiales; la ampliación de posibles cursos de acción (explicitando información implícita); y la estabilización de hábitos de acción externa y de simulación interna.

En el caso del resto de los animales, se ha defendido la posibilidad de que cierto tipo de representaciones relacionadas con (o basadas en) la acción ecológica pueda servir de base al tipo de estados modélicos y simuladores internos típicamente humanos (Magnani, 2007b), como los que se asumen en el caso de la arqueología cognitiva que veremos a continuación. Lo que se sostiene, específicamente, es un tipo de representaciones intrínsecamente entrelazadas con la percepción y ciertas habilidades kinestésicas. Sin embargo, es curiosa la forma cómo esto se tiende a justificar con interpretaciones claramente *ad hoc* de la evidencia empírica disponible. A modo de ejemplo, considérese el siguiente caso, donde la denominada abducción visual hace referencia al rol que las imágenes visuales (sensoriales y perceptuales) jugarían en la construcción de hipótesis o en la explicación de determinados hechos observados.

En el ámbito de la arqueología, Shelley (1996) ha estudiado la abducción visual que enfatiza la representación icónica, tanto de las formas como de las estructuras. El caso que ahora interesa hace alusión a cómo la “historia causal” de las formas pareciera estar determinada por un proceso imaginativo que sirve para explicar la forma actual de ciertos objetos. Las dos observaciones siguientes, presentadas por Shelley, se refieren a un tipo de piedras cortantes denominadas conocidas como lascas que se han ido encontrando en yacimientos paleolíticos:

1. Las lascas tenían dos formas distintas: ‘diestras’ y ‘zurdas’.
2. La proporción de lascas diestras y zurdas era de 57:43 por ciento. (p. 156)

Según una descripción más detallada, en la forma de una lasca diestra, el extremo derecho es más largo que el izquierdo, y la superficie superior (que corresponde a la cara externa

---

de la piedra original) es suave, en contraste a la inferior (“diseñada” por el tallador) que es más áspera. La lasca izquierda es un espejo de la diestra.

Se puede aceptar que no resulta inmediatamente obvio por qué debieran haberse producido dos tipos de lascas, o por qué los términos “diestras” y “zurdas” son apropiados. Sin embargo, lo que debiera resultar interesante en esto es el *experimento de pensamiento visual* al que se alude para explicar las razones pertinentes. Según este experimento, los científicos aludidos por Shelley reportan algo como lo que sigue: si se atiende a lo que típicamente haría un fabricante de herramientas diestro a través de golpes reiterados, se apreciaría que dicho individuo tiende a tomar la piedra que usa para martillar con la mano derecha dominante, y aquella otra que se quiere tallar en la mano izquierda. La mano izquierda cumple básicamente la función de torno, orientando apropiadamente la posición de la piedra que se talla, con el fin de facilitar golpes sucesivos. Adicionalmente, hay una tendencia a girar la mano izquierda en el sentido de las manecillas del reloj, en la medida que se desprenden las lascas. Y así sucesivamente. Luego, una lasca zurda debiera ser fabricada por un tallador zurdo, siguiendo el mismo procedimiento, pero intercambiando el rol de cada mano.

Shelley razona que también puede haber sido el caso de que las denominadas lascas zurdas se partieran por donde la piedra tallada tiene fisuras naturales, pero un procedimiento de tallado predominantemente diestro también puede explicar la mayor cantidad de lascas diestras. Para confirmar esto, se ha recurrido a experimentos donde se lleva a cabo este tipo de procedimiento de tallado, sin recurrir a al procedimiento zurdo, hallándose un proporción de lascas diestras y zurdas muy similar a las encontradas en los yacimientos señalados (56:44 por ciento).

Lo que se destaca como significativo en este ejemplo es que ilustraría cómo las formas de ciertos objetos puede tener una importancia teórica en el contexto de la investigación científica, y cómo es que imaginar una “historia causal” de aquellas formas (en este caso el método de fabricación de lascas) *puede sugerir explicaciones* de las mismas. La parte que destaco en cursiva en la oración anterior me parece de interés, puesto que ejemplifica una manera recurrente de hacer referencia a la participación de aspectos externalistas como constitutivos de los procesos de razonamiento inferenciales responsables de la explicación. Sin embargo, no me parece muy especulativo sostener que dicho rol constitutivo resulta poco obvio, o difícil de aceptar, especialmente cuando no se visualiza claramente la naturaleza de los procesos inferenciales relevantes. Dicho de otra manera, aún cuando se aceptara que, en último término, la explicación final descansa en inferencias realizadas sobre la base de estructuras lingüiformes más tradicionales, no quedaría claro o bien cómo se vinculan constitutivamente las representaciones modélicas a las tradicionales, o bien cómo podría funcionar una inferencia en base a modelos independientemente de las inferencias tradicionales.

Al igual que en el caso de abducción visual presentado, otro ejemplo de *experimento de pensamiento visual*, en el mismo campo de la arqueología reconstructiva, tiene que ver con el desafío de abducir la características que deben haber tenido las estructuras superiores de aquellas edificaciones de las cuales sólo se han desenterrado sus restos inferiores. En este caso, se recurre a la construcción y destrucción experimental de posibles réplicas arquitectónicas, con el doble fin de comparar los restos destruidos con las ruinas originales y de considerar alguna posible evidencia aún desatendida en áreas de excavación circundantes. Esto último tiene la importancia de facilitar el potencial hallazgo de estructuras superiores desprendidas de la edificación original, muy difíciles de identificar a no ser por el estudio de aquellas otras “ruinas” producidas experimentalmente.

Esta manera de dar cuenta de la relevancia de la denominada abducción visual en la formación de hipótesis científicas es recurrente en los enfoques externalistas. A partir de estos casos, complementariamente, también se suelen proponer diversos esquemas representacionales para implementar simulaciones computacionales, como son el caso de los trabajos de Layton (1989) y Thagard et al. (1997), por mencionar algunos. La visión optimista de Shelley (1996) se resume así:

***“Un mejor entendimiento de la abducción visual tiene amplias implicancias. Promete proporcionar una teoría mejorada de los procesos cognitivos inherentes a la conducta científica [...]” (p. 158).***

Lo que Magnani (2007b) rescata de la abducción visual, así entendida, es que se estaría proporcionando mayor evidencia acerca de la manera en que los animales (incluido el humano) usan artefactos, no como meros instrumentos, sino como “mediadores epistémicos” que cumplen un rol fundamental en proceso de “construcción de nichos cognitivos”. Sobre estas bases, se sostiene que la generación de hipótesis, y la cognición en general, son fenómenos cuyo origen no es dependiente del lenguaje humano, sino más bien dependiente de facultades instintivas y no-lingüísticas. El rol que le cabe al lenguaje humano en esta visión es aquel propuesto por la paleo-antropología cognitiva, donde se hipotetiza que el lenguaje hace posible la integración, en ambientes sociales, de módulos de dominio específico preexistentes pero independientes en homínidos pre-lingüísticos. Estos módulos deben entenderse, básicamente, como complejas habilidades motoras aprendidas inicialmente por imitación o por procesos creativos independientes. La interconexión mencionada es la que habría propiciado la capacidad de fabricar herramientas y artefactos a través del proceso de “descorporalización de la mente”, ya descrito en la sección 2.1.1.

De lo anterior, se hipotetiza la existencia de estructuras representacionales que no necesitan ser lingüiformes, sino más bien meramente basadas en modelos, específicamente relacionados con la acción, y de base perceptivo-kinestética<sup>65</sup>. Esto no desconoce el posterior desarrollo de representaciones lingüiformes, pero tampoco específica de qué forma podrían surgir estructuras lingüiformes a partir de representaciones modélicas en tanto estructuras pre-lingüísticas. No obstante lo anterior, el “*interplay* productivo” vinculado al proceso de descorporalización de la mente mencionado es lo que está en el centro de un sistema iniciador de una actividad ecológica. Dicho sistema es caracterizado como una “externalización semiótica” (Magnani, 2007a), donde la selección o modificación apropiada de elementos es considerada como un caso de abducción selectiva y/o creativa, marcado por un proceso continuo de “extracción” y “creación” de *affordances* más o menos disponibles en el ambiente, el que estaría constreñido por una tendencia general de “afinación ecológica” entre el organismo y su entorno. Esta última noción, implícita o explícitamente, parece ser una de las más ubicuas en las investigaciones de RBM externalista, y permite evitar la necesidad de dar cuenta del por qué cierto agente “desearía” realizar, inicialmente, tal o cual acción.

La afinación ecológica en cuestión habría tenido un giro evolutivo especial a partir del desarrollo y enriquecimiento de la cultura material, especialmente con el uso de artefactos manufacturados artificialmente. Este giro evolutivo no sólo sería significativo por el rol de mediador epistémico que normalmente se le asocia a los artefactos que potencian algún tipo de “pensamiento instrumental” imposible de ser capturado desde una perspectiva minimalista de la cognición animal, o alguna exclusivamente basada en reflejos. Más bien, esta afinación tiene un valor epistémico superior en el contexto de una mayor amplitud de

<sup>65</sup> Estos estados serían distintos, a su vez, a los que se puedan atribuir a los casos clásicos de tropismo y condicionamiento, puesto que éstos últimos resultan invariantes desde el momento que son registrados por primera vez.

las oportunidades cognitivas (perceptuales, kinestéticamente corporalizadas, simulativas, etc.), donde dicha ampliación de posibles cursos de acción aumenta la “sensibilidad hacia sus consecuencias”. Esta sensibilidad toma la forma de una *mayor sensibilidad al carácter dinámico de los fenómenos y a la detección de fenómenos anómalos*. Luego, mientras que los reflejos condicionados son el resultado de un estímulo condicionado que lo refuerza, el afinamiento ecológico es el resultado de una operatividad instrumental de la mediación epistémica que refuerza la conducta del agente. En este sentido, dadas las diferencias corporales en organismos de distintas especies, las *affordances* con cierto valor epistémico e instrumental que ofrece un mismo ambiente son distintas para distintos organismos. Y esta es la misma razón por la cual distintos organismos fabrican “nichos cognitivos” conforme a al afinamiento ecológico que le sea apropiado. A partir de esta situación, se hipotetiza que las *affordances* artificialmente construidas permiten una conducta creadora de nichos más flexible que la determinada por el ambiente natural, lo que daría como resultado continuo la ampliación de posibles cursos de acción, y la estabilización de hábitos de acción externa y simulaciones internas.

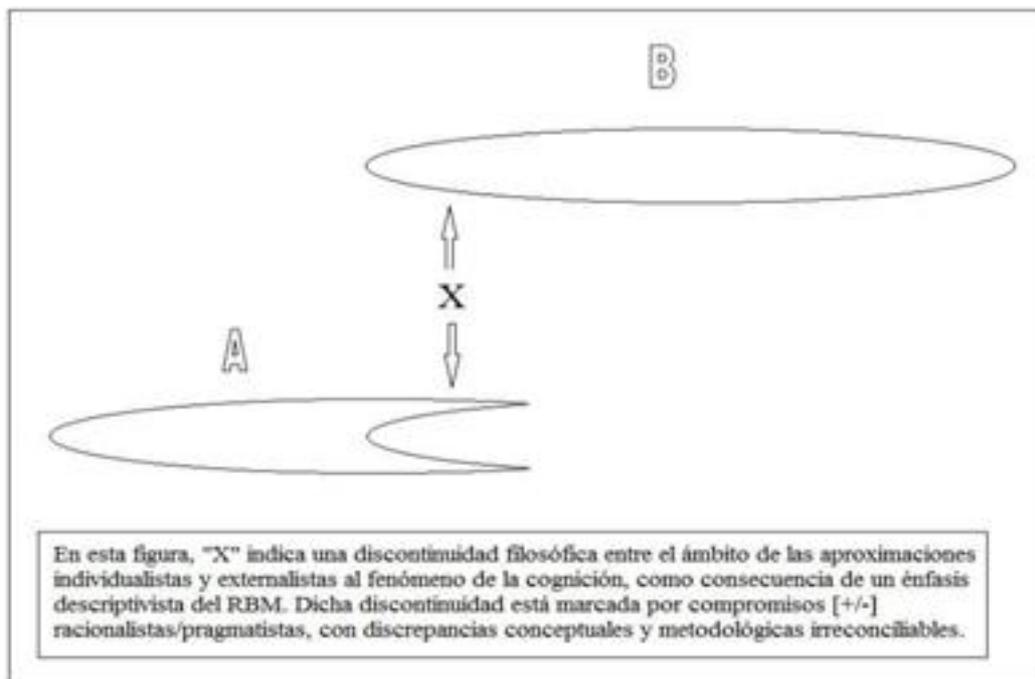
#### 4.1.2.3. Discusión sobre el estatus equívoco de los criterios de constreñimientos

Antes de pasar a la última sección de este trabajo, es relevante hacer una intermisión crítica con respecto a los criterios de constreñimientos propuestos. Estos criterios se han planteado a partir del caso de los denominados experimentos de pensamiento visual porque me ha parecido un caso paradigmático de la manera en que convergen múltiples enfoques conceptuales y experimentales del RBM, y del estudio de las “prácticas cognitivas” en ambientes reales. No obstante lo anterior, hay una pregunta que surge a raíz de dicha convergencia, marcada por posturas claramente descriptivas, como si un énfasis en tal postura pudiera desvincularse cómodamente de un enfoque filosófico normativo<sup>66</sup> más tradicional. Como sea que quiera formularse esta pregunta, lo que aquí interesa es plantear de qué manera, en el ámbito del RBM en cuestión, serían distintas las expresiones *abducción manipulativa* y “*manipulación abductiva*”. Ciertamente, la literatura que hemos examinado parece darle al segundo término el nombre del primero.

La distinción no parece trivial por cuanto podría estar escondiendo aquello que no es inmediatamente visible en aquella descripción didáctica que presentáramos en la figura 1 de la sección 1.1.3. La presente discusión simplemente pretende hacer patente una discontinuidad marcada por compromisos filosóficos que especifican discrepancias conceptuales irreconciliables. Según como lo ilustra la figura 3, creo que es pertinente relacionar un trasfondo filosófico de corte pragmatista con la noción de abducción en tanto *inferencia de base racional subyacente a cualquier acción provisional que permita aumentar oportunidades de conjetura o hipotetización productiva* (asumiendo que algo así exista). Sin embargo, tal noción tiene, como se describió acerca del segundo criterio de constreñimiento, una versión más radical asociada a un tipo de manipulación epistémica que pretende sentar las bases de una cierta “virtud epistémica funcional” en supuestos procesos de “afinación eco-cognitiva”. Una razón para considerar dichos criterios como ambiguos o equívocos es precisamente lo que pretende sondear la interrogante planteada en el párrafo anterior. Si es el caso de que la noción pragmatista más radical que hemos parafraseado como “manipulación abductiva” representa mejor la “búsqueda de

<sup>66</sup> Consideremos que incluso las aproximaciones más experimentales al estudio de la cognición están sujetas a marcos normativos, como podría ser el caso las hipótesis (y sus supuestas consecuencias) enmarcadas en la teoría de la evolución por selección natural, en tanto posibles mecanismos causales reguladores de la investigación empírica en las ciencias cognitivas actuales.

oportunidades” de la manipulación epistémica, entonces tendríamos que llegar a aceptar incluso la posibilidad de que cualquier descubrimientos científico serendípico, por ejemplo, es el resultado de un proceso racional determinado por una tendencia extra-mental a cometer errores por parte del agente cognitivo. Esto último se desprendería de una tendencia conductual observada por los defensores del RBM externalista a manipular cosas *como si* estuviera conjeturando posibles cursos de acción, imposibles de inferir sin recurrir a dichas prácticas supuestamente epistémicas. La expresión correspondiente (por paradójal que suene) sería que agentes como nosotros tienden a equivocarse a intencionadamente, porque a partir de prácticas manipulativas con “consecuencias epistémicas” se podría estabilizar un hábito o regla de acción, ineluctablemente conducente al desarrollo de una mayor sensibilidad hacia las consecuencias de determinadas plantillas [*templates*] de conducta. En la figura 3, cada una de las dos nociones pragmatistas de abducción señaladas se ubicaría en los extremos de B, siendo la más radical aquella que conlleva mayores problemas de justificación conceptual y experimental, por cuanto también tendría que dar explicación al hecho de que razonar es sólo una facultad más de la “agencia eco-cognitiva” propuesta. Este problema, según me interesa sostener, es insalvable porque la discontinuidad que separa los ámbitos de A-B y B, en la figura 3, es tan filosóficamente irreconciliable como lo son las aproximación a la inferencia de hipótesis explicativas desde una perspectiva pragmatista y otra racionalista, en sus versiones canónicas (i.e. donde la primera está regulada por la suma de sus consecuencias prácticas de aquello que se hipotetiza, y la segunda por una lógica formal).



*Figura 3: Aproximaciones irreconciliables a la inferencia de hipótesis explicativas, como resultado de una discontinuidad conceptual de trasfondo.*

Conceptualmente, podríamos señalar que el desarrollo de una mayor “sensibilidad hacia las consecuencias” de alguna determinada conducta estaría determinada por lo que anteriormente se llamó la condición mínima del razonamiento práctico (i.e. *que una persona p necesite estar convencida de que hacer A causará G*). En la disyuntiva de que dicha condición mínima pueda ser necesaria, suficiente o ambas, claramente tendríamos que aceptar que, no sólo tendría que tratarse de una condición únicamente

necesaria, sino que también debiera tratarse de una condición necesaria *en determinadas circunstancias*. Como se comentó recientemente, estas circunstancias son especificadas como un contexto de uso donde se manifestarían regularidades que requieren ser previamente adoptadas por una comunidad de determinada de científicos, a partir de ciertas contribuciones individuales. Esto más bien parece incorporar una contradicción, puesto que dicha “regularidad previamente adoptada” es precisamente lo que se estaría tratando de describir y dilucidar como característica epistémica de las prácticas interpretadas como productivas.

Metodológicamente, a su turno, una posible teoría del RBM podría salvaguardarse de ese riesgo a través de lo que hemos delineado como un razonamiento regulado por un “plan de acción” que afina reglas de acción eco-cognitivamente. Sin embargo, el problema no parece distinto en este caso, toda vez que, tanto al intentar describir prácticas científicas productivas, como al describir posibles conductas extra-mentales ecológicamente determinadas, lo que se está haciendo es mantener el constreñimiento de “virtud epistémica”, resultante de una inversión metodológica de base pragmatista y anti-representacionista.

Claramente, el énfasis empírico-descriptivista de la investigación acerca de un RBM es sólo compatible con una posición metodológica de inspiración pragmatista que soslaya la existencia de algún sistema de representación mental, implementado a través de vehículos representacionales estructurados, a favor de representaciones constructivísticamente *reificadas*. Esto último implica aceptar una visión conceptual subjetiva y no realista de tales estructuras representacionales, y, como *efecto* de aquello, aceptar también compromisos metodológicos como el de una individuación mecanicista de procedimientos epistémicos, tal como esta noción se caracterizó en la sección 4.1.

Por el contrario, si se abrazan convicciones realistas acerca de un sistema de vehículos representacionales de la cognición, tales procedimientos de individuación mecanicista no pueden menos que ser considerados arbitrarios y equívocos, en el sentido de que no es posible determinar con claridad aquellos mecanismos epistémicos esenciales en el tipo de descripciones propuestas acerca de las prácticas científicas como actividad cognitiva. Mientras no se pueda identificar aquel conjunto de rasgos compartidos en las “prácticas productivas” sin los cuales la expansión del conocimiento no podría ocurrir, la tendencia a incorporar aspectos de algún ambiente temporo-espacial determinado parece incluir una teorización de base empírica *ad hoc*— algo que también se puede reconocer en la diversidad de líneas de investigación y explicaciones mecanicistas de las ciencias cognitivas actuales. Bechtel (2010) expresa una visión similar de la siguiente manera:

***“Cómo las personas representan y simulan procesos en el mundo son temas de investigación dentro de ciencia cognitiva, y los filósofos que buscan entender explicaciones mecanicistas se benefician de esta investigación. Una característica importante de la investigación mecanicista es que usualmente se enfocan en sistemas de modelos [...], a partir de los cuales los investigadores pueden posteriormente intentar generalizaciones hacia otros sistemas. Pero los sistemas de modelos no son comúnmente elegidos por su tipicidad, sino porque favorecen la investigación. La generalización no consiste en tratar como universal aquello que se encuentra en casos particulares (como sería en el caso de las explicaciones inductivas simplistas), sino en la identificación de similitudes y diferencias entre los mecanismos”.*** (p. 364)

Dada esta evaluación del estudio mecanicista de las prácticas científicas en ambientes reales, parece difícil llegar a cumplir con alguna condición de adecuación que permita caracterizar el proyecto general de una teoría de RBM, dentro del marco del análisis integrado de mecanismos (o principios) cognitivos supuestamente comunes a las diversas líneas de investigación relevantes. Si esto es así, entonces tampoco habría una buena argumentación a favor de la manipulación de estructuras modélicas en tanto “agentes epistémicos autónomos”. Este es precisamente el punto que trataremos de reforzar y extender la siguiente y última sección del presente trabajo.

### 4.1.3. Condiciones de adecuación para una teoría del surgimiento de hipótesis por RBM

---

Para delinear algún posible tipo de condiciones de adecuación de la generación de hipótesis por RBM, y establecer las razones por las que aquí se considera inviable su satisfacción, presentaremos una narración unificada de los aspectos centrales de las distintas versiones examinadas a lo largo de este trabajo. Esto facilitará la identificación de los problemas que se quieren resaltar, sin asumir por esto que la integración programática de las distintas propuestas es algo claramente viable. Precisamente, trataremos de mostrar que tal integración no es (aún) factible, en la medida que las condiciones de adecuación ensayadas aquí no alcanzan a ser satisfechas sobre la base de mucho de los planteamientos conceptuales y metodológicos discutidos en las secciones anteriores.

El entendimiento del surgimiento de nuevas de hipótesis explicativas que se propone en el campo del RBM está regulado por prescripciones pragmatistas. Esta perspectiva se desarrolla como una reacción a las visiones filosóficas tradicionales que caracterizan el proceso en cuestión como irracional, ilusorio y oscuro. En general, los planteamientos alternativos caracterizan la *creatividad* y el *descubrimiento* como procesos racionales, en la medida que son constructivístamente (y experiencialistamente) entendidos como una relación compleja entre diferentes etapas inferenciales, susceptibles de ser claramente analizadas e identificadas en términos no tradicionales. Esto implica una multiplicidad de análisis que se enfocan en el rol de los denominados *mediadores epistémicos externos*, en la *abducción manipulativa*, o en la *abducción basada en modelos*. El primer tipo de abducción mencionado busca dar cuenta del carácter dinámico de la información (explícita e implícita) y de la manera en que ésta se materializa en segmentos variados de conocimiento, sujetos a una interacción causal continua entre estructuras representacionales internas y externas. El segundo tipo, a su vez, se funda en la premisa de que gran parte de nuestra actividad mental o cognitiva se realiza en base a modelos. Esta premisa adquiere su peculiar relevancia creativa cuando se aloja en procesos abductivos, a partir de los cuales se puede individuar una abducción basada en modelos, como ocurriría en el caso de la explicación de procesos superiores de cambio conceptual (Magnani, 1999a).

Dado que los procesos de creatividad y descubrimiento no pueden ser asumidos a priori (según las prescripciones pragmatistas), una línea correcta de investigación tendría que estar direccionada hacia el estudio de los efectos y las consecuencias que puedan tener las prácticas científicas asociadas a casos de descubrimiento. Lo que finalmente se entienda por creatividad y/o descubrimiento es algo que debiera ser informado por la suma de las consecuencias prácticas de las rutinas sistemáticas de expansión de conocimiento que la investigación sobre RBM logre identificar y describir. Esto pone el énfasis en el estudio de mecanismos epistémicos activados en diversos niveles de análisis,

especialmente donde se interprete que tales mecanismos cumplen un rol *explicitador* de constreñimientos implícitos. Luego, con el fin de establecer las bases de un proceso de generación de hipótesis apropiado, se sostiene que tal estudio se sustenta en la acción creadora de nichos cognitivos del agente activo. Esto permite describir regularidades en la acción epistémica consistentes con los desarrollos actuales del estudio experimental de la cognición en ambientes reales, específicamente, con las concepciones de la cognición distribuida.

Los casos típicos de un RBM se dan en torno al método histórico-cognitivo de propuesto por Nersessian (1999), donde se pretende re-enmarcar el problema de la formación y el cambio conceptual en ciencia. En este contexto, se pone atención a las prácticas empleadas por agentes humanos reales en la construcción, la comunicación y el reemplazo de las representaciones de un dominio específico. Los casos paradigmáticos de RBM serían la construcción y manipulación de representaciones visuales, los experimentos de pensamiento, el razonamiento analógico, entre otros. Si embargo, este enfoque estaría incompleto, razón por la cual, dentro de la comunidad de investigadores del RBM, se incluyen líneas de investigación en torno a procesos como el de la abducción manipulativa y las concepciones artefactualistas de la noción de modelo.

La incorporación de estas últimas líneas de investigación tendrían el propósito común de explorar maneras en que los agentes cognitivos pueden acceder a información tácita, o no disponible inmediatamente para procesos abductivos más tradicionales. Aquí, habría que entender (falazmente, según lo discutido en la sección 4.1.1.1.) que por limitaciones de procesos “tradicionales” se alude a limitaciones de procesos de tipo intracraneales. Sin embargo, lo que importa para una posible teoría de un RBM es que cierta conducta cognitiva extra-teórica (ni sentencial, ni basada en modelos) sería consistente con las tareas habituales y eficaces que agentes como nosotros realizan diariamente, sin la posibilidad inmediata de recurrir a su explicación conceptual (como se ilustraría en Hutchins, 1995). Esta línea de investigación dedicada al estudio de aspectos extra-teóricos no sólo ha servido para delinear prácticas abductivas multimodales, sino también para potenciar el desarrollo de programas computacionalmente adaptados para el descubrimiento o re-descubrimiento, por ejemplo, de teoremas matemáticos o hipótesis científicas (Magnani et al. 2005).

Lo que se propone en la base de un RBM es la existencia de plantillas de conducta que exhiben ciertas regularidades, vinculables a cierta actividad creativa característica de la participación de agentes científicos en sistemas inferenciales distribuidos. Estas plantillas son consideradas hipótesis de conductas que no tiene un carácter explicativo sino que meramente conjetural. Esto requiere aceptar que dichas plantillas de conducta [e.g. *construals*] pueden estar cognitivamente corporalizadas en el sistema constituido por el cuerpo y la mente del científico (como la internalización de la acción). Dichas plantillas pueden estar almacenadas para ser seleccionadas por el agente activo, en un caso similar al proceso de “abducción selectiva” de orden superior o teórico, o pueden ser construidas y afinadas a través de una *negociación epistémica* entre el agente y aspectos del ambiente, en un caso similar al proceso de “abducción creativa” de orden superior o teórico. Las regularidades que pueda expresar esta batería de plantillas no viene dada por sí sola, sino que se basa en un proceso doble que implica, a un nivel inferior, una afinación eco-cognitiva que tiene como efecto el desarrollo de una mayor “sensibilidad hacia las consecuencias”<sup>67</sup>

<sup>67</sup> Es pertinente destacar que esta sensibilidad parece entenderse como si se tratara de una “inferencia” de causa y efecto atribuida a dos cosas o eventos que tienden a ocurrir conjuntamente. En este sentido, “inferir” debiera entenderse como “esperar

de algún conjunto de prácticas potencialmente epistémicas; y, a un nivel superior, la comunicación y la atribución de un valor epistémico vinculado a los usos contextuales de un modelo (artefactual) que se adopten como prácticas habituales en una comunidad científica determinada cuyos miembros interactúan entre sí regularmente. A partir de estos procesos de acción o negociación epistémica continua, se gatillarían o facilitarían procesos de generación de hipótesis teóricos o más tradicionales, lo que completaría una teoría del surgimiento de hipótesis por RBM externalista (o distribuido). Hasta aquí el resumen referencial propuesto.

En la sección anterior, intentamos la identificación de dos criterios de constreñimientos para un RBM como el descrito anteriormente. Dichos criterios fueron discutidos y finalmente evaluados como equívocos. A continuación, se proponen tres condiciones de adecuación que, sin constituir un listado exhaustivo, debieran ser satisfechas para que la narración anterior tuviera cierta viabilidad epistémica como la mejor teoría disponible para el surgimiento de hipótesis nuevas por medio de un RBM:

1. Los usos contextuales atribuidos a un modelo externo, adoptados por una comunidad de científicos determinada, son una condición de adecuación para el desarrollo de plantillas de conducta (en tanto “hipótesis de conducta” corporalizadas)
2. Las plantillas de conducta, basadas en “reglas de acción” afinadas ecocognitivamente. Sobre el desarrollo de plantillas de conducta en base a la noción de “reglas de acción” estabilizadas en ambientes cognitivos apropiados, ver Pino, 2010., son condición de adecuación para el desarrollo de una sensibilidad hacia las consecuencias de la acción epistémica
3. La manipulación abductiva (según como la “abducción manipulativa” se caracterizó en la sección anterior) es condición de adecuación para explicitar información tácita útil para la inferencia abductiva de hipótesis explicativas por RBM.

Para reforzar las razones ya discutidas respecto a los problemas incluidos en estas condiciones, habría que insistir en el claro contraste entre dos aproximación pragmatistas generales del RBM, donde se abordan nociones de modelos conceptualmente irreconciliables. Estas dos aproximaciones pueden denominarse “instrumentalista” y “artefactualista”, respectivamente. La primera, ubicada hacia el extremo izquierdo de B, en la figura 3, está mejor representada por la visión de Giere, según la cual la *forma* de un modelo es su condición de adecuación para la representación de modelos (en función de un “buen calce” con la realidad), y la *función* del modelo es la condición de adecuación para los propósitos del científico. La segunda, ubicada hacia el extremo derecho de B, está mejor representada por los planteamientos de Knuuttila. Según esta visión, un RBM no requiere contar una noción representacionalista de modelo, porque es erróneo asumir que los modelos representan aspectos del mundo. La concepción de modelo más apropiada es la de un artefacto epistémico.

Entre las dos aproximaciones mencionadas, como ejemplo de cuánto varían los enfoques del RBM, el valor epistémico que se les atribuye a los modelos difiere sustancialmente en uno y otro caso. Mientras que en la primera los modelos sirven a los propósitos del agente intencional potenciando o acomodándose a procesamientos de orden intracraneal, en la segunda se asume cierta agencia cognitiva autónoma relativa a la materialidad del modelo. En otras palabras, el modelo artefactual tendría una virtud epistémica que radica tanto en su diseño intencional como en los usos que su propia materialidad constriñe, lo que sería capturado por la noción de *affordance*. Entre estas dos

que ocurra algo a partir de otra cosa”, por lo que uno podría sostener que detrás de dicha “sensibilidad” se esconde una confusión entre *causalidad* y *expectativas*.

---

aproximaciones, se ubican otras como las de Nersessian y de Magnani. Ambas contribuyen con dimensiones adicionales distintas a la multiplicidad de análisis que puede tener el estudio del RBM, pero no son reticentes a una unificación programática que no logra resolver las discrepancias conceptuales discutidas a lo largo de este último capítulo y contenidas a lo largo del ámbito de B (fig. 3).

Dado que todas las aproximaciones mencionadas comparten algunos o todos los planteamientos alternativos de la cognición distribuida, corporalizada, socio-ambientalmente situada, eco-cognitivamente determinada, etc., parece sensato exigir que, asumiendo que las condiciones especificadas anteriormente son correctas, una teoría unificada del RBM debería cumplir con todas las condiciones de adecuación. Es claro que no es así, pero también es claro que las condiciones de adecuación señaladas no son independientes unas de otras. No satisfacer alguna crea problemas con la satisfacción de alguna de las demás.

Se dijo en la sección anterior que son aquellas visiones más radicalmente externalistas las que tienen mayores desafíos para una unificación programática, básicamente porque incorporan una cantidad mayor y más heterogéneas de niveles de análisis. Pero también debido a la aceptación ilimitada de formatos modélicos en tanto eso potenciaría su valor como portadores de algún tipo de conocimiento implícito. Esto podría ser una ventaja para las posturas que aún sostienen la centralidad del agente intencional en los procesos cognitivos donde los modelos juegan un rol importante, como es el caso de Giere. No obstante aquello, y al menos con relación al tipo de RBM que se ha querido examinar críticamente en este trabajo, no es el caso que pueda vislumbrarse una teoría unificada mientras no haya luces de criterios de constreñimientos inequívocos, ampliamente compartidos por los investigadores de la *MBR Community*. Lo mismo puede decirse de las condiciones de adecuación que puedan ser satisfechas para la manipulación de “supuestos agentes epistémicos autónomos” sin formatos definidos, en las diversas líneas de investigación como, por ejemplo, el estudio del razonamiento diagnóstico, visual, espacial, analógico o temporal.

## CONCLUSIONES

El presente trabajo de tesis ha pretendido constituir una exploración aclaratoria del rol del denominado Razonamiento Basado en Modelos (RBM) en la generación de hipótesis y el razonamiento en ciencia. En términos generales, dicho rol podría ser caracterizado como explicitador de constreñimientos implícitos necesarios para conformar una teoría alternativa de generación de hipótesis y descubrimiento científicos. En ese sentido, nuestra investigación se plantea como un examen crítico de aproximaciones que intentan defender un rol cognitivo de los modelos como representaciones artefactuales y, al mismo tiempo, desafiar un estatus representacionalista tradicional de las representaciones mentales, con el fin de configurar un RBM en tanto método de inferencia productivo (o ampliador de conocimiento).

La noción de RBM como método de inferencia se funda en la idea de un sistema inferencial distribuido, cuyas características principales han podido especificarse en el contexto de una interacción entre Ciencia Cognitiva y Filosofía. En este trabajo hemos intentado argumentar a favor de la tesis de que tal interacción establece un antecedente inicuo para el desarrollo de la noción de sistema inferencial distribuido. Mucho de la inviabilidad de esta noción ha sido evaluado críticamente en términos de tal antecedente general, y, específicamente, con respecto al enfoque pragmatista que inspira gran parte del quehacer y la reflexión de base experimental en ambos ámbitos de la interacción.

Empíricamente, el RBM se funda en el estudio de las prácticas científicas en ambiente real. Dicho estudio tiene un compromiso variable con aproximaciones metodológicas que defienden el desborde oportunista y anti-individualista de la agencia cognitiva, y por lo tanto, una demarcación poco clara de los procesos que subyacen a la conducta inteligente. Con el propósito de contribuir a la identificación de criterios sistemáticos e inequívocos de constreñimientos para un RBM externalista, así como de ciertas condiciones de adecuación mínimas para el surgimiento de hipótesis por RBM, se postuló la siguiente hipótesis de investigación: **si es posible identificar criterios sistemáticos e inequívocos de constreñimientos para el RBM en tanto *sistema inferencial distribuido*, y satisfacer condiciones de adecuación mínimas para la manipulación de representaciones externas sin formato definido, entonces los modelos como *agentes epistémicos autónomos* constituye la concepción unificadora disponible más adecuada para dar cuenta de una teoría del RBM externalista.**

Como se especificará a continuación, el tratamiento de los temas de cada capítulo ha permitido dotar de un grado de plausibilidad a una tesis central que cruza toda esta investigación: la dificultad para identificar una sistematización de criterios de constreñimientos y condiciones mínimas de adecuación especificadas en la hipótesis anterior, está sesgada por los efectos de una tensión filosófica subyacente entre Pragmatismo y Racionalismo. Esta tensión se manifiesta en términos de un tipo de inversión pragmatistas de orden metodológico-conceptual, a través de la cual se intenta no sólo desafiar los tratamientos tradicionales del surgimiento de hipótesis y el razonamiento creativo en ciencia, sino que también sustituir supuestos y preguntas de investigación básicas. En este contexto, se ha terminado por impugnar la hipótesis de investigación donde – para expresarla en términos equivalentes a los planteados anteriormente –, el

---

rol de los “agentes epistémicos autónomos” entendidos como noción modélica unificadora del RBM externalista, depende de la identificación de constreñimientos inequívocos de tal razonamiento en tanto sistema inferencial distribuido, y de la satisfacción de condiciones mínimas para una manipulación epistémica de modelos artefactuales sin formato definido.

En el capítulo I, se presentaron algunas nociones y distinciones metodológicas preliminares para comenzar una indagación crítica del RBM externalista. Asumiendo que el estudio del quehacer científico está regulado por principios filosófico-experimentales de corte pragmatista, se propuso un esquema de transición de la externalización del RBM. El propósito de este esquema fue ilustrar el tipo de antecedentes que desde las ciencias cognitivas actuales motivan concepciones de razonamiento alternativas a la concepción racionalista tradicional, en términos de una serie de transiciones relativas a la supuesta interacción causal entre los componentes de la triada cerebro-cuerpo-ambiente que regula los argumentos de cognición activa. Con este esquema de trasfondo de la externalización cognitiva del RBM quedó estipulado, provisionalmente, el lugar que ocuparía el denominado “agente epistémico autónomo” pertinente a la hipótesis central de esta investigación.

En el capítulo II, se realizó un examen de diversas concepciones modélicas relacionadas a un RBM, con el fin de dar cuenta del posible desarrollo conceptual que ha dado lugar a la visión más externalista de *representación modélica*, caracterizada como un tipo especial de re/presentación artefactualista. En primer lugar, se estableció la equivalencia de esta concepción con la idea de una *agencia epistémica autónoma*, noción que terminará siendo impugnada a los largo del capítulo IV. Posteriormente, y como un efecto de la influencia ecléctica de los enfoques de la cognición alternativas al entendimiento hegemónico tradicional del agente psicológico intencional de la Ciencia Cognitiva clásica (presentados en el capítulo I), las concepciones de representación modélica fueron tratadas en términos de los siguientes contrastes.

Por un lado, se abordaron las características de una representación modélica pertinente a un RBM, en contraste a los modelos mentales propuestos por Johnson-Laird. En este caso, se hizo hincapié en que los compromisos que el RBM adquiere con las descripciones de los mecanismos cognitivos provenientes de las actuales ciencias cognitivas implican el abandono de la tesis de modelos mentales amodales fundados en métodos de inferencia tradicionales. Esto se resume de la siguiente manera. Johnson-Laird plantea que su teoría de modelos mentales introduce un nuevo tipo de representación mental, sin alterar el supuesto de que estas nuevas representaciones se relacionan con el razonamiento por medio de un sistema capaz de derivar inferencias deductivas e inductivas a partir de representaciones internas. En otras palabras, lo que diferencia un tipo de razonamiento llevado a cabo a través de la manipulación de representaciones de estructuración proposicional y un posible tipo de razonamiento basado en la manipulación de modelos amodales no es la presencia o ausencia de relaciones lógicas. Más bien, la diferencia radica en el tipo de representaciones sobre el cual operarían dichas relaciones lógicas. Alternativamente, los enfoques de RBM externalistas tienden a abandonar más o menos radicalmente la centralidad de un razonamiento de base lógica, y también el requisito de que las representaciones mentales modélicas tengan un carácter modal.

Por otro lado, se lleva a cabo un contraste entre representaciones modélicas desde la perspectiva de un ‘razonamiento basado en el agente’ y un ‘razonamiento basado en modelos’. Esto implicó abordar los rasgos centrales de la *noción de modelo como representaciones modales extra-mentales*, con el propósito de especificar la medida en que los enfoques de RBM externalistas se pueden entender como el resultado de una inversión empírico-conceptual de inspiración pragmatista.

Una de las conclusiones de este segundo capítulo es que el hecho de que el RBM acepte nociones vagas de representaciones modélicas caracterizadas por no poseer un formato definido es reflejo de la centralidad que la idea de representación ha tenido en la explicación de procesos y mecanismos cognitivos. Esto no justifica el hecho controversial de que bajo el rótulo de RBM converjan nociones que sugieran la posibilidad de una grave incompatibilidad entre concepciones modélicas diversas, así como también entre unidades de análisis asociadas a tales concepciones. Sin embargo, puede ser indicativo de las razones detrás de la persistencia en el uso del término 'representación', aún cuando lo que se busque es enfatizar un supuesto rol epistémico de objetos extra-mentales o aspectos físicos del ambiente externo al agente. En este contexto, se estipuló que, de ser cierto un RBM externalista, la concepción modélica pragmatista más representativa del mismo se puede delinear como a una *re/presentacionista* constructivista reificada, cuyo formato oportunista está determinado por el formato que posea aquello que se considere como el tipo de información participante del procesamiento cognitivo, regulado por el propósito de guiar la acción exitosa de un agente cognitivo en el contexto situacional de una tarea determinada. Ya que esta información se define laxamente como cualquier tipo de *datos* que incremente la base de conocimiento del agente, *la forma de la información es la forma de la re/presentación*, pero esa es una aseveración que uno puede interpretar, más bien, como la indefinición de algún formato para la noción de re/presentación. Esta concepción vaga - más bien un esbozo general - de representación modélica de un RBM externalista, estaría a la base de los enfoques más radicales que proponen subsumir los hallazgos de un campo de investigación ampliamente diversificado.

Según esta concepción de representación externa, los modelos no actúan como una imagen [*picture*] de alguna parte de la realidad, sino más bien como una suerte de *realidades artificiales paralelas*. Y cuando esta última noción se toma en un sentido real o literal (y no sólo metafórico, por ejemplo), se puede entender por qué los modelos artefactuales suelen ser considerados un tipo de *agente epistémico autónomo*. De aquí surge la línea de investigación que le atribuye un rol cognitivo a los instrumentos investigativos, y a su participación en sistemas distribuidos, que es donde tendrían lugar ciertos procesos capaces de transformar tareas cognitivas. Esto implica aceptar la existencia de un sistema en el que tanto los investigadores como los instrumentos de investigación actúan como socios causales en los procesos responsables de posibles hallazgos científicos. Lo anterior, nuevamente, fue considerado en esta investigación como la inversión de una visión de razonamiento tradicional, toda vez que la intuición de que los individuos aprenden de o a partir de los modelos es sustituida por la idea de que los individuos aprenden *con* los modelos. Según nuestra conclusión, aquellos planteamientos acerca de la expansión de conocimiento como parte de un tipo de aprendizaje extra-cerebralmente distribuido.

El capítulo III examina dos temas de interés dentro de la problemática de la formación de hipótesis por RBM en tanto método de inferencia. Estos temas dicen relación, por un lado, con dos aspectos de un razonamiento creativo distintos al tradicional: la abducción basada en modelos y la abducción manipulativas, y por otro lado, con el escrutinio crítico de la noción de *delegación cognitiva* subyacente a los roles epistémicos de objetos y procedimientos extra-mentales.

Asumiendo que el tratamiento de la abducción sobre la base de representaciones de estructuración supuestamente proposicional es insatisfactorio, los tipos de razonamientos abductivos señalados suponen que o bien no tienen la centralidad de los tratamientos proposicionales asumida originalmente a partir de las propuestas de Pierce, o bien podrían

ser complementarios (de una manera poco interesante para los investigadores en RBM) a los aspectos alternativos del razonamiento creativo. El énfasis de este capítulo no está en la abducción basada en modelos, cuyos defensores la sitúan como un proceso intracraneal - y por lo tanto, principalmente centrado en el agente intencional -, sino más bien en el surgimiento de un tipo de planteamientos que considera dicha abducción como una explicación incompleta del razonamiento creativo. En este sentido, se dio cuenta de la abducción manipulativas como la incorporación de un aspecto supuestamente soslayado por la abducción basada en modelo. Este aspecto dice relación con casos de razonamiento científico donde lo central del proceso creativo se explicaría en términos de una explotación crucial del ambiente, caracterizada por una interacción semiótico-cognitiva entre agente y ambiente. La idea central que se ponderó en este capítulo es aquella donde, a partir de dicha interacción, tiene lugar la creación de mediadores epistémicos en tanto agentes artefactuales que cumplen el rol de “interpretantes explicitadores de información tácita”, los que eventualmente gatillarían procesos de inferencias más tradicionales para la posterior generación y selección de hipótesis.

La primera conclusión surgida del examen de tales cuestiones dependen de una posible confusión entre asumir cierta ontología del formato de las representaciones internas y la adscripción de cierta semánticidad a formatos que sólo pueden ser dependientes de la capacidad de las personas para pensar en *cómo algo podría lucir en tales o cuales formatos*. Relacionado a lo anterior, se mantiene la permanente oscuridad con respecto a una posible *complementariedad* entre aspectos exclusivamente internos del razonamiento creativo (i.e. modélico y proposicional) y aspectos radicalmente externalistas del RBM. Según se destacó en la conclusión de la primera parte del capítulo en cuestión, mientras que las representaciones internas están dotadas de un carácter más o menos representacionalista (con matices que pueden hacer crisis al incorporar las representaciones corporalizadas), las externas tienden a abandonar un compromiso representacionalista, a favor de lo que se ha caracterizado como una re/presentación epistémica artificial. De aquí la dificultad de concebir una interacción literal entre concepciones representacionalistas y anti-representacionalistas de los vehículos de la cognición. En este caso, nos pronunciamos sosteniendo que dicha dificultad tiene que ver con dos problemas relacionados entre sí. Por un lado, tal complementariedad no cuenta con una lógica sobre la cual dependan prescripciones normativas que responda, por ejemplo, a “cómo se debiera proceder al enfrentarse a un proceso inferencial de tales características”. Por otro lado, la inversión pragmatista que hemos mencionado con anterioridad, según la cual el pensamiento y los procesos razonamiento creativo se entienden como la internalización de la acción de un agente cognitivo, en vez de entender a la acción del agente cognitivo como la externalización del pensamiento. Finalmente, se concluyó que tales apreciaciones permean la propuesta de un RBM en tanto sistema inferencial distribuido, toda vez que se corre un serio riesgo de trivialización de los supuestos constituyentes de un sistema inferencial de ese tipo. La trivialización en cuestión se manifiesta en el hecho de que lo que normalmente realiza un investigador del RBM externalista es la descripción de las circunstancias en las que ha tenido lugar un proceso inferencial exitoso. Sin embargo, de aquí no se sigue un entendimiento sustantivo de cómo funciona y qué constituye exactamente tal sistema inferencial.

Finalmente, en el capítulo IV se presentaron y discutieron los antecedentes que se estimaron relevantes para establecer criterios de constreñimientos para un RBM externalista, así como ciertas condiciones de adecuación que debiera satisfacer una posible teoría de manipulación epistémica para la generación de hipótesis. Específicamente, se enunciaron dos criterios de constreñimientos (**de funcionalidad epistémica contextual,**

**y de afinación eco-cognitiva)** y tres condiciones de adecuación interrelacionadas. Los criterios enunciados resultaron tener un carácter controversial e inequívoco, y ninguna de las condiciones explicitadas en el último capítulo parece estar cercana a satisfacerse por los enfoques de RBM externalistas estudiados. Sobre la base de estas razones específicas, se puede ratificar que la hipótesis de investigación no ha sido confirmada.

Si se quisiera dar continuidad a la presente investigación, sería pertinente ahondar en alguno de los temas que aquí no se pudieron clarificar mejor, especialmente con respecto a la tesis de la *complementariedad* entre representaciones modélicas internas y las externas, especialmente en el contexto de una suerte de “trivialización pragmatista” de la explicación del RBM a nivel de mecanismos cognitivos. En este contexto, también sería atractivo atender con mayor detalle la denominada *hipótesis del continuo* según la cual el razonamiento científico es explicado como una mera extensión del razonamiento ordinario, y cómo podría manifestarse coherentemente la presencia o ausencia de determinadas estrategias cognitivas. Finalmente, será necesario indagar en el tipo de formato de representación modélica para un RBM, ya que si la revolución cognitiva y la contra-revolución cognitiva tienen ambas algo de verdad, el formato, en tanto vehículo representacional, es importante.

---

# REFERENCIAS

- Adams, F. y Aizawa, K. (2001). "The bounds of cognition." *Philosophical Psychology* 14, N° 1: 43-64
- \_\_\_\_\_. (2008). *The Bounds of Cognition*. Malden, MA y Oxford: Blackwell Publishing
- \_\_\_\_\_. (2009). "Why the mind is still in the head." En: P. Robbins y M. Aydede (eds.), *The Cambridge Handbook for Situated Cognition* (pp. 78-95). New York: Cambridge University Press
- Armstrong, D.M. (1968). *A Materialistic Theory of the Mind*. London: RKP
- Bailer-Jones, D.M. (1999). *Tracing the development of models in the philosophy of science*. En: N.J. Nersessian, L. Magnani, P. Thagard (Eds.), *Model-based Reasoning in Scientific Discovery*, pages 23-40, New York: Kluwer Academic/Plenum Publishers
- Barsalou, L. W., (1999). "Perceptual symbol systems." *Behav Brain Sci* 22: 577–660
- Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence* 47: 139-159
- Artemov, S. (2008). "The logic of justification." *The Review of Symbolic Logic* 1(4), 477-513
- Bechtel, W. (2008). *Mental mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. London: Routledge
- Bechtel, W. (2010). "How can philosophy be a true cognitive science discipline?" *Topics in Cognitive Science* 2: 357-366
- Bechtel, W., y Graham, G. (1998). *A Companion to Cognitive Science*. Oxford: Basil Blackwell
- Burge, T. (1986). "Individualism and psychology." *Philosophical Review* 95: 3-45
- Clark, A. (1997). *Being There : Putting the Brain, Body, and World Together again*. Cambridge, MA: MIT Press
- Clark, A. (2001). *Mindware: an introduction to the philosophy of cognitive science*. New York : Oxford University Press
- Clark, A. (2003). *Natural-born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford: Oxford University Press
- Clark A. (2007). "Curing cognitive hiccups: A defense of the extended mind." *J Philos* 104(4):163–192
- Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press
- Clark, A. y D. Chalmers. (1998). The extended mind. *Analysis* 58, no 1: 7-19. Reimpreso en *The philosopher's annual*, Vol. 21, ed. P. Grim, 1998, y en *Philosophy*

- of mind: Classical and contemporary readings*, ed. D. Chalmers. New York: Oxford University Press, 2002
- Craik, K. (1943). *The Nature of Explanation*. Cambridge: Cambridge University Press
- Di Paolo E. (2009). "Extended life." *Topoi* 28, no 1: 9-21
- Dourish, P. (2001). *Where the action is: The foundations of embodied interaction*. Cambridge, MA: MIT Press
- Dretske, F.I. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press
- \_\_\_\_\_. (1988). *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press
- Fodor, J. A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press
- Fodor, J. A. (1980). "Methodological solipsism considered as a research strategy in cognitive psychology." Reprinted in his *Representations*. Sussex: Harvester Press, 1981
- Fodor, J. (1984). "Observation reconsidered." *Philosophy of Science* 51: 23-43.
- \_\_\_\_\_. (1987). *Psychosemantics*, Cambridge, MA: The MIT Press/A Bradford Book
- \_\_\_\_\_. (1998). *Concepts: Where Cognitive Science Went Wrong*, New York: Oxford University Press
- \_\_\_\_\_. (2000). *The mind doesn't work that way*. Cambridge, MA: The MIT Press
- \_\_\_\_\_. (2008). *LOT 2: the Language of Thought Revisited*. (Oxford: Oxford University Press)
- Gabbay, D. M. y Woods, J. (2005). "Advice on Abductive Logic." *Logic Journal of the IGPL*, 14(2), 189–219
- \_\_\_\_\_. (2006). "A formal Model of Abduction." En: L. Magnani (ed.), *Model-Based Reasoning in Science and Engineering*, Individual Authors and College Publications, London, UK, pp. 301-309
- Gatti, A. y L. Magnani. (2005). "On the Representational Role of Environment and on the Cognitive Nature of Manipulation." En: L. Magnani y R. Dossena, eds., *Computing, Philosophy and Cognition*, Individual Authors and College Publications, London, UK, pp. 227-242
- Gibson, J.J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin
- \_\_\_\_\_. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA
- Giere, Ronald, N. (1988). *Explaining Science: A Cognitive Approach*. Chicago y London: The University of Chicago Press
- \_\_\_\_\_. (1992). *Cognitive Models of Science*. Minneapolis, MN: University of Minnesota Press
- \_\_\_\_\_. (2002). "Models as Parts of Distributed Cognitive Systems" [en línea] <<http://www.tc.umn.edu/~giere/MPDCS.pdf>> [consulta: 12 septiembre 2010]

- \_\_\_\_\_. (2004). "How Models are used to Represent Reality." *Philosophy of Science* 71, pp. 742-752
- Gooding, D. (1990). *Experiment and the Making of Meaning*. Dordrecht: Kluwer Academic Publisher
- Gooding, D. y T. R. Addis. (1999). "A simulation of model-based reasoning about disparate phenomena." En Magnani, L., N. Nersessian, y P. Thagards (eds.), *Model-Based Reasoning in Scientific Discovery*. New York: Kluwer Academic/Plenum Publishers, pp. 103-123.
- Gooding, D.C y T. R. Addis. (2004). "Simulation Methods for an Abductive System in Science." En Magnani, L. (ed.), *MBR'04: Proceedings of the Fourth International Conference on Model-Based Reasoning*, Pavia, Italy, December 2004.
- Greeno, J.G. (1998). "The situativity of knowing, learning, and research." *American Psychologist* 53: 5-24
- Grush, R. (2002). "The philosophy of cognitive science." En: Machamer y Silberstein (eds.), *Blackwell Guide to the Philosophy of Science*. Basil Blackwell.
- Hanson, N. R. (1958). "The Logic of Discovery" *The Journal of Philosophy*, vol. LV N° 25
- \_\_\_\_\_. (1960). "More on the Logic of Discovery" *The journal of Philosophy*, vol. LVII N°6
- Harman G. (1986) *Change in View: Principles of Reasoning*. Cambridge, MA: MIT Press
- \_\_\_\_\_. (1966). "The Inference to the Best Explanation" *Philosophical Review*, 74: 88-95
- Hoyningen-Huene, P. (1987). "Context of Discovery and Context of Justification." *Studies in History and Philosophy of Science Part A* 18 (4):501-515
- Hurtshorne, C. y Weiss, P. (eds.) (1958). *Collected Papers of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press. Volumes 7-8 edited by A. Burks.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press
- Hutchins, E. (1999) Cognitive artifacts. En: R.A.Wilson & F.C. Keil, (Eds.), *Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press (pp. 126-7)
- Johnson-Laird, P.N. (1983). *Mental Models*. Cambridge: Cambridge University Press. Cambridge, Mass.: Harvard University Press
- \_\_\_\_\_. (2001). "Mental models and deduction." *Trends in Cognitive Sciences* 5(10):434-442
- \_\_\_\_\_. (2006). "Mental models, sentential reasoning, and illusory inferences." En: Held, C., Vosgerau, G., y Knauff, M. (Eds.) *Mental Models and the Mind*. New York: Elsevier. Pp. 27-52
- Kirsh, D. y Paul Maglio. (1994). "On distinguishing epistemic from pragmatic action." *Cognitive Science* 18, 4, 513-549
- Knuuttila, T. (2005). "Models as Epistemic Artefacts: Toward a Non-Representationalist Account of Scientific Representation" [en línea] <<http://ethesis.helsinki.fi/julkaisut/hum/filos/vk/knuuttila/modelsas.pdf>> [consulta: el 30 junio 2011]

- Knuuttila, T. y T. Honkela. (2005). "Questioning External and Internal Representation." En L. Magnani y R. Dossena, eds., *Computing, Philosophy and Cognition*, Individual Authors and College Publications, London, UK, pp. 209-226.
- Koch, C. y F. Crick. (2001). "On the zombie within." *Nature* 411, p. 893 (21 June 2001)
- Leyton, M. (1989). "Inferring causal history from shape." *Cognitive Science* 13: 357-387
- Lewis, David. (1980). "Mad Pain and Martian Pain". In Block (1980a) Vol. 1, pp. 216–222
- Lewis, R. L. (1999). *Cognitive Modeling, Symbolic*. En: Wilson, R. y Keil, F. (eds.), *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: The MIT Press
- Li, P. y D. Li. (2006). "Scientific cognition as model-based reasoning." En: L. Magnani (ed.), *Model-Based Reasoning in Science and Engineering*, Individual Authors and College Publications, London, UK, pp. 51-65
- Magnani, L. (1999a). "Model-based creative abduction." En Magnani, L., N. Nersessian, y P. Thagards (eds.), *Model –Based Reasoning in Scientific Discovery*. New York: Kluwer Academic/Plenum Publishers, pp. 219-238
- \_\_\_\_\_. (1999b). Visual abduction in mathematical discovery, Technical Report. Department of Philosophy, University of Pavia
- \_\_\_\_\_. (2001a). *Abduction, Reason, and Science. Processes of Discovery and Explanation*. New York: Kluwer Academic/Plenum Publishers
- \_\_\_\_\_. (2001b). *Philosophy and Geometry. Theoretical and Historical Issues*. Dordrecht: Kluwer Academic Publisher
- \_\_\_\_\_. (2002). "Conjectures and manipulations: external representations in scientific reasoning." *Mind & Society* 3: 9-31
- \_\_\_\_\_. (2006a). "Multimodal abduction. External semiotic anchors and hybrid representations," *Logic Journal of the IGPS* 14(1): 107-136
- \_\_\_\_\_. (2006b). "Disembodying Minds, Externalizing Minds." En Magnani, L. (ed.), *Model Based Reasoning in Science and Engineering. Cognitive Science, Epistemology, Logic*. London : College Publications
- \_\_\_\_\_. (2007a). "Semiotic brains and artificial minds. How brains make up material cognitive systems," En: R. Gudwin y J. Queiroz, eds., *Semiotics and Intelligent Systems Development*, Idea Group Inc., Hershey, PA, pp. 1-41
- \_\_\_\_\_. (2007b). "Animal Abduction. From Mindless Organisms to Artificial Mediators." En L. Magnani y P. Li (eds.), *Model-Based Reasoning in Science, Technology, and Medicine*, Series "Studies in Computational Intelligence", Vol. 64, Springer, Berlin/New York, pp. 3-37
- \_\_\_\_\_. (2008). "Creating Chances through Cognitive Niche Construction. The role of affordances" En I. Lovrek, R. J. Howlett, y L. C. Lakhmi (eds.), *Knowledge-Based Intelligent Information and Engineering Systems*, 11th International Conference, KES 2007, Vietri sul Mare, Italy, September 12-14, 2007, Proceedings, Part II, Series: Lecture Notes in Computer Science LNCS/LNAI 4693, pp. 917-926
- \_\_\_\_\_. (2009a). "Creative abduction and hypothesis withdrawal in science." En: J. Meheus y T. Nickles, eds., *Models of Discovery and Creativity*, Springer, Berlin, pp. 95-126

- \_\_\_\_\_. (2009b). *Abductive Cognition: The Epistemological and Eco-Cognitive Dimension of Hypothetical Reasoning*. Springer, Berlin/Heidelberg
- Magnani L, y Dossena R (2005) "Perceiving the infinite and then infinitesimal world: unveiling and optical diagrams in the construction of mathematical concepts." *Found Sci* 10:7-23
- Magnani, L., N. J. Nersessian, y P. Thagard (eds.) (1999). *Model-Based Reasoning in Scientific Discovery*, Kluwer Academic/Plenum Publishers, New York
- Magnani, L. y N. J. Nersessian (eds.) (2002). *Model-Based Reasoning. Scientific Discovery, Technological Innovation, Values*, Kluwer Academic/Plenum Publishers, New York
- Magnani, L. y P. Li (eds.) (2007), *Model-Based Reasoning in Science, Technology, and Medicine*, Series: Stud Comp Intell, Vol. 64, Springer, Berlin/New York
- Millikan, R.G. (1984) *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press
- \_\_\_\_\_. (1993) *White Queen Psychology and Other Essays for Alice*. Cambridge, MA: MIT Press.
- Miller, G.A. (2003). "The cognitive revolution: a historical perspective." *TRENDS in Cognitive Sciences* Vol.7 No.3 March
- Morgan, M.S. (1999) "Models of mediating instruments." En: Morgan, Mary S. y Morrison, Margaret, (eds.) *Models as mediators: perspectives on natural and social science*. Cambridge University Press, Cambridge, UK, pp. 10-37
- Nagel, Ernest (1961). *The Structure of Science*. New York: Harcourt, Brace & World
- Nersessian, N.J. (1999). "Model-based reasoning in conceptual change." En: N.J. Nersessian, L. Magnani, y P. Thagard (Eds.), *Model-based Reasoning in Scientific Discovery*, pages 5-22, New York: Kluwer Academic/Plenum Publishers.
- \_\_\_\_\_. (1999). "Model-based reasoning in conceptual change." En: Magnani, L., N. Nersessian, y P. Thagards (eds.), *Model-Based Reasoning in Scientific Discovery*. New York: Kluwer Academic/Plenum Publishers, pp. 5-22
- \_\_\_\_\_. (2002). "Maxwell and 'the method of physical analogy': model-based reasoning, generic abstraction, and conceptual change." En: D. Malament (ed.) *Essay in the History and Philosophy of Science and Mathematics*, p. 129-166, Lasalle, IL, 2002 Open Court
- \_\_\_\_\_. (2008). *Creating Scientific Concepts*. Cambridge, MA: The MIT Press.
- \_\_\_\_\_. (2009). "How Do Engineering Scientists Think? Model-Based Simulation in Biomedical Engineering Research Laboratories." En: *Topics in Cognitive Science*, 1:731-757
- Norman, D. (1988). *The Psychology of Everyday Things*. New York : Basic Books
- Perice, C.S. (1931-1958). *Collected Papers of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press, vols. 1-6, Hartshorne, C. y Weiss, P. (eds.); vols. 7-8, Burks, A.W. (ed.)
- Peirce, C. S. (1960). *Collected Papers*. Cambridge, MA, Harvard University Press.

- Peirce, C.S. (1878), "How to Make Our Ideas Clear", *Collected Papers (CP)*, 5.388-5.410
- Pylyshyn, Z. (2002) "Mental imagery: In search of a theory." *Behavioral and Brain Sciences* 25, 157–238
- Pinker, S. (1997). *How the Mind Works*. New York: W.W. Norton
- Pino, B. (2010). "Re-assessing ecology of tool transparency in epistemic practices." *Mind & Society: Volume 9, Issue 1*, Springer Berlin / Heidelberg, Pages 85-110. Disponible online (DOI 10.1007/s11299-010-0071-4)
- Pohjola, P. (2007). "Ontology, artefacts, and Models of Reasoning." En: Magnani, L., y P. Li (eds.), *Model-Based Reasoning in Science, Technology, and Medicine*. Series: Studies in Computational Intelligence, Vol. 64, Berlin/New York: Springer.
- Polanyi, M. (1996). *The Tacit Dimension*. London: Routledge & Kegan Paul
- Prendergast, C. (2000). *The Triangle of Representation*. New York: Columbia University Press
- Putnam, H. (1960). "Minds and machines." En: S. Hook (ed.), *Dimensions of Mind*, Collier Books
- Putnam, H. (1975). "The meaning of 'meaning'." Reprinted in his *Mind, Language, and Reality*. New York: Cambridge University Press
- Rupert, R. (2004). "Challenges to the Hypothesis of Extended Cognition." *Journal of Philosophy* 101 (8):389-428
- Searle, J. (1980). "Minds, brains and programs." *Behavioral and Brain Sciences* 3: 417-424
- \_\_\_\_\_. (1995). *The Construction of Social Reality*. London: Penguin Books.
- Shelley, C. (1996). "Visual abductive reasoning in archeology." *Philosophy of Science* 63: 278-301
- Smolensky, P. (1989). "Connectionist modeling: Neural computation/mental connections." En L. Nadel (Ed.), P. Culicover, L. A. Cooper, R. M. Harnish (Assoc. Eds.), *Neural connections, mental computation*. Cambridge, MA: MIT Press/Bradford. 49–67. [Re-impreso en J. Haugeland, (Ed.). 1997. *Mind Design II: Philosophy, Psychology, Artificial Intelligence*, MIT Press/Bradford Books.]
- Sterelny, K. (2004). "Epistemic Artefact and the Extended Mind." En: Schantz, Richard (ed.), *The Externalist Challenge*. Berlin: Walter de Gruyter
- Suárez, M. (2002). An inferential conception of scientific representation. Presentado como PSA 2002 [en línea] <<http://philsci-archive.pitt.edu/>> [consulta: 16 septiembre 2009]
- Stich, S. (1978). "Autonomous psychology and the belief-desirethesis." *Monist* 61: 573–591
- Suppes, P. (1961). "A comparison of the meaning and uses of models in mathematics and the empirical sciences". En H. Freudenthal (ed.), *The Concepts and the Role of the Model in Mathematics and Natural and Social Sciences*, Dordrecht: Reidel, pp. 163-177

- 
- Thagard, P. (1992). *Conceptual Revolutions*. Princeton: Princeton University Press.
- Thagard, P. y Shelley, C. (1997). "Abductive reasoning: Logic, visual thinking, and coherence." En: M.L. D. Chiara, K. Doets, D. Mundici, y J. van Benthem (eds.), *Logic and Scientific methods*. Dordrecht: Kluwer Academic Publisher, p. 413-427
- Tooby, J., y De Vore, I. (1987). "The reconstruction of hominid behavioral evolution through strategic modeling." En: Kinzey. W.G., (eds), *Primate Models of Hominid Behavior*. Albany: Suny Press, pp. 183-237
- Vallejos, G. (2008). *Conceptos y Ciencia Cognitiva*. Bravo y Allende editores.
- van Fraasen, B. (1980). *The Scientific Image*. Oxford: Clarendon Press
- van Gelder, T. y R. Port. (1995). "It's About Time: An Overview of the Dynamical Approach to Cognition." En: Port y van Gelder, (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*. MIT/Bradford
- van Gelder, T. (1995). "What might cognition be if not computation?" *J Philos*, 91: 345-381
- van Gelder, T. J. (1998). "The dynamical hypothesis in cognitive science." *Behavioral and Brain Sciences*, 21, 1-14
- \_\_\_\_\_. (1999) "Revisiting the Dynamical Hypothesis." Preprint No. 2/99, University of Melbourne, Department of Philosophy
- Vermass, P. y Houkes, W. (2006). "Technical functions: a drawbridge between the intentional and structural natures of technical artefacts." *Studies in History and Philosophy of Science* 37: 5-18
- von Wright, G.H. (1963). "Practical inference." *The Philosophical Review* 72: 159-179
- \_\_\_\_\_. (1971). *Explanation and understanding*. London: Routledge and Kegan Paul
- Wason P. (1966). "Reasoning." En: Foss B (ed.), *New Horizons in Psychology*. Harmondsworth, UK: Penguin
- Wilson, M. (2002). "Six Views of Embodied Cognition." University of California, Santa Cruz. *Psychonomic bulletin & review*, vol. 9, no4, pp. 625-636
- Wilson, R. (2005). "Meaning Making in the Mind of the Externalist" [en línea] <<http://www.arts.ualberta.ca/~raw/recpapers.htm>> [consulta: 30 junio 2011]
- Woods, J. (2007). "The concept of Fallacy is Empty. A Resource –Bound Approach to Error." En: Magnani, L y Li, P., (Eds.), *Model-Based Reasoning in Science, Technology, and Medicine*. Springer-Verlag Berlin Heidelberg