



UNIVERSIDAD DE CHILE
FACULTAD DE FILOSOFÍA Y HUMANIDADES
DEPARTAMENTO DE FILOSOFÍA

UN ANÁLISIS DEL CAMBIO CONCEPTUAL EN CIENCIA COGNITIVA: EL
CONCEPTO DE REPRESENTACIÓN

Tesis para optar al grado de MAGÍSTER EN ESTUDIOS COGNITIVOS

PENÉLOPE PEREIRA PERDOMO

PROFESOR GUÍA: GUIDO VALLEJOS

SANTIAGO DE CHILE

2021

*A Luis Perdomo Folch, mi abuelo, porque
el castillo de libros y medallas de oro
sigue en expansión.*

La presente tesis fue realizada con el financiamiento de CONICYT Magíster Nacional 2018, folio 22181186.

TABLA DE CONTENIDO

RESUMEN	v
INTRODUCCIÓN	1
CAPÍTULO 1: CONCEPTOS Y CAMBIO CONCEPTUAL	6
1. Introducción	6
2. La aproximación clásica del cambio conceptual.....	8
3. Proceso de cambio conceptual	28
4. Tipos de cambio conceptual.....	49
5. Discusión	56
6. Conclusiones	60
CAPÍTULO 2: CAMBIO CONCEPTUAL EN CIENCIA COGNITIVA.....	62
1. Introducción	62
2. La ciencia cognitiva: supuestos, interdisciplinariedad e inmadurez	63
3. Antecedentes de cambio conceptual en ciencia cognitiva.....	73
4. Discusión	92
5. Conclusiones	93
CAPÍTULO 3: ANÁLISIS DEL CAMBIO CONCEPTUAL EN EL CONCEPTO DE REPRESENTACIÓN.....	95
1. Introducción	95
2. Caracterización de los sistemas conceptuales sucesivos	97
3. Demostración de existencia de cambio conceptual y determinación del tipo ..	136
4. Mecanismos que posibilitan el cambio conceptual	140
5. Discusión	143
6. Conclusiones	145
CONCLUSIONES GENERALES.....	148
REFERENCIAS BIBLIOGRÁFICAS.....	150

RESUMEN

En la presente tesis realizaré un análisis del cambio conceptual (CC) en la noción de REPRESENTACIÓN en ciencia cognitiva. Por medio de una investigación documental, sostendré que su proceso principal de discontinuidad apunta a la incorporación de distribución y de multimodalidad. El primer punto se encuentra en los debates entre la Teoría Representacional de la Mente (TRM) (Fodor, 1975, 1983, 1987, 1997a; Fodor y Pylyshyn, 1988) y el Tratamiento Adecuado del Conexionismo (TAC) (Smolensky, 1987, 1988, 1989). El segundo punto se observa en las críticas planteadas desde el Sistema de Símbolos Perceptuales (SSP) (Barsalou, 1999). En el primer capítulo me centraré en las características del proceso de cambio y la metodología para su estudio, elementos que serán extraídos de la aproximación clásica del CC. En el segundo capítulo identificaré los rasgos de la ciencia cognitiva relevantes para el análisis (sus supuestos sustantivos, interdisciplinariedad e inmadurez), y revisaré antecedentes de investigaciones sobre CC en ella. En el tercer capítulo examinaré las transformaciones conceptuales en REPRESENTACIÓN a partir del abordaje de la TRM, el TAC y el SSP. A modo de conclusión general, sistematizaré los resultados, estableceré algunas limitaciones y futuras directrices.

INTRODUCCIÓN

Los conceptos poseen una centralidad notable tanto a nivel individual como a nivel colectivo, dado que permiten la representación del mundo y la comprensión del lenguaje. La presente tesis tiene como objetivo general la realización de un análisis del cambio conceptual (en adelante, CC) en la noción de REPRESENTACIÓN en ciencia cognitiva. Por medio de una investigación documental, sostendré que el proceso principal de transformación en esta noción apunta a la incorporación de un carácter distribucional y la incorporación del formato multimodal de las representaciones. El primer punto se puede encontrar en los debates entre la Teoría Representacional de la Mente (TRM) (Fodor, 1975, 1983, 1987, 1997a; Fodor y Pylyshyn, 1988) y el Tratamiento Adecuado del Conexionismo (TAC) (Smolensky, 1987, 1988, 1989). El segundo punto se observa en las críticas realizadas desde el Sistema de Símbolos Perceptuales (SSP) (Barsalou, 1999). Para llevar a cabo este objetivo general, me enfocaré en los siguientes objetivos específicos: (i) determinación de las características de los CC y la metodología apropiada para su estudio; (ii) identificación de los rasgos de la ciencia cognitiva que son relevantes para el análisis de CC en ella; (iii) identificación de antecedentes de estudio de CC en ciencia cognitiva, y (iv) análisis de la noción de REPRESENTACIÓN en la TRM, el TAC y el SSP. A continuación, ahondaré en este esquema general.

En el primer capítulo, abordaré el marco teórico de la aproximación clásica del CC para llevar a cabo el primer objetivo específico. El área de investigación de las discontinuidades conceptuales corresponde a un campo interdisciplinario constituido principalmente por tres disciplinas: historia y filosofía de las ciencias, enseñanza de las ciencias y psicología del desarrollo. A su vez, su aproximación clásica se caracteriza por tres consideraciones: la apelación al historicismo kuhniano, la afirmación de que el conocimiento presenta un marco coherente unificado que exhibe rasgos propios de una teoría, y la hipótesis del continuo.

Por un lado, el estudio de los CC fue iniciado por Thomas Kuhn (1970) para indicar la transformación en el significado de los conceptos contenidos en un paradigma científico tras el cambio de este, producto de una revolución. El autor sostiene que los paradigmas son inconmensurables entre sí, es decir, poseen disparidades conceptuales, observacionales y metodológicas que imposibilitan la comparación a través de un

parámetro neutro. No obstante, en trabajos posteriores, Kuhn (1982) restringe la noción de inconmensurabilidad a una versión local, según la cual existe solo un cambio de significado en un conjunto limitado de términos.

La intersección entre las ideas kuhnianas sobre los CC y la teoría del desarrollo cognitivo de Jean Piaget (1950, 1976) configura una visión del conocimiento como un marco coherente unificado que exhibe rasgos propios de una teoría (es decir, permite realizar explicaciones y predicciones que constriñen las creencias e inferencias del sujeto). Entre los exponentes que sostienen esta visión destacan Posner y colaboradores (Posner, Strike, Hewson y Gertzog, 1982; Strike y Posner, 1992) en el área de la enseñanza de las ciencias, y Susan Carey (1985, 1999a, 1999b, 2009) desde la psicología del desarrollo.

Por último, la convergencia entre el historicismo kuhniano y el giro cognitivo en historia y filosofía de las ciencias condujo al planteamiento de la hipótesis del continuo, que afirma que la cognición científica es una extensión sofisticada de la cognición humana común. La filosofía computacional de la ciencia de Paul Thagard (1992, 2011, 2014) y la historia cognitiva de la ciencia de Nancy Nersessian (1992, 1995, 2008) son exponentes de esta vertiente. Sumando los trabajos de Posner *et al.* y Carey, la hipótesis del continuo justifica la conformación de una cuádruple analogía entre los cambios en teorías científicas, en los científicos como sujetos epistémicos, en los estudiantes de ciencias y en los niños.

El abordaje del marco teórico de la aproximación clásica del CC me permitirá enfatizar características relevantes del proceso de transformación conceptual, donde se incluyen la noción de ecología conceptual (Posner *et al.*, 1982; Strike y Posner, 1992), el *bootstrapping* —en particular, las analogías— como mecanismo central de cambio, y la plausibilidad de describir el resultado del proceso como inconmensurable con la etapa inicial. Sostendré que una metodología de estudio de los CC que puede ser extraída de este marco teórico consta de: (a) caracterización de los sistemas conceptuales sucesivos; (b) individuación de los conceptos inexpresables en el vocabulario del sistema anterior o bien modificados con respecto a este, de acuerdo con el rol que cumplen en la teoría; (c) demostración de existencia de un cambio entre los sistemas conceptuales sucesivos junto con la determinación del tipo de cambio, y (d) estudio de los mecanismos que lo posibilitaron.

En el segundo capítulo, realizaré el segundo y tercer objetivos específicos planteados. Mostraré que los supuestos sustantivos (computacionalidad y

representacionalidad), la interdisciplinariedad y la inmadurez del área son factores que constriñen los análisis de los CC allí. Dado que las problematizaciones de los supuestos sustantivos cumplen un rol central en el desarrollo de una ciencia, y dado que REPRESENTACIÓN se halla involucrada en ellos, el proceso de transformación conceptual en ella es complejo. Además, la interdisciplinariedad introduce un factor de variabilidad en el rol que cumple esta noción en dependencia de la subdisciplina, y la inmadurez implica que los CC se encuentran en curso.

Sumado a lo anterior, analizaré algunos estudios que han individualizado CC en el área, a saber, Thagard (1992, 2011), Akagi (2017) y Keijzer (2002). Estos se han enfocado en nociones principales como PROCESAMIENTO DE INFORMACIÓN, CONCEPTO, COGNICIÓN y REPRESENTACIÓN. En conjunto, sugieren la existencia de tres focos de transformación de conceptos: en la denominada *revolución cognitiva*, en las discusiones sobre los límites de la cognición, y en la recientemente considerada *revolución neurocognitiva*. La revolución cognitiva da paso a los enfoques clásicos (simbólico y conexionista) en ciencia cognitiva, que introducen los supuestos sustantivos de representacionalidad y computacionalidad. Las discusiones sobre los límites de la cognición son originadas por enfoques no clásicos (Cognición 4EA) que problematizan dichos supuestos. La revolución neurocognitiva reivindica la importancia de la neurociencia en la ciencia cognitiva e integra explicaciones que aluden a estructuras y procesos neurales.

En el tercer capítulo, cumpliré con el cuarto objetivo específico. Para ello, seguiré los pasos de la metodología de estudio de los CC planteada con anterioridad. Primeramente, caracterizaré los sistemas conceptuales sucesivos: la TRM de Jerry Fodor (1975, 1983, 1987, 1997a; Fodor y Pylyshyn, 1988), el TAC de Paul Smolensky (1987, 1988, 1989) y el SSP de Lawrence Barsalou (1999). En los tres casos, me detendré en los compromisos epistemológicos y los problemas implicados en las posturas que originan los debates, que son elementos importantes en la ecología conceptual. Para facilitar la comparación entre posturas, me concentraré en tres puntos: la relación entre representaciones conceptuales y representaciones perceptuales, entre representaciones mentales y palabras del lenguaje natural, y entre representaciones mentales y representaciones neurales.

Luego, demostraré la existencia de transformación conceptual y determinaré el tipo de cambio. Sostendré que las nociones de REPRESENTACIÓN que se desprenden apuntan a REPRESENTACIÓN AMODAL, REPRESENTACIÓN DISTRIBUIDA y REPRESENTACIÓN

MULTIMODAL respectivamente. En ellas, se observan CC que corresponden a reorganizaciones básicas plasmadas en la incorporación de distribución y de multimodalidad, que no son plausiblemente entendidas en términos de inconmensurabilidad kuhniana. Posteriormente, me enfocaré en las analogías como un mecanismo que influye en estas reorganizaciones —la metáfora del computador, la analogía con el funcionamiento neuronal y la problematización de la analogía con los sistemas de grabación.

Vincularé los resultados del análisis con dos puntos de la ecología conceptual que subyacen a las distinciones conceptuales detectadas: la definición de cognición y la adopción o rechazo de la realizabilidad múltiple. Por último, revisaré los resultados a la luz de los tres rasgos de la ciencia cognitiva que constriñen el estudio en ella (sus supuestos sustantivos, su interdisciplinariedad y su inmadurez). A modo de conclusión general, sistematizaré los resultados, estableceré algunas limitaciones y futuras directrices.

Para efectos de la presente tesis, adoptaré una noción de CC basada principalmente en las definiciones de Thagard (1992) y Nersessian (2008) que incluye dos puntos relevantes: las acciones o eventos centrales del proceso de cambio y la definición de CONCEPTO involucrada. Con respecto al primer punto, el CC indica el proceso de creación, alteración y comunicación de conceptos individuales en la historia del desarrollo de una disciplina. La creación es la introducción de conceptos novedosos por medio del empleo del conocimiento disponible y los recursos cognitivos del sujeto epistémico; la alteración corresponde a la reestructuración ocurrida en los conceptos de un individuo o una ciencia producto del aprendizaje o el desarrollo científico, y la comunicación apunta a la presentación del concepto novedoso a los pares, usualmente a través de argumentos, lo que eventualmente puede desembocar en la alteración de la estructura conceptual del receptor.

En relación con el segundo punto, los conceptos son entendidos como un tipo de representación mental. La especificidad de esta caracterización varía en dependencia del enfoque cognitivo en consideración. Carey sostiene que las representaciones “son estados del sistema nervioso que tienen contenido, que se refieren a entidades concretas o abstractas (o incluso a entidades ficcionales), propiedades y eventos” (2009, p. 5). Thagard pertenece al enfoque simbólico en ciencia cognitiva (véase pp. 65-66), por lo que sostiene que los conceptos corresponden a estructuras mentales similares a las palabras. Nersessian, en cambio, es cercana a posturas alternativas tales como la

cognición corporizada (véase pp. 82-83), donde un concepto es una estructura neural relacionada con el sistema sensoriomotor. En el segundo capítulo mostraré que CONCEPTO es uno de los casos de CC en ciencia cognitiva, por lo que la discusión sobre su definición se encuentra abierta. Además, dado que los conceptos son un tipo de representación, el análisis de REPRESENTACIÓN en el tercer capítulo está directamente relacionado con este debate.

La relevancia del tema de la presente tesis puede ser observada al considerar que las discontinuidades conceptuales usualmente toman la forma de cambio de teoría. Por ende, los CC han sido analizados principalmente dentro del marco de las llamadas *revoluciones científicas* (referidas al cambio paradigmático producido luego de un periodo de crisis en una ciencia). Las investigaciones se desarrollan en ciencias como la física y la química, consideradas maduras en sentido kuhniano —es decir, regidas por un paradigma—. Sin embargo, la ciencia cognitiva es inmadura. De acuerdo con Barbara von Eckardt (1995), “lo que todos los científicos cognitivos comparten a lo largo del tiempo es un compromiso con algo parecido a un enfoque del estudio de la mente en lugar de un conjunto específico de teorías, explicaciones o leyes” (von Eckardt, 1995, p. 15). Tanto las propiedades de los fenómenos que conforman su dominio de investigación —esto es, las capacidades cognitivas humanas— como las preguntas básicas planteadas con respecto a él son conceptualizadas preteóricamente, puesto que se vinculan con lo que el sentido común conoce sobre ellos. Por otro lado, las teorías que su investigación genera son tentativas, por lo que “está todavía en la búsqueda de un conjunto de compromisos teóricos más permanentes” (von Eckardt, 1995, p. 15). Dado que sus teorías son tentativas, parece difícil localizar aquí un proceso de cambio teórico.

La presente tesis responderá afirmativamente a la posibilidad de identificación de CC en ciencia cognitiva por medio del abordaje de los antecedentes expuestos en el segundo capítulo y la realización del análisis del tercer capítulo. En cuanto ninguna de las teorías revisadas allí ha constituido un reemplazo del enfoque simbólico, las transformaciones observadas en REPRESENTACIÓN parecen indicar que no se requiere un cambio teórico para sostener la existencia de CC.

CAPÍTULO 1: CONCEPTOS Y CAMBIO CONCEPTUAL

1. Introducción

El estudio de los CC configuró un área interdisciplinaria que comprende tres disciplinas: historia y filosofía de las ciencias, psicología del desarrollo y enseñanza de las ciencias. Esta área, a su vez, estableció una cuádruple analogía no exenta de críticas entre: (i) los CC ocurridos en la historia de la ciencia, (ii) aquellos que experimentan los científicos como sujetos epistémicos que investigan fenómenos y crean conceptos, (iii) los estudiantes o novatos en su proceso de aprendizaje de una ciencia, y (iv) los niños en su desarrollo cognitivo. Cada arista enfatiza definiciones, tipos o mecanismos distintos, por lo que no existe una noción unívoca del fenómeno. Por ejemplo, en enseñanza de las ciencias, Posner, Strike, Hewson y Gertzog (1982; Strike y Posner, 1992) describen el CC como un proceso de acomodación: la alteración de concepciones centrales durante el aprendizaje, cuya etapa inicial es la asimilación y su mecanismo central radica en la operación de la ecología conceptual (véase pp. 29-31). En psicología del desarrollo, Susan Carey (1985, 1999a, 1999b, 2009) afirma que se trata de una transformación en conceptos individuales en el marco de la reestructuración del conocimiento durante el desarrollo cognitivo, cuya etapa inicial engloba la cognición núcleo y las teorías intuitivas, y cuyo mecanismo central es el *bootstrapping quineano* (véase pp. 32-39). En historia y filosofía de las ciencias, Paul Thagard (1992, 2011, 2014) y Nancy Nersessian (1992, 1995, 2008) sostienen que el CC se define como la creación, alteración y comunicación de conceptos en el marco de las revoluciones científicas, donde la etapa inicial comprende el contexto del sujeto enfrentado a la resolución de problemas. Para Thagard, el mecanismo central es la coherencia explicativa (véase pp. 16-17), mientras que para Nersessian es el *bootstrapping*, con un énfasis en los razonamientos basados en modelo (pp. 33-38). Este conjunto de enfoques configura lo que se denomina *aproximación clásica del CC*.

En el presente capítulo me concentraré en la determinación de una metodología de estudio de los CC que considere rasgos relevantes del proceso y que sea útil para su aplicación en el abordaje de teorías científicas. Sostendré que esta metodología consiste en el siguiente conjunto de pasos: (a) caracterización de los sistemas conceptuales sucesivos; (b) individuación de los conceptos inexpresables en el vocabulario del sistema

anterior o bien modificados con respecto a este, de acuerdo con el rol que cumplen en la teoría; (c) demostración de existencia de un cambio entre los sistemas conceptuales sucesivos, junto con la determinación del tipo de cambio, y (d) estudio de los mecanismos que posibilitaron el CC. Algunas características centrales a considerar son las influencias teóricas y problemas presentes en los debates entre los investigadores —lo que se desprende de la noción de ecología conceptual (Posner *et al.*, 1982; Strike y Posner, 1992) —; el *bootstrapping* (en particular, las analogías) como mecanismo central, y la plausibilidad de describir el resultado del proceso como inconmensurable con la etapa inicial.

Para cumplir con el objetivo planteado, analizaré las propuestas fundacionales dentro de la aproximación clásica mencionada anteriormente. Su relevancia para el objetivo principal de la tesis —esto es, el análisis del CC en ciencia cognitiva— se encuentra en su apelación y complemento de ideas provenientes de la historia y filosofía de las ciencias¹. Los rasgos definitorios extraídos constituirán la base para analizar las transformaciones sucesivas en REPRESENTACIÓN. Así, en el apartado siguiente realizaré una caracterización de la aproximación clásica por medio de sus tres consideraciones principales: la influencia del historicismo kuhniano, la afirmación de que el conocimiento presenta un marco coherente unificado que exhibe rasgos propios de una teoría, y la hipótesis del continuo. En el tercer apartado abordaré el proceso del CC y, en particular, su contexto, mecanismo central y resultados. En el cuarto apartado observaré los tipos o grados de CC que se han planteado en la literatura. En la sección de discusión sistematizaré los rasgos que serán relevantes para el análisis realizado en el tercer capítulo. A modo de conclusión, mencionaré algunas consideraciones preliminares sobre la posibilidad de la existencia de CC sin un cambio de teoría, a propósito del vínculo estrecho entre discontinuidades conceptuales y revoluciones científicas. Ello abre las puertas a su estudio en una ciencia inmadura como la ciencia cognitiva, tema que abordaré en el segundo y tercer capítulo.

¹ Esta relevancia contrasta con enfoques alternativos, que abandonan la similitud entre el proceso de aprendizaje o el desarrollo cognitivo infantil y el trabajo científico (Özdemir y Clark, 2007, p. 355). En estos casos no es posible extraer rasgos de los CC que sean relevantes para el análisis de las discontinuidades conceptuales en una ciencia.

2. La aproximación clásica del cambio conceptual

La aproximación clásica del CC puede ser caracterizada por medio de tres consideraciones: la apelación al historicismo kuhniano, la afirmación de que el conocimiento presenta un marco coherente unificado que exhibe rasgos propios de una teoría, y la hipótesis del continuo. A continuación, abordaré cada uno de estos puntos, que permitirán posteriormente extraer algunos rasgos de los CC relevantes para el estudio propuesto en el tercer capítulo de la presente tesis.

2.1. Historicismo kuhniano

La primera consideración que caracteriza la aproximación clásica corresponde al legado del historicismo, particularmente el historicismo kuhniano y sus nociones de ciencia normal, paradigma, revolución, cambio paradigmático y cambio gestáltico. Thomas Kuhn presentó los antecedentes del CC con la publicación de *The Structure of Scientific Revolutions* (cuya primera edición data de 1962) y trabajos posteriores, al interesarse por “la inconmensurabilidad y la naturaleza de la división conceptual entre las etapas de desarrollo separadas por (...) ‘revoluciones científicas’” (Kuhn, 2000c, p. 228). En dicha obra, el autor presenta su teoría de los cambios paradigmáticos en ciencias, según la cual el avance y desarrollo científico no se producen por un proceso gradual de acumulación de conocimiento —descubrimientos individuales e inventos—, sino por revolución. Para entender el desarrollo de un campo científico distingue, de manera esquemática, entre un periodo pre-paradigmático y un periodo post-paradigmático. En un estadio inicial, existe una competencia entre diferentes escuelas por el dominio en el campo². La diferencia entre ellas radica en “sus maneras inconmensurables de ver el mundo y de practicar la ciencia en él” (1970, p. 4). Frente a la ocurrencia de un logro científico notable, el número

² Para Kuhn, el desarrollo temprano de la ciencia se caracteriza por la carencia de estándares y consenso, las disputas sobre normas de investigación y la recolección aleatoria de datos. Todo ello se debe a la ausencia de un tipo de paradigma cuya naturaleza identifique problemas desafiantes y sus posibles soluciones (1970, p. 179). Estas divergencias iniciales suelen desaparecer tras el triunfo de una de las escuelas que, “por sus propias creencias y preconcepciones características, enfatizó solo alguna parte especial del conjunto de información demasiado grande y rudimentaria” (1970, p. 17). Las otras escuelas comienzan gradualmente a desaparecer, en gran parte por la conversión de los investigadores al grupo ganador. Quienes no realicen esta conversión son considerados como externos a la profesión, y su trabajo no es estimado.

de escuelas se reduce (frecuentemente, a una). Comienza así la operación de la ciencia normal, guiada por un paradigma. Ante la presentación de anomalías y los intentos infructíferos por abordarlas, la ciencia normal entra en un periodo de crisis que desemboca eventualmente en el cambio paradigmático. En este marco, el CC se entiende como la transformación en el significado de los conceptos dado el cambio del marco teórico en el que se encuentran y del que obtienen dicho significado, producto de la ocurrencia de una revolución. Por su parte, la noción de CONCEPTO involucrada tiene en vista conceptos científicos tales como FUERZA y MASA. Estos se entienden como entidades históricas, es decir, elementos propuestos en el marco de una teoría que cambian junto con esta, frente a lo cual es “imposible definir todos los términos de una teoría en el vocabulario de la otra” (Kuhn, 1982, p. 669). A pesar de que *The Structure* no emplea la noción de CC directamente, la teoría de los cambios paradigmáticos en ciencias sienta las bases de la noción, que será empleada ocasionalmente por el autor en obras posteriores.

La etapa inicial del proceso de transformación conceptual en el esquema kuhniano es la ciencia normal y el paradigma, puesto que allí se establecen los sistemas conceptuales embebidos en teorías que, con la revolución, sufrirán un cambio. La ciencia normal es una “investigación firmemente basada en uno o varios logros científicos pasados, logros que una comunidad científica particular reconoce por un tiempo como la base de su práctica posterior” (1970, p. 10). En otras palabras, es la vigencia de un paradigma. La noción de paradigma determina las entidades existentes e inexistentes y su comportamiento, además de proveer teorías y estándares para el funcionamiento de la ciencia. Una de las primeras definiciones que el autor entrega afirma que son “logros científicos universalmente reconocidos que, por un tiempo, proveen problemas y soluciones modelo a una comunidad de profesionales” (1970, p. viii). Los logros que son tomados como modelos incluyen teorías, leyes y aplicaciones, y su estudio es fundamental para formar parte de una comunidad científica determinada. En esta primera definición se apela a una circularidad: un paradigma es aquello compartido por una comunidad científica, y la comunidad científica madura es definida por la posesión de un paradigma.

En el Epílogo, escrito siete años después de la primera edición de *The Structure*, Kuhn identifica dos sentidos principales de paradigma. Por una parte, en sentido sociológico, corresponde a la *matriz disciplinar*: la constelación de creencias, técnicas, y valores compartidos por la comunidad. Incluye generalizaciones simbólicas, creencias en

modelos ontológicos o heurísticos —que permiten establecer analogías para determinar qué se entiende por explicación o solución adecuada— y valores como la consistencia y la simplicidad, que son importantes ante una crisis o en la elección entre paradigmas. El segundo sentido, en cambio, es más estrecho³ y corresponde al cuarto elemento en dicha constelación: los ejemplares de soluciones a los problemas en la ciencia normal. El tipo de instrucción que reciben los futuros científicos contempla el aprendizaje de teorías y leyes por medio de sus aplicaciones. En particular, se basa en la enseñanza de soluciones estándar a problemas seleccionados. Estos sirven, principalmente, para desarrollar la habilidad de *ver* una variedad de situaciones de manera similar a los científicos de cuya comunidad forman parte, y a agruparlas en conjuntos de similitud. Para ello, los ejemplares suponen la existencia de una solución y de reglas (compromisos teóricos, conceptuales, metodológicos e instrumentales) que determinen su naturaleza y los pasos para conseguirla. Cabe notar, no obstante, que los paradigmas son anteriores y más completos que cualquier conjunto de reglas, y guían la investigación sin ellas (1970, p. 42). Por tanto, dichos ejemplares no guardan en común reglas explícitas⁴, sino un parecido de familia, en términos wittgensteinianos⁵. El conocimiento que condensan es tácito, esto es, no puede articularse explícitamente. Según esta idea, los estudiantes no aprenden mediante una teoría y un conjunto de reglas, sino por la resolución de ejemplares.

La posesión de un paradigma de esta naturaleza es un signo de madurez del campo científico. Por medio de él, la labor científica puede ser sistemática, focalizada y con una delimitación más rígida en comparación con el periodo pre-paradigmático. No es necesaria la reconstrucción, justificación o reexamen de los fundamentos del campo y sus

³ En *The Essential Tension* (1977), Kuhn sostiene que la noción de paradigma como solución de problemas ejemplares corresponde a su noción originaria. El procedimiento de aprendizaje que instaura es similar al observado en la enseñanza de la conjugación de verbos y declinación de sustantivos y adjetivos, denominado *paradigma*, y a ello debe su nombre. Sin embargo, el concepto terminó expandiéndose para englobar los libros de textos en los que aparecen, y los compromisos compartidos por los miembros de la comunidad (1977, p. xix).

⁴ De acuerdo con el autor, las reglas cobran relevancia cuando el paradigma es insuficiente para la resolución de problemas, siendo frecuentes en el periodo pre-paradigmático y en las revoluciones científicas. Por otro lado, dada la diversidad de especialidades científicas en un área, un grupo amplio puede compartir reglas explícitas, pero teniendo paradigmas distintos, en cuanto se educan por medio de diferentes modelos de logros científicos.

⁵ Kuhn recurre a las críticas que realiza Wittgenstein a la noción clásica de concepto, según la cual existen condiciones necesarias y suficientes para proveer una definición. Wittgenstein sostiene que es imposible la distinción entre propiedades esenciales y accidentales, por lo que no se puede establecer la necesidad en la inclusión de una propiedad determinada a la definición. Lo que comparten entre sí las instancias de algunos conceptos como JUEGO son parecidos de familia, y no una lista de condiciones necesarias y suficientes para formar parte de la categoría.

conceptos en cada trabajo. La investigación científica se centra en desarrollar el paradigma: determinar y precisar los hechos que este indica como significativos, buscar coincidencia entre los hechos y la teoría mediante su comparación con las predicciones, y articular la teoría, resolviendo problemas y ambigüedades residuales. En suma, se busca una coincidencia de la naturaleza con los esquemas provistos por el paradigma. Por tanto, no ocurren descubrimientos fácticos o teóricos, pues “no busca novedades de hecho o de teoría y, cuando tiene éxito, no encuentra ninguna” (1970, p. 52). Como resultado, la ciencia normal es una etapa de acumulación de conocimiento. Existe en ella un progreso que es visto como un proceso “unidireccional e irreversible” (1970, p. 206). No se entiende como un progreso teleológico: una teoría científica no es superior a su predecesora por ofrecer una mejor representación de lo que la naturaleza es realmente. Se trata de una mejora gradual en la capacidad de resolver problemas en comparación con las anteriores.

Por otro lado, el mecanismo de cambio entre paradigmas en una ciencia madura corresponde a la revolución. Las revoluciones son episodios de la historia de una ciencia en los que se generan cambios en los compromisos profesionales de la comunidad y en el aparato conceptual y metodológico del estándar de ciencia vigente. Se rechaza una teoría anteriormente reconocida y se adopta otra incompatible con ella; se reemplazan los modelos de problemas y soluciones, y ocurre una transformación en el modo de ver el mundo (véase pp. 39-40). Este periodo de transición no es un proceso acumulativo, sino que requiere la reconstrucción de los fundamentos del campo⁶. Por ejemplo, un descubrimiento como el del oxígeno no implica solamente la adición de una entidad, sino que acarrea una reevaluación de los compromisos sostenidos por la comunidad. No se trata de una mera transformación en la interpretación de las observaciones de los datos que el científico realiza: los datos recogidos son distintos, transformando una porción de la experiencia para vincularla con el nuevo paradigma. Las mediciones y operaciones son determinadas por el paradigma por lo que, al cambiar este, algunas se vuelven irrelevantes y son reemplazadas. Incluso el lenguaje observacional empleado presupone “un mundo ya subdividido perceptual y conceptualmente de un cierto modo” (1970, p. 129) que cambia en el proceso. Así, a pesar de que la ciencia puede mantener gran parte de sus términos e instrumentos luego de la revolución, estos capturan un aspecto de la naturaleza

⁶ Esta radicalidad no implica que solo cuenten como tales las mayores revoluciones de las ciencias, como las asociadas con Copérnico, Newton o Einstein. Dada la diversidad de especializaciones en la ciencia normal, es posible que ocurra una revolución en una tradición y una comunidad específica sin afectar otras.

distinto. En suma, la aceptación de una nueva teoría implica un cambio en las reglas de la práctica de la ciencia normal, lo que explica su característica no acumulativa.

La antesala de las revoluciones es la percepción de una anomalía, es decir, el reconocimiento de que la naturaleza de alguna manera ha violado las expectativas inducidas por el paradigma que gobierna la ciencia normal. Luego continúa con una exploración más o menos extensa del área de la anomalía. Y acaba solo cuando la teoría del paradigma ha sido ajustada tal que lo anómalo se ha convertido en lo esperado (1970, pp. 52-53).

Las anomalías surgen en el marco de las expectativas teóricas y prácticas provistas por el paradigma, y son una ocurrencia que no encaja con el orden de los fenómenos instaurado por este. La presencia de una mera discrepancia no es sinónimo de crisis: la ciencia normal se enfrenta frecuentemente a dificultades que el paradigma no resuelve, por lo que constituye una empresa de resolución de problemas. Por tanto, para que la crisis ocurra se debe reconocer que la discrepancia es distinta a dichos problemas y que, por ejemplo, cuestiona generalizaciones centrales o inhibe aplicaciones importantes.

El periodo de crisis es provocado por la falla en la resolución de problemas en la ciencia normal, y acarrea modificaciones mayores en los estándares de problemas, técnicas y soluciones. Se caracteriza por una labor de investigación semejante a la llevada a cabo en el periodo pre-paradigmático, aunque con una mayor focalización y definición producto de la existencia de un paradigma. La comunidad científica, en lugar de abandonar el paradigma que generó la crisis, realizará numerosos esfuerzos para alinear la anomalía con las expectativas⁷. La aislará y estructurará, aplicará más estrictamente las reglas de la ciencia normal y generará numerosas teorías especulativas. A raíz de ello, el paradigma pierde enfoque y se aflojan los criterios de trabajo. Si la actividad normal fracasa, comienza un periodo extraordinario de investigación donde la anomalía pasa a convertirse en el tema central del quehacer científico. Este periodo se encuentra marcado por “la proliferación de articulaciones en competencia, la disposición a intentar lo que sea, la expresión de un descontento explícito, el recurso a la filosofía y al debate sobre los fundamentos” (1970, p. 91). Eventualmente, aparecerá un nuevo candidato a paradigma,

⁷ El proceso de reemplazo de un paradigma por otro comienza con algunos científicos enfocados en el problema gatillante de la crisis, y usualmente menos comprometidos con la visión de mundo y las reglas del paradigma anterior. Sin embargo, la resistencia al cambio es inevitable y se produce por la seguridad de que el viejo paradigma eventualmente resolverá sus problemas.

que ya puede figurar de manera incipiente antes del desarrollo o reconocimiento de la crisis⁸.

El trabajo de Kuhn se sitúa dentro del marco del denominado *giro historicista* en filosofía de las ciencias, donde *The Structure* es considerado un manifiesto principal. En términos generales, este giro constituye una respuesta al giro lingüístico en filosofía de las ciencias, propio de la aproximación formal del positivismo lógico y del racionalismo crítico de Popper. El positivismo consideraba los conceptos de manera separada de su sentido psicológico, limitándolos a su naturaleza de signos con significado determinado por un lenguaje. Realizaba una separación tajante entre el componente fáctico y el componente racional de las teorías científicas, y las analizaba con las herramientas de la lógica, excluyendo el rol del sujeto epistémico que las plantea. El historicismo, en cambio, traslada el enfoque desde la estructura lógico-sintáctica de los enunciados de la ciencia hacia sus prácticas y procesos históricos reales⁹.

⁸ Un candidato a paradigma es aceptado al cumplirse dos condiciones: que parezca resolver problemas no resueltos y que prometa preservar buena parte de las habilidades de resolución de problemas que la ciencia acumuló con el paradigma anterior. La elección no depende de los criterios de la ciencia normal, porque esta se guía precisamente por el paradigma que se encuentra en cuestionamiento. Los paradigmas en competencia difieren en sus estándares con respecto a los problemas y sus soluciones, por lo que los criterios a satisfacer serán distintos, y sus defensas apelarán a los elementos del paradigma mismo en el que se enmarcan (es decir, son circulares). Además, como ninguno resuelve todos los problemas, las prioridades varían. Por ello, la discusión de sus valores debe recurrir a criterios externos a la ciencia normal. Las disparidades entre paradigmas son profundas: no suelen radicar en la estructura lógica, sino que se hallan en la postulación de distintas entidades constituyentes del universo y la redefinición de los problemas relevantes y sus soluciones. Por tanto, la lógica y la experimentación no son suficientes para la elección. Ello se resume en la noción de incommensurabilidad entre tradiciones científicas, que será abordada más adelante (véase pp. 39-43). En algunos casos, las características del nuevo paradigma mantienen parte de las características de la investigación extraordinaria que lo antecedió, mientras que en otros el paradigma nuevo emerge de repente (1970, pp. 89-90).

⁹ El giro lingüístico en filosofía de las ciencias condujo a la consideración de la estructura conceptual como un lenguaje: expresiones lingüísticas que forman enunciados, que a su vez forman teorías. Los conceptos que se encuentran en ellas son entendidos, de acuerdo con Carnap, como “signos, por ejemplo, palabras, expresiones compuestas por palabras, símbolos artificiales, etcétera, por supuesto con el significado que tienen en el lenguaje en cuestión. No nos referimos a ‘concepto’ en su sentido psicológico, esto es, a las imágenes o pensamientos relacionados de alguna manera con una palabra; ello no pertenece a la lógica” (1981, p. 118). El proyecto positivista lógico buscaba la unificación de la ciencia mediante el ordenamiento de los conceptos en sistemas jerárquicos, donde fuera posible su reducción gradual a otros conceptos hasta llegar a un nivel inferior que refiere directamente a la experiencia (Arabatzis y Kindi, 2008, p. 346). Así, se analizaban las relaciones entre enunciados por medio de la lógica. El CC corresponde allí a un fenómeno continuo y acumulativo, donde los conceptos nuevos son extensiones lógicas de los anteriores: pueden ser definidos en términos del vocabulario anterior, o bien determinados por oraciones de reducción (oraciones que describen las condiciones experimentales requeridas para que se aplique el concepto). No obstante, la visión del trabajo científico que se deriva de aquí es artificial y alejada del trabajo real de las comunidades científicas. Un escenario similar presenta el racionalismo crítico de Karl Popper que, si bien se distancia del positivismo lógico en cuanto a la noción del método científico —rechazando el inductivismo y

Posteriormente, diversos autores ampliaron y problematizaron el estudio de las revoluciones científicas dentro de una tendencia a la integración de enfoques cognitivos en la producción del conocimiento. Este giro cognitivo mantuvo el rechazo al positivismo lógico: reintroduce el contexto de descubrimiento, esto es, el proceso de generación de nuevas ideas o hipótesis que incluye el dominio descriptivo de la psicología, historia y sociología. Ello se observa en la consideración de los conceptos como representaciones mentales en lugar de signos con significado determinado por un lenguaje. Además, sostiene la insuficiencia del enfoque historicista para dar cuenta de las prácticas científicas. Según esta crítica, muchos de los recursos conceptuales propuestos por Kuhn y otros historicistas son necesarios pero no suficientes para estudiar la estructura compleja del conocimiento científico. Dos de los exponentes que se enmarcan en esta tendencia corresponden a la filosofía computacional de la ciencia de Paul Thagard (1988, 1992, 2014) y la historia cognitiva de la ciencia de Nancy Nersessian (1992, 1995, 2008).

En *Conceptual Revolutions* (1992), Thagard provee una teoría de las revoluciones conceptuales que integra “aproximaciones filosóficas, psicológicas y computacionales” (1992, p. 10). Las revoluciones científicas son entendidas como sucesos que involucran transformaciones mayores en sistemas conceptuales y proposicionales. Los CC son caracterizados como “la creación y alteración de representaciones mentales que corresponden aproximadamente a palabras” (2014, p. 1532), mientras que los cambios proposicionales son transformaciones en “estructuras mentales que representan lo que las oraciones representan” (1992, p. 21). Estas definiciones se enmarcan en el enfoque simbólico en ciencia cognitiva. Allí, los conceptos son representaciones mentales análogas a estructuras de datos en los computadores. Esquemáticamente, estas estructuras incluyen la determinación de sus tipos, subtipos, partes, sinónimos, antónimos, reglas e instancias. Los conceptos cumplen diversos roles psicológicos, entre los cuales se cuentan la categorización, explicación, generalización, solución de problemas, aprendizaje, memoria, realización de inferencias, comprensión y producción de lenguaje.

Los conceptos individuales forman sistemas conceptuales, que son “conceptos organizados en jerarquías de tipo y jerarquías de parte, y vinculados entre sí por reglas” (1992, p. 30). Estas jerarquías son modos de estructurar el sistema por medio de la apelación a relaciones entre tipos (por ejemplo, un tocororo es un tipo de ave, que es un

adoptando el falsacionismo—, mantiene la exclusión del psicologismo y la recurrencia a la lógica para estudiar el conocimiento en la ciencia.

tipo de animal) y entre partes (por ejemplo, el tálamo es parte del diencéfalo, que es parte del sistema nervioso). Estas relaciones son importantes por, al menos, dos razones. En primer lugar, poseen un poder organizativo de las jerarquías que forman. En segundo lugar —y ello constituye su rol fundamental en el sistema conceptual—, especifican los constituyentes del mundo, es decir, la ontología. Con la revolución, se produce un cambio principalmente en relaciones de tipo o relaciones de parte.

Thagard emplea herramientas de inteligencia artificial que permiten describir y analizar la estructura, cambios y mecanismos cognitivos de estos sistemas conceptuales¹⁰. Los sistemas se analizan como una red de nodos, donde los nodos son los conceptos y las líneas de la red son los enlaces entre ellos. Entre los tipos de enlace se cuentan aquellos que indican que un concepto es un tipo o instancia de otro, aquellos que indican propiedades o partes, y reglas que expresan relaciones generales entre conceptos. En este marco pueden distinguirse mecanismos de CC ocurridos a nivel de descubrimiento o de instrucción científica tanto en el caso del desarrollo conceptual como en el reemplazo de sistemas. El desarrollo es entendido como la realización de un sistema conceptual nuevo. Cuando ocurre por descubrimiento, cuenta con mecanismos tales como la generalización empírica¹¹ desde datos experimentales y la abducción —esto es, la formación de hipótesis explicativas—. Esta última aplica a los conceptos teóricos tales como OXÍGENO o FUERZA, que se producen por combinación conceptual (desde partes de conceptos anteriores) y no por descripciones observables, puesto que involucran inobservables. En el caso de la instrucción, se incluye la enseñanza de hipótesis y experimentos a los científicos de la comunidad, quienes los someten a evaluación antes de aceptarlos.

Por su parte, el reemplazo de sistemas conceptuales es el proceso de sustitución de un sistema existente por otro. El reemplazo, que puede producirse también por

¹⁰ Por ejemplo, en *Computational Philosophy of Science*, Thagard (1988) rescata elementos de la noción kuhniana de paradigma tales como los procesos de solución de problemas basados en ejemplares y el rol organizador de las estructuras conceptuales en el pensamiento y percepción de los fenómenos. Estos elementos son capturados por el programa de inteligencia artificial *Processes of Induction* (PI), un solucionador de problemas escrito en LISP que desarrolla una aproximación computacional al descubrimiento y justificación científica. El programa, dado un conjunto de condiciones iniciales y de metas, usa reglas de tipo *si-entonces* para averiguar cómo cumplir las metas. Integra partes deductivas, estadísticas, analógicas y esquemáticas de la explicación. Por ejemplo, aplica reglas como el *modus ponens*, utiliza razonamiento analógico para implementar episodios pasados de solución de problemas en el caso presente y, si es satisfactorio, esquematiza los casos en una caracterización general para emplear en problemas futuros.

¹¹ Las generalizaciones empíricas son enunciados universales expresados por reglas en los sistemas de inteligencia artificial.

descubrimiento o por instrucción, tiene como mecanismo la coherencia explicativa abordada en la *teoría computacional de coherencia explicativa* (TEC)¹². Esta sostiene que la sustitución se realiza si la hipótesis de la nueva teoría posee mayor coherencia que la teoría actual. Así, el científico descubre que una teoría posee mayor coherencia explicativa que la anterior, o bien instruye a otros científicos presentando argumentos a favor de la mayor coherencia. La coherencia explicativa es primariamente una relación entre dos proposiciones: P y Q son coherentes si existe una relación explicativa entre ellas, es decir, si se cumple que una es parte de la explicación de la otra, si juntas son partes de la explicación de otra proposición R o si ambas son análogas en la explicación que respectivamente dan a otras proposiciones R y S (1992, p. 65). Por otro lado, P y Q son incoherentes si se contradicen entre sí u ofrecen explicaciones competitivas.

La apreciación de mayor coherencia explicativa actuará como gatillante para promover el cambio de teorías, entendiendo por *teoría* una colección coherente y estructurada de hipótesis, relaciones conceptuales y éxitos en soluciones de problemas pasados. La transición ocurre en las siguientes etapas. Tras la consideración de la existencia de una teoría que compite con la teoría presente (estado inicial), el científico, aunque escéptico, comienza a aprender más acerca de la nueva teoría y gradualmente comprende sus afirmaciones (1992, p. 104). Ello conduce a la apreciación de la mayor coherencia explicativa de la nueva teoría en comparación con la anterior. Por ello, los CC son vistos como un proceso de solución de problemas que recurre a subprocesos deductivos, estadísticos, esquemáticos, analógicos, causales y lingüísticos.

¹² En cuanto el interés de Thagard radica en el desarrollo de una filosofía computacional de la ciencia, realiza una implementación de TEC por medio de ECHO, un programa computacional con arquitectura híbrida que aplica algoritmos conexionistas y lenguaje de programación LISP para mostrar la computabilidad de las relaciones de coherencia. ECHO contiene parámetros manipulables que modulan los principios propios de las relaciones de coherencia. Por ejemplo, el parámetro de excitación de dato manipula el principio de prioridad de datos, y el parámetro de simplicidad disminuye mayormente los pesos en presencia de cohipótesis. El programa muestra la operación de un algoritmo que convierte las relaciones de coherencia en juicios de aceptabilidad en un sistema de proposiciones. Las proposiciones son representadas como unidades; la coherencia se representa con el establecimiento de enlaces excitatorios con valor positivo, y la incoherencia con enlaces inhibitorios con valor negativo; la aceptabilidad, por su parte, se representa por medio de la activación. El programa establece influencias entre proposiciones para mostrar relaciones explicativas, y permite distinguir entre hipótesis fuertes y débiles. Para ello, toma en consideración propiedades del razonamiento científico, como la simplicidad, la analogía y la importancia de la amplitud explicativa. Por ejemplo, las unidades que se activan suelen ser aquellas que representan hipótesis que explican más, que proveen explicaciones más simples o que son análogas a otras hipótesis explicativas.

En distintas revoluciones se pueden apreciar distintos procesos de descubrimiento. Algunos métodos son generalizados desde las observaciones y experiencia (basados en datos), otros son abductivos (basados en la explicación) y otros formados para superar contradicciones (basados en la coherencia). La amplitud explicativa es la principal contribución en la mayoría de los casos, pero también influyen otros factores como la simplicidad, la contradicción y la analogía. A pesar de los cambios producidos con la revolución (véase pp. 45-46), se mantiene una continuidad por medio de la permanencia de enlaces a otros conceptos en la red. El reemplazo puede ser plausiblemente explicado por competencias entre las fuerzas de las reglas de los sistemas: su empleo frecuente y exitoso en la solución de problemas aumenta la fuerza de las reglas hasta que superen las del sistema antiguo. Estas últimas no desaparecen, sino que se vuelven suficientemente débiles y dejan de figurar en los procesos de pensamiento (1992, p. 58). En el proceso, se va conformando un sistema conceptual nuevo que reemplaza el anterior. No obstante, ambos sistemas se mantienen conectados a un marco de conceptos relativamente estable que no cambia.

Por último, la obra de Nancy Nersessian profundiza en las dimensiones cognitivas que subyacen a la innovación conceptual propia de las revoluciones, pero también de la cognición humana común. El CC es definido por la autora como el problema de entender:

cómo los científicos, trabajando solos o en colaboración, combinan sus capacidades cognitivas humanas con los recursos conceptuales disponibles para ellos como miembros de comunidades científicas y contextos sociales más amplios para crear y comunicar nuevas representaciones científicas de un dominio (1995, p. 204).

Por tanto, el CC corresponde principalmente a la creación y comunicación de nuevas representaciones en el proceso de solución de problemas. Las estructuras conceptuales nuevas surgen tras un razonamiento sistemático para solucionar un problema específico. La dinámica del CC es conformada por aspectos individuales (sensoriomotores), comunitarios y contextuales de la práctica de la ciencia. El individuo, ante una dificultad, emplea sus capacidades cognitivas y herramientas derivadas de un contexto sociocultural determinado. Ello es congruente con los enfoques alternativos en ciencia cognitiva, los cuales extienden los límites de la cognición desde el sujeto hacia su ambiente sociocultural y material (véase pp. 82-83). Para controlar su entorno, el humano desarrolla herramientas mentales por medio de la selección biológica y sociocultural. La ciencia se sitúa en el extremo de este desarrollo: forma parte de las estrategias cognitivas implementadas por los sujetos para entender su entorno y resolver problemas surgidos en

él. Por ende, la cognición científica involucra la historia evolutiva humana, el desarrollo cognitivo infantil y los medios socioculturales en los que se lleva a cabo. La innovación conceptual es un problema representacional, dado que se busca determinar “cómo representar la información conocida para permitir inferencias satisfactorias que van más allá de la información *target* a mano y que conduzcan a nuevas hipótesis para su posterior investigación” (2008, p. xii).

Para la autora, los conceptos son las unidades básicas que permiten representar el mundo. En el nivel del individuo cumplen diversos propósitos, como la categorización de las experiencias, la comprensión del lenguaje, la solución de problemas, entre otros. En el nivel de la ciencia, proveen representaciones sistemáticas para la explicación y predicción de los fenómenos. Su formación y cambio tiene como mecanismo central el *bootstrapping* (véase pp. 32-39). Nersessian toma distancia de las teorías clásicas de representación en ciencia cognitiva —a las que adhiere Thagard—, donde los conceptos se representan por estructuras similares al lenguaje y el razonamiento se realiza por la manipulación con base en reglas lógicas o específicas de dominio. Se acerca a la perspectiva de la cognición corporizada, según la cual un concepto es una estructura neural relacionada con el sistema sensoriomotor. Por ello, la representación involucra aspectos modales, es decir, análogos a los estados perceptuales desde los que se extraen, sin limitarse a los aspectos amodales (esto es, no perceptuales) que caracterizan a las teorías clásicas (véase pp. 125-126).

Los procesos creativos que subyacen al descubrimiento científico conforman el contexto de desarrollo, donde los clásicos contextos de descubrimiento (generación de nuevas ideas en las ciencias) y de justificación (criterios que se requieren para que estas sean aceptadas) se integran. El contexto de desarrollo aborda los procesos de articulación de teorías desde ideas vagas, su comunicación y su reemplazo. La dinámica del CC incluye tanto capacidades individuales como herramientas comunitarias y contextuales. Dada la relación entre el nivel individual y el nivel colectivo, el proceso ocurre primariamente a nivel del científico individual, pero no es exclusivo de este. En efecto, también ocurren cambios representacionales dentro de la comunidad científica. Los cambios se vinculan con el razonamiento basado en modelos que permite la creación de nuevos conceptos, y en el que me detendré más adelante (pp. 33-38).

En suma, la primera consideración por medio de la cual se caracteriza la aproximación clásica del CC consiste en la apelación y problematización de nociones del historicismo

kuhniano. Kuhn estudia los cambios paradigmáticos de manera general, mientras que Thagard y Nersessian se enfocan en procesos cognitivos que subyacen al cambio. No obstante, esta influencia kuhniana no se limita a la historia y filosofía de las ciencias, sino que tiene aplicaciones en otras disciplinas. Ello se relaciona con la segunda consideración de la aproximación clásica: la tesis de que el conocimiento posee rasgos propios de las teorías.

2.2. Conocimiento como teoría

La aplicabilidad de las ideas kuhnianas sobre el desarrollo de la ciencia a otras áreas de estudio es explicada por el mismo autor en *The Structure*:

en la medida que el libro representa el desarrollo científico como una sucesión de periodos atados a la tradición interrumpidos por cortes no acumulativos, sus tesis son indudablemente de amplia aplicabilidad. Pero deben serlo, porque son prestadas de otros campos (1970, p. 208).

Dos de los campos a los que se refiere en la cita corresponden a la psicología de la Gestalt y la psicología del desarrollo. De la primera extrae la noción de cambio gestáltico, esto es, una transformación abrupta en la percepción. Con ella caracteriza, como se verá después (pp. 39-40), el cambio ocurrido en la transición entre paradigmas inconmensurables. Con respecto a la psicología del desarrollo, Kuhn refiere a los experimentos por los cuales Jean Piaget estudia “los distintos mundos del niño en crecimiento y el proceso de transición de uno a otro” (1970, p. vi). Por lo tanto, adopta nociones propias de la psicología para aplicarlas a la historia y filosofía de las ciencias y, a su vez, tanto la psicología como la enseñanza de las ciencias recurren posteriormente a nociones kuhnianas para emplearlas en el estudio del CC en el desarrollo cognitivo humano. Esta relación entre disciplinas justificó en gran medida las analogías en las que se basa la aproximación clásica del CC¹³.

En particular, el trabajo de Kuhn y su intersección con la teoría del desarrollo cognitivo de Jean Piaget (1950, 1976) dan paso a la segunda consideración central en la aproximación clásica: la visión del conocimiento como un marco coherente unificado que

¹³ Kuhn no es explícito con respecto a las influencias exactas de Piaget en su obra. Este punto es problemático en la medida en que dicha influencia es considerada central para la aplicación de la aproximación kuhniana al aprendizaje humano. Además, la relación entre ambos y las analogías que permite han sido criticadas por autores que, como Greiffenhagen y Sherman (2008), sostienen que no existe una compatibilidad entre sus aproximaciones al CC.

exhibe rasgos propios de teoría, denominada *knowledge-as-theory perspectives*, de acuerdo con Özdemir y Clark (2007). Según esta postura, el conocimiento se representa como estructuras conceptuales con rasgos de una teoría, es decir, conforma un conjunto coherente de ideas que generan explicaciones y predicciones que constriñen las creencias y las inferencias que realizan los individuos. Con *conocimiento* se refiere a proposiciones mentalmente representadas que se relacionan con un dominio específico de fenómenos. Por otro lado, *teoría* corresponde a un cuerpo coherente de dominio específico, con una ontología, causalidad y principios explicativos propios, que juega un rol similar a los paradigmas kuhnianos en la determinación de las creencias del sujeto sobre el dominio de fenómenos en cuestión. No se trata, por ende, de una teoría explícita y compartida (Vosniadou, 2008, p. xv). Dos exponentes de esta postura son Posner y colaboradores (1982; Strike y Posner 1992) en la enseñanza de las ciencias, y Susan Carey (1985, 1999a, 1999b, 2009) en psicología del desarrollo.

En el área de la enseñanza de las ciencias¹⁴, la postura anteriormente mencionada representa el enfoque clásico (Vosniadou 2008). Posner y colaboradores definen el CC como la alteración de concepciones¹⁵ organizadoras o generativas en el aprendizaje y el pensamiento durante la enseñanza. Plantearon una teoría del aprendizaje de concepciones científicas como ENERGÍA y FUERZA, cuyo objetivo consiste en determinar “cómo los aprendices hacen una transición desde una concepción, C1, a una concepción sucesora, C2” (Strike y Posner, 1992, p. 148) que es incompatible con la primera, producto de la instrucción. Los autores asociaron, respectivamente, las etapas piagetianas de asimilación y acomodación con las nociones kuhnianas de ciencia normal y cambio paradigmático.

¹⁴ El impacto que tuvo la obra de Kuhn en la enseñanza de las ciencias fue notable. Un ejemplo de ello se puede observar en el análisis de citas al autor llevado a cabo por Loving y Cobern (2000), según el cual la mayor parte de las referencias a sus textos que aparecen en *Science Education* y *The Journal of Research in Science Teaching* entre 1985 y 1998 apoyan notablemente posiciones kuhnianas con respecto al CC.

¹⁵ La noción de concepto que manejan los autores es poco clara. En la teoría inicial planteada en el artículo *Accommodation of a Scientific Conception: Toward a Theory of Conceptual Change* (1982), Posner y colaboradores utilizan indistintamente las nociones de compromiso, concepto y concepción. Reconocen que las dos últimas refieren a distintos niveles de conceptualización (1982, p. 212), pero no explicitan las diferencias entre ellos. Por otro lado, en la revisión de la teoría presente en el texto *A Revisionist Theory of Conceptual Change* (1992), Strike y Posner aclaran que optaron principalmente por CONCEPCIÓN en lugar de CONCEPTO en cuanto pretendían “puntualizar la pluralidad y complejidad interna de estos objetos de cambio, y distinguirlos del término *concepto* como es usado en el discurso normal” (Strike y Posner, 1992, p. 148). No obstante, la elección es problemática, porque en la literatura se ha entendido a los conceptos como aquellas representaciones que refieren a entidades en el mundo, y las concepciones como las creencias acerca de esas entidades (Carey, 2009, p. 489). Los autores se limitan a explicitar su enfoque exclusivo en las conceptualizaciones que juegan un papel generativo u organizador en el pensamiento. Estas constriñen la relación del individuo con la nueva información, e influyen en la manera en que perciben el mundo.

La asimilación es la integración de experiencia ambiental a la estructura cognitiva del sujeto y sus patrones de comportamiento (Piaget, 1950, p. 7), y posibilita la generación de conocimiento. De acuerdo con Piaget,

Ningún comportamiento, incluso si es nuevo para el individuo, constituye un absoluto comienzo. Siempre se injerta en esquemas anteriores y, por tanto, equivale a asimilar nuevos elementos a estructuras ya construidas (innatas, como las reflejas, o previamente adquiridas) (Piaget, 1976, p. 17).

Por tanto, permite tanto la continuidad de las estructuras como la incorporación de nuevos elementos en ella.

La asimilación es el punto de partida o primera fase del CC, y se relaciona con la ciencia normal o vigencia del paradigma. Los autores son fuertemente influenciados por los postulados kuhnianos sobre la racionalidad científica, según los cuales el razonamiento científico se enmarca en un conjunto de creencias sustantivas, técnicas, valores y modelos de solución de problemas que cambian frente a dificultades importantes. Así, aplican la idea de que los problemas y soluciones dependen de las concepciones actuales y compromisos compartidos por la comunidad al aprendizaje de los estudiantes¹⁶. Este aprendizaje se comprende como una actividad racional similar a una investigación, en cuanto las concepciones de los estudiantes con respecto a un tópico cambian a la luz de ideas que son comprendidas y aceptadas por ser vistas como racionales (Posner *et al.*, 1982, p. 212). Además, se atribuyen concepciones previas al aprendiz que constriñen su relación con la nueva información y juegan un papel organizador análogo a los paradigmas kuhnianos. En la asimilación, los estudiantes abordan fenómenos o ideas nuevas empleando sus concepciones previas, y no desde cero.

En concreto, las concepciones previas exhiben tres características análogas a los paradigmas. En primer lugar, son altamente resistentes a la instrucción. De manera similar a como sostiene Kuhn, se interconectan sintáctica y semánticamente con otros conceptos, por lo que el cambio en uno requiere la modificación de otros, siendo difícil la alteración.

¹⁶ Los autores sostienen que la consideración de la literatura de la historia y filosofía de las ciencias otorga una naturaleza epistemológica a sus planteamientos (entendiendo por *epistemología* la determinación de criterios lógicos de las creencias racionales) y normativa (dada la premisa de que el aprendizaje debe ser una actividad racional). Por ello, no “describe el funcionamiento típico de las mentes de los estudiantes o alguna ley del aprendizaje. Tampoco predice ningún resultado de aprendizaje muy específico” (Strike y Posner 1992, p. 155). Dicho de otro modo, no se trata de una teoría psicológica que determina los procesos cognitivos involucrados en el aprendizaje de una ciencia. La racionalidad científica se aplica al aprendizaje estudiantil tomando en consideración la inmadurez del estudiante y dejando de lado los aspectos sociológicos de las comunidades científicas (Strike y Posner, 1992, p. 152).

En segundo lugar, pueden ser vistas como herramientas del pensamiento que permiten determinar cómo se articulan los fenómenos a considerar, qué cuenta como evidencia, etcétera. Por último, funcionan como categorías perceptuales: los fenómenos son percibidos desde las concepciones poseídas, por lo que un cambio en las concepciones implica un cambio en la percepción del mundo (véase pp. 39-40).

Sin embargo, en la modificación de la teoría presente en Strike y Posner (1992), se puntualiza una característica importante que marca la diferencia con los paradigmas. Las concepciones previas no siempre se encuentran totalmente articuladas antes de las concepciones científicas impartidas durante la instrucción. En ocasiones no poseen una formulación completa o son formuladas de manera no simbólica (icónica o enactiva). Por ejemplo, las personas pueden tener “imágenes de cómo funciona o ‘lenguaje corporal’ de cómo funciona” (1992, p. 156) un objeto determinado. Los paradigmas, en cambio, presentan un nivel de expresión simbólica más sofisticado.

El cambio en el conocimiento en la enseñanza de las ciencias, por su parte, se entiende a través de la acomodación. Según Piaget, esta es “toda modificación de un esquema o estructura asimilatoria por los elementos que asimila” (1976, p. 18). En otras palabras, corresponde al proceso de modificación de las estructuras en el desarrollo cognitivo. La acomodación es vinculada por Posner y colaboradores con la fase de la revolución científica, que se caracteriza por la modificación de los compromisos compartidos por la comunidad tras la aparición de anomalías. En esta fase, los estudiantes deben reorganizar sus conceptos centrales frente a la incapacidad para comprender un nuevo fenómeno, y se considera como un CC radical. El gatillante principal es el conflicto cognitivo. La presentación de experiencias nuevas o inconsistentes con las concepciones previas genera un desequilibrio y una insatisfacción con la concepción actual del sujeto que, a su vez, provoca una reestructuración en las estructuras cognitivas. Volveré a este punto en el tercer apartado.

Por su parte, desde la psicología del desarrollo, Susan Carey (1985, 1999a, 1999b, 2009) estudia los CC presentados en el desarrollo cognitivo infantil. Entiende por CC una transformación a nivel de conceptos individuales en sistemas conceptuales sucesivos (C1 y C2) que comparten el mismo dominio de fenómenos. El cambio involucra a todo el sistema dadas las interrelaciones conceptuales establecidas en este. C1 y C2 son inconmensurables entre sí (véase pp. 43-45) porque, dado que se modifican los primitivos conceptuales que representan y explican los fenómenos, los conceptos nuevos no son expresables en el vocabulario anterior (Carey, 2009, p. 365). La autora adhiere a una

versión de la Teoría-Teoría del desarrollo conceptual infantil de acuerdo con la cual buena parte de la adquisición del conocimiento puede ser analizada en términos de reestructuraciones de conocimientos iniciales. Las reestructuraciones se enmarcan en cambios teóricos que poseen similitudes importantes con los cambios en nociones centrales de teorías científicas sucesivas en la historia de la ciencia (2009, p. 362).

El estado inicial del CC incluye el repertorio representacional innato¹⁷ y estructuras de conocimiento explícito lingüísticamente codificado, dentro de las que destacan las teorías intuitivas. En sus trabajos iniciales (1985, 1999b), Carey describe el repertorio innato como aquel que incluye un conjunto de dominios de contenido del mundo (dominios núcleo) y un conjunto reducido de estructuras conceptuales (conocimiento núcleo). El conocimiento núcleo determina algunas distinciones ontológicas y atribuciones causales, y su adquisición está guiada por principios innatos de dominio específico y más cercanos a los datos perceptuales. Por su parte, las teorías intuitivas juegan roles similares a las teorías científicas, como la determinación de características esenciales de un concepto que ayuda a la categorización, la representación de conocimiento causal y explicativo, y la realización de inferencias explicativas y predicciones. Se destacan una teoría intuitiva de los objetos físicos y una psicología intuitiva ya presentes a los 12 meses (1999b, p. 311). La teoría intuitiva de los objetos físicos tiene a *objeto físico* como clase ontológica central, y *causalidad directa* como principio causal básico; en el caso de la psicología intuitiva, lo son *persona* y *causalidad intencional* respectivamente.

El desarrollo cognitivo corresponde, en buena parte, a la emergencia de nuevos dominios teóricos desde estos dominios iniciales, junto con “la reestructuración concomitante de los conceptos ontológicamente importantes y la emergencia de nuevas nociones explicativas” (1985, p. 14). Un ejemplo figura en *Conceptual Change in Childhood* (1985), donde Carey se enfoca en el CC presentado en el conocimiento biológico durante los 4 y los 10 años. Evalúa nociones tales como ANIMAL, PERSONA, conceptos de plantas, animales y procesos corporales, realizando una comparación entre niños y adultos. La autora determina que, en ese rango de edad, los infantes presentan una

¹⁷ En este contexto, Carey emplea *innato* como *no aprendido*, y no como *presente desde el nacimiento*. Es decir, las representaciones que son innatas son aquellas que no resultan de procesos de aprendizaje. Estos procesos de aprendizaje son entendidos como “aquellos que construyen representaciones del mundo sobre la base de computaciones en un *input* que es en sí mismo representacional” (2009, p. 12). La teoría que sostiene la autora critica la reducción del repertorio representacional inicial a representaciones sensoriales.

reestructuración que “puede ser concebida como la emergencia de una nueva teoría (una biología intuitiva) desde su teoría madre (una teoría intuitiva de comportamiento animal)” (1985, p. 8). El núcleo de esta emergencia es el concepto de VIDA, que se desarrolla por medio de la adquisición y reorganización del conocimiento biológico (véase pp. 51-52).

Para Carey, los conceptos, las creencias y las teorías son estructuras mentalmente representadas.

Los conceptos son unidades de representación mental, algo así como los constituyentes de los *ítems* léxicos individuales, tales como *objeto*, *materia* y *peso*. Las creencias son proposiciones mentalmente representadas que el creyente considera verdaderas, tales como *El aire no está hecho de materia*. Los conceptos son los constituyentes de las creencias, esto es, las proposiciones son representadas por estructuras de conceptos. Las teorías corresponden a estructuras mentales complejas que consisten en un dominio de fenómenos representados mentalmente y un principio explicativo que da cuenta de ellos (1999a, pp. 459-460).

En otras palabras, los conceptos son unidades de pensamiento, símbolos mentales que constituyen las creencias y las teorías, y forman parte del conjunto de representaciones mentales que posee el humano. Estas representaciones son estados del sistema nervioso con contenido. Pueden referir a objetos, propiedades o eventos, ya sean concretos, abstractos o ficcionales (2009, p. 5).

En *The Origin of Concepts* (2009), la autora profundiza en la distinción entre las representaciones conceptuales pertenecientes a los sistemas de cognición núcleo y aquellas que figuran en las teorías intuitivas. La noción de cognición núcleo proviene del *conocimiento núcleo* empleado anteriormente (el cambio de terminología puntualiza que las representaciones que pertenecen a la cognición núcleo no necesariamente son verdícas, por lo que no constituyen conocimiento bajo la noción clásica de *conocimiento* como creencia verdadera justificada).

Por una parte, la cognición núcleo es el “fundamento del desarrollo de la comprensión conceptual humana” (2009, p. 11). Corresponde a una fuente de representaciones que incluye los objetos de tamaño y de distancia media, los números, entre otras. Los sistemas de cognición núcleo poseen seis características principales. En primer lugar, presentan contenido conceptual¹⁸, por lo que no se reducen a representaciones sensoriales, perceptuales o sensoriomotoras. Juegan un rol inferencial

¹⁸ Para Carey, el contenido conceptual se determina por un enlace causal entre ciertas entidades en el mundo y las instanciaciones de los símbolos en la mente, y dicha relación se establece por selección natural. Esta postura es denominada *teoría causal de la determinación del contenido*.

importante para el pensamiento. En segundo lugar, sus representaciones son adquiridas por dispositivos de *inputs* perceptuales que son modulares. La modularidad implica que se trata de mecanismos innatos, automáticos, obligatorios, de dominio específico y encapsulados (véase pp. 101-104). La operación de estos dispositivos es determinada por procesos evolutivos que garantizan “que las representaciones relevantes refieran a aspectos del entorno que son importantes para la supervivencia” (2009, p. 29). Así, la selección natural ha definido las conexiones causales entre las entidades en el mundo y las representaciones en la mente (estados del sistema nervioso). En tercer lugar, los dispositivos poseen una historia de desarrollo, por lo que operan a lo largo de la vida. En cuarto lugar, los sistemas de la cognición núcleo son dispositivos de aprendizaje: permiten procesos de aprendizaje individual que crean nuevos recursos representacionales con información útil del entorno. En quinto lugar, algunos sistemas de cognición núcleo no son exclusivos del humano, sino que se comparten con otras especies (por ejemplo, primates no humanos). Por último, el formato de sus representaciones es probablemente icónico en lugar de estructurado en términos de símbolos explícitos (como aquellos del lenguaje o las matemáticas). La diferencia entre ambos formatos radica en que el formato icónico es analógico, es decir, las partes del símbolo representan las partes correspondientes a la entidad representada. Ello contrasta con las representaciones simbólicas similares al lenguaje, que no contienen información sobre dichas partes.

Las teorías intuitivas son distintas a la cognición núcleo en diversas aristas. En primer lugar, no son comandadas por dispositivos de *input* innatos y modulares, por lo que sus conceptos no son innatos ni causalmente conectados a las entidades que representan. Los mecanismos que establecen relaciones entre ellas y las entidades que representan incluyen procesos sociales. En segundo lugar, no son continuas durante el desarrollo conceptual, sino que están sujetas a cambio por medio de mecanismos tales como el *bootstrapping quineano*, que abordaré más adelante (pp. 32-39). En tercer lugar, no dependen de dispositivos de aprendizaje de dominio específico, y tienen un rol inferencial mayor. Por último, su formato de representación no suele ser icónico, sino que se representa por símbolos explícitos como los del lenguaje y las matemáticas. Las denominadas teorías marco (*framework theories*) son especialmente relevantes, porque dictan los compromisos ontológicos y los principios explicativos más fundamentales con los cuales el sujeto entiende el mundo.

A pesar de estas diferencias, existe una superposición entre la cognición núcleo y las teorías intuitivas. Dado que la cognición núcleo es un dispositivo de aprendizaje,

permite el desarrollo de las teorías intuitivas. En otras palabras, “el *output* de los sistemas de cognición núcleo es parte del *input* de la construcción de teorías” (2009, p. 24). Además, ambas estructuras constriñen las creencias del sujeto con respecto a un dominio específico de fenómenos. Este constreñimiento adquiere la forma de una determinación de las propiedades esenciales de un concepto para pertenecer a una categoría¹⁹, una representación del conocimiento explicativo y causal, un soporte de las inferencias y compromisos ontológicos. Por ello, “la cognición núcleo de los objetos se superpone con el conocimiento de la física intuitiva, y la cognición núcleo de los agentes intencionales se superpone con una teoría intuitiva de la mente” (2009, p. 24).

A modo de sumario, la segunda consideración central para la caracterización de la aproximación clásica del CC sostiene que el conocimiento posee una estructura similar a las teorías, es decir, permite realizar explicaciones y predicciones que constriñen las creencias e inferencias del sujeto. Tanto los trabajos de Posner y colaboradores como los de Carey se enmarcan en esta propuesta, aplicando rasgos generales del estudio kuhniano de los cambios paradigmáticos al proceso de desarrollo y aprendizaje estudiantil e infantil respectivamente. A continuación, me detendré en la tercera y última consideración de la aproximación clásica: la hipótesis del continuo entre la cognición científica y la cognición humana común, que de una u otra forma subyace a todas las posturas vistas.

2.3. Hipótesis del continuo

La tercera y última consideración que caracteriza la aproximación clásica del CC corresponde a la hipótesis del continuo, según la cual la cognición científica es una extensión sofisticada de la cognición humana común. La hipótesis subyace, con algunas diferencias, a las elaboraciones teóricas de cada uno de los exponentes vistos anteriormente, puesto que justifica la aplicación de nociones propias de la historia y filosofía de las ciencias al desarrollo cognitivo humano y, a su vez, la aplicación de herramientas psicológicas a la historia y filosofía de las ciencias. Como mencioné con anterioridad, la aproximación clásica configura una cuádruple analogía entre los cambios de teorías en ciencias, los científicos como sujetos epistémicos, los estudiantes de ciencias

¹⁹ Carey (1985) sostiene que el sistema conceptual humano se articula por un conjunto de categorías ontológicamente básicas como, por ejemplo, *ser vivo*. La realización de una mayor cantidad de distinciones ontológicas forma parte del desarrollo cognitivo.

y los niños²⁰. Así, Posner y colaboradores mantienen una similitud entre estudiantes y científicos en relación con la posesión de concepciones previas similares a los paradigmas y los rasgos de sus procesos de cambio; Carey establece una relación entre los CC ocurridos en el desarrollo conceptual infantil y los cambios en nociones centrales de teorías científicas sucesivas, donde el uso del *bootstrapping* es común para ambos (véase pp. 32-39); Thagard vincula los CC en ciencias con aquellos que presentan los científicos involucrados en el proceso de solución de problemas, y con los estudiantes en su proceso de instrucción (para este segundo punto, véase p. 57).

La hipótesis del continuo encuentra una exposición más completa en el análisis histórico-cognitivo de Nersessian, quien se centra en los CC en ciencias. La autora estudia cómo las estructuras conceptuales son desarrolladas y transformadas por medio de este método pluralista que combina (i) exámenes de casos históricos de prácticas teóricas y experimentales científicas, (ii) teorías sobre las capacidades cognitivas humanas provenientes de la ciencia cognitiva, e (iii) investigaciones epistemológicas y de filosofía de las ciencias. La dimensión histórica investiga archivos históricos (diarios, correspondencia, publicaciones, entre otros) para obtener una visión de los logros de los agentes creativos y su integración en la comunidad y contexto sociocultural. La dimensión cognitiva, por su parte, considera que tales prácticas tienen como base capacidades cognitivas humanas comunes, y las interpreta desde las teorías cognitivas disponibles en la literatura²¹. La hipótesis del continuo es el supuesto básico de este método y sostiene que:

las estrategias de solución de problemas que los científicos han creado y las prácticas representacionales que han desarrollado en el transcurso de la historia de la ciencia son una evolución muy sofisticada y refinada de los procesos ordinarios de razonamiento y representación (Nersessian, 1992, p. 5).

²⁰ Como mencionaré en la sección de discusión, algunos autores problematizan ciertas aristas de esta cuádruple analogía. Por ejemplo, Thagard critica su aplicación al desarrollo cognitivo infantil, mientras que Carey apunta a la baja evidencia en el caso de los estudiantes. No obstante, en términos generales, se mantienen las analogías entre los científicos/las teorías científicas y los niños/los estudiantes.

²¹ Como tal, es considerado un método naturalista: sostiene la realización de una filosofía de la ciencia informada por las mejores teorías científicas y las prácticas reales de construcción de conocimiento utilizadas por los científicos. Además, apela al planteamiento y testeo de hipótesis filosóficas por medio de métodos empíricos —como el análisis histórico, los experimentos psicológicos y la modelación computacional.

Por tanto, la congruencia entre las prácticas históricas y los análisis cognitivos sirven tanto para estudiar las actividades científicas como para enriquecer los modelos de cognición desarrollados en psicología cognitiva.

En el presente apartado abordé la aproximación clásica del CC. Sostuve que esta puede caracterizarse mediante tres consideraciones principales: la apelación al historicismo kuhniano, la postura según la cual el conocimiento posee rasgos propios de las teorías y la hipótesis del continuo. En el próximo apartado me concentraré en tres puntos principales vinculados con el análisis de las discontinuidades conceptuales que se desprenden de las posturas anteriormente revisadas: el contexto del CC, tal y como puede ser entendido desde la noción de ecología conceptual de Posner y colaboradores; el *bootstrapping* como mecanismo central de cambio, estudiado por Carey y Nersessian, y la comparativa entre el estado inicial y el estado final del proceso, por medio de la indagación de la existencia o inexistencia de inconmensurabilidad. Estos aspectos serán posteriormente de utilidad para el abordaje de los CC en la noción de REPRESENTACIÓN en ciencia cognitiva.

3. Proceso de cambio conceptual

Dentro de la aproximación clásica del CC, las discontinuidades conceptuales usualmente toman la forma de cambio de teoría. Esta concepción tiene como fuente la consideración del historicismo kuhniano. Como mencioné anteriormente, Kuhn sostiene que los conceptos se encuentran embebidos en teorías y sufren una transformación tras el proceso de cambio teórico. Ahora bien, ¿cómo se desarrolla el CC? En el presente apartado me enfocaré en algunos rasgos relevantes del proceso que se derivan de las propuestas vistas en el apartado anterior: el contexto del cambio, un mecanismo central y su resultado. Con respecto al primer punto, abordaré la noción de ecología conceptual (Posner *et al.*, 1982; Strike y Posner, 1992). En relación con el segundo punto, explicaré el *bootstrapping* (Carey, 2009; Nersessian, 2008). Por último, indagaré sobre las nociones de inconmensurabilidad y cambio gestáltico dada la comparativa entre el estado inicial y el estado final del proceso.

3.1. Contexto de los cambios conceptuales: la ecología conceptual

Los conceptos no cambian en el vacío: sus transformaciones se producen dentro de un sistema que da soporte y coherencia a sus desarrollos y reestructuraciones. Una noción útil desde la cual se aborda esta característica corresponde a la *ecología conceptual*, planteada por Posner y colaboradores. Los autores sostienen que, durante el proceso de CC, existen conceptos centrales que el sujeto mantiene y que sirven de guía para este. Tales conceptos conforman la ecología conceptual, el contexto por medio del cual el sujeto concibe los conceptos nuevos. Sobre la base de ella se entienden tanto las concepciones previas como las concepciones que son candidatas para reemplazarlas.

Para que ocurra el proceso de CC, se requiere la presentación de cuatro condiciones. Primeramente, debe ocurrir una insatisfacción con los conceptos actuales provocada por la acumulación de dificultades que estos no pueden resolver²². Ello ocurre, por ejemplo, con estudiantes que poseen conceptos newtonianos de ESPACIO y TIEMPO y se enfrentan a un curso de relatividad especial. Los aprendices necesitan entender por qué los resultados experimentales que se les presentan constituyen una anomalía, así como tener la disposición de reconciliarlos con sus concepciones actuales y de reducir las inconsistencias entre sus propias creencias. En última instancia, fracasan en el intento de asimilar dichos resultados, lo que constituye una motivación para gatillar el proceso de acomodación: “las personas no acomodan mientras la asimilación sea aún razonable” (Strike y Posner, 1992, p. 149). Frente a la presencia de dificultades, tenderán a realizar cambios menores en su esquema conceptual, hasta que se haga aparente la necesidad de una transformación más profunda. Ello es análogo al comportamiento presentado por las comunidades científicas frente a la crisis desencadenada por la presencia de anomalías, quienes evitan renunciar en primera instancia al paradigma que condujo a la crisis y realizan modificaciones *ad hoc* con el afán de eliminar el conflicto. La segunda condición para que ocurra el CC es la existencia de una nueva concepción que sea inteligible, esto es, que el sujeto pueda poseer al menos un sentido de cómo se estructuraría la realidad de ser verdadera. De lo contrario, no será comprendida y, por lo tanto, no será candidata para reemplazar a la concepción anterior. La tercera condición apunta a la plausibilidad de dicha concepción, es decir, que exhiba capacidad de resolver las anomalías. En último

²² El objetivo principal en la enseñanza, bajo esta perspectiva, es crear un conflicto cognitivo para generar la insatisfacción.

lugar, debe ser fructífera, al posibilitar la extensión de las indagaciones a otros campos y al ser una herramienta de pensamiento productiva.

La ecología conceptual incluye al menos cinco características o artefactos cognitivos en la primera formulación de la teoría de los autores (Posner *et al.*, 1982). En primer lugar, las anomalías gatillan la insatisfacción del sujeto hacia sus conceptos actuales. En segundo lugar, las analogías permiten la inteligibilidad de las concepciones nuevas. En tercer lugar, los compromisos epistemológicos (ideales explicativos y visiones acerca del carácter del conocimiento) permiten determinar un modelo de explicación satisfactoria y de conocimiento adecuado. En cuarto lugar, las creencias metafísicas acerca de la ciencia (como, por ejemplo, la concepción del universo como un todo ordenado) y los conceptos metafísicos de esta (como los conceptos que apuntan a la naturaleza última del universo) pueden jugar un rol en la preferencia por un tipo de explicación determinado. Por último, también forma parte de la ecología el conocimiento sobre otras áreas y sobre los conceptos en competencia.

En la revisión de la teoría planteada por Strike y Posner (1992) se amplía el rango de factores a considerar, incluyendo las concepciones científicas actuales, las concepciones previas erróneas, motivaciones y metas. Todos los elementos se mantienen en constante interacción. De las características anteriormente mencionadas, las anomalías, compromisos epistemológicos y creencias metafísicas funcionan como guías fundamentales del CC. Las anomalías surgen frente a la incapacidad de asimilar algo que es asimilable, y pueden provocar un conflicto cognitivo similar a la crisis en sentido kuhniano, que “prepara la ecología conceptual del estudiante para una acomodación” (Posner *et al.*, 1982, p. 224). Por su parte, los compromisos epistemológicos y creencias metafísicas constituyen los supuestos fundamentales sobre el conocimiento y la ciencia, formando una base sobre la cual se juzga el nuevo conocimiento. Por ejemplo, la creencia en la existencia de un espacio y tiempo absolutos conlleva a que los estudiantes consideren como contraintuitiva la teoría de la relatividad especial de Einstein. Para que adopten satisfactoriamente la interpretación del espacio y el tiempo como relativos a cualquier sistema inercial, requieren transitar por el proceso de acomodación.

Como puede observarse, Posner y colaboradores mantienen sus influencias kuhnianas en la elaboración de la noción de ecología conceptual: de manera general, tiene una similitud con los paradigmas y, de manera específica, incluye elementos tales como las anomalías. Por tanto, si bien esta noción es formulada en el marco de la enseñanza de

la ciencia para caracterizar las transformaciones ocurridas en los estudiantes, es cercana a estudios relacionados con la historia y la filosofía de las ciencias²³.

Una de las ventajas de utilizar la noción de ecología conceptual en lugar de la noción kuhniana de paradigma para el estudio del CC en REPRESENTACIÓN en ciencia cognitiva radica en su énfasis en elementos tales como las analogías, que forman parte de los mecanismos clave para el proceso. Por ello, optaré por *ecología conceptual* en el análisis. En el tercer capítulo me concentraré principalmente en tres elementos englobados en la ecología: los compromisos epistemológicos, las anomalías y las analogías. Los compromisos epistemológicos serán abordados por medio de la especificación de algunos componentes centrales para la caracterización de la noción de REPRESENTACIÓN en cada propuesta vista: la relación entre representaciones perceptuales y conceptuales, la relación entre representaciones mentales y palabras del lenguaje natural, y la relación entre representaciones mentales y representaciones neurales. Eventualmente realizaré algunos alcances sobre implicancias ontológicas relacionadas. Las anomalías se observarán a través de las discusiones que gatillan los cambios. No obstante, optaré por nociones neutrales como *problema*, puesto que el concepto de *anomalía* parece encontrarse estrechamente ligado al proceso revolucionario de ciencias consideradas como maduras en sentido kuhniano. Por último, las analogías serán consideradas como un mecanismo principal que gatilla la creación, cambio y comunicación de conceptos. En el próximo subapartado me detendré en la explicación del proceso del cual forman parte las analogías, es decir, el *bootstrapping*.

²³ Por ejemplo, el contexto del problema al que se enfrenta el sujeto epistémico ha sido ampliamente abordado por Nersessian. Dado que la transformación conceptual es principalmente un proceso de solución de problemas, la primera etapa incluye los recursos cognitivos, el contexto sociocultural y material del sujeto epistémico. Los recursos cognitivos incorporan elementos tales como los conocimientos de dominio específico y la experiencia en la solución de problemas. En el contexto sociocultural se cuentan las comunidades de las que forma parte y los factores culturales que ejercen influencia sobre él (como la religión que profesa). Por último, el contexto material se vincula con recursos e instrumentos enmarcados en las prácticas científicas. Un ejemplo de la operación de los elementos mencionados se encuentra en el estudio histórico-cognitivo de Michael Faraday, quien desarrolló un concepto de campo para fuerzas eléctricas y magnéticas. El estado inicial que propició esta formulación incluye la capacidad del científico de razonar a partir de la creación de representaciones visuales (una representación visual de líneas de fuerza con movimientos diversos, y una representación visual de curvas entrelazadas que representan la relación entre las líneas eléctricas y magnéticas); la religión sandemania del científico, que influyó en su creencia de que las fuerzas de la naturaleza se unifican por medio de dichos movimientos; su trabajo junto a Humphrey Davy y los recursos materiales involucrados en sus experimentos sobre electromagnetismo.

3.2. Un mecanismo de producción del cambio conceptual: el *bootstrapping*

En la literatura de la aproximación clásica del CC se han planteado diversas explicaciones sobre los elementos que gatillan las discontinuidades conceptuales. Por ejemplo, en la enseñanza de las ciencias, Posner y colaboradores (1982; Strike y Posner, 1992) sostienen que el factor principal es la ecología conceptual expuesta en el subapartado anterior; mientras que, para Thagard (1992), la coherencia explicativa cumple ese rol (véase pp. 16-17). Sin embargo, estos elementos no explican los detalles del proceso de innovación conceptual a nivel cognitivo, sino que muestran algunos factores generales que impulsan o motivan dicha innovación. En cambio, un recurso principal que subyace a la creación, alteración y comunicación de conceptos es el *bootstrapping*, estudiado por Carey (2009) y Nersessian (2008). La justificación de mi énfasis en el *bootstrapping* radica en su importancia para el CC en las ciencias y, en particular, en la ciencia cognitiva. Este mecanismo se encuentra a la base del razonamiento analógico y, como mencionaré en el tercer capítulo (pp. 140-143), las analogías constituyen un punto importante en las transformaciones propias de REPRESENTACIÓN. En el presente subapartado, me concentraré en la descripción general de las dos etapas del *bootstrapping* propuesta por Carey, y el razonamiento basado en modelos estudiado por Nersessian como punto principal de la segunda etapa.

El *bootstrapping* es un mecanismo que permite la creación de nuevos recursos representacionales que no se reducen a las representaciones previas. Su operación subyace tanto a la cognición científica como a la cognición humana común²⁴. Carey (2009) sostiene que este mecanismo permite explicar las discontinuidades conceptuales surgidas en teorías intuitivas con respecto a los sistemas de cognición núcleo y en dominios particulares de conocimiento en el desarrollo cognitivo infantil. Nersessian (2008), por su parte, se enfoca en su aporte para la innovación conceptual propia de los cambios de teorías en la historia de las ciencias, y afirma que es un elemento central tanto para los CC débiles como para los CC revolucionarios en sentido kuhniano (2008, p. 187).

²⁴ A propósito de este punto, Nersessian muestra la operación de este mecanismo en la reconceptualización del dominio del electromagnetismo de Maxwell y en la integración de la noción TORSIÓN en el concepto de resorte de un individuo durante la aplicación de un protocolo *Think-Aloud*. En el primer caso, surge una idea fundamentalmente nueva en la historia de la ciencia. En el segundo, surge una idea nueva —ya disponible culturalmente— para un individuo determinado.

Para ambas autoras, se crean nuevos recursos representacionales que “permiten pensamientos que no podíamos considerar anteriormente” (Carey, 2009, p. 333).

La complejidad de este proceso es ilustrada por Carey, quien acuña el término *bootstrapping quineano* en honor a las metáforas que planteó Quine para su descripción. Dentro de ellas, el bote de Neurath es especialmente famosa: asimila la construcción de teoría en casos de discontinuidad conceptual con construir un bote mientras se está sobre él en medio del océano. En otras palabras, muestra la dificultad de la creación de representaciones nuevas que trascienden las representaciones antecesoras, puesto que, en un inicio, la persona no posee los conceptos en el dominio *target* (el dominio de interés) que mapeará en los conceptos relevantes en el dominio fuente (el dominio conocido).

Carey identifica dos etapas centrales en el desarrollo del *bootstrapping*: el empleo de marcadores (*placeholders*) y la elaboración de fundamento conceptual para dichos marcadores. Con respecto a la primera etapa, se emplean símbolos explícitos nuevos (del lenguaje escrito, de notación matemática, diagramas, entre otros) que sirven como marcadores, esto es, estructuras con información incompleta que será posteriormente completada. Por ejemplo, Kepler empleó el marcador *vis motrix* en la elaboración de su hipótesis de que el sol causa el movimiento de los planetas, y la fórmula $D = W/V$ (la densidad es igual al peso dividido por el volumen) funciona como un marcador en el proceso de distinción entre peso y densidad en los niños. El significado de este sistema de marcadores se limita a la interrelación entre los símbolos. El individuo interpretará estos símbolos inicialmente solo de manera parcial, y en términos de los conceptos que ya posee, puesto que todavía no cuenta con la capacidad de formular los conceptos que los símbolos representarán. Entre científicos y niños existen diferencias en la fuente de los marcadores: mientras los primeros postulan los símbolos como ideas tentativas, los segundos los adquieren mediante la enseñanza explícita.

En relación con la segunda etapa, el *bootstrapping* utiliza razonamientos basados en modelo para otorgar fundamento conceptual a los marcadores. Los marcadores introducen símbolos mentales nuevos, y los procesos de modelación les otorgan contenido. Carey recurre a la obra de Nersessian (2008) sobre la capacidad humana de pensamiento simulativo por medio de la modelación mental para el abordaje de este tipo de razonamiento, por lo que me remitiré a dicha obra a continuación.

Nersessian estudia la capacidad de representar situaciones reales o imaginarias desde la memoria y de realizar inferencias a partir de ellas en ausencia de la situación física que es objeto de razonamiento. Para la autora, los procesos de solución de problemas

usualmente involucran la creación de un modelo icónico de las situaciones, procesos y eventos que puede ser manipulado. Un modelo mental es “una representación analógica estructural, conductual o funcional de una situación, evento o proceso del mundo real o imaginario” (2008, p. 93). En otras palabras, son unidades que representan rasgos de un sistema físico determinado que son pertinentes para el problema que se está considerando, y que son empleadas en tareas cognitivas como la resolución de problemas²⁵. Un razonamiento basado en modelos, por su parte, es el proceso de elaboración, manipulación y revisión de modelos mentales en la memoria de trabajo que son estructural, conductual o funcionalmente análogos al fenómeno *target*. En él interactúan representaciones internas y externas que apuntan a la conformación de un sistema acoplado. Así, si bien parte del razonamiento ocurre en el medio interno (en la cabeza), requiere emplear representaciones externas de varios formatos —lingüísticos, físicos, matemáticos o gestos, entre otros—. El entorno forma parte del sistema cognitivo al proveer constreñimientos (*constraints*) y posibilidades efectivas (*affordances*). Mientras que los primeros son regularidades en un dominio que restringen ciertas acciones en cuanto dependen de condiciones específicas, las segundas son recursos que ayudan a la solución de problemas. Entre ellos se cuentan aspectos espaciales, temporales, funcionales, causales entre otros. Esta interacción con representaciones externas permite a la autora sostener que los modelos poseen aspectos multimodales²⁶, es decir, mantienen características perceptuales.

La segunda etapa (y etapa central) del *bootstrapping* consiste en “ciclos de construcción, simulación, evaluación y adaptación de modelos que sirven como interpretaciones provisionales del problema *target*” (Nersessian, 2008, p. 184). Se integran sistemas representacionales diferentes, puesto que los recursos conceptuales de

²⁵ Los modelos mentales poseen al menos tres características relevantes. En primer lugar, son analógicos: mantienen constreñimientos propios del sistema representado, entendiendo por *constreñimiento* los aspectos del sistema que restringen las acciones o interacciones posibles con él. En segundo lugar, son abstracciones idealizadas y esquemáticas: no representan la totalidad del sistema físico, sino que se concentran en propiedades, relaciones u objetos que son relevantes para el razonamiento que se está llevando a cabo (por ejemplo, relaciones espaciotemporales, causales, entre otras). En tercer lugar, y en vínculo con los rasgos anteriores, poseen una naturaleza icónica: mantienen similitudes con el sistema físico que representan. Las similitudes pueden presentar distintos grados, siendo el isomorfismo el grado máximo. Las representaciones icónicas incluyen información de diversos formatos, tales como fórmulas, formato lingüístico, visual, auditivo, etcétera. Existen distintos tipos de modelos: computacionales, matemáticos, físicos o conceptuales, y son considerados como las unidades básicas de los científicos cuando trabajan con teorías (entendidas como sistemas representacionales).

²⁶ Me detendré en la noción de multimodalidad en el tercer capítulo (véase pp. 129-130) a propósito de los sistemas de símbolos perceptuales de Lawrence Barsalou, uno de los autores en quienes Nersessian se basa.

un dominio se proyectan como fuente de la estructura conceptual de otro. El individuo inicia la construcción del modelo con una comprensión básica del dominio *target*. La comprensión básica distingue rasgos generales del fenómeno en cuestión, y de ella se extraen los constreñimientos iniciales que, junto con el objetivo perseguido, determinan qué dominio analógico fuente se empleará y cuáles son sus constreñimientos relevantes. En la selección se pueden usar procesos simulativos o imaginativos.

Posteriormente, se crea un modelo inicial por medio de procesos de abstracción²⁷, que posibilitan el enfoque selectivo en rasgos relevantes. En este modelo se integran los constreñimientos del dominio *target* y el dominio fuente. Tras la construcción, se utiliza la simulación para manipular el modelo, examinar la relación de los constreñimientos y realizar predicciones. La simulación es definida como “un esfuerzo epistémico cuyo propósito es fomentar la comprensión e inferencia mediante la predicción de estados futuros de los modelos y el examen de las consecuencias de estos” (2008, p. 199). Esta manipulación posiblemente destaca constreñimientos no considerados con anterioridad. Luego, se realiza una evaluación en la que se comparan los rasgos del modelo y el *target* y se determina si el modelo representa todos sus constreñimientos relevantes. De ser así, la solución al problema en el modelo se proyecta al *target* y se evalúa su rendimiento frente a los requerimientos del problema. De lo contrario, se lleva a cabo la adaptación. En todo el proceso, la selectividad es clave, en cuanto permite enfocar la atención en rasgos relevantes y desechar los irrelevantes. En suma, en el *bootstrapping* “cada dominio suministra constreñimientos que pueden considerarse como uno de los cordones, los modelos híbridos intermedios son los cruces de los cordones, y cada cruce apoya o contribuye a una mayor construcción del modelo y a una mejor comprensión del *target*” (2008, p. 133).

Algunos tipos de razonamiento basados en modelo que son centrales para el cambio representacional son el razonamiento analógico, el razonamiento imaginista y los experimentos mentales. Los razonamientos analógicos consisten en el empleo de similitudes relacionales entre los dos dominios (*target* y fuente) para realizar inferencias. La solución del problema en el dominio fuente se proyecta al dominio *target*. Para ello,

²⁷ Entre estos procesos se incluyen las idealizaciones —es decir, la suposición de hechos del fenómeno que, en estricto rigor, no se cumplen, pero que ayudan a la comprensión del modelo y a su relación con representaciones matemáticas—; las generalizaciones —donde las propiedades comunes de las instancias se formulan como enunciados universales de tipo *Todo A es B*—, y las abstracciones genéricas —un proceso en el cual se suprimen los detalles de la instancia del objeto o fenómeno para enfocarse en el caso genérico y hacer inferencias sobre él.

usualmente se requiere que los dominios se representen de forma más abstracta: se aíslan las propiedades o relaciones relevantes para establecer la analogía, y se consideran de manera separada al resto de sus cualidades. Por ejemplo, el modelo atómico clásicamente atribuido a Rutherford establece una analogía entre el sistema solar y los átomos, donde ambos son representados como entidades localizadas al centro con otras entidades que giran en torno a ellas.

No obstante, en ocasiones no es posible realizar proyecciones de un dominio a otro de manera directa. En esos casos, es necesario realizar un modelo intermedio que permita dichas proyecciones. La fuente provee constreñimientos que se conjugan con los constreñimientos del *target* para conformar este modelo intermedio, que es híbrido porque consiste en la conjugación de constreñimientos del dominio fuente, del dominio *target*, y de sus propios constreñimientos. La solución al problema se logra por medio del examen de las interacciones entre los tres, luego de lo cual se puede proyectar al problema *target* original. Por ejemplo, Maxwell realizó una analogía entre el electromagnetismo y la mecánica de medios continuos en la que representó las fuerzas magnéticas como movimientos de vórtices en un medio fluido elástico, lo que le permitió extraer representaciones matemáticas propias de los fenómenos continuos para emplearlas en la representación del electromagnetismo. Aquí, el dominio fuente es la mecánica de medios continuos y la mecánica de máquina, y el dominio *target* es el electromagnetismo. Como el dominio fuente no permitía la proyección directa, Maxwell construyó una serie de modelos híbridos que conjugaban los constreñimientos de los dominios: el modelo de fluidos de vórtices y el modelo de ruedas pasivas.

El razonamiento imaginista, por su parte, corresponde a la construcción y manipulación de representaciones esquemáticas que se emplean para analizar y predecir. Estas representaciones son principalmente perceptuales, y se destacan aquellas que pertenecen a la modalidad visual. Permiten realizar inferencias perceptuales inmediatas y procesos de simulación, así como comunicar nuevas representaciones a otros sujetos por medio de la instrucción. Ello se debe a la accesibilidad con la que muestran las partes relevantes del fenómeno en investigación (en comparación con textos o fórmulas). Un ejemplo corresponde a la creación de visualizaciones esquemáticas externas (en papel), como los diagramas. Maxwell realizó diagramas que expresan su modelo de ruedas pasivas, y estas representaciones visuales en dos dimensiones fueron un soporte para el análisis matemático y permitieron comunicar sus avances a los otros científicos

(ayudándolos, a su vez, a construir sus propios modelos mentales para entender el modelo) (Nersessian, 2008, p. 259).

Por último, los experimentos mentales constituyen un tipo de razonamiento simulativo que involucra la construcción y cambio de un modelo mental mediante la imaginación de casos futuros. Un modelo de experimento mental es “un sistema conceptual que representa el sistema físico sobre el que se razona” (Nersessian, 2008, p. 179). Por medio de la manipulación del modelo se determina qué sucede si se realizan operaciones específicas. Sus formulaciones pueden adquirir la forma de historias (donde la narrativa describe una secuencia de eventos que llama a la imaginación del receptor) o de preguntas contrafácticas de tipo *¿Qué pasaría si...?* Poseen supuestos explícitos e implícitos de la representación que se investiga, por lo que son útiles en la exposición de inconsistencias u otros problemas que surgen al intentar aplicar partes de dicha representación a una situación específica. Además, su narrativa constituye una guía para que el receptor construya un modelo de la situación y realice inferencias a partir de ello. Un ejemplo de experimento mental es la paradoja del gato de Schrödinger, en la cual se muestran las contradicciones derivadas de la aceptación de la teoría de la indeterminación de la física cuántica.

Cabe notar que la separación entre las tres formas de razonamiento no es tajante. Nersessian sostiene que las prácticas científicas indican un uso en conjunto y una superposición frecuente entre los procesos analógicos, los procesos imaginistas y los experimentos mentales. Por ejemplo, la analogía newtoniana entre las órbitas planetarias y lunares y el movimiento de un proyectil lanzado con velocidad sucesivamente mayor desde una montaña con gran altura en la superficie terrestre también funciona como un experimento mental: la analogía utiliza un bosquejo relacionado con un razonamiento simulativo donde, por medio de la imaginación, se representan los pasos sucesivos que realizaría el proyectil de la montaña a la Tierra, y luego la velocidad de escape con la que orbitaría la Tierra por efecto de la fuerza centrípeta (Nersessian, 2008, p. 134).

El *bootstrapping* permite el CC porque posibilita la emergencia de soluciones que son combinaciones novedosas de constreñimientos. Estas soluciones superan los recursos de representación de los dominios fuentes de los constreñimientos, por lo que no pueden representarse en términos de las estructuras conceptuales anteriores. De este modo, el CC surge “de procesos iterativos de construcción, manipulación, evaluación y revisión de modelos analógicos para satisfacer constreñimientos” (2008, p. 128). Por ejemplo, las analogías permiten formar conceptos novedosos mediante la abstracción; las

representaciones imaginistas lo posibilitan por procesos tales como la creación de una visualización externa a partir de una imaginación interna (como en el caso de los diagramas), y los experimentos mentales promueven CC en una comunidad a través de su narrativa, puesto que instan al razonamiento acerca de los problemas que derivan de las representaciones. Con respecto a su importancia para la cognición humana común, se trata de una herramienta evolutivamente valiosa para la adaptación al entorno, en cuanto simula resultados potenciales de acciones en este y permite resolver problemas significativos para la supervivencia.

En suma, las aproximaciones de Carey y Nersessian son complementarias en el tratamiento del *bootstrapping*, a pesar de que el énfasis de la primera se encuentre en el desarrollo cognitivo y la segunda se concentre en la innovación conceptual científica. Partiendo desde la base de su descripción como un mecanismo que permite crear representaciones novedosas tanto para la cognición científica como para la cognición humana común, Carey describe el proceso general y Nersessian se enfoca en su núcleo: los razonamientos basados en modelos. Un punto interesante que surge de la conjunción entre ambas se encuentra en la relación entre *bootstrapping* y razonamiento basado en modelos. Esta relación parece involucrar una circularidad: la segunda etapa del *bootstrapping* corresponde a la realización de razonamientos basados en modelos, y los razonamientos basados en modelos involucran a su vez un proceso de *bootstrapping*. En ambos casos, se trata de la construcción y manipulación de modelos que permiten interpretar el fenómeno *target*. Esta circularidad podría explicarse si se considera que el término *bootstrapping* es una metáfora para concebir la capacidad de creación de nuevos recursos representacionales que trascienden las representaciones que uno posee. El término significa *levantarse a sí mismo tirando de los cordones de sus propios zapatos* (Carey, 2009, p. 20). Aparentemente, la metáfora caracteriza tanto el mecanismo general de creación de conceptos nuevos como los razonamientos particulares que permiten dicha creación, lo que explicaría la circularidad. El mecanismo general es entendido como la conformación de conceptos novedosos con los recursos que se adquieren a medida que se desarrollan las etapas del proceso. El razonamiento basado en modelos, por su parte, se explica por la relación entre constreñimientos de los dominios *target* y fuente, que son entendidos como cordones de zapatos cuyos cruces representan los modelos creados (Nersessian, 2008, p. 133).

Luego de caracterizar este mecanismo central, en el siguiente subapartado me concentraré en la etapa final (C2) del proceso de CC, reflexionando acerca de la

posibilidad de caracterizarla en términos de inconmensurabilidad y cambio gestáltico en relación con la etapa inicial (C1).

3.3. Resultado de los cambios conceptuales: ¿inconmensurabilidad y cambio gestáltico?

En los subapartados anteriores, revisé el contexto de los CC y un mecanismo central, el *bootstrapping*, que permite su desarrollo. En este subapartado abordaré una problemática recurrente en el estudio del estado final del proceso: la existencia o inexistencia de inconmensurabilidad y cambio gestáltico entre C1 y C2. Este problema es relevante dado que constituye un aspecto central de análisis y discusión en la aproximación clásica del CC por la influencia del historicismo kuhniano en esta. A continuación, desarrollaré la postura de Kuhn y su adopción o problematización por parte de los exponentes de dicha aproximación.

Para Kuhn, se producen al menos dos diferencias centrales entre un paradigma nuevo y el anterior. En primer lugar, surgen diferencias conceptuales y metodológicas: se establecen relaciones nuevas entre conceptos e instrumentos heredados del paradigma viejo, se adoptan otros conceptos e instrumentos nuevos, y cambian los estándares de la ciencia (es decir, los problemas a resolver y la metodología apropiada). En segundo lugar, y de manera más radical, ocurre un cambio en la visión de mundo, un cambio gestáltico. En este punto, Kuhn exhibe influencias de la psicología de la Gestalt, una corriente psicológica que estudia principalmente la organización perceptual humana²⁸. La aplicación de los principios de la Gestalt al estudio de las ciencias lleva al autor a sostener una relación entre paradigmas y percepción. Los paradigmas incluirían un rol perceptual por lo que, con la revolución, ocurre un cambio brusco de percepción: dado que el acceso al mundo “es a través de lo que ven y hacen, podríamos decir que tras una revolución los científicos responden a un mundo diferente” (1970, p. 111). Además, para la percepción

²⁸ El principio fundamental de la psicología de la Gestalt sostiene que la totalidad estructural de la percepción es mayor que la suma de los estímulos parciales que la conforman. Ante los mismos estímulos se obtiene una percepción distinta en un contexto diferente porque los objetos se perciben como un todo organizado y no como elementos aislados. En esta organización opera la ley de la pregnancia: la experiencia perceptiva organiza los elementos de la mejor forma posible en consideración con la totalidad de los factores presentes (Saal, 1975, p. 274). La mejor forma posible suele ser aquella que tiene características como mayor simplicidad y simetría, y los factores considerados comprenden la profundidad, volumen, distancia, entre otros. Las ilusiones ópticas como el pato-conejo o el cubo de Necker son ejemplos de la operación de la organización y su cambio.

misma se requeriría un paradigma: “lo que un hombre ve depende de lo que mira y también de lo que su experiencia visual-conceptual previa le ha enseñado a ver” (1970, p. 113)²⁹.

Las diferencias entre el paradigma viejo y su sucesor los convierte en inconmensurables entre sí. La inconmensurabilidad en *The Structure* es entendida como las disparidades conceptuales, observacionales y metodológicas que imposibilitan la comparación a través de un parámetro neutro. En el caso de la revolución, esta transición entre inconmensurables acarrea un cambio radical, es decir, un cambio que no se desarrolla gradualmente, sino que “como un cambio gestáltico, debe ocurrir todo de una vez (aunque no necesariamente en un instante) o no ocurrir en absoluto” (1970, p. 150). Este concepto inicial de inconmensurabilidad fue objeto de diversas críticas, principalmente porque parecía implicar que la comunicación entre científicos de paradigmas distintos era imposible, y que los conceptos científicos pasados no podían ser explicados de manera comprensible en términos actuales (lo que, efectivamente, realiza Kuhn en su libro). Por ende, desde la escritura del Epílogo de *The Structure*, el autor realizó una revisión de esta noción.

La inconmensurabilidad es reformulada en términos de pertenencia a comunidades lingüísticas distintas, donde los problemas de comunicación pueden ser vistos como problemas de traducción. Los miembros de una comunidad científica comparten un lenguaje común que los diferencian de otros grupos. Junto con este lenguaje, adquieren un conjunto de compromisos cognitivos que son “una consecuencia de los modos en que los términos, frases y oraciones del lenguaje son aplicados a la naturaleza” (1977, p. xxii). Entender el punto de vista de un hablante de otro lenguaje se encuentra limitado por las imperfecciones del proceso de traducción y la determinación de referencia (1977, p. xxiii). Dos personas que emplean el mismo vocabulario, pero que mantienen dos modos de percibir distintos presentan dificultades en la comunicación. En ambos casos, las palabras que usan fueron aprendidas, en parte, mediante ejemplares. Dado que comparten los mismos estímulos, una programación neural similar y, por tanto, gran parte de su mundo científico y cotidiano, se reconocen como pertenecientes a comunidades lingüísticas distintas.

²⁹ El componente perceptual de los paradigmas se aprecia también en la visión de los futuros científicos durante su aprendizaje, por ejemplo, “al observar una fotografía de una cámara de burbujas, el estudiante ve líneas confusas y discontinuas, el físico un registro de eventos subnucleares familiares” (1970, p. 111).

Las locuciones no problemáticas se traducen de manera homófona, es decir, conservando el sonido de las palabras sin agregar elementos idiosincráticos. Los términos que son foco de problema se aíslan, luego de lo cual se puede recurrir a los vocabularios compartidos para descubrir lo que el otro ve frente a un estímulo. La traducción permite experimentar indirectamente puntos de vista ajenos, así como establecer comparaciones, por lo que es una herramienta que eventualmente permite la persuasión y la conversión. Por ejemplo, en la comunicación de investigaciones dentro de la comunidad científica, algunos resultados son descriptibles en oraciones que son entendidas por dos grupos distintos. Si el nuevo punto de vista demuestra ser fructífero, aumentarán los resultados verbalizables en este modo y su poder de persuasión para otros científicos. Sin embargo, traducir una teoría al lenguaje propio no es lo mismo que hacerla propia. Para lograr lo segundo, el sujeto debe pensar y trabajar en el otro lenguaje. Quien emplee una teoría sin internalizarla de esta forma carece del modo de ver de los nativos, y solo puede usarla como extranjero. Nuevamente se muestra aquí la dependencia que posee la visión de mundo con respecto a un paradigma.

En un artículo posterior titulado *Commensurability, Comparability, Communicability* (1982), Kuhn formula una noción modesta, denominada *inconmensurabilidad local*, que ahonda en la reformulación lingüística presente en el Epílogo de *The Structure*. Esta inconmensurabilidad local es una relación entre los lenguajes en los que se encuentran formuladas las teorías y los conceptos que las articulan. De acuerdo con ella, “no existe un lenguaje, neutro o no, al que ambas teorías, concebidas como conjuntos de oraciones, puedan traducirse sin residuos ni pérdidas” (1982, p. 670). En otras palabras, los elementos (explicaciones, conceptos, leyes) expresados en la terminología de una teoría no pueden ser expresados en su sucesora. Cabe notar, en primer lugar, que los problemas de traducción surgen en un grupo reducido de términos, usualmente interdefinidos, y las oraciones que lo contienen. El resto de los términos en común entre ambas teorías mantienen su significado, funcionan de la misma forma y se traducen de manera homófona. Ello provee una base de comparación y discusión, e incluso de exploración del significado de los términos que son inconmensurables. En segundo lugar, los conceptos y oraciones que no poseen traducción pueden estar sujetos a la interpretación. Este proceso es diferente a la traducción. El traductor conoce ambos lenguajes (L1 y L2) y produce un texto equivalente por medio del reemplazo de los términos de L1 por los de L2 que compartan el mismo significado. El intérprete puede inicialmente dominar solo L2, y descubrir el uso de los términos de L1 por la observación

del comportamiento lingüístico y las circunstancias que rodean la producción del texto. El traductor no cambia el significado de los términos y oraciones —cuya existencia es anterior a la traducción—, y la traducción no incluye las glosas o prefacios del traductor. En cambio, el intérprete aprende y enseña un nuevo lenguaje.

La inconmensurabilidad surge dado que la fijación de las referencias (la relación entre un símbolo y las entidades que simboliza) es un proceso complejo. El mero reemplazo de un concepto por una expresión que comparta su referencia no garantiza una traducción. Las traducciones que solo preservan las referencias desembocan en textos incoherentes. En las traducciones entre dos textos inconmensurables, se reemplazaría un término de L1 con un término de L2 en algunos contextos, y con otro término de L2 en otros contextos. Es posible que no exista un término o descripción de término en L2 que comparta referencia con uno en L1, ya que este forma parte de un grupo de términos interrelacionados, algunos de los cuales deben ser aprendidos juntos, y que aluden a la conceptualización de un dominio de fenómenos que es incompatible con aquella sostenida por L2. Ello sucede porque las conexiones entre entidades en el mundo y los significados de las palabras se encuentran mediadas por diversos criterios que varían entre individuos³⁰, algunos de los cuales pueden ser teóricos. Las comunidades de hablantes aprenden un grupo de términos en conjunto que describen fenómenos del mundo de manera distinta a otros conjuntos de otras lenguas. La dependencia que mantienen los significados de estos términos con la teoría de la que forman parte es congruente con la afirmación de Kuhn según la cual “diferentes lenguajes imponen diferentes estructuras del mundo” (1982, p. 682). A su vez, este punto sigue la línea de las diferencias constatables bajo distintos paradigmas. Por tanto, no se debe preservar solo la referencia, sino también el sentido o la intensión. Cuando ello no ocurre, aparece la inconmensurabilidad.

Dado que la inconmensurabilidad surge en un conjunto de términos y las oraciones en las que figuran, el resto de los términos en común garantiza que entre teorías inconmensurables exista comunicación y comparabilidad. La comunicación entre hablantes no requiere compartir los mismos criterios de identificación de los referentes, pero sí necesita mantener una homología de la estructura léxica, es decir, una semejanza en las interrelaciones de palabras que conforman un lenguaje (de las cuales dependen sus

³⁰ Según Kuhn, la coincidencia usual en la relación término-referencia, que ocurre a pesar de la diferencia en los criterios empleados por los hablantes, se debe a que el mundo al cual se adecúa dicho lenguaje favorece tales identificaciones.

significados). La similitud entre estructuras taxonómicas es central para la comunicación porque esta estructura refleja los aspectos del mundo que describe, y sus diferencias desembocan en diferencias de estructuras de mundo.

Las nociones kuhnianas de inconmensurabilidad y cambio gestáltico han sido adoptadas o problematizadas por los autores de la aproximación clásica del CC revisados anteriormente. Mientras que existen opiniones divididas sobre la adecuación de la noción de inconmensurabilidad, se aprecia un consenso en las críticas al cambio gestáltico como forma de describir el proceso entre C1 y C2. A continuación, indagaré en ambos puntos.

Posner y colaboradores sostienen que la concepción actual de los estudiantes y aquella que adquieren por medio de la enseñanza de las ciencias son incompatibles. Sin embargo, no ahondan en esta incompatibilidad y, en particular, no explicitan si tienen en vista una inconmensurabilidad entre C1 y C2. Por otro lado, toman postura con respecto al problema de caracterizar los CC como cambios gestálticos. A pesar de que sostienen que una de las similitudes entre las concepciones previas y los paradigmas reside en su rol como categorías perceptuales (véase p. 22), el cambio paradigmático no se debe entender como un cambio gestáltico, porque no ocurre de manera abrupta:

La acomodación, sobre todo para los novatos, se concibe mejor como un ajuste gradual en las concepciones propias, en el que cada nuevo ajuste sienta las bases para otros ajustes, pero donde el resultado final es una reorganización o un cambio sustancial en los conceptos centrales propios (1982, p. 223).

Los trabajos de Carey, en cambio, han recurrido sistemáticamente a la noción de inconmensurabilidad como una característica central del CC. La autora identifica dos tipos de discontinuidades que ocurren en el desarrollo conceptual entre las estructuras conceptuales infantiles (C1) y las adultas (C2): un mayor poder expresivo con respecto a la estructura conceptual anterior, producto de los procesos de modelación implicados en el *bootstrapping quineano*, y una inconmensurabilidad entre sistemas sucesivos. Este segundo fenómeno es definido como “la relación entre sistemas conceptuales de tal manera que uno contiene conceptos que no están simplemente ausentes en el otro, sino que en realidad son incoherentes desde el punto de vista del otro” (2009, p. 359). Las creencias y conceptos de un sistema no pueden ser formulados en términos de las creencias y conceptos del otro. Para Carey:

Cuando nos preguntamos si el lenguaje de los niños (L1) y el sistema conceptual que expresa (C1) a veces pueden ser inconmensurables con el lenguaje (L2) y con el sistema conceptual (C2) de los adultos, donde C1 y C2 abarcan el mismo dominio de la naturaleza, estamos preguntando si existe un conjunto de conceptos en el núcleo

de C1 que no puede ser expresado en términos de C2, y viceversa. Nos preguntamos si L1 puede ser traducido a L2 sin glosa de traductor. (...) Los conceptos indiferenciados de C1 ya no pueden desempeñar ningún rol en C2, y los conceptos coalescentes de C2 no pueden jugar un papel en C1 (1999a, p. 465).

Dicho de otro modo, la inconmensurabilidad es una relación entre dos sistemas conceptuales y sus correspondientes lenguajes que ocurre cuando existe una incompatibilidad en el potencial referencial: los métodos de fijación de referencia para un término difieren, a pesar de que se superpongan en cuanto a contenido. Por ejemplo, un término asociado a una única entidad en una teoría se asocia a dos entidades en la teoría sucesiva, como ocurre con el concepto indiferenciado de PESO y DENSIDAD que poseen los infantes en términos de GRADO DE PESADEZ. La autora se refiere a la inconmensurabilidad local kuhniana, que surge solo en un conjunto de conceptos que articulan las teorías, usualmente interdefinidos y aprendidos en conjunto por la comunidad de hablantes (2009, p. 365).

Los niños adquieren el lenguaje por medio de procesos sociales y aprendizaje explícito de los adultos. No obstante, deben asignar los términos aprendidos a los conceptos propios de su sistema conceptual, construido a partir de los recursos de la cognición núcleo y las teorías intuitivas (véase pp. 24-26). Así, “el lenguaje que construyen está condicionado tanto por el idioma que están escuchando como por la conceptualización del mundo que ya han construido” (1999a, p. 467). Volviendo al ejemplo del concepto indiferenciado de PESO/DENSIDAD, la teoría intuitiva de los objetos físicos propia del niño no posee los recursos conceptuales suficientes para formular las creencias, conceptos y explicaciones de la teoría intuitiva de las entidades materiales que domina el adulto. La indiferenciación de PESO/DENSIDAD es incoherente desde el sistema adulto por lo que, para entender el concepto GRADO DE PESADEZ, se debe recurrir a su relación con otros conceptos en el sistema del niño, como DENSIDAD y SUSTANCIA, que se interrelacionan de un modo distinto. En el CC, se trazan nuevas distinciones ontológicas sobre la base de las cuales los conceptos son subsumidos en categorías ontológicas nuevas, reasignados a nuevas ramas de una jerarquía conceptual, fusionados o diferenciados. El proceso corresponde a una transformación gradual que requiere un periodo de años —incluso siglos en el caso de la historia de la ciencia— para llevarse a cabo. Por tanto, no puede corresponder a un cambio gestáltico.

Siguiendo a Kuhn, la autora señala que la inconmensurabilidad no anula la posibilidad de comunicación e interpretación (2009, p. 378). La primera se realiza sobre

la base de los términos compartidos que mantienen su significado, lo que ayuda a fijar los referentes de algunos términos no compartidos. La segunda consiste en un proceso de *bootstrapping* en reversa. El psicólogo o el historiador de la ciencia crea un sistema conceptual que es incompatible con el propio mediante el aprendizaje de un conjunto de términos que lo articulan y que son, al inicio, parcialmente interpretados en términos de los conceptos que posee. Analiza los usos interrelacionados de los términos y cómo se vinculan a los fenómenos, y utiliza procesos de modelación para construir una teoría al respecto (2009, p. 370). Por ejemplo, las glosas de traductor forman parte del proceso en el que los conceptos de L1 son enseñados a los hablantes de L2.

Por su parte, Thagard sostiene cuatro tipos de relación posible entre C1 y C2 que problematizan la inconmensurabilidad y el cambio gestáltico kuhniano. En primer lugar, puede ocurrir una incorporación, donde C2 absorbe completamente a C1. Una segunda opción es la sublación, en la cual C2 incorpora parcialmente C1 mientras rechaza algunos de sus aspectos. En otras palabras, se mantienen las hipótesis explicativas exitosas de C1, pero los conceptos sufren eliminaciones o reorganizaciones. La suplantación, en cambio, implica que C2 rechaza de manera casi total a C1: las hipótesis explicativas y sus conceptos. Por último, puede ocurrir una indiferencia, donde la adopción de C2 no toma en consideración a C1, por lo que no existe ninguna relación entre ellas. Las diferencias entre estas relaciones se encuentran en el nivel de acumulabilidad conceptual, metodológica o de evidencias observado en las revoluciones, aunque la adopción de una teoría nueva siempre involucra al menos el rechazo de algunas afirmaciones teóricas de la anterior. Mientras que la mayor parte de los desarrollos no revolucionarios en la ciencia son vistos como incorporaciones, las revoluciones tienden a presentar suplantaciones y sublaciones. Bajo esta perspectiva, la visión kuhniana de la ciencia es limitada, en cuanto sostiene que una nueva teoría suplantada o ignora a la otra, sin considerar casos como la incorporación y la sublación.

En el proceso de la revolución se va conformando un sistema conceptual nuevo que reemplaza el anterior, pero donde se mantiene una continuidad. Ambos permanecen conectados a un marco de conceptos relativamente estable que no cambia. Esta idea parece similar a la ecología conceptual planteada por Posner y colaboradores, porque se incluyen conceptos que se mantienen y guían el cambio. También es parecida a los términos con el mismo significado que son compartidos entre los sistemas conceptuales de niños y adultos mencionados por Carey.

La continuidad permite sostener que la adopción de nuevas teorías es racional en general, lo que aplica para muchos casos históricos (1992, p. 112). Con ello, Thagard pretende superar la irracionalidad del cambio científico que puede desprenderse de la descripción kuhniana de la revolución como un episodio de cambio brusco que imposibilita la evaluación entre posturas por estándares externos³¹, lo que corresponde a la noción de inconmensurabilidad desarrollada en *The Structure*. También critica su versión local y, en particular, la comparación que realiza entre la adquisición de una nueva teoría y un segundo lenguaje. Para Thagard, el aprendizaje de un lenguaje consiste en la adquisición de representaciones mentales de la gramática y el léxico, y su traducción depende de la capacidad de uso de estas representaciones para la transformación de términos y oraciones de un lenguaje en términos y oraciones de otro (1992, p. 114). El sistema conceptual científico no requiere aprender un conjunto de relaciones sintácticas tan amplias y variadas como la gramática y el vocabulario, por lo que es gramaticalmente más sencillo de adquirir que un lenguaje. No obstante, léxicamente es a la vez más fácil y más difícil: más simple por la superposición léxica entre marcos conceptuales distintos, y más complejo por las reorganizaciones jerárquicas, redefiniciones jerárquicas y diferencias ontológicas, que hacen que el aprendizaje del lenguaje natural requiera menos cambios radicales. Por tanto, la comparación entre una nueva teoría y un segundo lenguaje es imprecisa. Si bien la inconmensurabilidad local de Kuhn se encuentra en lo correcto al afirmar que no existen traducciones sin residuos entre marcos conceptuales distintos, Thagard sostiene que una traducción completa no es necesaria para establecer juicios acerca de la coherencia explicativa. Basta con conocer la teoría nueva, sus explicaciones y qué cuenta como datos para ella. Aunque dos sujetos estén en desacuerdo sobre este último punto, solo se requiere una apreciación de los experimentos del otro, dado que sus resultados, si bien no están desligados de consideraciones teóricas, son guiados por el mundo. En conclusión,

adquirir un nuevo sistema conceptual es algo así como adquirir un nuevo idioma, pero las personas lo hacen todo el tiempo, y la adquisición y traducción del lenguaje son solo impedimentos temporales para la aplicación racional de consideraciones de coherencia explicativa a la evaluación de la teoría (Thagard, 1992, p. 117).

³¹ Sin embargo, Kuhn ha respondido a las críticas que le acusan de proveer una imagen irracional de la ciencia. Por ejemplo, en *Reflections on my Critics*, el autor señala que el hecho de que la lógica y la observación no cumplan un rol decisivo en la elección de teorías no implica descartar ambas o sugerir que no existen buenas razones que apoyen la elección de una teoría por sobre la otra (2000b, p. 126).

Ello también podría aplicarse a la inconmensurabilidad que estudia Carey en el caso del desarrollo cognitivo infantil: si bien los niños deben atravesar un proceso de asignación de términos del lenguaje aprendido a los conceptos de su sistema conceptual, se logra normalmente sin mayores inconvenientes.

Esta crítica parece no capturar el núcleo de la noción de inconmensurabilidad. La factibilidad de adquirir un nuevo sistema conceptual, en el caso del niño y del científico, no implica que el proceso sea sencillo. En el caso del desarrollo cognitivo, Carey (2009) muestra la dificultad del cambio entre la teoría física infantil, que no diferencia entre PESO y DENSIDAD, y la teoría adulta. A pesar de ser parte de la enseñanza en los colegios, un gran porcentaje de estudiantes secundarios fallan en realizar el CC correspondiente. Un escenario similar sucede con el aprendizaje de fracciones y decimales, donde la mitad de los estudiantes estadounidenses de secundaria que toman el examen de admisión universitaria SAT aún no posee una comprensión de ellos (2009, p. 470). Por otro lado, es cierto que no se requiere una traducción completa para establecer un análisis entre teorías sucesivas, pero ello no es contrario a la inconmensurabilidad local kuhniana: los términos que mantienen su significado en ambas teorías permiten la comparación, discusión y exploración de los términos inconmensurables. No obstante, es posible que el nivel de continuidad entre C1 y C2 sea aún mayor que los términos con mismo significado a los que se refiere Kuhn, por lo que se justificaría la problematización general de la noción de inconmensurabilidad.

Por último, Nersessian sostiene que el problema de la inconmensurabilidad y de la racionalidad del cambio son “artefactos del planteamiento del cambio, de una manera que solo compara los puntos finales de un proceso largo, y no toma en consideración la fina estructura intermedia” (2008, p. x). De manera similar a la crítica de Thagard, la autora sostiene que el análisis propuesto por Kuhn es históricamente inadecuado porque el empleo de la metáfora del cambio gestáltico enfatiza la etapa final del proceso de CC, reflejando un cambio abrupto y discontinuo donde el resto del proceso queda inanalizable. Así, “la metáfora ha obstruido el desarrollo de la epistemología historicista que es defendida” (1992, p. 7). Es implausible que los cambios en el conocimiento se caractericen por una serie de cambios gestálticos inconexos. El CC es producido por prácticas de solución de problemas que son creativas, y que permiten articular, comunicar y reemplazar representaciones de un dominio.

El estudio del razonamiento basado en modelos como herramienta de innovación conceptual (véase pp. 33-38) permite a Nersessian reformular dos planteamientos de la

propuesta de Kuhn: el problema de la acumulatividad en el cambio científico y el cambio de percepción asociado con un cambio de paradigma. En el primer caso, la autora sostiene que los cambios en la ciencia poseen a la vez un carácter continuo y no acumulativo: las estructuras conceptuales nuevas se desarrollan a partir de las estructuras anteriores, pero no son una extensión de estas. El razonamiento basado en modelos posibilita la creación de soluciones que no pueden representarse en los mismos términos de las estructuras conceptuales anteriores, dado que superan los recursos representacionales de los dominios fuentes. Sin embargo, son desarrolladas a partir de constreñimientos de estos dominios. Allí, la metáfora del cambio gestáltico puede describir solo la última parte del proceso de CC, pero no debe ser empleada para caracterizar el proceso completo. El segundo punto, no obstante, parece contradecir su crítica a la inconmensurabilidad y cambio gestáltico. Para la autora, el entendimiento y percepción de mundo podría ser inconmensurable con la etapa prerrevolucionaria en cuanto “si negociamos el mundo por medio de la construcción de modelos mentales, los científicos prerrevolucionarios y los posrevolucionarios construirían modelos mentales diferentes y, por tanto, tendrían realmente experiencias diferentes del mundo” (1992, p. 36). En otras palabras, los procesos creativos de solución de problemas que son empleados por los sujetos epistémicos son diferentes, con lo cual también difieren sus percepciones sobre el mundo. En obras posteriores (2009), Nersessian parece abandonar este punto.

En resumen, la crítica principal planteada desde la aproximación clásica del CC a la comparativa kuhniana entre C1 y C2 se concentra mayormente en la noción de cambio gestáltico. Según esta crítica, la analogía oscurece los microprocesos que conforman el proceso de discontinuidad conceptual porque aparenta que existe un salto abrupto entre la etapa inicial y la etapa final³². Con respecto a la inconmensurabilidad, Carey sostiene su versión local para el desarrollo cognitivo y la identifica como un rasgo propio de los CC —a diferencia del enriquecimiento de conocimiento, como mencionaré en el próximo apartado—; mientras que Thagard y Nersessian problematizan la noción a partir de la constatación de continuidades entre C1 y C2. En el tercer capítulo examinaré su adecuación a los CC de REPRESENTACIÓN en ciencia cognitiva.

³² Kuhn mismo admite esta crítica, sosteniendo que la transferencia del término *cambio gestáltico* desde los individuos hasta las comunidades es incorrecta, dado que reduce a un instante un proceso que es extendido e impide la posibilidad de su estudio. Sin embargo, afirma que su trabajo abre sendas para la exploración de este (Kuhn, 2000a, pp. 88-89).

En el tercer apartado me he concentrado en el contexto, mecanismos y resultado de los CC. El contexto fue descrito a partir de la noción de ecología conceptual, que incluye elementos clave como los compromisos epistemológicos, los problemas que suscitan debates y las analogías. A propósito de este último punto, me enfoqué en el *bootstrapping* como mecanismo que permite el desarrollo de representaciones novedosas. Por último, en cuando al resultado del CC, analicé las nociones de inconmensurabilidad y cambio gestáltico. Luego de describir los detalles de este fenómeno que serán relevantes para el análisis de las transformaciones conceptuales en REPRESENTACIÓN, en el próximo apartado me resta identificar algunos tipos de CC que se han estudiado en la literatura de la aproximación clásica.

4. Tipos de cambio conceptual

En la aproximación clásica del CC, es posible encontrar distintos tipos o grados de discontinuidades conceptuales tanto en historia y filosofía de las ciencias como en el desarrollo cognitivo humano. Previo a la clasificación de estos, los autores usualmente realizan una diferencia entre CC y revisiones de creencia o enriquecimiento de conocimiento. Mientras que las primeras corresponden a discontinuidades conceptuales, las segundas son meras adiciones o eliminaciones relativamente triviales en el concepto que no implican una transformación significativa de este ni de sus relaciones con otros conceptos. Esta diferencia se observa tanto en enseñanza de las ciencias como en psicología del desarrollo e historia y filosofía de las ciencias. Por ejemplo, Posner y colaboradores afirman la disimilitud entre una revisión de creencias y una transformación radical de conceptos. Mientras que el CC afecta conceptos centrales en el aprendizaje y el pensamiento, el segundo no:

Un aprendiz que es capaz de sustituir una visión newtoniana o aristotélica del movimiento por la de Einstein ha experimentado el tipo de cambio conceptual al que nos referimos. Alguien que aprende que está lloviendo afuera o que aprende que el aire caliente sube, no lo ha hecho. Esta distinción se mantiene incluso si estos últimos casos de aprendizaje involucran el reemplazo de una creencia o concepto por otro (1992, p. 148).

En otras palabras, en el aprendizaje existen casos en los cuales se produce un cambio entre conceptos sin corresponder a un CC, por lo que cuenta como revisión de creencias. Una distinción similar radica en el planteamiento de Thagard, para quien tanto el

aprendizaje humano como el descubrimiento científico no pueden explicarse como una simple revisión de creencia. El autor afirma que:

Contaré los cambios conceptuales como revolucionarios si ellos involucran el reemplazo de un sistema completo de conceptos y reglas por un nuevo sistema (...). Simplemente agregar un nuevo conjunto de ideas no plantea problemas especiales, y la sustitución de un solo concepto o regla debería ser un proceso simple (1992, p. 6).

Por su parte, Carey realiza dos distinciones al respecto: entre CC y enriquecimiento de conceptos (Carey, 2009) y entre reestructuraciones fuertes y reestructuraciones débiles de conocimiento (Carey, 1985). La primera distinción se entiende desde la presencia o ausencia de inconmensurabilidad. En los CC existe inconmensurabilidad: se involucra la creación de conceptos nuevos no expresables en términos de los conceptos anteriores. Esta inconmensurabilidad muestra que el cambio no puede ser entendido meramente en términos de un enriquecimiento de las estructuras iniciales. La transformación abarca compromisos ontológicos y explicaciones de un dominio determinado: el fenómeno en cuestión, sus mecanismos explicativos y los conceptos que articulan sus representaciones. En cuanto el sistema conceptual se conforma de conceptos interrelacionados de manera específica, existe una coherencia dentro del sistema (es decir, la comprensión de un aspecto constriñe la comprensión de otros) que es modificada con la reestructuración. Algunos ejemplos corresponden a los CC en conceptos de dominio núcleo como la mecánica y psicología intuitiva. En el enriquecimiento de conocimientos, en cambio, no existe inconmensurabilidad, puesto que no existe discordancia en el potencial referencial o en los roles conceptuales que determinan su significado. Se adquieren nuevas creencias que se mantienen representadas en términos de los conceptos anteriores, o se añaden nuevos conceptos que no acarrearán una revisión de los antecedentes. No exhibe coherencia necesariamente, dado que los nuevos hechos se añaden de manera independiente. Por ejemplo, la diferencia de PERRO en conceptos subordinados tales como POODLE o COCKER SPANIEL no implica CC, porque el concepto mantiene la misma extensión y rol conceptual.

En relación con la segunda distinción, el CC se entiende como una reestructuración fuerte del conocimiento donde se presentan diferencias en el dominio de fenómenos considerados por la teoría, las estructuras explicativas planteadas y los conceptos individuales entre teorías sucesivas. Esta reestructuración es análoga a las diferencias ocurridas en las nociones centrales en teorías científicas expuestas por Kuhn en *The*

Structure. En ciencias pueden ocurrir dos tipos de cambio de teoría. El primero corresponde al cambio verdadero de teoría (*true theory change*), referido a la transformación por medio de la revolución de un dominio de estudio establecido. Un ejemplo de ello radica en la transición entre la teoría pre-galileana del ímpetu y la mecánica newtoniana. El segundo tipo es el surgimiento de un nuevo dominio desde otro u otros, es decir, la emergencia de teoría (*theory emergence*). La biología evolutiva, lógica, psicología y química del s. XIX se incluyen en este tipo. La emergencia cuenta como cambio de teoría en cuanto “algunos de los fenómenos que motivan el nuevo dominio están ya estudiados en las tradiciones científicas predecesoras” (1985, p. 7), por lo que este nuevo dominio se superpone con teorías anteriores.

La autora sostiene que los cambios presentados en los niños son descritos de mejor manera en términos similares al desarrollo científico³³. En particular, el cambio entre la teoría intuitiva del comportamiento y la biología intuitiva es un caso de emergencia de teoría. Los niños menores de 10 años presentan una teoría intuitiva que es animista³⁴: atribuyen estados intencionales tanto a animales como a objetos inanimados. Además, otorgan a los animales funciones biológicas como comer, respirar y crecer a partir de la similitud con PERSONA. Con la adquisición de la biología intuitiva, se forman conceptos superordinados como SER VIVO (donde se integran plantas y animales) y ANIMAL (donde se incluyen las personas). Por medio del empleo de este conocimiento, entienden que las funciones biológicas forman parte del ciclo de la vida, por lo que se abandona PERSONA como ejemplar que permite la proyección inductiva de estas propiedades.

La relación entre la teoría intuitiva del comportamiento y la biología intuitiva cumple con los tres requisitos que caracterizan las reestructuraciones fuertes del conocimiento. En primer lugar, los dominios de ambas teorías se superponen (los

³³ Ello contrasta con la postura piagetiana, quien describe el desarrollo de los niños en términos de un conjunto fijo de estadios cognitivos, desde el pensamiento pre-operacional a las operaciones concretas y luego a las operaciones formales.

³⁴ El animismo ocurre porque los niños incurren en una confusión semántica o porque no saben suficiente biología. En el primer caso, al juzgar si un objeto está vivo o no, el niño puede interpretar que debe establecer una diferencia entre *vivo* y *muerto* y no entre *vivo* e *inanimado*. Así, en cuanto no entiende la muerte como un cese de las funciones corporales, integra otros contrastes tales como *real* e *imaginario* o *existente* e *inexistente*. En el segundo caso, las diferencias conceptuales entre niños y adultos hacen que las palabras *vivo* y *no vivo* no se mapeen en los conceptos similares a los conceptos biológicos VIVO y NO VIVO de los adultos. Los niños pequeños no representan ningún concepto con la extensión adulta del concepto superordinado de SER VIVO, donde se integran animales y plantas en una misma categoría. Aplican su conocimiento biológico a la pregunta acerca de qué está vivo y qué no lo está, pero no es suficiente para lograr establecer las distinciones que establecen los adultos.

conceptos de PERSONA y ANIMAL son ontológicamente centrales a las dos, y diversos fenómenos son compartidos, como comer o respirar), pero existen cambios en ellos. En segundo lugar, aparecen diferencias en los mecanismos explicativos que se observan en las disparidades de los patrones de atribución de propiedades biológicas. Por ejemplo, los niños pequeños recurren a una motivación (tener hambre) para explicar la conducta de comer como parte de un comportamiento humano, lo que radica en una estructura explicativa psicológica y social. En cambio, la explicación biológica de la conducta de comer por necesidad biológica no es intencional, sino que apela a su rol en el mantenimiento de la vida. Los desarrollos en la comprensión de los procesos biológicos se interconectan: “ninguno es lógicamente requerido para el otro, y todos se apoyan mutuamente” (Carey, 1985, p. 70). Siguiendo a Kuhn, la autora afirma que este rasgo es propio de la emergencia de una teoría. En tercer y último lugar, existe CC en los conceptos individuales, como diferenciaciones entre dos sentidos de NO VIVO (muerto e inanimado) o coalescencias entre ANIMAL y PLANTA. Desarrollaré tales tipos de cambio más adelante (véase p. 53).

Por su parte, la reestructuración débil se encuentra involucrada en el paso de novato a experto, es decir, las transformaciones ocurridas en el proceso de adquisición de experticia por parte de un principiante en un dominio determinado, como las ciencias físicas o las matemáticas. Allí tiene lugar una representación de diferentes relaciones entre conceptos y la creación de nuevos conceptos y esquemas abstractos sobre los patrones entre estas nuevas relaciones. Se mantienen diversos conceptos en común; por ejemplo, FUERZA y ENERGÍA son compartidos por el sistema del novato y del experto en el dominio de la física, y sus términos son idénticos o traducibles fácilmente. Por ende, no se halla CC al nivel de los conceptos individuales. La diferencia entre los dos tipos de reestructuración (fuerte y débil) son sintetizados por la autora de la siguiente manera:

Dos sistemas conceptuales sucesivos son estructuralmente diferentes en el sentido débil si el último representa diferentes relaciones entre los conceptos que el anterior, y si los patrones de estas relaciones motivan conceptos superordinados en el sistema posterior que no son representados en el primero. El cambio de novato a experto en ajedrez provee un ejemplo paradigmático. Dos sistemas conceptuales sucesivos son estructuralmente diferentes en el sentido fuerte si la transición entre ambos involucra cambio conceptual; la presencia frente a la ausencia de cambio conceptual es la diferencia esencial entre los dos tipos de reestructuración. El cambio de la mecánica aristotélica a la galileana provee un ejemplo paradigmático de reestructuración fuerte (1985, p. 186).

En otras palabras, una reestructuración en sentido débil contempla la adición de nuevos conceptos o relaciones entre conceptos, mientras que una reestructuración en sentido fuerte involucra un CC: cambios en el dominio de fenómenos, estructuras explicativas y conceptos individuales.

Luego de la diferenciación general entre CC y revisiones de creencia, el número de tipos específicos de transformaciones que se pueden encontrar en el desarrollo cognitivo o científico varía considerablemente. Los trabajos de Carey afirman la existencia de, al menos, tres tipos de reestructuración fuerte del conocimiento (que pueden encontrarse de manera conjunta), a saber:

1. Diferenciación: un concepto inicial que permanece indiferenciado en C1 es dividido en dos o más conceptos nuevos en C2, donde el concepto indiferenciado inicial no juega ningún rol y no tiene sentido. En el desarrollo cognitivo se presenta, por ejemplo, el concepto de NO VIVO en el sentido de *muerto* y en el sentido de *inanimado*, o la diferenciación entre PESO y DENSIDAD como dos tipos de magnitudes físicas diferentes. En ciencias, un ejemplo se encuentra en la diferenciación entre CALOR y TEMPERATURA llevada a cabo por Black en la historia de la química.
2. Coalescencia: entidades consideradas como ontológicamente distintas en C1 y que jugaban roles conceptuales incompatibles son subsumidas en un concepto en C2, como ocurre con ANIMAL y PLANTA en el concepto superordinado de SER VIVO durante el desarrollo cognitivo, o la coalescencia, por parte de Galileo, de los términos aristotélicos de MOVIMIENTO NATURAL y MOVIMIENTO VIOLENTO en ciencias.
3. Reanálisis de la estructura básica de un concepto: es posible que ocurran cambios en el núcleo y en el tipo. Propiedades tomadas como fundamentales se convierten en relativas a otras, como ocurre en la transición entre considerar esenciales rasgos de un bebé tales como ser pequeño y desvalido, a ser vistos como derivados. Una propiedad puede, además, ser reconceptualizada en términos de relaciones, como sucede en el reanálisis newtoniano de PESO desde su concepción como una propiedad del objeto a su consideración como una relación entre objetos.

Thagard (1992), por su parte, distingue entre nueve grados de CC, ordenados en nivel de radicalidad:

1. Adición o supresión de una instancia: corresponde a un cambio relativamente trivial, como la adición de YORKSHIRE TERRIER en el concepto de PERRO.

2. Adición o supresión de una regla débil: adición de una regla (es decir, una relación general entre conceptos) con menor importancia pragmática para la solución de problemas, como ocurre al aprender que Kiribati se encuentra en Oceanía.
3. Adición o supresión de una regla robusta: adición de una regla con mayor importancia pragmática para la solución de problemas. Por ejemplo, aprender que el chocolate es nocivo para los gatos es potencialmente más valioso en términos pragmáticos que aprender que Kiribati está en Oceanía.
4. Adición o supresión de una relación de partes (descomposición): Un ejemplo de este tipo de cambio se encuentra en el concepto de ÁTOMO tras el descubrimiento de que contiene electrones, realizado por Thomson.
5. Adición o supresión de una relación de tipos (coalescencia o diferenciación): Thagard se remite a la coalescencia y la diferenciación como tipos introducidos inicialmente por Carey (1985), por lo que mantiene la caracterización de la autora.
6. Adición o supresión de un concepto: Un ejemplo de este tipo de cambio radica en la adición de ELECTROMAGNETISMO, producido por una coalescencia entre electricidad y magnetismo.
7. Colapso de parte de una jerarquía de tipo, abandonando la distinción previa: corresponde al abandono o la atenuación de distinciones antes realizadas. Por ejemplo, la atenuación en la distinción entre especies y variedades por parte de Darwin.
8. Reorganización jerárquica por un desplazamiento de concepto de una rama taxonómica a otra (*branch jumping*): involucra el rechazo de relaciones de tipo o de parte anteriores, así como la adición de nuevas relaciones. Ello la distingue de reorganizaciones simples tales como la diferenciación, coalescencia y descomposición, donde solo aparece una extensión de relaciones ya existentes. Un ejemplo corresponde a la reclasificación de TIERRA como PLANETA tras Copérnico.
9. Redefinición jerárquica (*tree switching*): es el cambio en el principio organizador de un árbol jerárquico. Corresponde al CC más radical, dado que no solo existe una reclasificación, sino también una revisión en la manera de clasificar. Darwin cambió el significado de TIPO, que pasó de basarse solamente en la similitud a constituir primariamente *ser de un ancestro común*. Thagard (2014) sostiene que un mejor término para este tipo de cambio corresponde a *metaclasificación*.

Los tres primeros tipos de CC podrían ser interpretados en términos de revisión de creencias. Ello no ocurre con el resto, dado que involucran jerarquías conceptuales formadas por relaciones de tipo y de parte. Estos cambios son más dramáticos —y menos frecuentes— que la mera adición o eliminación de nodos o enlaces, en cuanto dichas jerarquías proveen un andamiaje (*scaffold*) que organiza otros conceptos. En general, los nueve grados pueden agruparse en los siguientes grupos: adición, supresión, reorganización simple (diferenciación, coalescencia y descomposición), reorganización revisionista (reorganización jerárquica) y redefinición jerárquica o metaclasificación.

Por último, Nersessian identifica al menos cuatro tipos de cambio que pueden ser observados en las revoluciones científicas, a saber:

1. Creación de nuevos conceptos (por ejemplo, SPIN en mecánica cuántica).
2. Eliminación de conceptos anteriores (por ejemplo, FLOGISTO en química)
3. Descendencia conceptual de conceptos sobre la base de otros existentes (por ejemplo, MASA y CAMPO en relatividad).
4. Absorción de conceptos anteriores por otros (por ejemplo, la absorción de ÉTER por CAMPO o ESPACIO-TIEMPO).

A pesar de las discrepancias en el número de tipos o grados específicos de CC observados, las clasificaciones no son incompatibles entre sí, sino que responden a distintos grados de especificidad. Es patente que la propuesta de clasificación presentada por Thagard (1992) es más detallada, y los tipos identificados por Carey y Nersessian se incluyen en su taxonomía. La diferenciación y la coalescencia de Carey corresponderían a tipos de reorganización simple, mientras que el reanálisis de la estructura básica de un concepto puede corresponder a una reorganización simple o revisionista, en dependencia de su radicalidad. En el caso de la propuesta de Nersessian, la creación y eliminación de conceptos también son consideradas por Thagard, y tanto la descendencia como la absorción conceptual pueden contarse como reorganización simple o revisionista. Además, otras absorciones pueden ser sinónimo de eliminación, por ejemplo, en casos en los cuales el nuevo concepto no mantiene ningún rasgo central del concepto anterior.

En este apartado y el anterior, analicé detalles del proceso de las discontinuidades conceptuales y sus distintos tipos que se derivan de las propuestas enmarcadas en la aproximación clásica del CC. En la sección de discusión me centraré en la sistematización de los rasgos extraídos, así como el conjunto de pasos mediante los cuales abordaré el CC de REPRESENTACIÓN en ciencia cognitiva.

5. Discusión

La aproximación clásica del CC, tal como ha sido entendida en la presente tesis, se encuentra configurada por la cuádruple analogía establecida entre las teorías científicas en la historia de la ciencia, los científicos como sujetos epistémicos, los estudiantes de ciencias y los niños en su desarrollo cognitivo. Esta aproximación fue caracterizada por medio de tres consideraciones. En primer lugar, la apelación al historicismo kuhniano aborda nociones tales como ciencia normal, paradigma, revolución científica e inconmensurabilidad. En segundo lugar, la afirmación de que el conocimiento presenta un marco coherente unificado que exhibe rasgos propios de las teorías (sostenida en enseñanza de las ciencias y psicología del desarrollo por autores como Posner y colaboradores y Carey) considera estructuras conceptuales organizadas que, entre otras funciones, constriñen las creencias y las inferencias que realizan los individuos. Por último, la hipótesis del continuo entre la cognición científica y la cognición humana común sostiene que, dado que ambas capacidades son continuas, puede ocurrir CC en razonamientos no científicos que son análogos a los observados en las ciencias.

La exposición anteriormente presentada no implica que los elementos que conforman esta cuádruple analogía sean iguales en todos los aspectos: las teorías, los científicos, los estudiantes y los niños mantienen diversas diferencias. Este punto es resaltado, por ejemplo, por Posner y colaboradores:

Las aproximaciones epistemológicas y aquellas de la psicología cognitiva y la ciencia cognitiva representan diferentes tradiciones. Sus problemas dominantes son diferentes, sus vocabularios son diferentes, y sus enfoques investigativos son diferentes. Se requiere una investigación cuidadosa del significado de esas diferencias (Strike y Posner, 1992, pp. 172-173).

El proceso de aprendizaje y el desarrollo de teorías científicas difieren en elementos tales como la radicalidad, el rol de la intención y la reflexión metacognitiva. Con respecto a la radicalidad, Carey (1985) sostiene que el cambio de novato a experto estudiado en el contexto de la enseñanza de la ciencia es una reestructuración en sentido débil. En este tipo de reestructuración se representan relaciones diferentes entre conceptos y se introducen conceptos nuevos, pero no existe un CC. Ello parece ser contrario a la tesis sostenida por autores como Posner y colaboradores (1982), para quienes la acomodación constituye una reestructuración de estructuras cognitivas análoga a las revoluciones científicas kuhnianas. Carey reconoce el trabajo de investigadores como McCloskey

(1983), quienes comparan el aprendizaje de la física con el paso histórico de la teoría del ímpetu a la mecánica galileana, y luego a la mecánica newtoniana. Sin embargo, afirma que los datos que respaldan dichas afirmaciones no son conclusivos y requieren mayor evaluación (Carey, 1985, p. 7).

Thagard (1992), en cambio, es positivo sobre la existencia de reorganización conceptual en el aprendizaje de la física similar a las revisiones conceptuales mayores en ciencias, mientras que cuestiona la falta de evidencia empírica que respalde la ocurrencia de revoluciones en el desarrollo cognitivo infantil. Critica particularmente el trabajo de Carey (1985). A su juicio, a pesar de que la autora sostiene que los niños reemplazan su teoría animista por teorías biológicas dada la mayor coherencia explicativa, muestra ejemplos que son compatibles con una visión según la cual el conocimiento biológico no compite ni es comparado con la teoría anterior. En lugar de un reemplazo, las teorías iniciales simplemente son olvidadas o relegadas a un empleo en contextos acotados³⁵. Por ejemplo, el concepto superordinado de SER VIVO no sería similar al colapso en la distinción aristotélica entre MOVIMIENTO NATURAL y MOVIMIENTO VIOLENTO realizada por Galileo, dado que la coalescencia posee menor grado de radicalidad que el colapso. En los conceptos ANIMAL y PERSONA parecen radicar meras adiciones de creencia, sin abandono de creencias previas. Por tanto, “comparado con las revoluciones científicas, el cambio conceptual descrito por Carey es modesto: una cierta reorganización conceptual no revisionista sin mucho rechazo de las creencias previamente sostenidas” (1992, p. 254). La plausibilidad de la existencia de CC parecidos a los presentados en la historia de la ciencia y de mecanismos de cambio similares a los vistos en los científicos requiere mayor investigación.

³⁵ Esta crítica en contra de la existencia de comparación sistemática y reemplazo teórico en niños se relaciona con una perspectiva que considera que el conocimiento infantil o estudiantil se describe de mejor manera por medio de una ecología que cuenta con elementos cuasi independientes y donde no existe un proceso de reemplazo teórico. Así, problematiza la fuerte apelación a la racionalidad y la coherencia adoptada por la aproximación clásica. Uno de sus mayores exponentes es diSessa (1993), quien describe el desarrollo del conocimiento físico de sentido común en términos de la perspectiva del conocimiento por piezas (*knowledge in pieces*). Las estructuras que lo conforman son denominadas primitivos fenomenológicos (*p-prims*), estructuras mentales atómicas débilmente organizadas que se originan en interpretaciones superficiales de la realidad, permiten interpretar la experiencia y usualmente son autoexplicativas —un fenómeno X sucede porque “así es como son las cosas” (diSessa, 1993, p. 112)—. No pueden ser considerados como teorías en cuanto no exhiben la organización y consistencia de ellas. En esta perspectiva, el cambio producido por la enseñanza de la ciencia opera por medio del cambio en las prioridades y los contextos de activación de los *p-prims*, que se reorganizan en dependencia de los roles de los elementos en el sistema de conocimiento de física.

Con respecto a la intención y reflexión metacognitiva, se puede afirmar preliminarmente que existe una diferencia importante en el papel que cumplen en las reestructuraciones producidas durante el aprendizaje y durante la producción científica. El segundo caso se trata de un proceso de solución de problemas, donde los sujetos epistémicos utilizan sus capacidades cognitivas y metacognitivas en razonamientos sistemáticos para resolver problemáticas surgidas en su campo. Por su parte, durante el aprendizaje y el desarrollo cognitivo, pareciera que las intenciones operan en elementos tales como la motivación para resolver contradicciones o fallos en predicciones, pero a menor escala³⁶. Muchos estudiantes no realizan el esfuerzo cognitivo necesario para provocar la reestructuración en sus conceptos, y los niños no siempre exhiben un nivel de conciencia y control cognitivo tal en el proceso de aprendizaje. Además, niños y estudiantes no poseen el grado de conciencia metacognitiva observado en los científicos. Por ejemplo, no realizan modelos formales o experimentación sistemática (Carey, 2009), y su empleo del *bootstrapping* no es explícito (Nersessian, 2008).

Las diferencias anteriores problematizan la adecuación de la aproximación clásica del estudio del CC en niños y estudiantes en algunos aspectos, pero su relevancia para el análisis del fenómeno en historia y filosofía de las ciencias se mantiene. En otras palabras, no afecta el empleo de recursos psicológicos en el área. Si bien los rasgos que se desprenden de ellas enfatizan aspectos distintos desde los recursos teóricos y los objetivos propios de cada disciplina, existen puntos de convergencia que posibilitan obtener una visión general. En particular, es posible establecer un conjunto de pasos para el análisis del CC en una ciencia basado parcialmente de los estudios de las discontinuidades conceptuales llevados a cabo por los autores estudiados:

1. Caracterización de los sistemas conceptuales sucesivos: en el abordaje de los sistemas conceptuales, se deben analizar factores tales como los compromisos epistemológicos y los problemas que gatillan la creación, cambio y comunicación

³⁶ Sin embargo, Sinatra y Pintrich (2003) argumentan en contra de esta idea. Los autores han estudiado el CC intencional que, de acuerdo con ellos, “se caracteriza por la iniciación y regulación consciente y dirigida a metas de los procesos cognitivos, metacognitivos y motivacionales para producir un cambio en el conocimiento” (2003, p. 6). Bajo esta perspectiva, el aprendizaje intencional media entre los procesos psicológicos del aprendiz y los factores externos, tales como el formato de la instrucción. El aprendiz inicia y monitorea su aprendizaje por medio de mecanismos como la autorregulación y el control metacognitivo. Por ejemplo, al poseer la intención de aprender cierta materia, el sujeto puede focalizarse en procesar una información mientras desatiende otra.

de conceptos, lo que se desprende de la noción de ecología conceptual de Posner y colaboradores (1982; Strike y Posner, 1992).

2. Individuación de los conceptos inexpresables en el vocabulario del sistema anterior, o bien que presenten modificaciones, de acuerdo con el rol que cumplen en la teoría: en el caso presente me enfocaré en REPRESENTACIÓN, por lo que este paso se cumple de antemano.
3. Demostración de existencia de un cambio entre los sistemas conceptuales sucesivos, y determinación del tipo de cambio: la comparación entre los sistemas incluye el examen de la plausibilidad de caracterizar estas discontinuidades en términos de inconmensurabilidad. Los tipos o grados principales del CC se observan en la propuesta de Thagard (1992), que posee mayor nivel de especificidad y es compatible con las otras propuestas planteadas. Por tanto, la taxonomía puede reducirse a los siguientes elementos:
 - a. Creación o adición de conceptos.
 - b. Eliminación de conceptos.
 - c. Reorganización básica: incluye diferenciaciones, coalescencias, descomposiciones, reanálisis de estructuras básicas con menor nivel de radicalidad, descendencias y absorciones conceptuales con menor nivel de radicalidad.
 - d. Reorganización compleja: incluye reorganización jerárquica, reanálisis de estructuras básicas con mayor nivel de radicalidad, descendencias y absorciones conceptuales con mayor nivel de radicalidad.
 - e. Metaclasificación.

Como mencionaré en el tercer capítulo (véase p. 139), los CC en REPRESENTACIÓN son mayoritariamente reorganizaciones básicas.

4. Estudio de los mecanismos que posibilitaron el CC: el proceso de *bootstrapping* planteado por Carey (2009) y Nersessian (2008) es especialmente relevante para la innovación en ciencias. Como mencionaré en el tercer capítulo (véase pp. 140-143), el CC en REPRESENTACIÓN ha sido propiciado por tres analogías: la analogía con el computador, con el funcionamiento neuronal y la problematización de los sistemas de grabación.

Este conjunto de pasos que se desprende del análisis de las propuestas vistas en el capítulo constituirá el marco por medio del cual estudiaré una línea de CC en la noción de REPRESENTACIÓN en ciencia cognitiva.

6. Conclusiones

El objetivo del presente capítulo consistió en la determinación de rasgos característicos de los CC relevantes para su análisis. Para ello, abordé obras fundacionales dentro de las tres disciplinas que conforman la aproximación clásica del CC: historia y filosofía de las ciencias, enseñanza de las ciencias y psicología del desarrollo (Kuhn 1970, 1982; Posner *et al.*, 1982; Strike y Posner, 1992; Carey, 1985, 1999a, 1999b, 2009; Thagard, 1988, 1992, 2011, 2014; Nersessian, 1992, 1995, 2008). Caractericé esta aproximación a través de tres consideraciones: (i) la apelación al historicismo kuhniano, (ii) la afirmación de que el conocimiento presenta un marco coherente unificado que exhibe rasgos propios de una teoría, y (iii) la hipótesis del continuo. La metodología para el análisis que se desprende de ello considera la caracterización de los sistemas conceptuales sucesivos, la individuación de los conceptos inexpresables en el vocabulario del sistema anterior o que presentan modificaciones, la demostración de existencia de un cambio entre los sistemas conceptuales junto con la determinación del tipo de cambio, y el estudio de los mecanismos que posibilitaron la discontinuidad. Algunos elementos de la ecología conceptual, el *bootstrapping* como método y la evaluación de existencia de inconmensurabilidad son elementos importantes en el estudio.

A partir de la información abordada en el capítulo, el CC puede entenderse como un proceso de solución de problemas tanto histórico como cognitivo: ocurre en las transformaciones sucesivas en una ciencia particular y en el proceso cognitivo de modificación de representaciones presente en el sujeto epistémico. En el caso de la producción de conocimiento de las ciencias y su relación con el quehacer del científico, ello apunta a la interrelación entre el componente fáctico y el componente racional mencionado en la introducción al capítulo (véase p. 13). Por otro lado, con respecto a su relación con la cognición humana en general, favorece la hipótesis del continuo. Dado que es un proceso que se desarrolla en un periodo de meses o años, la noción de cambio gestáltico es inapropiada para su descripción.

Como mencioné anteriormente, el CC se ha asociado a una transición entre teorías, siendo clásicamente analizado dentro del marco de las revoluciones científicas o sus análogos cognitivos (en el desarrollo cognitivo y la enseñanza). Las investigaciones se sitúan mayormente en disciplinas como la física y la química, consideradas maduras en sentido kuhniano. Con base en lo anterior, cabe plantear la interrogante sobre la

ocurrencia de CC en ciencias inmaduras tales como la ciencia cognitiva. Como mencioné en la introducción, de acuerdo con Barbara von Eckardt (1995), “lo que todos los científicos cognitivos comparten a lo largo del tiempo es un compromiso con algo parecido a un enfoque del estudio de la mente en lugar de un conjunto específico de teorías, explicaciones o leyes” (von Eckardt, 1995, p. 15). Las propiedades de las capacidades cognitivas (su dominio³⁷ de investigación) y las preguntas que se plantean sobre ellas son preteóricas. Por otro lado, las teorías que su investigación puede generar son tentativas. Dado que son tentativas, no existe un proceso de cambio teórico como los ocurridos en ciencias maduras en sentido kuhniano. Ello parece constituir un problema porque, como mencioné con anterioridad, el CC ha sido clásicamente asociado con el cambio de teoría.

Preliminarmente, algunos de los detalles abordados en este capítulo parecen apuntar a la posibilidad de CC sin cambio de teoría. En primer lugar, la identificación que realiza Carey de la emergencia de teorías como un segundo tipo de cambio teórico amplía el análisis tanto en la psicología del desarrollo como en la filosofía de las ciencias. En el primer caso, este sentido permite a la autora sostener que la reestructuración corresponde a la emergencia de una nueva teoría desde teorías intuitivas previas. En el segundo caso, posibilita plantear la interrogante sobre la ocurrencia del fenómeno en ciencias inmaduras. En segundo lugar, la crítica realizada por Thagard (1992) a Carey (1985) indica que, si bien el cambio de teoría es relevante en la mayor parte de las aproximaciones de CC, no es una condición necesaria: es posible que no exista una comparación entre teorías, o que la teoría anterior quede relegada a su uso en situaciones más concretas. En el próximo capítulo responderé afirmativamente a la posibilidad de CC sin cambio de teoría sobre la base del análisis de los antecedentes de identificación de algunas transformaciones conceptuales en ciencia cognitiva.

³⁷ Con *dominio* la autora se refiere al conjunto de fenómenos considerados por el marco de investigación (1995, p. 19), que existen con independencia del aparato conceptual o teórico.

CAPÍTULO 2: CAMBIO CONCEPTUAL EN CIENCIA COGNITIVA

1. Introducción

En el capítulo anterior expuse la aproximación clásica del CC y los rasgos del proceso que pueden extraerse de ella. En el presente capítulo acotaré el estudio a los antecedentes de CC encontrados en la literatura especializada en ciencia cognitiva. Su revisión es relevante para los objetivos de la presente tesis por dos razones. En primer lugar, permite observar el estado del arte de las discusiones acerca de la posibilidad de la existencia de reconceptualizaciones en el área. En segundo lugar, ofrece un punto de partida desde el cual abordar posteriormente las variaciones en REPRESENTACIÓN.

En el apartado siguiente me centraré en tres características de la ciencia cognitiva clásica relevantes para el primer paso del análisis del CC que he propuesto en el quinto apartado del capítulo precedente, a saber, la caracterización de los sistemas conceptuales sucesivos. Estos rasgos de la ciencia cognitiva clásica son (i) sus supuestos sustantivos —computacional y representacional—, (ii) su interdisciplinariedad y (iii) su inmadurez. Las problematizaciones de los supuestos sustantivos cumplen un rol central en las discusiones donde posiblemente se localicen transformaciones importantes de conceptos. La interdisciplinariedad introduce un factor de variabilidad en el rol que cumplen algunas nociones —dentro de las cuales se cuenta REPRESENTACIÓN— en ciertas subdisciplinas que componen la ciencia cognitiva. Por último, la inmadurez implica que los debates entre teorías continúan abiertos, por lo que el proceso de transformación de conceptos también se encuentra en curso. Más precisamente, la actualidad de las prácticas y estudios en ciencia cognitiva muestra una variabilidad de enfoques que divergen entre sí. La determinación del proceso de CC debe considerar este hecho.

En el tercer apartado analizaré los estudios que han individualizado CC en el área, a saber, Thagard (1992, 2011), Akagi (2017) y Keijzer (2002). Estos se han enfocado en nociones centrales como PROCESAMIENTO DE INFORMACIÓN, CONCEPTO, COGNICIÓN y REPRESENTACIÓN. En la sección de discusión, sostendré que los análisis son, en términos generales, congruentes entre sí en la identificación de discontinuidades conceptuales ocurridas en las etapas de cambio relevantes. En conjunto, sugieren la existencia de tres

momentos de transformación conceptual: en la denominada *revolución cognitiva* (dando paso a los enfoques simbólicos y conexionistas); en las discusiones de los límites de la cognición, y en la recientemente considerada *revolución neurocognitiva*. En cada una de ellas se observan problematizaciones y concepciones con respecto a las representaciones. A modo de conclusión, afirmaré que los antecedentes presentados, en la medida en que sean correctos, responden afirmativamente a la posibilidad de CC en una ciencia inmadura como la ciencia cognitiva.

2. La ciencia cognitiva: supuestos, interdisciplinariedad e inmadurez

La historia de la ciencia cognitiva, entendida como el estudio interdisciplinario de los procesos cognitivos³⁸, data de la segunda mitad del siglo XX, entre las décadas de 1950 y 1960. Si bien las disciplinas que la componen (psicología, neurociencia, inteligencia artificial, lingüística, filosofía y antropología) poseen una historia más amplia y consolidada, la ciencia cognitiva como tal corresponde a un área nueva de investigación. A pesar de este hecho, autores como Barbara von Eckardt sostienen que su comunidad de investigadores mantiene una unidad y coherencia en el campo. En *What is Cognitive Science?* (1995) von Eckardt explica estas características en términos de la posesión de un marco transdisciplinario de compromisos compartidos, que funciona como marco de investigación. El componente central del programa de investigación es el estudio de la cognición del adulto normal típico, un punto de referencia para otros enfoques tales como el estudio del desarrollo o el deterioro cognitivo. El marco de investigación es conformado por un esquema de cuatro componentes: las especificaciones preteóricas de dominio, las preguntas empíricas básicas, los supuestos sustantivos y los supuestos metodológicos³⁹.

La caracterización de los detalles propios del marco de investigación de la ciencia cognitiva se escapa de los objetivos de la presente tesis. Sin embargo, este posee al menos tres rasgos que son especialmente relevantes para el primer paso del análisis del CC en

³⁸ La definición de cognición es un problema conceptual central en la ciencia cognitiva. Como se verá (pp. 83-86), uno de los análisis sobre CC, llevado a cabo por Akagi (2017), precisamente involucra esta noción.

³⁹ Los supuestos sustantivos especifican la aproximación adoptada para el estudio de la cognición del adulto normal típico y restringen las respuestas satisfactorias a las preguntas empíricas básicas que direccionan la investigación. Esta aproximación entiende las capacidades cognitivas como un sistema de capacidades computacionales y representacionales, por lo que reformulan las preguntas en términos de procesamiento de información.

esta ciencia, a saber, (i) sus supuestos sustantivos —representacional y computacional— ; (ii) su interdisciplinariedad y (iii) su inmadurez. En los próximos tres subapartados abordaré cada una de estas características, basándome de manera no estricta en la descripción otorgada por von Eckardt (1995, 2006). Sostendré que imponen restricciones al análisis del CC de una noción central como REPRESENTACIÓN.

2.1. Computacionalidad y representacionalidad

El mérito históricamente atribuible a la ciencia cognitiva radicó en la postulación de una batería teórica y metodológica nueva para estudiar científicamente la mente/cerebro. Previo a su surgimiento, la metodología imperante correspondía a aquella postulada por el conductismo. Entre 1920 y 1960, autores como Watson y Skinner limitaban el objeto de estudio de la psicología al comportamiento observable, rechazando la definición mentalista del objeto de estudio y el método introspeccionista⁴⁰ por el cual se abordaba. Para conservar un estatus científico, se eliminó la mente (entendida como experiencia consciente) como tema de investigación y, en su lugar, se desarrollaron leyes y explicaciones en torno al esquema estímulo-respuesta. En contraste con este panorama, la estrategia explicativa clásicamente adoptada por la ciencia cognitiva consiste en la explicación de las capacidades cognitivas por medio de su reconceptualización en términos de un sistema de capacidades computacionales y representacionales, es decir, capacidades de procesamiento de información. Según von Eckardt, estas capacidades poseen cinco características importantes. En primer lugar, son intencionales: involucran transformaciones de representaciones que poseen contenido o son acerca de un objeto o estado de cosas en el mundo. En segundo lugar, son evaluables pragmáticamente, ya que pueden ser ejercidas correcta o incorrectamente. En tercer lugar, son coherentes, dado que exhiben una relación de satisfacción entre los estados inicial y final cuando son ejercidas correctamente, en virtud de los contenidos de los estados involucrados. En cuarto lugar,

⁴⁰ La reconstrucción clásica de la historia de la psicología sostiene que los psicólogos experimentales de finales del siglo XIX y principios del siglo XX, como Wilhelm Wundt, confiaban en el introspeccionismo como método para estudiar la mente; posteriormente, fue reemplazado por el conductismo de John Watson y otros, quienes desplazaron el objeto de estudio a la conducta observable; por último, el conductismo fue superado por la revolución cognitiva. Esta historia contiene diversas distorsiones. Por ejemplo, es posible hallar enfoques conductistas entre los investigadores de los primeros laboratorios de psicología y la introspección como método fue cuestionada por Wundt mismo (Gallagher y Zahavi, 2008, p. 3).

son productivas: manifestables de maneras novedosas. Por último, son confiables: suelen ser ejercitadas con éxito.

La computacionalidad y representacionalidad ocupan el puesto de supuestos sustantivos en el área. El supuesto computacional afirma que la mente/cerebro corresponde, en gran medida⁴¹, a un dispositivo computacional. Un computador es entendido como un

dispositivo capaz de ingresar, almacenar, manipular y generar automáticamente información en virtud del ingreso, almacenamiento, manipulación y generación de representaciones de dicha información (...) en concordancia con un conjunto finito de reglas que son efectivas y que están, en algún sentido, en la máquina misma (von Eckardt, 1995, p. 50).

En otras palabras, se trata de dispositivos que procesan información por medio del procesamiento de las representaciones de esta. Este concepto es deudor del avance tecnológico de los computadores electrónicos y los trabajos en computabilidad de autores como Alan Turing, que permitieron posteriormente estudiar el funcionamiento de la mente/cerebro humana y su estructura a través del desarrollo de arquitecturas cognitivas.

Una arquitectura cognitiva corresponde a “una propuesta general acerca de las representaciones y procesos que producen el pensamiento inteligente” (Thagard, 2012 p. 50). Se refiere tanto a los modelos de cognición como a sus instanciaciones en los *software* desarrollados en inteligencia artificial. Incluye descripciones de las capacidades, operaciones y recursos del sistema. En términos generales, las arquitecturas se pueden dividir en aquellas que implementan las funciones en máquinas convencionales y aquellas que recurren a máquinas conexionistas. Los computadores convencionales consisten en dispositivos de procesamiento comúnmente seriales⁴², es decir, capaces de ejecutar una operación a la vez. Las arquitecturas convencionales esencialmente contemplan los siguientes componentes principales: dispositivos de *input*, encargados de la percepción; dispositivos efectores; unidad de procesamiento central (que incluye una unidad de control y unidades lógicas/aritméticas), más diversas unidades de memoria asociadas a los dos primeros componentes (Stillings *et al.*, 1995, p. 18). Los dispositivos también

⁴¹ La autora sostiene que la mente/cerebro corresponde en gran medida a un dispositivo computacional porque incluye transducción sensorial y de movimiento motor que puede no ser estrictamente computacional.

⁴² Existen algunas máquinas convencionales que pueden poseer un procesamiento parcialmente paralelo denominado *procesamiento concurrente*, donde algunos componentes funcionan de manera serial y otros en paralelo.

cuentan con un conjunto de operaciones básicas⁴³ que se encuentran definidas sobre representaciones o estructuras de datos. Las funciones de las unidades de *input* y *output*, que permiten el ingreso y salida de información del computador, se computan por medio de la ejecución de estas operaciones básicas de acuerdo con un programa: un conjunto de instrucciones explícitas que pueden estar almacenadas en la memoria o bien ser fijas (creadas por la interconexión de cables en el *hardware*). El orden de ejecución de las instrucciones es determinado por la unidad de control, y están escritas en código de máquina. El código de máquina es un lenguaje que interpreta directamente el computador y que resulta complejo de entender y manipular por el humano. Por ello, se han desarrollado lenguajes de nivel superior que codifican las instrucciones de manera similar al lenguaje natural, donde cada comando básico equivale a una subrutina en el código de máquina. En otras palabras, es reducible al conjunto de instrucciones efectivas básicas del código de máquina.

Por su parte, las arquitecturas conexionistas o de procesamiento paralelo distribuido están vagamente inspiradas en las redes neuronales del cerebro, capaces de ejecutar varias operaciones a la vez. El diseño de la arquitectura consiste en redes conformadas por un conjunto de unidades de procesamiento análogas a neuronas abstractas, que presentan patrones de conectividad con pesos (fuerzas numéricas). Cada unidad posee un valor numérico de activación que es enviado a otras unidades por medio de conexiones excitatorias o inhibitorias. Al llegar la señal de *input*, las unidades que la reciben calculan su valor de activación en dependencia de la naturaleza excitatoria o inhibitoria de las señales, los pesos de los enlaces entre unidades y el número de señales de *input* que reciben. De acuerdo con ello, computan *outputs*: una unidad con valor de activación y peso positivo establece una influencia excitatoria, mientras que una unidad con valor de activación y peso negativo establece una conexión inhibitoria. Los valores de activación de las unidades de *input* corresponden al *input* del modelo, que propaga dicha activación a través de las conexiones hasta las unidades de *output*, las que codifican el *output* del modelo. Además, entre ambas existen unidades ocultas (*hidden units*), que no se asocian con la computación de *inputs* ni *outputs*, sino que pueden desarrollar representaciones internas asociadas con el procesamiento de información a niveles más complejos. Los pesos o fuerzas numéricas de conexión codifican el conocimiento del sistema —la

⁴³ El conjunto de operaciones básicas varía en dependencia del tipo de computador, pero usualmente incluye operaciones aritméticas, lógicas, operaciones de movimientos de datos, entre otras.

memoria y el procesamiento son acarreados por el mismo conjunto de unidades—. Por ende, cumplen un rol análogo a los programas en los computadores convencionales. La diferencia con ellos radica en que los patrones de conectividad de las redes son capaces de modificar sus pesos para realizar una computación en específico mediante el aprendizaje. Este aprendizaje se logra por un entrenamiento donde se presentan pares de *input/output* de la función que debe computar (Smolensky, 1988, p. 1). El comportamiento de la red depende del estado de activación inicial y los pesos de conexión.

El supuesto computacional se relaciona con el segundo supuesto, esto es, el supuesto representacional, dado que un computador es un dispositivo que procesa información por medio del procesamiento de representaciones. Este segundo supuesto sostiene que la mente/cerebro es un dispositivo representacional, un “dispositivo que posee estados o que contiene dentro de sí entidades que son representacionales” (von Eckardt, 1995, p. 50).

Las representaciones son clásicamente consideradas como entidades mentales que tienen propiedades semánticas, es decir, poseen un contenido vinculado a fenómenos, propiedades o estados de cosas en el mundo. Sin embargo, a pesar de ser clásicamente constitutiva del estudio de la mente y la conducta, no posee una noción clara o estándar. Frente a ello, von Eckardt (1995, 2006) sistematiza los componentes relevantes de la definición general por medio de la consideración de la teoría de las representaciones propuesta por Charles Sanders Peirce. Para el autor, la representación involucra una relación triádica entre un signo o *representamen*, un objeto y un interpretante. Algo es un signo (una cualidad, evento o tipo) solo si es signo de un objeto⁴⁴ (no mental) con respecto a un interpretante (efecto mental en el intérprete). von Eckardt rescata de manera parcial cuatro elementos centrales de la teoría peirciana de las representaciones para elaborar su descripción. En primer lugar, el portador de la representación (*representation bearer*) corresponde a estructuras computacionales que dependen del tipo de computador a considerar. En segundo lugar, el objeto representacional es aquel o aquellos objetos con los que las representaciones mantienen relaciones semánticas. En otras palabras, son los

⁴⁴ La noción de objeto es un poco más compleja en la teoría de Peirce. El autor realiza una distinción entre objeto inmediato y objeto dinámico. El objeto inmediato es el aspecto particular del objeto representado por el signo, mientras que el objeto dinámico es el objeto mismo, independiente de la mente y de la relación de representación. Una instancia de representación puede involucrar ambos. En este caso, estarían relacionados entre sí, pero no serían idénticos (von Eckardt, 1995, p. 147).

objetos que constituyen su significado o contenido. Entre ellos se incluyen propiedades, eventos, estados de cosas, objetos concretos, entre otros. En tercer lugar, el fundamento (*grounding*) de la representación corresponde a las propiedades o relaciones que determinan sus propiedades semánticas (el contenido)⁴⁵. Por último, el interpretante de una representación corresponde al conjunto de procesos computacionales posibles involucrados en la posesión de la representación por parte de un sujeto. El interpretante es producido por la representación y permite que su contenido posea eficacia psicológica sobre el sujeto que la posee, es decir, que influya en sus estados mentales y comportamientos. Al final del tercer capítulo (pp. 146-147) indicaré qué elementos son entendidos como portador, fundamento e interpretante para las máquinas convencionales, conexionistas y los sistemas de símbolos perceptuales.

En suma, las representaciones mentales son estructuras o estados en la mente que poseen propiedades semánticas y que cuentan con cuatro aspectos centrales: están realizadas por un portador, representan uno o varios objetos representacionales, se encuentran fundamentadas (*grounded*) y son interpretadas por (o funcionan como representación para) un sujeto o intérprete. Volveré a las características de las representaciones en el tercer capítulo, donde consideraré el proceso de CC en esta noción.

Los supuestos sustantivos computacional y representacional como características principales de la ciencia cognitiva constituyen un primer rasgo relevante a considerar en el análisis del CC en esta ciencia. Las problematizaciones de los supuestos sustantivos cumplen un rol central en el desarrollo de una disciplina, sobre todo si es inmadura en sentido kuhniano. Por tanto, la caracterización de los sistemas conceptuales sucesivos y los fenómenos que pretenden explicar debe considerar estos elementos para dilucidar el proceso de transformación de los conceptos involucrados. En particular, el hecho de que REPRESENTACIÓN figure dentro de los supuestos sustantivos indica, preliminarmente, la complejidad de dicho proceso. A continuación, abordaré la segunda característica relevante para el estudio de CC en ciencia cognitiva: su interdisciplinariedad.

⁴⁵ Este fundamento debe ser naturalista, es decir, no involucrar propiedades o relaciones semánticas ni intencionales (para no caer en una petición de principio, a saber, la explicación del contenido de la representación en términos de contenido).

2.2. La interdisciplinariedad de la ciencia cognitiva

Un segundo rasgo relevante para el análisis del CC en ciencia cognitiva corresponde a su interdisciplinariedad. Como mencioné anteriormente, esta área se encuentra conformada por diversas ciencias, donde clásicamente se cuenta la psicología, filosofía, inteligencia artificial, neurociencias, lingüística y antropología. Ello justifica la existencia de niveles de descripción y de explicación distintos en ella. Dos ejemplos de este punto radican en las distinciones entre niveles propuestas por Marr (1982) y Dennett (1986), que han sido utilizadas recurrentemente en la literatura. Si bien el primero se refiere específicamente al tratamiento computacional de la percepción visual, su división refleja la necesidad de apelar a disciplinas distintas en el estudio de los sistemas cognitivos.

Marr (1982) identifica tres niveles de organización en un dispositivo de procesamiento de información. El nivel superior refiere a la teoría computacional, donde se caracteriza el desempeño del dispositivo —entendido como un mapeo de una clase de información a otra—. En él, se determinan las propiedades del mapeo y la meta de la computación. El segundo nivel pertenece a la representación y algoritmo, que describe las estructuras representacionales y procesos computacionales que permiten que el dispositivo ejecute la función. Allí se elige la representación para el *input* y el *output*, así como el algoritmo que transforma una en la otra. Por último, el nivel de la implementación describe la realización física de los algoritmos y las representaciones, desarrollando la arquitectura de forma detallada. Marr establece que, si bien los tres niveles se encuentran relacionados, existe una relativa independencia. Esta independencia se debe a dos hechos: las elecciones disponibles en cada uno son amplias (qué tipo de algoritmo se empleará en el segundo nivel, qué tipo de implementación física en el tercer nivel, etcétera), y algunos elementos explicativos específicos en cada uno son independientes. Por ello, existen fenómenos que pueden ser explicados solo en uno de estos niveles: por ejemplo, la neuroanatomía pertenece al tercer nivel, mientras que la psicofísica al segundo nivel⁴⁶. Los niveles de Marr han sido, en ocasiones, renombrados como nivel de contenido, forma y medio en algunos textos de psicología cognitiva, o como nivel semántico, sintáctico y físico.

⁴⁶ Marr (1982) sugiere que esta división entre niveles ayuda en la evaluación de objeciones clásicas como la diferencia entre el cerebro y un computador serial. Según el autor, en cuanto la distinción entre procesamientos seriales y paralelos pertenece al nivel representacional y algorítmico, no excluye la posibilidad de programar un computador para realizar las mismas tareas que un cerebro.

Por su parte, en *Content and Consciousness* (1986) —con una primera edición de 1969—, Dennett realiza una distinción entre el nivel personal y subpersonal de explicación. El primero de ellos se refiere a las personas y sus acciones, mientras que el segundo versa sobre los eventos propios del cerebro y el sistema nervioso. El autor introduce la diferencia durante su análisis del fenómeno mental del dolor. A nivel personal, las sensaciones dolorosas pueden ser discriminadas, localizadas y gatillar una reacción de una manera no mecánica: no existe un proceso específico por el cual el individuo realiza la discriminación o localización, sino que simplemente puede realizarlas. El dolor, por tanto, corresponde a una cualidad inanalizable. En el nivel subpersonal, en cambio, el fenómeno se explica en términos mecánicos, por medio de la transmisión del estímulo desde el área afectada a través de una red neural específica. A pesar de la diferencia, ambos niveles de explicación se encuentran relacionados. Con respecto al dolor, “no puede ser el caso de que no exista una relación entre los dolores y los impulsos neurales” (1986, p. 95). Esta distinción ha sido utilizada recurrentemente en filosofía de la mente y ciencia cognitiva⁴⁷. Por ejemplo, como mencionaré en el tercer apartado, Keijzer (2002) basa parte de su crítica sobre el concepto de REPRESENTACIÓN en el hecho de que la ciencia cognitiva clásica ha utilizado la explicación sobre las representaciones a nivel personal como guía para explicaciones de nivel subpersonal (pp. 87-88).

La interdisciplinariedad de la ciencia cognitiva y la multiplicidad de niveles de explicación y descripción impone una dificultad al estudio del CC de una noción central como es el caso de REPRESENTACIÓN: sus características y roles varían en dependencia de la subdisciplina. En psicología cognitiva, por ejemplo, es considerada como un supuesto básico aún en la actualidad. En el *Tratado de neuropsicología clínica* (segunda edición, 2018) se define la psicología cognitiva como “la rama de la psicología que estudia cómo la mente se representa internamente el mundo externo y lleva a cabo la computación mental requerida para todos los aspectos del pensamiento” (Saavedra, 2018, p. 3). El panorama es distinto en otra subdisciplina como la inteligencia artificial donde, para la conformación de dispositivos inteligentes, no se requiere forzosamente la apelación de representaciones. En *The Cambridge Handbook of Artificial Intelligence* (2014) se identifica como uno de los temas principales de la inteligencia artificial el

⁴⁷ Cabe notar que la distinción no es propia de Dennett. El autor señala que puede ser localizada en trabajos anteriores de autores como Ryle y Wittgenstein.

cuestionamiento a la necesidad de que exista una representación o modelo del mundo en el sistema (en forma de reglas, símbolos u otros). De acuerdo con Franklin (2014), “surgió una controversia sobre la cantidad de modelización que realmente se necesitaba. Algunos afirmaron que se podía hacer mucho sin necesidad de una amplia modelización interna” (p. 16). Uno de los principales impulsores de este cuestionamiento fue Rodney Brooks (1999) quien, desde la robótica situada, sostiene que “la representación es la unidad de abstracción errónea en la construcción de las mayores partes de los sistemas inteligentes” (Brooks, 1999, p. 81). Cuestiona la necesidad de representaciones centrales explícitas en el diseño de robots autónomos: en niveles muy simples de inteligencia, la complejidad de dichas representaciones ralentiza su procesamiento, lo que dificulta la reacción en tiempo real del robot en un ambiente dinámico. Las diferencias vistas entre disciplinas como la psicología cognitiva y la inteligencia artificial instauran la posibilidad de que las variaciones en la noción de REPRESENTACIÓN dependan del área o las áreas involucradas directamente en un enfoque específico. En el próximo subapartado me detendré en el último punto relevante para el estudio de discontinuidades conceptuales en la ciencia cognitiva: su inmadurez.

2.3. La inmadurez de la ciencia cognitiva

Un tercer rasgo de la ciencia cognitiva a considerar en el análisis de CC es su inmadurez. Esta característica apunta a la posesión de un marco de investigación inmaduro, es decir, un marco sin un antecesor científico del cual haya evolucionado, cuyas especificaciones de dominio se encuentran ligadas a las concepciones de sentido común. Ello implica la ausencia de un grupo estable de teorías, leyes o explicaciones que no varíen de manera importante en relación con el proceso de investigación científica. Para von Eckardt, “lo que todos los científicos cognitivos comparten en el transcurso del tiempo es un compromiso con algo parecido a un enfoque del estudio de la mente” (von Eckardt, 1995, p. 15). A pesar de ello, su actividad de investigación exhibe coherencia, producto de la existencia de compromisos compartidos.

La diferencia entre una ciencia madura y una ciencia inmadura es, según la autora, una semejanza genética. En sus inicios, la actividad científica se funda en preguntas surgidas de la experiencia ordinaria. Por ende, en las ciencias inmaduras, el marco de investigación se extrae de concepciones ordinarias y sin antecedentes científicos del mundo, de las cuales surgen las especificaciones de dominio (los tipos de fenómenos y

las propiedades que lo determinan) y las preguntas básicas. La autora puntualiza, en particular, cuatro diferencias entre las ciencias maduras e inmaduras. En primer lugar, existe una menor confianza con respecto a la concepción del dominio de estudio que posee una ciencia inmadura. Ello sucede porque, durante el progreso de la investigación, se generan revisiones de los supuestos de dominio y preguntas básicas —surgidos inicialmente del sentido común— producto de los cambios conceptuales generados por las actividades científicas. En segundo lugar, los constreñimientos existentes en la investigación desarrollada bajo un marco inmaduro poseen el estatus de supuestos de trabajo y no de hipótesis aceptadas. En cuanto no cuentan con las referencias a descubrimientos científicos y conceptualizaciones de marcos anteriores, dichos constreñimientos son menores en comparación con una ciencia madura. En tercer lugar y en relación con la tecnología disponible, esta es típicamente menos avanzada que en el caso de una ciencia madura, de modo que “un marco inmaduro a menudo necesita inventar su tecnología desde cero” (von Eckardt, 1995, p. 32). Esto se debe a la falta de una herencia de equipamientos y métodos típicamente otorgada por marcos anteriores. Por último, la comunidad científica inmadura no posee una identidad profesional desarrollada —por ejemplo, mostrada a través de revistas especializadas o recursos de fondos.

Durante el transcurso del desarrollo científico, el trabajo realizado dentro de un marco de investigación puede conducir a otro marco de investigación. Un ejemplo se encuentra en la evolución de la biología, donde las preguntas formuladas en términos del gen en genética mendeliana avanzaron hacia la conformación del marco de la genética molecular. Así, en contraste con las ciencias inmaduras, las ciencias maduras poseen un marco producto de la evolución de otros marcos, y la herencia aportada por estos permite que la comunidad posea mayor confianza en la concepción que tiene sobre su dominio de estudio. Los resultados empíricos y teóricos anteriores ocasionan un constreñimiento mayor de las investigaciones. Además, cuentan con mayor avance tecnológico y una comunidad más afianzada en relación con la identidad profesional, por heredar parte de la estructura sociológica de la comunidad anterior. En suma, las diferencias mencionadas sobre los marcos que caracterizan ambos tipos de ciencias son expresadas por la autora de la siguiente forma:

La distinción entre un marco de investigación científico inmaduro y uno maduro es claramente de grado. Un marco de investigación enraizado en nuestra concepción ordinaria del mundo y sin ningún antecesor científico es claramente inmaduro. Un marco que ha evolucionado a partir de una larga lista de otros marcos científicos de

investigación es claramente maduro. Un marco con cierto linaje científico, pero no mucho, será un caso poco claro (1995, p. 31).

La ciencia cognitiva cumple con todos los rasgos de una ciencia con un marco de investigación inmaduro. En esta área, tanto las propiedades de los fenómenos que conforman su dominio de investigación —esto es, las capacidades cognitivas humanas— como las preguntas básicas planteadas con respecto a él son conceptualizadas preteóricamente, puesto que corresponden a lo que el sentido común conoce sobre ellos. Por otro lado, las teorías que su investigación puede generar son tentativas, por lo que la ciencia cognitiva “se encuentra todavía en la búsqueda de un conjunto de compromisos teóricos más permanentes” (von Eckardt, 1995, p. 15).

La inmadurez de la ciencia cognitiva impone una dificultad al análisis del CC allí. En el primer capítulo sostuve que este fenómeno se ha asociado sistemáticamente con un cambio de teoría (véase p. 28). Ello no excluye la posibilidad de que exista transformación de conceptos en el área —como mencioné antes, autores como Susan Carey y Paul Thagard otorgan antecedentes que permiten sostener su existencia (p. 61)—. Sin embargo, esto vuelve problemática la identificación del proceso del CC si este se entiende en términos radicales. El panorama actual de la disciplina muestra un conjunto de aproximaciones que, junto con problematizar uno o más supuestos sustantivos, postulan una batería de enfoques teóricos y metodológicos distintos entre sí. Por tanto, la noción central de REPRESENTACIÓN muestra divergencias entre teorías, y las discusiones continúan abiertas. Ello puede implicar que el proceso de transformación de conceptos también se encuentra en curso.

En el presente apartado describí los tres rasgos de la ciencia cognitiva que son relevantes para la caracterización de los sistemas conceptuales sucesivos—sus supuestos, su interdisciplinariedad y su inmadurez—. A continuación, revisaré los antecedentes de estudio que han identificado CC en la disciplina.

3. Antecedentes de cambio conceptual en ciencia cognitiva

Existen escasos antecedentes en la literatura especializada que responden afirmativamente a la pregunta sobre la posibilidad de CC en ciencia cognitiva. Las investigaciones se han enfocado en nociones principales en el área, como

PROCESAMIENTO DE INFORMACIÓN, CONCEPTO, COGNICIÓN y REPRESENTACIÓN. Analizaré la línea argumental de tres investigaciones: Thagard (1992, 2011), Akagi (2017) y Keijzer (2002).

3.1. Thagard sobre revolución cognitiva y revolución neurocognitiva

Un primer estudio fue llevado a cabo por Thagard (1992) en el marco de la denominada *revolución cognitiva* ocurrida en psicología. Como mencioné en el primer capítulo, el autor sostiene que el mecanismo central que opera en las revoluciones conceptuales es la coherencia explicativa, una relación entre proposiciones modulada por un conjunto de principios que posibilitan la evaluación de su aceptabilidad en un sistema explicativo. En ciencias naturales tales como la física, la biología o la química, las revoluciones provocan cambios mayores en las jerarquías conceptuales. Ello no ocurre en el caso de la psicología. Si bien el autor admite la existencia de CC allí, sostiene que mantiene diferencias importantes con las ciencias antes mencionadas. A grandes rasgos, la reconstrucción argumental de Thagard puede ser expuesta de la siguiente forma:

- a. Las ciencias naturales presentan CC en sentido fuerte, esto es, revoluciones conceptuales.
- b. Las revoluciones conceptuales en ciencias naturales se producen por la aceptación de una teoría que posee mayor coherencia explicativa que la anterior.
- c. Las ciencias sociales tales como la psicología no poseen un cuerpo de teorías que puedan ser reemplazadas por otras con mayor coherencia explicativa, sino que solo cuentan con aproximaciones en competencia.
- d. Por lo tanto, la psicología no presenta revoluciones conceptuales.
- e. La psicología exhibe CC en sentido débil, producido por comparaciones entre metodologías o expectativas de una coherencia explicativa futura.

Para el autor, el paso del conductismo al cognitivismo corresponde a la competencia entre dos aproximaciones, no entre dos teorías explicativas. La diferencia entre ambas nociones es central para el argumento. Mientras que una teoría es una colección coherente de hipótesis que explican una porción amplia de fenómenos, una aproximación es entendida como un conjunto de estilos de explicación y de metodologías que es más general y difusa que una teoría. Las aproximaciones poseen marcos de investigación, es decir, enfoques distintos en el tratamiento del objeto de estudio que son más específicos que las aproximaciones, pero más vagos que las teorías. La relación entre estas tres es

resumida por Thagard de la siguiente manera: “una teoría puede formar parte de un marco que asume una aproximación” (1992, p. 226)⁴⁸. Bajo este contexto, el cognitivismo es una aproximación caracterizada por la postulación de representaciones mentales y procesos computacionales, que son análogos a las estructuras de datos y los algoritmos de los programas computacionales. Así, el comportamiento inteligente es explicado mediante la posesión de estas representaciones y procesos que operan sobre ellas. A través de la analogía con el computador, el desarrollo de las ciencias de la computación otorgó a dicha aproximación la posibilidad de explicar potencialmente una mayor cantidad de hechos en comparación con su competidor, el conductismo, demostrando mayor expectativa de coherencia explicativa.

El esquema conductista explicaba el comportamiento humano en términos de estímulos y respuestas. Conceptualmente, estaba caracterizado por la eliminación de conceptos mentalistas de sentido común incorporados en el introspeccionismo tales como CONCIENCIA o VOLUNTAD y algunas reorganizaciones revisionistas, como la reinterpretación de LENGUAJE en términos de comportamiento verbal. Su simplicidad excesiva limitaba las explicaciones de la conducta humana y animal al comportamiento observable. En contraste, las herramientas de las ciencias de la computación permitieron al cognitivismo aumentar la estimación de éxito en comparación con el conductismo. El proceso de reemplazo entre ellos fue un proceso de sublación, es decir, una incorporación parcial de conceptos conductistas mientras se rechazaron sus explicaciones del comportamiento. El cognitivismo integró parcialmente tanto el aparataje conceptual conductista como el de la psicología de sentido común. Para Thagard, este tipo de evaluaciones conducentes a la elección entre aproximaciones posee una forma subjetiva e inexacta como “parece que la aproximación A1 será capaz de producir una teoría para explicar E, pero la aproximación A2 no parece ser lo suficientemente robusta como para producir tal teoría” (1992, p. 240). Siguiendo la clasificación de tipos de CC propuesta en el capítulo anterior, se pueden identificar las siguientes transformaciones:

1. Creación o adición: se han introducido conceptos computacionales: ESQUEMA, REGLA, PLAN, entre otros, como parte de la consideración de la analogía con el computador.

⁴⁸ En presente tesis no sigo las diferencias entre marco y aproximación tal y como son planteadas por Thagard (1992), sino que utilizo ambos términos en sentido amplio.

2. Reorganización básica: Existe una diferenciación de INFORMACIÓN como una noción matemática, como ligada a un procesador y como ecológica. En el primer caso, corresponde a una propiedad de una señal expresada por la fórmula matemática $I(s) = \log 1/p(s)$, donde la señal conduce mayor información (en *bits*) por un canal de comunicación si es menos probable el mensaje. Esta noción fue desarrollada por Shannon y Weaver (1949) en teoría de la comunicación. En el segundo caso, INFORMACIÓN refiere a estructuras computacionales que se encuentran en la mente. Esta noción se vincula con los trabajos de Miller (1956) y Newell, Shaw y Simon (1958). Miller (1956) sostuvo que la limitación que posee la memoria a corto plazo (siete elementos, más o menos dos) puede ser superada por la recodificación de los estímulos en fragmentos o *chunks* de información en lugar de *bits*. Newell, Shaw y Simon (1958) desarrollaron la aproximación de procesamiento de información que establecía la analogía directa entre el pensamiento y los procesos computacionales. En el tercer caso, INFORMACIÓN es una propiedad que se encuentra en el entorno, y se origina en la aproximación ecológica al estudio de la cognición propuesta por Neisser (1976) e inspirada en la aproximación ecológica a la percepción de Gibson. Además de la diferenciación de INFORMACIÓN, se puede hallar una coalescencia de los conceptos MENTE y COMPUTADORA en el concepto de PROCESAMIENTO DE INFORMACIÓN. Ello ocurre porque la mente es entendida como un sistema de capacidades computacionales y representacionales análogo a los computadores. Por último, existe una absorción de conceptos conductistas en términos cognitivos, como CONDICIONAMIENTO, que transita de ser considerado como una asociación de patrones de comportamiento frente a estímulos determinados (Watson, 1959; Skinner, 1953) a ser entendido como un aprendizaje de relaciones entre eventos que depende de expectativas y representaciones internas, sostenido por autores como Rescorla (1988).
3. Reorganización compleja: se produce una reorganización jerárquica por reclasificación de MENTE y PENSAMIENTO. Estas nociones, antes consideradas desde una perspectiva dualista como distintas a lo físico, se reclasificaron como procesos computacionales.

Con respecto a las transformaciones mencionadas, Thagard sostiene que:

Desde un punto de vista puramente conceptual, el cognitivismo posee un carácter levemente revolucionario, dado que diversos conceptos fueron añadidos y

reorganizados, a pesar de que el cognitivismo involucre una baja eliminación de conceptos de sentido común y conductistas (1992, p. 238).

El apelativo de *carácter levemente revolucionario* probablemente se encuentre dado por la comparación entre la radicalidad del cambio entre el conductismo y el cognitivismo y la radicalidad del cambio entre el introspeccionismo y el conductismo. Como mencioné más arriba, el esquema conductista tendió a la eliminación de toda una batería de conceptos. El cognitivismo, en cambio se caracterizó por la incorporación de los conceptos anteriores (de sentido común y conductista), aunque reorganizándolos.

Dentro de la aproximación cognitivista, el conexionismo corresponde a uno de los marcos principales. Anteriormente mencioné algunas de sus características: se trata de modelos computacionales que postulan un conjunto de unidades de procesamiento similares a neuronas abstractas, con umbrales de activación e inhibición. En cuanto marco perteneciente a la aproximación cognitivista, el conexionismo mantiene la postulación de representaciones mentales y procesos computacionales, aunque el aparataje conceptual difiere del marco simbólico. Thagard menciona tres conceptos principales. El primero de ellos es APRENDIZAJE CONEXIONISTA, que apunta al resultado de la modificación de los patrones de conectividad en función de la experiencia. Ello contrasta con la noción de APRENDIZAJE en el enfoque simbólico, entendida como la adquisición, almacenamiento y recuperación de información en un dispositivo convencional. El segundo concepto es REPRESENTACIÓN DISTRIBUIDA. En términos generales, las unidades individuales no representan conceptos completos, sino que estos se distribuyen en diversos nodos. En cambio, REPRESENTACIÓN en el enfoque simbólico se encuentra ligada a una estructura de datos o un símbolo que puede representar un concepto. Por último, el concepto de SATISFACCIÓN DE CONSTREÑIMIENTOS PARALELOS se vincula con la capacidad de los modelos de satisfacer múltiples constreñimientos simultáneos, a diferencia del procesamiento serial. Estas diferencias serán profundizadas en el tercer capítulo.

A pesar de las innovaciones conceptuales y otras novedades en términos de implementación, Thagard aclara que el conexionismo no constituye una revolución. La naturaleza de los marcos permite que no se tenga que elegir uno en desmedro de otro, sino que “somos libres de construir una síntesis de conceptos valiosos a partir de ellos” (1992, p. 242), lo que no sucede cuando se trata de teorías en competencia. Además, el estado de desarrollo aún incipiente que presentaba el conexionismo a la fecha de publicación de su libro *Conceptual Revolutions* (1992) llevó al autor a afirmar que se

requieren mayores avances para hacer frente al éxito del programa simbólico dentro de la ciencia cognitiva.

En suma, el análisis de Thagard arroja que el cognitivismo presenta un carácter revolucionario débil caracterizado por el rechazo de la explicación conductista a todo tipo de comportamiento en términos de estímulo y respuesta, así como la postulación de que el pensamiento es procesamiento de información —es decir, es computacional y representacional—. Los CC hallados son la adición de conceptos y algunas reorganizaciones.

Un panorama ligeramente diferente es descrito en un texto publicado casi veinte años después (Thagard, 2011), donde el autor describe el reciente desarrollo de una *revolución del cerebro* en ciencia cognitiva. La línea argumental puede ser reconstruida como sigue:

- a. Las revoluciones conceptuales en ciencias naturales introducen nuevos conceptos, nuevas organizaciones conceptuales, nuevas hipótesis, nuevos esquemas explicativos y suelen depender del desarrollo de tecnología.
- b. Las revoluciones conceptuales en ciencias naturales involucran un reemplazo dramático de las teorías anteriores.
- c. La revolución del cerebro ha introducido nuevos conceptos y organizaciones conceptuales en términos neurales; nuevas hipótesis, como la identificación de emociones con procesos cerebrales, y nuevos esquemas explicativos neurales y neuroquímicos. Además, dependió en gran medida del desarrollo de técnicas de neuroimagen.
- d. La revolución del cerebro no involucra un reemplazo dramático de las teorías anteriores, sino una modificación lenta en curso.
- e. Por lo tanto, la revolución del cerebro corresponde a una revolución conceptual en curso.

La revolución del cerebro se caracteriza por la aplicación de explicaciones neurales en las áreas de la psicología a raíz del desarrollo de tecnologías tales como la imagen por resonancia magnética funcional para observar la actividad cerebral durante tareas cognitivas. Además, el progreso de los modelos conexionistas influyó en la conformación de aproximaciones neuralmente inspiradas, en comparación con las clásicas estructuras de datos y algoritmos del enfoque simbólico. Por ende, se transita desde las explicaciones en términos de procesamiento de información a las explicaciones en términos de

estructuras y procesos neurales: las neuronas se organizan en poblaciones por sus conexiones sinápticas. Algunos CC identificados por el autor son:

1. Creación o adición: introducción de conceptos nuevos tales como PATRÓN DE ACTIVACIÓN NEURAL, dada la integración de explicaciones neurales.
2. Reorganización compleja: ocurrió una reclasificación de los estados mentales como procesos cerebrales. Un ejemplo de ello es la reclasificación de CONCEPTO. Esta noción era considerada por los conductistas como una ficción dada su inobservabilidad, y posteriormente se entendió como estructuras de datos en programas computacionales en el marco del cognitivismo. En el esquema explicativo de la neurociencia, en cambio, corresponde a un patrón de activación cerebral que ocurre en el tiempo. Este patrón tiene la ventaja explicativa de implementar varias propuestas sobre la naturaleza de los conceptos planteadas desde la filosofía y la psicología, entre las que se cuentan: la noción clásica, según la cual los conceptos tienen una estructura definida en términos de condiciones necesarias y suficientes; la noción prototípica, según la cual su estructura es probabilística y depende de los rasgos típicos de una clase determinada, y la noción de ejemplar, donde los conceptos constituyen el almacenamiento de representaciones de ejemplares de la clase. La implementación dependería del entrenamiento de la red y el comportamiento de las poblaciones neuronales y sus conexiones. A modo de ejemplificación, el entrenamiento con múltiples ejemplos permite la formación de conexiones que almacenan sus rasgos y dan paso a la implementación de ejemplares. Otra reorganización se encuentra en EMOCIÓN en términos de “patrones complejos de actividad neural que integran valoraciones cognitivas sobre la relevancia de la situación actual para los objetivos de un agente con las percepciones de los estados fisiológicos internos” (Thagard, 2011, pp. 69-70).

Al igual que ocurrió con el cognitivismo, la revolución del cerebro no implica una eliminación de la batería previa de conceptos. La evidencia que sostiene esta revolución es variada. Thagard (2011) coloca el ejemplo de recursos como la psicofarmacología: explicaciones neuroquímicas del efecto de las drogas en los estados psicológicos, que apoyan la consideración de que los estados mentales son procesos neurales (químicos y eléctricos) en lugar de solo computacionales. No obstante, el autor no se explaya en la evidencia principal de la disponibilidad de explicaciones neurales de fenómenos cognitivos, a saber, los modelos propuestos desde la neurociencia para explicar los

correlatos neuronales de la conducta. El problema con esta omisión radica en que, en neurociencias, existen disputas teóricas centrales al respecto. Cada modelo planteado descansa en la descripción de circuitos y cambios de actividad, pero difieren en el correlato neuronal que atribuyen a los estados, conductas o percepciones particulares. Por ejemplo, Maldonado (2007) distingue entre tres modelos principales que han guiado las investigaciones. El primero de ellos se denomina *modelo de convergencia jerárquica*, y se relaciona con la llamada *doctrina de la neurona*. De acuerdo con él, el correlato neuronal de una conducta o estado cognitivo se encuentra en un cambio en la tasa de descarga neuronal en un lugar específico del sistema nervioso. El segundo modelo se denomina *representación distribuida* o *código de poblaciones*, y parece ser el que Thagard considera al afirmar la influencia del conexionismo. Según este, el correlato neuronal se halla en un patrón distribuido de actividad neuronal en distintas estructuras del sistema nervioso. Por último, el *modelo de correlación temporal* o *poblaciones sincrónicas* sostiene que el correlato neuronal corresponde a ensambles neuronales que están distribuidos (tal como señala el segundo modelo), pero que se definen por una sincronía y, usualmente, por una actividad oscilatoria. La diferencia entre estos tres modelos permite observar que la afirmación de la existencia de un CC inducido por los avances en neurociencia es inexacta, en cuanto pasa por alto la complejidad del panorama de las neurociencias. De todos modos, es cierto que se observa un fenómeno similar a la denominada *revolución del cerebro* en la tendencia a la integración de la neurociencia en la ciencia cognitiva.

En suma, la revolución del cerebro se caracteriza conceptualmente por la adición de conceptos y reorganizaciones conceptuales complejas. No obstante, la realización de una comparación con la revolución cognitiva permite observar que el autor varía los criterios de consideración. Como puede notarse en la reconstrucción general del argumento presentado en *Conceptual Revolutions* (1992), se apoya fuertemente en una comparación entre la psicología cognitiva y las ciencias naturales. En sus palabras,

La psicología aún no ha alcanzado una teoría unificada comparable con la mecánica de Newton, la teoría del oxígeno de Lavoisier o la teoría de la evolución por selección natural de Darwin. A pesar de los innegables avances en la comprensión del pensamiento humano que se han producido en este siglo, todavía nos falta una teoría unificada que explique un rango amplio de fenómenos. La consideración de este hecho es crucial para la comprensión del desarrollo conceptual de la psicología moderna (...) (1992, p. 225).

La consideración del desarrollo conceptual en ciencias tales como la física o la química como modelo de revolución conceptual conduce al autor a sostener que las innovaciones en psicología son débiles. Esta comparación puede resultar contraproducente para evaluar la profundidad de los CC ocurridos en otras ciencias, por las notorias diferencias entre disciplinas. Tanto los objetos de estudio como las metodologías para dar cuenta de ellos difieren, por lo que podría especularse que la naturaleza débil o revolucionaria de un CC también varía en dependencia del área en cuestión. La recurrencia a esta comparación se mantiene en Thagard (2011), pero su repercusión cambia en la afirmación del carácter revolucionario del cognitivismo y de la revolución del cerebro. Ambas revoluciones se diferencian de los desarrollos en ciencias maduras en sentido kuhniano en la medida en que no constituyen un rechazo completo de las ideas previas. Sin embargo, con respecto a la revolución cognitiva, el autor no duda en sostener que “los avances en psicología en la década de 1950 y 1960 son revolucionarios” (2011, p. 67). En particular, ese carácter viene dado por el rechazo de que todo el comportamiento puede explicarse en términos conductistas de estímulo-respuesta, y el empleo de la metáfora del computador para sostener la existencia de representaciones mentales, procesos computacionales y una reclasificación del pensamiento como procesamiento de información. Por otra parte, la *revolución del cerebro* presenta una transformación importante en las explicaciones de las capacidades cognitivas. Al respecto, según el autor,

A diferencia de las revoluciones de Copérnico y Pasteur, la Revolución del Cerebro no es un reemplazo drástico de las teorías científicas viejas por otras nuevas, sino más bien una modificación lenta de hipótesis y explicaciones psicológicas a medida que surgen nuevas ideas teóricas en neurociencias (2011, p. 76).

Por tanto, a pesar de mantenerse la comparación, esta no desemboca en la consideración de las transformaciones conceptuales en psicología como débiles. En conclusión, los dos análisis que ofrece Thagard corresponden a dos ejemplos de CC en una ciencia inmadura en sentido kuhniano.

3.2. Akagi sobre reconceptualizaciones en teorías alternativas

Un segundo foco de análisis de CC en ciencia cognitiva se concentra en el proceso de reconceptualización ocurrido a raíz de las denominadas *teorías alternativas* a los enfoques simbólico y conexionista. Este apelativo engloba un conjunto de

aproximaciones que critican alguno de los supuestos sustantivos de la disciplina (computacional y representacional). Volviendo al análisis que entrega von Eckardt (1995), la estrategia explicativa de la ciencia cognitiva consiste en la reconceptualización de las capacidades cognitivas en términos de capacidades de procesamiento de información. En otras palabras, los procesos cognitivos consisten en computaciones sobre representaciones mentales —que, como abordaré en el tercer capítulo, son consideradas como amodales (véase pp. 125-126).

La mayoría de las posturas que han problematizado estos supuestos son actualmente agrupadas bajo el nombre de Cognición 4EA⁴⁹: cognición encarnada (corporizada), embebida (situada), enactiva, extendida (distribuida) y afectiva. De acuerdo con este conjunto, en primer lugar, la cognición se relaciona con una interconexión entre el cuerpo —principalmente, la experiencia sensoriomotora del organismo— y el ambiente. Es decir, incluye estructuras y procesos corporales que no son neurales. En segundo lugar, siempre se encuentra situada en un entorno (material y sociocultural) determinado que ofrece ciertas posibilidades efectivas (*affordances*) para su interacción con él. Su diseño, por tanto, está orientado para funcionar de manera apropiada en ese ambiente. En tercer lugar, la cognición se halla en un cuerpo que es biológico, un sistema vivo que posee determinadas características. Una de estas características es la acción y autorregulación del sistema para utilizar efectivamente los recursos del entorno. En cuarto lugar, se encuentra también distribuida entre artefactos y otros individuos que permiten extender los recursos cognitivos. Por último, los procesos cognitivos se integran con los procesos emocionales. La valencia afectiva es central para la operación en el ambiente y para la toma de decisiones: la información del medio externo se conjuga con la información del medio interno, vinculada con el procesamiento afectivo y emocional, lo que implica que los estímulos externos dejan de ser neutros para el individuo. En términos muy generales, algunas obras fundacionales incluyen la psicología ecológica propuesta en *The Ecological Approach to Visual Perception* (1979), de Gibson, el enactivismo de Varela, Thompson y Rosch en *The Embodied Mind* (1991), la robótica situada de Brooks en *Intelligence without Representation* (1991, recogido en

⁴⁹ El surgimiento del nombre *Cognición 4E* se sitúa en un *workshop* sobre cognición corporizada realizado en la Universidad de Cardiff en el año 2006, donde participaron autores tales como Shaun Gallagher, Richard Menary, Matthew Radcliffe, Mark Rowlands, entre otros (Newen, de Bruin y Gallagher, 2018, p. 4). El término se empleó posteriormente por Richard Menary en su edición del número especial de la revista *Phenomenology and the Cognitive Sciences* (2010). La *A* de *afectiva* fue añadida después.

1999), la cognición extendida de Clark y Chalmers en *The Extended Mind* (1998), la cognición distribuida de Hutchins en *Cognition in the Wild* (1995), el enfoque dinámico propuesto por van Gelder en *What Might Cognition Be, if Not Computation* (1995), y la cognición corporizada radical de Chemero en *Radical Embodied Cognitive Science* (2009). A pesar de las variaciones entre los autores mencionados, comparten una crítica a los supuestos sustantivos computacional o representacional, así como la discusión de los límites de la cognición.

En este marco, Akagi (2017) sostiene que COGNICIÓN está en un proceso de transformación progresiva. A continuación, realizaré una reconstrucción argumental del planteamiento del autor:

- a. El *problema de la cognición* (es decir, la carencia de una definición apropiada de esta) es central en ciencia cognitiva.
- b. El *problema de la cognición* está accidentalmente ligado con la resolución de las *guerras de la frontera de la cognición*, disputas teóricas y metodológicas llevadas a cabo por las posturas no clásicas y que muestran el *hecho del desacuerdo experto generalizado*.
- c. Las *guerras de la frontera de la cognición* han llevado a adoptar la búsqueda de la *marca de lo cognitivo* (criterio dicotómico de demarcación de los fenómenos cognitivos que excluye muchas fronteras anti-clásicas) como estrategia frente al *problema de la cognición*.
- d. Sin embargo, el *problema de la cognición* también es motivado por el CC en COGNICIÓN, que constituye un progreso.
- e. Por tanto, el *problema de la cognición* es mejor abordado por la adecuación extensional ecuménica, a saber, la visualización del CC de COGNICIÓN por medio de la descripción de los desacuerdos entre científicos cognitivos sobre qué es la cognición, lo que permite caracterizar su progreso a pesar de no resolver las preguntas sobre sus límites.

El problema de la carencia de una definición establecida y ampliamente aceptada de la cognición es central para la disciplina, puesto que se trata justamente de su objeto de estudio y su criterio principal de demarcación como ciencia. El interés en dicho problema se asoció fuertemente con el debate llevado a cabo por las posturas no clásicas acerca de los límites de la cognición. Las disputas vistas entre aproximaciones apuntan a un hecho principal de la ciencia cognitiva: el hecho del desacuerdo experto generalizado. En otras palabras, no ha sido posible llegar a un consenso sobre qué es la cognición. La

estrategia para abordar el problema se concentró en gran medida en la búsqueda de la *marca de lo cognitivo*, esto es, un criterio dicotómico de demarcación impulsado por autores como Fred Adams y Ken Aizawa (2001, 2008), que pretende distinguir entre los fenómenos que son cognitivos y los que no lo son.

Sin embargo, la búsqueda de resolución del problema de la cognición por medio de la determinación de sus límites es un accidente histórico (Akagi, 2017, p. 3567). Con *accidente* probablemente el autor se refiera a *contingencia*: el problema ha sido abordado así no por el hecho de que sea la vía única o más relevante, sino porque históricamente fue el caso. No obstante, existen otros fenómenos de interés por los cuales resolver el problema más allá de las guerras de la frontera de la cognición. Uno de ellos es el proceso de CC en el que se encuentra esta noción. Considerando los tipos de CC estudiados en el primer capítulo (p. 59), esta discontinuidad conceptual corresponde a:

1. Reorganización básica: se ha realizado una diferenciación entre dos sentidos de COGNICIÓN: cognición superior o *highfalutin* y cognición inclusiva. La primera de ellas es una noción restrictiva entendida como sinónimo de pensamiento racional. La segunda es una conceptualización reciente, en la que se considera la percepción, el control motor y el afecto como fenómenos cognitivos. Sumado a ello, se han incluido otros fenómenos naturales dentro de la noción de cognición inclusiva, tales como artefactos u otros individuos, como parte de las discusiones llevadas a cabo por los enfoques no clásicos. Al respecto, las disputas continúan abiertas.

Akagi (2017) sostiene la existencia de CC en COGNICIÓN por medio de dos evidencias: (i) la conceptualización de la percepción, las emociones, entre otras, como fenómenos cognitivos, y (ii) las variaciones del concepto producto de las disputas entre el enfoque clásico y las aproximaciones no clásicas acerca de los límites de la cognición. Las etapas que describen el cambio son las siguientes. En los inicios de la revolución cognitiva, la concepción de los procesos cognitivos de nivel personal —en el sentido de Dennett (1986)— era cercana a la psicología del sentido común (esto es, la adscripción de creencias y deseos, empleada por las personas como medio de explicación y predicción de la conducta de los congéneres). Desde la década de 1960 y sobre todo en la década de 1980, los resultados empíricos favorecieron una subdivisión de las clases de la psicología del sentido común en clases cognitivas más específicas. Por ejemplo, la memoria se dividió en distintos sistemas: en relación con la temporalidad, Atkinson y Shiffrin distinguieron entre memoria de corto y de largo plazo; en relación con el contenido,

Tulving diferenció entre memoria episódica y semántica, y Squire distinguió entre memoria declarativa y no declarativa. En la actualidad se considera un modelo que incluye memoria semántica, episódica, de trabajo, procedural y de representación sensorial (Schacter y Tulving, 2005), cada una de las cuales cuenta con procesos y circuitos separados. Además, investigaciones sobre sesgos cognitivos en la toma de decisiones y la sensibilidad al contexto de diversos procesos subpersonales ocasionaron una reducción de la sinonimia entre cognición y pensamiento racional. La pérdida de distinción entre la cognición superior y otros fenómenos cognitivos conllevaron el empleo de la noción inclusiva y motivaron las guerras de la frontera de la cognición. Estas disputas se encuentran orientadas a discutir si el concepto inclusivo de COGNICIÓN incluye otros fenómenos naturales, como partes del cuerpo distintas al cerebro (cognición corporizada), la relación entre objetos, individuos y artefactos (cognición distribuida), entre otras. Con ello, intentan influir en el CC que tiene lugar en la ciencia cognitiva. Estos cambios son descritos como un progreso, en cuanto “están motivados en gran parte por los esfuerzos para responder a la evidencia, o más generalmente para describir bien el mundo a la luz de la experiencia recalcitrante” (Akagi, 2017, p. 3554).

La búsqueda de la *marca de lo cognitivo* es congruente con la adopción de una estrategia sectaria de la explicación conceptual, esto es, la toma de postura en la guerra de las fronteras de la cognición y la subsecuente explicación de COGNICIÓN a través de esa postura. En otras palabras, la búsqueda de resolución del problema de la cognición por medio de la determinación de sus límites condujo a la búsqueda de un criterio de demarcación junto con la adopción de una postura determinada en el debate, con lo cual COGNICIÓN es descrita de una forma sesgada y controversial. Akagi (2017) sostiene que esta aproximación no es la más adecuada. La explicación de COGNICIÓN posiblemente no tome la forma de condiciones necesarias y suficientes para delimitar esta noción. En su lugar, “el hecho del desacuerdo experto generalizado invita al tratamiento en términos de estructuras conceptuales más complejas que vienen en grados, o tienen subtipos, o algo similar” (2017, p. 3563). Por otro lado, adoptar la estrategia de tomar partido en la disputa aún en curso recae en la elección de marcas de lo cognitivo que son especulativas y que ocultan el proceso de CC que está en desarrollo. Dado lo anterior, el autor propone la adecuación extensional ecuménica, un estándar que expresa el estado de acuerdo y de desacuerdo entre los científicos cognitivos con respecto a qué se entiende por COGNICIÓN. Este estándar de adecuación explicitaría los fenómenos que son generalmente

considerados como cognitivos, los que son considerados como no cognitivos y los casos generalmente considerados como problemáticos. Según Akagi,

Plausiblemente, una explicación ecuménicamente adecuada de la cognición consistirá en algún criterio (...) que articule los rasgos de la cognición y marque explícitamente algunas características como compartidas entre las concepciones de la cognición, y marque otras como objetos de disputa en curso (2017, p. 3561).

Por ello, este criterio produce una extensión no clásica, esto es, una extensión que incluye grados de acuerdo sobre la pertenencia de los *ítems* en la extensión del concepto. Akagi concluye que su estrategia “hace explícito el cambio conceptual, incluso a expensas de respuestas claras a las preguntas abiertas” (p. 3567).

En suma, el análisis llevado a cabo por Akagi (2017) identifica un CC en la noción de COGNICIÓN que se encuentra ligado a las discusiones entre el enfoque clásico computacional-representacional y las aproximaciones alternativas. Este CC es principalmente una reorganización básica. Ello implicaría que estas discusiones constituyen una fuente de estudio para la determinación de otras posibles discontinuidades conceptuales.

3.3. Keijzer sobre la eliminación de REPRESENTACIÓN

El tercer y último estudio que me interesa analizar se encuentra igualmente enmarcado en el proceso de reconceptualización ocurrido a raíz de los enfoques no clásicos. Se centra en las discusiones entre el enfoque clásico y el enfoque dinámico y corporizado de la cognición (abreviado como DEC). De acuerdo con este enfoque, la cognición es eminentemente corporizada: el cuerpo biológico (sobre todo, la experiencia sensoriomotora) y el entorno poseen interconexiones importantes que la constituyen. Además, se describe más apropiadamente por medio de sistemas dinámicos: sistemas de ecuaciones diferenciales cuyas variables muestran distintos estados en función del tiempo (van Gelder, 1995). En este marco, Keijzer (2002) realiza una discusión sobre el CC en REPRESENTACIÓN. Su estudio cuenta con la siguiente línea argumental:

- a. En el enfoque clásico en ciencia cognitiva, las representaciones a nivel subpersonal son entendidas como un tipo de modelación interna que produce las regularidades conductuales.
- b. En DEC, las entidades interpretables como representacionales son los parámetros de control interno (ICP).

- c. Los ICP no son el equivalente biológico de las representaciones en un contexto cognitivo.
- d. El caso biológico muestra que es plausible una interpretación no representacional.
- e. Por tanto, DEC involucra un CC en REPRESENTACIÓN en ciencia cognitiva.

En el enfoque clásico, las representaciones han jugado un rol central. La concepción según la cual las personas re-presentan los objetos que no se encuentran inmediatamente presentes es cercana a las explicaciones de la psicología del sentido común con respecto al comportamiento, por lo que brindan familiaridad. Sin embargo, este nivel explicativo—nivel personal, en la terminología de Dennett (1986)— debe ser distinguido del nivel subpersonal, que explica los mecanismos cerebrales que subyacen a las características del nivel personal. La crítica de Keijzer comienza afirmando que el enfoque clásico utiliza las explicaciones de nivel personal como guía para las explicaciones de nivel subpersonal. Este paso es denominado por el autor como Teoría del Agente (*Agent Theory* o AT), y consiste en sostener la hipótesis de que el cerebro opera de manera similar al agente tomado en su conjunto. En otras palabras, el cerebro se interpreta en términos de la psicología del sentido común. La AT oscurece el foco de las ciencias y neurociencias cognitivas, cuyo objeto de estudio comprende “los principios que nos permiten ser humanos” (2002, p. 277) en lugar de la psicología popular. El punto principal de este oscurecimiento radica en el empleo de la noción tradicional de representación a nivel subpersonal como descripción del funcionamiento cerebral. Cabe enfatizar que se refiere a la noción tradicional porque, como abordaré en el tercer capítulo (p. 108), la neurociencia utiliza el concepto de representación neural para indicar patrones de actividad neural que representan características ambientales.

Estas representaciones subpersonales poseen dos características. En primer lugar, se entienden como un tipo de modelación interna que es isomórfica con respecto a los procesos externos relevantes. Los modelos mentales contienen información importante para explicar, predecir, y anticipar eventos. La noción de modelo mental se origina en el trabajo de Kenneth Craik (1943), quien sostiene que el rasgo central del pensamiento radica en que provee un modelo del mundo externo. Este modelo es conseguido por tres procesos esenciales: la traducción de procesos externos a símbolos (palabras, números, etcétera), la obtención de otros símbolos mediante procesos de razonamiento o inferencia, y la re-traducción de ellos a procesos externos o el reconocimiento de su correspondencia con eventos externos. El modelo interno permite explicar la generación del comportamiento orientado a metas, en cuanto “actúa como un programa que instruye al

sistema motor de modo tal que las metas específicas y distales sean conseguidas” (Keijzer, 2002, p. 279). La concepción de las representaciones como modelación mental es empleada en la ciencia cognitiva en conjunción con el supuesto de la pobreza de estímulos: la insuficiencia y ambigüedad de la información sensorial (véase p. 102). Ello justifica que las representaciones dependan del modelo interno de la situación externa y no de las circunstancias externas mismas. La segunda característica de las representaciones subpersonales propias de la ciencia cognitiva clásica corresponde a su consideración como la fuente productora de regularidades conductuales. El sistema cognitivo interno conduce las tareas de planificación que son comunicadas al sistema motor para que este ejecute las acciones correspondientes. Sin el mando del sistema cognitivo interno, el sistema motor se mantiene pasivo.

Frente al enfoque clásico, las aproximaciones dinámicas y corporizadas de la cognición explican el comportamiento por medio de las interacciones dinámicas entre el cuerpo, el ambiente y el sistema neural. En primera instancia, las características dinámicas y de corporización fueron consideradas como suficientes para explicar dicho funcionamiento, por lo que no se requería recurrir a la noción de representación ni limitarse al sistema cognitivo interno. Un ejemplo de ello son los robots de Brooks, quien rechazó la representación ligada a un procesador central y a situaciones particulares. Sin embargo, este postulado fue problematizado principalmente por dos argumentos: (i) la apelación a una interpretación más amplia de REPRESENTACIÓN, dado que sus críticas se enfocan en una noción estrecha, y (ii) la afirmación de su rol en procesos cognitivos de nivel superior, puesto que las críticas se centran en procesos principalmente motores. Con ello, se consideró que los postulados anti-representacionistas eran demasiado radicales y, en su lugar, autores como Clark (1997a) y Bechtel (1998) concibieron el DEC como complementario al representacionalismo.

Contrario a lo anterior, Keijzer sostiene la inconveniencia de una postura que complemente el DEC con el representacionalismo⁵⁰:

Una visión DEC completa aún contendrá entidades similares a las representaciones.
Será posible interpretar esas entidades como representaciones. No obstante, estas

⁵⁰ Para Keijzer (2002), el interés por la adopción de una postura híbrida se encuentra en la necesidad de explicar conductas tales como el comportamiento anticipativo, en el que el acoplamiento presente entre el organismo y el entorno inmediato parece no ser suficiente. La integración de las representaciones permitiría explicar cómo el comportamiento se orienta a metas a largo plazo, para las que la interacción presente no provee suficiente instrucción. Sin embargo, el autor rechaza esta utilidad. Bajo su criterio, la propuesta híbrida mantendría el problema del origen en el nivel personal de explicación, y no explicaría en qué medida las instrucciones representacionales motoras transfieren el control al acoplamiento dinámico.

entidades similares a las representaciones tienen su origen en un contexto conceptual muy diferente y desempeñan un papel distinto en comparación con el uso ordinario de las representaciones en ciencia cognitiva. Los cambios son lo suficientemente importantes como para justificar la afirmación de que estas entidades no son representaciones, sino entidades que requieren una terminología diferente (2002, p. 277).

La visión DEC a la que alude el autor tiene como concepto fundamental el de autoorganización, esto es, la capacidad de coordinar y establecer patrones entre un conjunto de elementos inicialmente desordenados. Existen dos tipos de parámetros relevantes para el proceso. Por un lado, el parámetro de orden representa la organización macroscópica, que es producida por las interacciones entre partes que operan a menor escala y cuyo comportamiento es constreñido por dicho parámetro. Por otro lado, los parámetros de control son aquellos que determinan el estado del sistema, lo que a su vez es responsable del surgimiento y forma del parámetro de orden. En otras palabras, por medio de la manipulación de los parámetros de control, el sistema en su conjunto puede variar su orden. Los sistemas biológicos son capaces de manipular sus propios parámetros de control para alcanzar estados macroscópicos que son descritos como “parámetros de orden de baja dimensión” (Keijzer, 2002, p. 282).

A diferencia de los sistemas físicos, los sistemas biológicos cuentan con un factor que dirige los procesos de formación de patrones en direcciones específicas. Este factor es presumiblemente el ADN. Bajo esa interpretación, los genes corresponderían a un conjunto de parámetros de control que permiten modular procesos de autoorganización. Serían, de acuerdo con Keijzer, parámetros de control interno (*internal control parameter* o ICP), que pertenecen a un nivel microscópico y actúan iniciando o deteniendo procesos biológicos específicos de autoorganización, patrones macroscópicos particulares. Los ICP son a su vez dirigidos por otros factores, que pueden ser controlados por el sistema. Así, se conforma una red reguladora genética que “consiste en influencias ampliamente recíprocas y dispersas entre genes, proteínas y procesos metabólicos” (2002, p. 282). Esta red es también interpretada como procesos de autoorganización. En suma, el nivel microscópico de los ICP se autoorganiza en un patrón macroscópico que es descrito por un parámetro de orden.

Los ICP son anteriores e independientes de los patrones macroscópicos y de los procesos microscópicos. Ello se debe a que, dada su función como parámetros de control, su influencia debe variar para poder iniciar y detener procesos de autoorganización. El ADN, entendido como ICP, influencia en los procesos biomoleculares inmediatos al

codificar ciertas proteínas, pero no depende de dichos procesos, en cuanto sus codificaciones específicas se deben a la selección evolutiva. Los genes forman parte de una red de factores regulatorios y no parecen controlar las estructuras macroscópicas de manera unidireccional.

Para el autor, el caso biológico anteriormente planteado muestra que:

es posible imaginar un mecanismo de dirección a largo plazo para la cognición que no se base en la noción de representaciones subpersonales. Los estados internos —tales como los parámetros biológicos de control interno— podrían guiar los acoplamientos a corto plazo entre organismo y entorno en escalas de tiempo más largas hacia direcciones específicas. Estos estados internos desempeñarían una función que es equivalente a la que suelen realizar las representaciones en aproximaciones cognitivas comunes. Sin embargo, al mismo tiempo, estos estados internos son tan diferentes de la noción tradicional de representación de la ciencia cognitiva que es perfectamente razonable sostener que no son representaciones (2002, p. 283).

Para Keijzer, los ICP no deben ser entendidos de manera equivalente a las representaciones en el contexto cognitivo, si bien tienen similitudes funcionales. Las representaciones clásicas basadas en modelo derivan su representacionalidad del isomorfismo con ciertos aspectos del mundo externo, que las hace útiles como guía de la conducta de un sistema cognitivo. El usuario selecciona las representaciones y es un intermediario entre estas y lo representado, por lo que las representaciones tienen una existencia anterior. Son el componente principal de la explicación para el comportamiento inteligente, mientras que el resto (como el sistema motor) es visto como periférico. Los ICP, por su parte, no poseen un estatus representacional independiente, sino que mantienen un rol secundario como parámetros reguladores de procesos de autoorganización. Modulan la capacidad de autoorganización sin ser la fuente del orden. Enmarcándose en la periferia desde el punto de vista de las representaciones clásicas, son los recursos que posibilitan el comportamiento inteligente para el DEC. De acuerdo con Keijzer, esta diferencia es suficiente para sostener que “incluso cuando se considera útil insertar estados internos como ICP, esta inserción no implica de ninguna manera un compromiso directo con la representación interna” (2002, p. 284). La introducción de estados internos como los ICP configura la aproximación como no representacional.

Dado lo anterior, Keijzer sostiene una diferencia conceptual entre ICP y representación —entendiendo por REPRESENTACIÓN las representaciones subpersonales y basadas en modelo—, así como la plausibilidad de una interpretación no representacional. No es necesario incluir representaciones en los problemas que parecían

requerirlas —como el comportamiento anticipativo u orientado a metas, que requiere más que la interacción inmediata entre agente y entorno—. El autor no niega la posibilidad de existencia de representaciones subpersonales en adición a los ICP o que los ICP tengan potencial representacional por razones distintas a su operación como patrones de control. Tampoco excluye la posibilidad de que se demuestre empíricamente que la realización de la función de los ICP requiere que estos posean atajos representacionales para el entorno distal. Sin embargo, realizar una distinción conceptual es relevante. En efecto,

nombrar a los ICP como una forma de representación destacaría su relevancia conductual a largo plazo en lugar de su funcionamiento como parámetro de control inmediato dentro de los desordenados detalles de autoorganización de los que surgen los patrones de comportamiento. Esto es exactamente el mensaje equivocado desde la perspectiva del DEC (2002, p. 287).

Dicho de otro modo, igualar ICP y representación constituye un error. Por conclusión, la consideración de la diferencia entre ambos permite sostener que el DEC involucra un CC con respecto al enfoque clásico. Este CC correspondería a:

1. Eliminación de REPRESENTACIÓN: Dado que las entidades interpretables como representaciones en este enfoque poseen un rol distinto a las entidades clásicamente nombradas como representaciones, se necesitaría el empleo de una terminología diferente para ellas.

En suma, Keijzer apela al contexto conceptual en el que surgen los ICP y el rol que cumple allí para sostener su diferencia con las representaciones tradicionales —la importancia del contexto conceptual ya fue abordada en el primer capítulo a partir de la noción de ecología conceptual de Posner y colaboradores (1982, 1992)—. El análisis del CC se circunscribe en los debates entre enfoques clásicos y alternativos, por lo que sería una muestra del hecho del desacuerdo experto generalizado, en los términos de Akagi (2017). Su resultado es especialmente relevante para los objetivos de la presente tesis, puesto que muestra un antecedente de transformación en REPRESENTACIÓN. La fuerza de la conclusión anti-representacional que Keijzer (2002) sostiene con respecto al DEC depende en buena medida de que las características atribuidas a la noción tradicional de REPRESENTACIÓN —modelación mental que produce regularidades conductuales— sean correctas, y que los ICP sean las entidades interpretables como representacionales en el DEC. En otras palabras, depende de las dos primeras premisas en la reconstrucción argumental expuesta anteriormente (pp. 86-87). De ser verdaderas, el resultado del análisis del CC arrojaría una preferencia por la eliminación de este concepto. Sin

embargo, incluso bajo el supuesto de que ambas premisas son verdaderas, la conclusión mantiene abierta la posibilidad de interpretaciones representacionales de los ICP, y de que existan otro tipo de representaciones. Por tanto, el cambio que apunta a la eliminación de la noción de REPRESENTACIÓN parece una sugerencia de terminología en lugar de una discontinuidad conceptual constatable.

A modo de sumario, los trabajos de Thagard (1992, 2011), Akagi (2017) y Keijzer (2002) corresponden a antecedentes directos que responden afirmativamente a la posibilidad de CC en una ciencia inmadura como la ciencia cognitiva. A ellos subyace una serie de debates abiertos que pueden ser considerados como posible foco de CC para esta área. En el próximo apartado sistematizaré sus alcances y límites.

4. Discusión

En los apartados anteriores revisé algunos antecedentes de CC identificados en la ciencia cognitiva. Tomados en su conjunto, estos trabajos sugieren al menos tres momentos importantes de transformación de conceptos en ciencia cognitiva, a saber:

1. La revolución cognitiva.
2. La discusión de los límites de la cognición.
3. La revolución neurocognitiva.

Para Thagard (1992), la revolución cognitiva surgió de la competencia entre dos aproximaciones —el conductismo y el cognitismo—. El cognitismo explica el comportamiento inteligente por medio de la posesión de representaciones mentales y procesos computacionales que operan sobre ellas. A través de la metáfora del computador, muestra una mayor posibilidad de coherencia explicativa futura, es decir, posibilita eventualmente la explicación de una mayor cantidad de hechos que el conductismo. El conductismo, en cambio, se caracterizaba por la explicación del comportamiento en términos de estímulo y respuesta, lo que denotaba una simplicidad excesiva que limitaba su coherencia explicativa. El cognitismo integró parcialmente los conceptos conductistas y rechazó sus explicaciones limitadas. Por otro lado, el conexionismo es un marco dentro de la aproximación cognitivista que mantiene la postulación de representaciones mentales y procesos computacionales. A pesar de algunas innovaciones conceptuales y de implementación, este marco no constituye una revolución, sino que se encuentra en pleno desarrollo y no excluye la posibilidad de

síntesis con otros marcos. El CC llevado a cabo por la revolución cognitiva incluye la introducción de conceptos computacionales y algunas reorganizaciones conceptuales en nociones claves como INFORMACIÓN.

En segundo lugar, la discusión de los límites de la cognición llevada a cabo por las posturas no clásicas en ciencia cognitiva (Cognición 4EA) se posicionan críticamente frente a los supuestos sustantivos de la disciplina (computacional y representacional). Problematizan que el resto del cuerpo físico y el entorno queden relegados a ser fuentes de *inputs* que puedan explicarse por medio de la apelación a los estados internos. Para Akagi (2017), las discusiones sobre los límites de la cognición impulsan el CC en este concepto, como la reorganización ocurrida al diferenciar entre sentidos de COGNICIÓN. Además, este cambio es un progreso en el sentido de capacidad de respuesta a la evidencia. Por otro lado, según Keijzer (2002), la discusión específica entre el enfoque clásico y el enfoque dinámico y corporizado de la cognición apunta a la preferencia por la eliminación de REPRESENTACIÓN, puesto que las entidades que postula el DEC y que son interpretables como representaciones (es decir, los ICP) no son sus equivalentes biológicos en un contexto cognitivo. Por tanto, las disputas abiertas entre aproximaciones clásicas y no clásicas son un foco de CC para la ciencia cognitiva.

Por último, la revolución neurocognitiva reivindica la importancia de la neurociencia, que formaba parte del nivel de implementación según la división de Marr (1982). La revolución neurocognitiva estudiada por Thagard (2011) permite un tránsito desde las explicaciones en términos de procesamiento de información a las explicaciones que aluden a estructuras y procesos neurales. Por ende, se integra parte de la batería previa de conceptos, mientras que se producen CC tales como la adición de conceptos propios de la neurociencia y algunas reorganizaciones conceptuales.

Los estudios llevados a cabo por Thagard (1992, 2011), Akagi (2017) y Keijzer (2002) constituyen un marco teórico útil para el análisis del CC en REPRESENTACIÓN, labor que será realizada en el próximo capítulo.

5. Conclusiones

El presente capítulo tuvo un objetivo doble. En primer lugar, analicé las características propias de la ciencia cognitiva que son relevantes en el análisis del CC en ella, a saber: sus supuestos sustantivos computacional y representacional, su interdisciplinariedad y su inmadurez. En segundo lugar, revisé los antecedentes de estudios que individúan CC en

esta ciencia. Si bien las investigaciones revisadas (Thagard, 1992, 2011; Akagi, 2017; Keijzer, 2002) se concentran en transformaciones conceptuales distintas e identifican tipos de cambios diferentes, muestran un panorama congruente en relación con la posibilidad de CC en la ciencia cognitiva, y en los focos de transformación conceptual que pueden identificarse. Ello permite abordar, en el siguiente capítulo, algunas discontinuidades conceptuales significativas en REPRESENTACIÓN.

La conclusión más evidente que puede extraerse de los antecedentes presentados es la siguiente: en la medida en que los trabajos revisados sean correctos, es posible hallar CC en ciencias inmaduras en sentido kuhniano tales como la ciencia cognitiva. En el primer capítulo mencioné algunos aspectos preliminares que sugerían la posibilidad de CC sin cambio de teoría, a saber, la identificación de Carey (1985) de la emergencia de teorías como un segundo tipo de cambio teórico, y la crítica realizada por Thagard (1992) a la autora, que indica la posibilidad de reducción de uso de la teoría anterior sin un proceso de comparación y reemplazo con la teoría nueva. Los antecedentes presentados en este capítulo constituyen un reforzamiento directo a esta idea.

Sumado a lo anterior, la pregunta relevante a raíz del panorama histórico presentado en estos tres momentos de transformación de conceptos —revolución cognitiva, discusión de los límites de la cognición y revolución neurocognitiva— corresponde a si existe un CC en la noción de REPRESENTACIÓN involucrada en ellos. La respuesta, adelantada por el trabajo de Keijzer (2002), es positiva. Como abordaré en el tercer capítulo, el enfoque simbólico se encuentra ligado a las representaciones amodales, y el enfoque subsimbólico o conexionista desarrolla la noción de representación distribuida. Ambos son producto de la revolución cognitiva estudiada por Thagard (1992). Por otro lado, las discusiones sobre los límites de la cognición y el aporte de la revolución neurocognitiva permiten transitar hacia la incorporación de representaciones modales o multimodales, pero también hacia el anti-representacionalismo. Por supuesto, cabe destacar que esta clasificación de momentos de transformación dentro de la ciencia cognitiva es una reconstrucción para posibilitar el análisis. En el desarrollo histórico real del campo no existe una secuencialidad y limitación tan clara entre los hitos.

CAPÍTULO 3: ANÁLISIS DEL CAMBIO CONCEPTUAL EN EL CONCEPTO DE REPRESENTACIÓN

1. Introducción

Los dos capítulos anteriores han buscado transitar desde la generalidad del estudio del CC en la aproximación clásica hacia la particularidad de su análisis en ciencia cognitiva. Las ideas y teorías presentadas constituyen los antecedentes por medio de los cuales analizaré las transformaciones en el concepto de REPRESENTACIÓN en dicha ciencia. A modo de sumario, los argumentos principales que han sido presentados son los siguientes:

1. Un CC es el proceso de creación, alteración y comunicación de conceptos individuales, entendiendo por tales estructuras representacionales.
2. Los pasos para el análisis del CC incluyen: (i) caracterización de los sistemas conceptuales sucesivos, (ii) demostración de existencia de cambio y determinación del tipo de cambio, y (iii) mecanismos que posibilitan el cambio.
3. En ciencia cognitiva, el análisis del CC se encuentra constreñido por la consideración de los supuestos sustantivos (representacionalidad y computacionalidad), su interdisciplinariedad y su inmadurez.
4. En ciencia cognitiva es posible localizar al menos tres focos de transformación conceptual importantes, a saber: la revolución cognitiva, la discusión sobre los límites de la cognición y la revolución neurocognitiva.
5. Estos focos de transformación indican la posibilidad de la presentación de discontinuidades conceptuales sin un cambio de teoría en ciencias inmaduras.

El objetivo del presente capítulo corresponde a la realización del análisis del CC en el concepto de REPRESENTACIÓN en ciencia cognitiva. Esta noción es central para el área, y forma parte de uno de sus supuestos sustantivos. Como mencioné en el segundo capítulo, una representación ha sido entendida esquemáticamente como una entidad con propiedades semánticas que representa propiedades o estados de cosas. Un ejemplo de los intentos por proveer una caracterización operativa de esta noción se encuentra en von Eckardt (1995), quien recurre a la teoría de la representación de Peirce para extraer un concepto de representación en general, y luego analizar su aplicación en el caso de las representaciones mentales. Con base en ello, afirma que cuatro aspectos de la

representación son el portador, el objeto de representación, el fundamento o *grounding* y el interpretante (véase pp. 67-68). Así, sistematiza los aspectos centrales que han sido objeto de investigación por parte de filósofos y científicos cognitivos.

En el presente análisis sostendré, tras seguir los pasos propuestos para el abordaje de las discontinuidades conceptuales, que el proceso principal de cambio apunta a (i) la incorporación de un carácter distribucional y (ii) la incorporación del formato multimodal de las representaciones. Para ello, en el siguiente apartado caracterizaré los sistemas conceptuales propios de la exposición clásica de la Teoría Representacional de la Mente (TRM) de Jerry Fodor, el Tratamiento Adecuado del Conexionismo (TAC) de Paul Smolensky y el Sistema de Símbolos Perceptuales (SSP) de Lawrence Barsalou. En la caracterización, me detendré en los compromisos epistemológicos y los problemas implicados en las posturas que originan los debates, que son elementos importantes en la ecología conceptual. Para ello, me concentraré en tres puntos de las definiciones de REPRESENTACIÓN involucradas en tales aproximaciones teóricas:

- a. Relación entre representaciones conceptuales y representaciones perceptuales.
- b. Relación entre representaciones mentales y palabras del lenguaje natural.
- c. Relación entre representaciones mentales y representaciones neurales.

En el tercer apartado, abordaré la demostración de existencia de cambio y la determinación del tipo de cambio. Sostendré que las nociones de REPRESENTACIÓN que se desprenden apuntan a REPRESENTACIÓN AMODAL, REPRESENTACIÓN DISTRIBUIDA y REPRESENTACIÓN MULTIMODAL respectivamente. Los tipos de CC vistos en ellas corresponden principalmente a reorganizaciones básicas. Además, afirmaré que estos cambios no son plausiblemente entendidos como inconmensurables, puesto que no constituyen un cambio radical. En el cuarto apartado me referiré a las analogías como un mecanismo central que subyace a estas discontinuidades conceptuales. En el caso de la TRM destaca la metáfora del computador, en el conexionismo el funcionamiento neuronal y en el SSP la problematización de la analogía con los sistemas de grabación (*recording systems*). En la sección de discusión vincularé estos resultados con dos puntos de la ecología conceptual que subyacen a las distinciones conceptuales detectadas: la definición de cognición y la adopción o rechazo de la realizabilidad múltiple. A modo de conclusión, revisaré el análisis a la luz de los tres rasgos de la ciencia cognitiva que constriñen el estudio (sus supuestos sustantivos, su interdisciplinariedad y su inmadurez).

2. Caracterización de los sistemas conceptuales sucesivos

La centralidad de la noción de REPRESENTACIÓN en ciencia cognitiva conlleva una amplitud de uso en la literatura especializada en el área, junto con variaciones importantes en las características centrales propuestas. Por tanto, los CC presentados son diversos. En el presente análisis me concentraré en una línea de transformaciones conceptuales que puede ser considerada principal, a saber, el cambio localizado en los debates entre una versión del enfoque simbólico —específicamente, la Teoría Representacional de la Mente (TRM)— propuesta por Jerry Fodor, el modelo de procesamiento paralelo distribuido tal y como es expuesto por el Tratamiento Adecuado del Conexionismo (TAC) de Paul Smolensky, y el Sistema de Símbolos Perceptuales (SSP) de Lawrence Barsalou. Siguiendo el paso inicial de la metodología de análisis descrita en el primer capítulo, caracterizaré estos sistemas conceptuales a continuación.

2.1. Teoría Representacional de la Mente, de Jerry Fodor

Como mencioné en el segundo capítulo, la ciencia cognitiva clásicamente reconceptualiza las capacidades cognitivas en términos de capacidades de procesamiento de información, es decir, computacionales y representacionales. Uno de los exponentes principales del enfoque simbólico es Jerry Fodor (1975, 1983, 1987, 1997a), cuya exposición clásica⁵¹ de la TRM sostiene que el pensamiento es un proceso computacional cuyo medio es una serie de estructuras representacionales. En particular, su teoría involucra (i) la hipótesis del lenguaje del pensamiento, según la cual este sistema representacional adquiere la forma de un código interno que exhibe propiedades similares a los lenguajes naturales, y (ii) un compromiso ontológico con las actitudes proposicionales (estados mentales de sentido común, como creencias y deseos), entendidas como relaciones computacionales/funcionales entre un sujeto y una representación mental. A continuación, abordaré ambos aspectos de la teoría.

⁵¹ En obras posteriores, Jerry Fodor modificó algunos puntos de su exposición inicial sobre el lenguaje del pensamiento. Por ejemplo, en *LOT 2: The Language of Thought Revisited* (2008), el autor provee una actualización de las afirmaciones centrales de *The Language of Thought* (1975). La diferencia principal entre ambas obras se encuentra en temas vinculados con la tesis principal (como las críticas al pragmatismo o la defensa a una teoría atomista y referencialista de conceptos en *LOT 2*), pero la TRM se mantiene intacta. En la presente tesis abordé la exposición clásica, es decir, la presentada en el libro de 1975, puesto que mantiene su estatus fundacional y es el objeto de los debates principales a los que me referiré (sobre todo, los debates con Smolensky).

En primer lugar, la hipótesis del lenguaje del pensamiento se puede sistematizar en cuatro puntos:

- a. Los procesos cognitivos son procesos computacionales.
- b. Los procesos computacionales presuponen un sistema representacional que funciona como medio de las computaciones.
- c. El sistema representacional es entendido como un lenguaje del pensamiento o Mentalés.
- d. El lenguaje del pensamiento no es un lenguaje natural, pero proporciona el aparato representacional para aprender un lenguaje natural y realizar conductas inteligentes.

De acuerdo con Fodor (1975), las teorías de la cognición con mayor plausibilidad propuestas en psicología cognitiva conceptualizan los procesos cognitivos como procesos computacionales. Un proceso computacional es entendido como un proceso de manipulación de símbolos en virtud de su sintaxis y preservando su semántica. En particular, las computaciones son definidas preliminarmente como “aquellas relaciones causales entre símbolos que respetan de forma fiable las propiedades semánticas” (Fodor, 1997a, p. 832). La computación requiere un medio para computar, y este medio es provisto por el sistema representacional.

Las representaciones son símbolos (objetos físicos) que poseen tres características centrales. En primer lugar, tienen contenido o significado. Este contenido depende de las relaciones causales y nomológicas con las entidades con las que se vinculan. Para Fodor (1987), un concepto *C* significa la propiedad *P* porque una instanciación de *P* causa una instanciación de *C*. Por ejemplo, el concepto PERRO significa la propiedad *perro* porque la propiedad *perro* causa instanciaciones de PERRO⁵². De ello se desprende que las representaciones son evaluables semánticamente (pueden ser individualizadas por su contenido y examinadas de acuerdo con su relación con el mundo). En segundo lugar, poseen poderes causales, es decir, la capacidad de producir efectos. En tercer lugar, “muchas representaciones mentales tienen estructura constituyente (parte/todo), y muchos procesos mentales son sensibles a la estructura constituyente de las representaciones mentales a las que se aplican” (1997a, p. 832). Las representaciones poseen sintaxis y semántica combinatoria, lo que implica la existencia de fórmulas

⁵² Volveré a este punto en las conclusiones (p. 146), a propósito del fundamento o *grounding* de las representaciones.

complejas cuyo contenido se determina por el contenido de sus partes simples o subfórmulas. Las subfórmulas pueden aparecer en otras fórmulas (son transportables). Así, las representaciones pueden ser atómicas o moleculares, donde las representaciones moleculares se constituyen por otras representaciones que a su vez son atómicas o moleculares. La semántica de las representaciones moleculares depende de la semántica de sus partes sintácticas y de su estructura constituyente (Fodor y Pylyshyn, 1988, p. 12).

Los procesos son sensibles a dicha estructura, lo que permite explicar la coherencia semántica del pensamiento. Esta relación entre sintaxis y semántica es abordada en el enfoque simbólico por medio de la teoría de la prueba, una rama de la lógica que proporciona reglas de inferencia lógica (manipulación simbólica) que permiten explorar las relaciones semánticas de coherencia. Por ejemplo, la inferencia *Si Rover muere entonces algo muere*, formalizada como $Fa \rightarrow \exists x (Fx)$, es una generalización existencial que aplica por su forma sintáctica, pero que también preserva la semántica. Abordaré las características de la estructura de las representaciones, así como los argumentos en los que se sostiene su postulación, en el subapartado *Relación entre representaciones mentales y palabras del lenguaje natural* (pp. 104-108).

Las representaciones que poseen los organismos adquieren la forma de oraciones internas en un lenguaje del pensamiento donde se llevan a cabo las computaciones que subyacen al comportamiento. Este lenguaje interno no es un lenguaje natural⁵³, sino un sistema innato (es decir, no aprendido) que permite adquirir lenguajes naturales y realizar conductas inteligentes. Por una parte, aprender un lenguaje involucra un proceso de aprendizaje de las propiedades semánticas de sus predicados, lo que implica a su vez aprender generalizaciones que determinan sus extensiones⁵⁴. Para todo lo anterior se requiere poseer un lenguaje previo distinto al lenguaje que se está aprendiendo: un sistema capaz de representar las extensiones de los predicados de los lenguajes naturales que pueden ser aprendidos. El empleo racional de los recursos representacionales depende de variables tales como las exigencias de la tarea, la atención y la motivación. Por otra parte, los modelos relacionados con la toma de decisiones constituyen un ejemplo que muestra

⁵³ El lenguaje interno es privado, por lo que no puede ser un lenguaje natural. Además, la existencia de capacidad de pensamiento en organismos no verbales como animales o infantes preverbales constituye otro factor para trazar la diferencia con los lenguajes naturales.

⁵⁴ Las generalizaciones que determinan las extensiones de los predicados son reglas de verdad que estipulan las condiciones en las que el predicado es verdadero. Para ello, distinguen las representaciones que determinan las extensiones de los predicados elementales y señalan las reglas que definen los predicados complejos en términos de predicados elementares.

la relación del lenguaje del pensamiento con las conductas inteligentes. Las decisiones que toma un organismo frente a una situación determinada dependen del orden de preferencia que le asigna a las consecuencias probables estimadas para las opciones de comportamiento que cree poseer. Ello implica que tiene un medio para representar situaciones, comportamientos, consecuencias y probabilidades, a saber, el lenguaje del pensamiento.

El segundo punto relevante en la caracterización de la TRM corresponde al compromiso ontológico con las actitudes proposicionales, estados mentales de sentido común que tienen como objetos inmediatos a las representaciones (Fodor, 1987, pp. 16-17). Fodor sostiene que la TRM posee dos afirmaciones centrales relacionadas con este punto:

Afirmación 1 (la naturaleza de las actitudes proposicionales):

Para todo organismo *O*, y para toda actitud *A* hacia la proposición *P*, existe una relación *R* (computacional/funcional) y una representación mental *MP* tal que

MP significa que *P* y

O tiene *A* si y solo si *O* tiene *R* con *MP* (1987, p. 17).

Las actitudes proposicionales tienen un carácter relacional: son relaciones diádicas computacionales/funcionales entre un organismo y una representación mental. En una actitud proposicional como *Jerry cree que está lloviendo*, *cree que está lloviendo* es el estado mental, y el sujeto Jerry mantiene una relación computacional/ funcional con la representación mental *está lloviendo* que significa, a su vez, la proposición *está lloviendo*. Jerry posee esta actitud proposicional si y solo si tiene la relación de *creer que* con la representación mental *está lloviendo*. En otras palabras, por cada episodio de tener una actitud proposicional con un contenido determinado, existe un episodio de estar relacionado con un símbolo o representación con el mismo contenido.

El bicondicional presente en la *Afirmación 1* es luego mitigado por el autor. En efecto, es posible la existencia de actitudes proposicionales a las que no corresponda una representación mental, y representaciones mentales sin actitudes. Ello introduce una distinción entre casos centrales y casos derivados en la teoría. Los casos centrales se pueden determinar por medio de la *Afirmación 2*:

Afirmación 2 (la naturaleza de los procesos mentales):

Los procesos mentales son secuencias causales de ejemplificaciones de representaciones mentales (1987, p. 17).

Los estados mentales que están causalmente relacionados se presentan de manera sucesiva siguiendo principios computacionales que aplican formalmente a las representaciones. En la secuencia *Tengo prueba, estudiaré*, la ejemplificación *estudiaré* es causada por la ejemplificación *tengo prueba*, de modo que $EM1 \rightarrow EM2$. Los casos centrales para la teoría son aquellos donde los contenidos de las actitudes que figuran en los procesos mentales están representados explícitamente, lo que permite la causación intencional. Las actitudes proposicionales sin una representación mental correspondiente, así como las representaciones mentales sin actitudes asociadas, son casos derivados.

Una de las virtudes principales de la TRM radica en la sugerencia de un mecanismo que explique el paralelismo entre las relaciones causales entre estados mentales y las relaciones semánticas entre objetos proposicionales. Este mecanismo se funda en la conjunción de la postulación de representaciones mentales y la metáfora del computador (véase pp. 140-141). Dicha conjunción permite enlazar las propiedades semánticas con las relaciones causales por medio de las propiedades sintácticas de los símbolos.

En suma, las representaciones mentales son entendidas como símbolos con contenido, poderes causales y estructura. Ahora bien, ahondando en la noción de REPRESENTACIÓN fodoriana, es posible enfatizar tres rasgos: (i) la existencia de diferencias en el procesamiento de representaciones ligadas a los sistemas de *inputs* modulares y representaciones propias de los procesadores centrales (pp. 101-104); (ii) el vínculo entre las representaciones mentales y las palabras del lenguaje natural (pp. 104-108), y (iii) la relación entre las representaciones mentales y las denominadas *representaciones neurales* (pp. 108-110). En los próximos subapartados, abordaré cada uno de estos puntos.

Relación entre representaciones conceptuales y representaciones perceptuales

Fodor (1983, 1985) sostiene una diferencia en cuanto al procesamiento de tipos de representaciones (representaciones ligadas a los sistemas de *inputs* modulares y representaciones propias de los procesadores centrales) que ha sido entendida en términos de una distinción entre representaciones perceptuales y representaciones conceptuales (por ejemplo, Carey, 2009, pp. 9-10).

La modularidad fodoriana afirma la existencia de módulos cognitivos. Un módulo es entendido como “un sistema computacional encapsulado informacionalmente, un mecanismo de realización de inferencias cuyo acceso a la información *background* está

constreñido por características generales de la arquitectura cognitiva” (1985, p. 3). Sobre la base de ello, niega la continuidad entre percepción y cognición. Sostiene que la tesis de la continuidad entre ambas surge de la incapacidad de diferenciar entre las características inferenciales de los procesos de fijación de creencias en los procesadores centrales y los procesos encapsulados de los módulos perceptuales. Por un lado, los procesos perceptuales poseen una naturaleza inferencial. Ello se sostiene desde el argumento de la pobreza del estímulo, según el cual la integración perceptual debe incluir información aportada por el organismo, porque “por lo general, existe más información en una respuesta perceptual que en el estímulo proximal que provoca la respuesta” (1985, p. 2). Por tanto, no se puede reducir a una mera respuesta refleja. Por otro lado, los procesos perceptuales son encapsulados, es decir, son insensibles al sistema de creencias del organismo. Un ejemplo clásico radica en la persistencia en la observación de ilusiones ópticas tales como la diferencia en longitud de líneas en la figura de Mueller-Lyer, a pesar de aprender que ambas son iguales. La naturaleza encapsulada del procesamiento perceptual apunta a la tesis de la impenetrabilidad cognitiva de la percepción, la que sostiene que los módulos no son penetrables (no pueden ser alterados o influidos) por la información proveniente del sistema de creencias del procesador central. La información disponible para los procesos cognitivos se encuentra separada de aquella disponible para los procesos perceptuales. Si bien las inferencias perceptuales aportan información a las integraciones perceptuales, esta es de dominio específico: incluye información de su propia *base de datos* y especificaciones de los estímulos proximales. En esta línea, la tesis de la modularidad distingue sistemas modulares de sistemas no modulares.

Los sistemas modulares incluyen sistemas de *input* como los mecanismos de percepción y lenguaje, cuya función es llevar información desde los *outputs* de los transductores (que traducen la información del mundo al formato que sea procesable por el sistema) hacia los procesadores centrales. Además de ser encapsulados, estos sistemas son de dominio específico (se encuentran especializados para procesar *inputs* específicos), innatos, autónomos y asociados con mecanismos neuroanatómicos particulares. Suelen organizarse en un flujo de información *bottom-up*. También son obligatorios (su activación no puede detenerse una vez que se presente el estímulo adecuado), y poseen eficiencia en cuanto a la rapidez de su funcionamiento. Cabe destacar nuevamente que los sistemas modulares no son respuestas reflejas a los estímulos proximales, porque tienen una función de filtro de la variabilidad en las proyecciones proximales y de inferencia de disposiciones distales a partir de estimulaciones proximales

(1985, p. 4). El autor deja abierta la posibilidad de que los sistemas motores sean modulares.

Los sistemas no modulares, en cambio, incluyen procesadores centrales relacionados con la resolución de problemas y otras facultades cognitivas de nivel superior. Suelen poseer las características contrarias a los sistemas de *input*. No son encapsulados ni de dominio específico, sino que integran diversas informaciones de distintos dominios. Tampoco son obligatorios, ya que suelen someterse a control voluntario. Son más lentos, se asocian con estructuras neurológicas difusas y el flujo de la información no es específicamente *top-down* ni *bottom-up*⁵⁵.

La exposición anterior parece indicar, a primera vista, una asociación de los sistemas de *input* modulares con el procesamiento de representaciones perceptuales, y de los procesadores centrales con las representaciones conceptuales. No obstante, a pesar de la existencia de representaciones que cumplan con tal asociación, la relación es más compleja, lo que se observa en la existencia de grados de modularidad y en la falta de estudios sobre los procesadores centrales. Con respecto al primer aspecto, de acuerdo con Fodor, “el problema de evaluar el grado de encapsulamiento informativo de los sistemas de *input* no es (...) independiente del problema de determinar cómo se individualizan dichos sistemas y qué tipo de representaciones constituyen sus *outputs*” (1983, p. 74). Existen casos intermedios. Por ejemplo, la información de los sistemas de *input* puede ser corregida de manera *top-down* en casos en los que el estímulo no contiene suficiente información o en casos en los que “sabemos que el mundo no puede ser realmente como parece” (1983, p. 73). En relación con la falta de estudios sobre los procesadores centrales, su carácter no modular impone mayores dificultades en su investigación. Este punto puede constituir un factor que influya en la determinación de las representaciones asociadas a dichos procesadores.

A pesar de las dificultades anteriormente nombradas, es posible constatar una diferencia entre tipos de representaciones en cuanto a su procesamiento. Estas diferencias se enmarcan en la arquitectura cognitiva computacional-representacional. Fodor asume “sin discusión que los procesos mentales son computacionales en la medida en que son cognitivos, por lo que la función típica de los mecanismos cognitivos sea la transformación de las representaciones mentales” (1983, p. 13). A partir de allí, su tesis

⁵⁵ *Top-down* refiere al flujo de la información desde el nivel cognitivo superior (experiencias previas, metas, etcétera) hacia los aspectos perceptuales, mientras que *bottom-up* comienza con los aspectos perceptuales y avanza hacia el nivel cognitivo superior.

central es que existe “una distinción entre dos modos en los que los procesos computacionales pueden ser ‘inteligentes’” (1985, p. 1): por un lado, los procesos cognitivos son inferenciales y no encapsulados y, por otro, los procesos perceptuales son inferenciales y encapsulados. Como veré más adelante al abordar el SSP de Barsalou (1999), es posible problematizar esta diferencia de procesamiento al considerar los sistemas neurales compartidos entre percepción y cognición.

A continuación, me centraré en el segundo rasgo relevante de las representaciones mentales en la aproximación fodoriana, es decir, su vínculo con las palabras del lenguaje natural.

Relación entre representaciones mentales y palabras del lenguaje natural

Como se vio en la presentación general de la TRM, el lenguaje del pensamiento⁵⁶ exhibe varias de las características de los lenguajes naturales tanto a nivel sintáctico como a nivel semántico. Consta de una serie de conceptos (similares a *ítems* léxicos simples) y un conjunto finito de reglas de composición que operan con un carácter recursivo y cuyo poder expresivo es infinito. El punto medular de esta relación radica en la consideración de complejidad, es decir, la postulación de estructura constituyente para los estados representacionales. Los principales argumentos en favor de la estructura apelan a tres nociones fundamentales: productividad, sistematicidad y composicionalidad, rasgos de la cognición que están íntimamente relacionados.

En primer lugar, la productividad corresponde a la capacidad de poseer un conjunto potencialmente infinito de estados con distintos objetos intencionales y roles causales a partir de recursos finitos, por medio de la combinación y recursividad. El argumento clásico a su favor se basa en la competencia lingüística y su capacidad generativa, esto es, la capacidad de comprender y producir, en principio, una cantidad ilimitada de

⁵⁶ El código interno mantiene una analogía con el lenguaje natural que justifica el nombre de *lenguaje del pensamiento*. En los lenguajes naturales, la correspondencia entre la relación del sujeto con las fórmulas y sus actitudes proposicionales está mediada por las convenciones que gobiernan el uso coherente del lenguaje. Por ejemplo, un sujeto O emplea la fórmula *a es F* para representar que a es F si tiene la actitud proposicional de creer que a es F y acepta la fórmula *a es F*. Lo que media entre la creencia del sujeto y su empleo de dicha fórmula es una convención. En cambio, si *a es F* es una fórmula del lenguaje del pensamiento, el asentimiento a la fórmula es reemplazado por los determinantes de las relaciones computacionales entre esta y el sujeto y, además, la convención se reemplaza por lo nomológicamente necesario (1975, p. 78). Sin embargo, en ambos casos se mantiene una correspondencia entre la relación del sujeto con la fórmula y la actitud proposicional que posee.

representaciones novedosas sin límite de complejidad. Esta capacidad se debe lograr por medios finitos. En particular, es posible por la recursión, un recurso formal que permite la producción de oraciones indefinidamente largas sobre la base de la aplicación sucesiva de una regla o parte de una regla a una misma expresión. Por ejemplo, la regla del uso del *que* relativo permite construir oraciones tales como *Luis, que es ingeniero, que nació en Cuba, que luchó en la guerra...* El desempeño real se determinaría por la relación entre el conocimiento base del individuo y sus recursos computacionales finitos, y mejora manejando parámetros tales como los constreñimientos de tiempo o el nivel de motivación. La productividad se explica si los estados tienen estructura combinatoria, es decir, si se constituyen de partes cuyas combinaciones produzcan el conjunto potencialmente infinito de estados.

El argumento de la productividad se basa en la idealización de una competencia ilimitada que es discutida por su lejanía con la evidencia, por lo que Fodor (1975) y Fodor y Pylyshyn (1988) apelan a otro argumento, relacionado con la sistematicidad. La sistematicidad se refiere a una relación entre la capacidad de producir y entender algunas oraciones y la capacidad de producir y entender otras. Por ejemplo, quien entiende y profiere la oración *Juan ama a María* sabe el significado de *Juan*, de *ama* y de *María*, y puede entender y proferir la oración *María ama a Juan*. Los autores introducen la noción por medio de la comparación del aprendizaje de idiomas y la memorización de un libro de frases: es posible aprender cualquier parte del libro sin aprender el resto, pero el manejo de la sintaxis del lenguaje permite aprender a producir o entender oraciones relacionadas. Se trata de una característica propia de la capacidad lingüística, y depende de la semántica combinatoria de los lenguajes naturales. En ella, existen mecanismos que determinan el significado de las oraciones a partir de la especificación del significado de sus elementos léxicos y su estructura sintáctica. Las oraciones se forman con palabras y frases que pueden participar de diversas fórmulas. El hecho de que tengan semántica combinatoria depende de que tengan estructura constituyente: si todas las oraciones fueran atómicas, la producción y comprensión de unas no implicaría la producción y comprensión de otras. De la sistematicidad del lenguaje se deriva la sistematicidad del pensamiento: las capacidades cognitivas también son sistemáticas (al menos tan sistemáticas como las capacidades lingüísticas, puesto que el lenguaje tiene como función expresar el pensamiento): la capacidad de poseer un pensamiento determinado implica la capacidad de poseer otros que se encuentran relacionados. Ello se explica sosteniendo que poseer un pensamiento requiere estar en un estado representacional, y que deben existir

relaciones estructurales (es decir, composición con las mismas partes) entre las representaciones que corresponden a los pensamientos *Juan ama a María y María ama a Juan*. Por tanto, los pensamientos también deben tener una estructura constituyente. A diferencia de la productividad, la sistematicidad es un rasgo directo (no idealizado) de la cognición.

Otra fuente de sistematicidad radica en la coherencia inferencial, esto es, la posibilidad de realizar cualquier inferencia de un tipo determinado dada la capacidad de realizar algunas inferencias de ese tipo. Por ejemplo, “no existen mentes preparadas para inferir que *John fue a la tienda* a partir de *John y Mary y Susan y Sally fueron a la tienda* y a partir de *John y Mary fueron a la tienda*, pero no a partir de *John y Mary y Susan fueron a la tienda*” (Fodor y Pylyshyn, 1988, p. 48). De acuerdo con el enfoque simbólico, existe un paralelismo entre principios lógicos como *una conjunción de tipo P&Q implica sus constituyentes P y Q*, por un lado, y leyes psicológicas sobre que los pensamientos *P&Q* causan los pensamientos *P* y *Q*. Ambos casos se explican por la estructura constituyente: la semántica combinatoria es sensible a la sintaxis de las representaciones en el primer caso, y los procesos mentales operan de manera sensible a la estructura de las representaciones en el segundo caso. De ello se deriva un compromiso con la relación entre inferencias de forma lógica similar y roles similares en el pensamiento. Las inferencias similares se llevan a cabo por mecanismos psicológicos similares, de modo tal que, al inferir *P* a partir de *P&Q&R* y al inferir *P* a partir de *P&Q*,

las premisas de ambas inferencias son expresadas por representaciones mentales que satisfacen el mismo análisis sintáctico (...) y el proceso de elaboración de la inferencia corresponde, en ambos casos, a la misma operación formal de separar el constituyente que expresa la conclusión (1988, p. 47).

Dado que se reconoce la existencia de representaciones estructuradas y procesos sensibles a esta estructura, se predice que las representaciones con estructura parecida tendrán roles parecidos en el pensamiento.

En tercer lugar, la composicionalidad apunta a que el contenido de las representaciones complejas se determina por el contenido de las representaciones primitivas que la conforman junto con las reglas de composición o sintaxis. Se relaciona con la sistematicidad, de modo que “tal vez sea mejor considerarlas como aspectos de un único fenómeno” (1988, p. 41). Las oraciones relacionadas de manera sistemática se encuentran también relacionadas semánticamente, y para explicar este vínculo se requiere apelar al principio de composicionalidad, de acuerdo con el cual “en la medida en que un

lenguaje es sistemático, un *ítem* léxico debe aportar aproximadamente la misma contribución semántica a cada expresión en la que aparece” (1988, p. 42). Entre oraciones sistemáticamente relacionadas, la similitud de sus estructuras da cuenta de su relación semántica porque las propiedades semánticas de los elementos compartidos son independientes del contexto. En los modismos, los elementos no son considerados en su significado habitual, por lo que la expresión no es sistemática con respecto a oraciones con los mismos constituyentes pero tomados en su significado estándar. Por tanto, la sistematicidad depende de la composicionalidad, y la composicionalidad se requiere para explicar “cómo un lenguaje finitamente representable puede contener infinitas expresiones no sinónimas” (1988, p. 43). En este punto se reintroduce la productividad como un requisito formal.

Otro argumento en favor de la composicionalidad apela a las relaciones entre las estructuras sintácticas de las oraciones y sus vínculos de implicación. Por ejemplo, un predicado de tipo *es una vaca café* se relaciona semánticamente con el predicado *es café*. El primer predicado implica el segundo dado que el segundo es un constituyente suyo, la forma sintáctica tiene la fuerza semántica de una conjunción y la simplificación de la conjunción no afecta el valor semántico de *café*.

Nuevamente, la composicionalidad requiere la existencia de representaciones estructuradas, dado que aplica a las relaciones semánticas entre las palabras y las oraciones de las que son parte. Y, nuevamente, considerando la relación entre lenguaje y pensamiento, lo mismo aplica a las representaciones mentales. De la misma forma que la capacidad de emplear algunas oraciones se conecta con la capacidad de emplear otras semánticamente relacionadas, la capacidad para pensar algunos pensamientos (y de encontrarse en un estado representacional determinado) se encuentra conectada con la capacidad para pensar otros pensamientos (y de encontrarse en otros estados representacionales) semánticamente relacionados. El enfoque simbólico es capaz de considerar la relación entre sistematicidad y vínculo semántico, porque es una propiedad de la arquitectura cognitiva⁵⁷.

⁵⁷ De acuerdo con Fodor y Pylyshyn (1988), el conexionismo no es capaz de dar cuenta de la productividad, sistematicidad y composicionalidad. En relación con la primera, presenta problemas con su implementación. Una máquina convencional tiene una distinción funcional entre memoria y programa, por lo que puede implementar capacidades cognitivas productivas: los cambios en la memoria no afectan la estructura computacional. En una máquina conexionista, en cambio, la memoria se encuentra contenida en el patrón de activación, y su modificación conllevaría alteraciones en la estructura computacional. Con respecto a la sistematicidad de la coherencia inferencial, el conexionismo podría plantear modelos

En suma, el lenguaje del pensamiento mantiene un vínculo estrecho con el lenguaje natural. La complejidad es el núcleo de este vínculo, y permite dar cuenta de la productividad, sistematicidad y composicionalidad, rasgos exhibidos tanto por las capacidades lingüísticas como por el código interno. En el próximo subapartado me centraré en el último rasgo relevante de las representaciones mentales en la aproximación fodoriana, es decir, su vínculo con las representaciones neurales.

Relación entre representaciones mentales y representaciones neurales

En los subapartados anteriores me enfoqué en dos aspectos de la noción de representación propia de Fodor: la existencia de diferencias en el procesamiento de representaciones ligadas a los sistemas de *inputs* modulares y representaciones de los procesadores centrales, y el vínculo entre las representaciones mentales y las palabras del lenguaje natural. El tercer y último rasgo que abordaré corresponde a la distinción entre representaciones mentales y las denominadas *representaciones neurales*. En neurociencias, las representaciones neurales han sido comprendidas como “un patrón de actividad neural que representa alguna característica ambiental en el funcionamiento interno del cerebro” (Villarroya, 2017, p. 4). No obstante, el sentido en el cual las representaciones neurales *representan* algo varía en dependencia del autor y no necesariamente concuerda con el sentido en el cual las representaciones mentales, entendidas como estados con contenido intencional, representan estados de cosas. En neurociencias, representar puede ser entendido como designar, codificar, mapear, llevar información, entre otros verbos, a nivel de sistema nervioso. El establecimiento de identidad o diferencias entre representaciones mentales y representaciones neurales depende, en gran medida, de consideraciones ontológicas originadas en la filosofía de la mente. En términos generales, una postura reduccionista sostendría que las representaciones mentales son idénticas o reducibles a las representaciones neurales,

cognitivos donde inferencias de forma lógica similar son computadas por mecanismos diferentes o bien algunas inferencias son computadas y otras, de la misma forma lógica, no. Los nodos son atómicos, por lo que la falta de una noción de sintaxis lógica conduce a que las inferencias de forma lógica similar no tengan que ser ejecutadas por procesos computacionales sintácticamente similares. Por último, en lo tocante a la composicionalidad, tampoco explica la relación entre sistematicidad y vínculo semántico: las etiquetas de los nodos no poseen implicancias semánticas para las etiquetas de otros. Estas y otras falencias se derivarían, según Fodor y Pylyshyn, de no postular estructura constituyente para las representaciones. Abordaré la discusión entre el enfoque simbólico y el conexionismo más adelante (pp. 117-123).

mientras que una postura no reduccionista establecería vínculos entre ellas distintos a la identidad o la reducción.

En particular, Fodor sostiene un fisicalismo no reduccionista de acuerdo con el cual las ciencias especiales tales como la psicología no deben ser reducidas a las teorías físicas. Por tanto, rechaza el reduccionismo psicológico, es decir, la postura que afirma que las clases naturales psicológicas son idénticas o coextensivas con las clases naturales neurológicas. El interés de la psicología cognitiva se encuentra en “caracterizar la etiología del comportamiento en términos de una serie de transformaciones de información” (Fodor, 1975, p. 52). Aunque el comportamiento posea entre sus determinantes los eventos neuronales concomitantes, su explicación se construye de manera más amplia (incluyendo, por ejemplo, el pensamiento y el conocimiento).

En relación con el desarrollo de arquitecturas convencionales, una crítica recurrente formulada desde otros enfoques como el conexionismo radica en la falta de consideración de la neurociencia. Sobre este punto, Fodor establece una distinción entre el nivel cognitivo (el nivel de postulación de estados representacionales) y el nivel de implementación (computadores electrónicos, cerebro humano, entre otros). Según Fodor y Pylyshyn (1988), no es claro cómo datos tales como la importancia de las conexiones neuronales para los patrones de actividad cerebral o la sensibilidad de sus campos receptivos a un conjunto estrecho de valores sean relevantes para la arquitectura cognitiva. Los autores sostienen al menos dos argumentos relacionados con los aportes de la neurociencia. En primer lugar, considerando que existen distintos niveles de organización en un sistema, la estructura de niveles superiores en los que figuran procesos cognitivos abstractos no suele ser isomórfica con la estructura de niveles inferiores del sistema. Por ejemplo, el razonamiento no es definido como la difusión de activaciones en las redes. En segundo lugar, el compromiso con la inspiración neural condujo a la reaparición de teorías psicológicas que han mostrado limitaciones en el pasado. Entre estas se cuenta la psicología asociacionista relacionada con la noción del cerebro como una red neural, y modelos estadísticos de aprendizaje que surgieron a partir de “la idea de que la fuerza de una conexión entre neuronas es afectada por la frecuencia de su coactivación” (Fodor y Pylyshyn, 1988, p. 64). Por tanto, es necesario evaluar cuidadosamente las aplicaciones de datos neurocientíficos en el desarrollo de arquitecturas cognitivas.

En suma, es posible trazar una separación entre representaciones mentales y representaciones neurales sobre la base de dos puntos de la postura de Fodor: su

fisicalismo no reduccionista y su distinción entre un nivel cognitivo y un nivel de implementación. Ello permitiría sostener, preliminarmente, que las representaciones mentales se encuentran a un nivel cognitivo, mientras que las representaciones neurales pertenecen al nivel de implementación, y que una no se reduce a la otra. Sin embargo, ello depende de la definición de *representación neural* que se tome en consideración — en particular, qué se entiende por *representar* en ese contexto.

En el presente subapartado caractericé la TRM propuesta por Jerry Fodor, haciendo hincapié en tres puntos relacionados con la noción de REPRESENTACIÓN: la relación entre representaciones conceptuales y representaciones perceptuales, la relación entre representaciones mentales y palabras del lenguaje natural y la relación entre representaciones mentales y representaciones neurales. Considerando la ecología conceptual en la cual surge esta noción, es posible puntualizar influencias teóricas ya mencionadas tales como la arquitectura clásica de von Neumann y Turing, la teoría de la prueba, el fisicalismo no reduccionista, la distinción entre un nivel cognitivo y un nivel de implementación, y el compromiso ontológico con las actitudes proposicionales. Estos y otros factores influyen en el concepto de REPRESENTACIÓN resultante, entendido como símbolos con contenido, poderes causales y estructura. A continuación, abordaré el primer debate en el que se localiza un cambio en él: el modelo de procesamiento paralelo distribuido presentado por Paul Smolensky.

2.2. Tratamiento Adecuado del Conexionismo, de Paul Smolensky

El conexionismo corresponde a una de las grandes alternativas al enfoque presentado en el subapartado anterior, aunque mantiene el carácter computacional y representacional de los supuestos sustantivos. Uno de sus exponentes centrales es Paul Smolensky (1987, 1988, 1989), quien provee una articulación de la aproximación por medio de su Tratamiento Adecuado del Conexionismo (TAC). Como mencioné en el segundo capítulo, el conexionismo postula un modelo de redes compuesto por un conjunto de unidades de procesamiento análogas a neuronas abstractas, con valores numéricos de activación y patrones de conectividad entre ellas que poseen pesos. El modelo es denominado de procesamiento paralelo distribuido porque muchas unidades pueden computar al mismo tiempo, y la información se distribuye en las conexiones entre nodos. Ello constituye, de acuerdo con Smolensky, una ventaja con respecto a las máquinas

convencionales, cuyo carácter serial implicaría un colapso ante una lesión (ruptura de circuito). Las máquinas conexionistas pueden seguir ejecutando tareas aun cuando se remuevan algunas unidades, dado que cada unidad juega un rol pequeño. Esta característica se relaciona con la propiedad de completación de patrones, es decir, rellenar los elementos faltantes en un *input* incompleto para que la red genere un patrón completo. Lo anterior permite reconocer patrones mayoritariamente perceptuales. Además, se asocia con la degradación grácil, una capacidad de la máquina de mantener funcionalidad a pesar de la presencia de errores o daño. Con la falta de esta capacidad, un error en el reconocimiento o ejecución de una tarea implica detenerse y no realizar acción alguna⁵⁸.

La propuesta de Smolensky incluye los siguientes puntos importantes:

- a. La adopción de un marco de modelación cognitiva se relaciona con la toma de postura ante la *paradoja de la cognición*.
- b. La paradoja de la cognición es respondida por el marco de modelación conexionista sosteniendo que el sistema cognitivo es una máquina con constreñimientos blandos que, a nivel superior, muestra constreñimientos fuertes.
- c. El marco de modelación conexionista es un paradigma subsimbólico cuyo nivel de explicación preferente es subconceptual.
- d. El paradigma subsimbólico puede mantener una relación de complementariedad con el paradigma simbólico.

De acuerdo con Smolensky (1987, 1988, 1989), la justificación para la adopción de un marco de modelación cognitiva u otro incluye la toma de partido ante la denominada

⁵⁸ Este punto se vincula directamente con críticas al enfoque simbólico sobre las limitaciones de los computadores convencionales para estudiar la cognición humana. Entre ellas se cuentan restricciones de tiempo de procesamiento como el constreñimiento de los cien pasos, que apunta a la existencia de una asimetría entre el tiempo requerido para la ejecución de procesos computacionales (nanosegundos) y la rapidez neural (milisegundos): un programa serial neuralmente instanciado solo puede llevar a cabo 100 instrucciones en el tiempo en que una persona realiza tareas cognitivas. Otras restricciones refieren a la alta sensibilidad al daño y ruido —que puede originarse por la existencia de un control y memoria altamente localizados—. Dicha sensibilidad se relaciona con la falta de degradación grácil. Para Fodor y Pylyshyn (1988), estas y otras críticas se basan o bien en propiedades que no son intrínsecas de las arquitecturas convencionales o bien en propiedades que dependen de su implementación en los computadores actuales. Las críticas sobre los límites de velocidad o la falta de resistencia a daño y ruido son rasgos de implementación que solo apuntan a un hecho evidente: las arquitecturas no se implementan en los computadores de la misma forma que lo hacen en el cerebro. Ello no es problemático: si se acepta la diferencia entre el nivel cognitivo y el de implementación (véase p. 109), se sigue que los detalles de implementación son irrelevantes para la teoría psicológica.

paradoja de la cognición o dilema estructura/estadística. Esta paradoja consiste en la asimetría entre los aspectos duros y los aspectos blandos de la cognición: los sistemas cognitivos parecen estar gobernados por un sistema de reglas rígidas (constreñimientos fuertes) y, al mismo tiempo, presentan flexibilidad y fluidez (constreñimientos blandos). La competencia cognitiva de nivel superior se rige por reglas que indican representaciones y procesos simbólicos y estructurados, mientras que el desempeño real indica descripciones estadísticas. Según Smolensky, entre las posibles soluciones adoptadas frente a este dilema, el enfoque simbólico coloca el énfasis en los constreñimientos fuertes. Sus reglas e inferencias lógicas se aplican de forma serial e independiente de otras reglas⁵⁹. El conexionismo, en cambio, distingue entre dos niveles de análisis. En un nivel inferior, el sistema cognitivo es una máquina que realiza satisfacción paralela masiva de constreñimientos blandos. La red intenta establecer un balance entre todos los constreñimientos blandos, que son entendidos como restricciones que el sistema puede violar mientras realiza una tarea de forma apropiada. Estos se formalizan como inferencia estadística, un proceso donde la fuerza de la conexión entre dos unidades no depende solo de su actividad, sino que suele depender de las otras fuerzas. Por otra parte, en un nivel superior, el sistema exhibe constreñimientos fuertes: parece seguir reglas lógicas. En suma, los aspectos duros observados a nivel superior emergen de los aspectos blandos propios del nivel inferior, y “el sistema cognitivo es fundamentalmente una máquina blanda tan compleja que a veces parece dura cuando se ve a niveles superiores” (1987, pp. 139-140).

La arquitectura conexionista adoptada por Smolensky considera que la cognición debe ser entendida en el marco de un sistema dinámico, es decir, un sistema complejo gobernado por ecuaciones diferenciales. El sistema genera representaciones del entorno por medio de la extracción de información y su codificación en los pesos de los patrones de la red. Posee dos tipos de variables: valores numéricos de activación para cada unidad y fuerzas de conexión para sus enlaces. Sus estados evolucionan continuamente en el

⁵⁹ El enfoque simbólico ha sido criticado por su empleo de reglas representadas de forma explícita y operaciones de todo o nada, las cuales contrastarían con los constreñimientos blandos enfatizados por el conexionismo. Para Fodor y Pylyshyn (1988), esta incompatibilidad con los constreñimientos blandos no es una propiedad intrínseca de la arquitectura convencional, por lo que no afecta al nivel cognitivo. La arquitectura convencional no requiere la existencia de reglas explícitas ni la postulación de operaciones de todo o nada. Por ejemplo, los computadores de programa fijo pueden tener todas sus reglas implícitas, y la aplicación de reglas puede depender de magnitudes variables o influir en conjunto o por separado en el comportamiento.

tiempo de acuerdo con ecuaciones diferenciales: reglas de propagación de activación y reglas de modificación de fuerzas de conexión. Las primeras son entendidas como reglas de inferencia estadística, y las segundas son procedimientos de aprendizaje y almacenamiento, y se encargan de indicar los cambios en los pesos durante el entrenamiento.

Para entender la relación entre el enfoque tradicional visto en los subapartados anteriores, el enfoque conexionista y el cerebro, Smolensky realiza una distinción teórica entre tres sistemas cognitivos (paradigma simbólico, subsimbólico⁶⁰ y cerebro humano) y tres niveles de análisis (nivel conceptual, subconceptual y neural). El cerebro y su nivel neural de análisis, así como su relación con el nivel subconceptual del conexionismo, serán abordados en el subapartado *Relación entre representaciones mentales y representaciones neurales* (pp. 123-124). Por otra parte, el modelamiento cognitivo sostenido por autores como Fodor (1975, 1987; Fodor y Pylyshyn, 1988) es denominado *paradigma simbólico*. Su nivel preferente de análisis es el nivel conceptual. Desde un punto de vista sintáctico, las nociones propias de la psicología común como las creencias, deseos y conceptos son formalizadas como estructuras simbólicas que se apoyan en un sistema de símbolos físicos, esto es, “un dispositivo computacional físico para manipular símbolos” (Smolensky, 1989, p. 235). Las estructuras son operadas por manipulación simbólica. Se emplean instrucciones lingüísticas formalizadas en procedimientos efectivos al modo de Turing, e incluyen reglas formuladas lingüísticamente e interpretadas de manera secuencial. Con respecto a la semántica, los mismos símbolos que son operados por manipulación simbólica son interpretados en términos de los conceptos que el individuo utiliza para conceptualizar la tarea.

El modelamiento cognitivo del conexionismo, en cambio, corresponde al paradigma subsimbólico. Se usan subsímbolos, patrones de actividad de las unidades de las redes conexionistas. Estos mantienen diferencias con el paradigma simbólico tanto semántica como sintácticamente. Con respecto a la semántica, las entidades que son semánticamente interpretables —las representaciones— son entendidas como los patrones de activación. Estas variables numéricas son inferiores a los conceptos empleados de manera consciente en la descripción de la tarea. Su contenido depende del contexto (véase p. 120) y del modelo conexionista adoptado. En este sentido, según

⁶⁰ La referencia a los dos tipos de modelación (simbólica y conexionista) en términos de *paradigma* es propia de Smolensky.

Smolensky “cada modelo subsimbólico individual adopta procedimientos particulares para relacionar los patrones de actividad —vectores de actividad— con las descripciones a nivel conceptual de *inputs* y *outputs* que definen la tarea del modelo” (1988, p. 7). En relación con la sintaxis, en lugar de ser operados por manipulación simbólica, son operados por reglas de propagación de activación. Una operación discreta del paradigma simbólico corresponde aquí a diversas operaciones numéricas. En la modelación y predicción del comportamiento de nivel superior, la implementación de las unidades no se considera relevante, puesto que pertenece a un nivel inferior al nivel subconceptual, y es en el nivel subconceptual donde se puede realizar una predicción conductual exacta (Smolensky, 1988, p. 11).

Este paradigma posee un nivel de análisis inferior a aquel sostenido por la aproximación simbólica: en lugar de emplear el nivel conceptual, adopta un nivel de análisis preferentemente subconceptual. Sin embargo, un sistema conexionista también puede ser descrito a nivel conceptual. De acuerdo con Smolensky, “si adoptamos un nivel más alto de descripción de lo que ocurre en estos sistemas subsimbólicos (y eso implica, en un grado significativo, una aproximación) entonces obtenemos descripciones que son aproximadamente como las descripciones simbólicas” (1989, p. 238). Los conceptos simbólicos propios de la psicología del sentido común proveen descripciones que son válidas de manera aproximada y permiten entender los sistemas complejos. El carácter aproximado se debe a que estas descripciones excluyen aspectos relevantes del procesamiento, describen el comportamiento en términos cualitativos o por idealizaciones alejadas del desempeño real. Un ejemplo de descripción a nivel superior se encuentra en la reformulación de las conexiones estadísticas en términos de principio de mejor ajuste (*best fit principle*), según el cual el sistema conexionista produce como *output* un conjunto de inferencias que, tomadas en su conjunto, corresponden al mejor ajuste con respecto al *input* recibido “en un sentido estadístico definido por el conocimiento estadístico almacenado en las conexiones del sistema” (1988, p. 19).

A nivel sintáctico, un computador von Neumann puede simular una red conexionista y viceversa. La diferencia principal entre ambos paradigmas figura, por tanto, a nivel semántico. El paradigma simbólico provee explicaciones en las que los procesos tienen la misma semántica de nivel conceptual que el comportamiento a explicar. En el paradigma subsimbólico, en cambio, las unidades no poseen la misma semántica del comportamiento original. Las disposiciones de los sistemas cognitivos son explicadas “como regularidades aproximadas de nivel superior que surgen de leyes

cuantitativas que operan a un nivel más fundamental con una semántica diferente” (1988, p. 11). En suma, la diferencia principal entre ambos paradigmas es resumida por el autor de la siguiente manera: “los modelos subsimbólicos describen con precisión la microestructura de la cognición, mientras que los modelos simbólicos proporcionan una descripción aproximada de la macroestructura” (1988, p. 12). En este sentido, Smolensky señala que el paradigma subsimbólico no busca reemplazar al paradigma simbólico, sino explicarlo, refinarlo y complementarlo mediante el enfoque en un nivel inferior de análisis que permita entender “cómo los constructos teóricos de nivel conceptual emergen desde la computación subconceptual” (1988, p. 22).

Tras la caracterización general de los sistemas de procesamiento paralelo distribuido sistematizada por Smolensky, abordaré los detalles de la noción de representación involucrada por medio del examen de los tres rasgos determinados anteriormente: (i) la relación entre representaciones conceptuales y representaciones perceptuales (pp. 115-116); (ii) la relación entre las representaciones mentales y las palabras del lenguaje natural (pp. 116-123), y (iii) la distinción entre las representaciones mentales y las representaciones neurales (pp. 123-124).

Relación entre representaciones conceptuales y representaciones perceptuales

El conexionismo mantiene una demarcación entre percepción y cognición, y tiene interés en la modelación tanto de procesos perceptuales de nivel inferior como de procesos de nivel superior. La percepción —en particular, procesos perceptuales como la categorización— es considerada como fundamento de la inteligencia (Smolensky, 1987, p. 140), lo que justifica en parte la prioridad otorgada a los aspectos blandos en la respuesta a la paradoja de la cognición (pp. 111-112). Los aspectos fuertes serían excepcionales en el desempeño cognitivo, y se desarrollarían más tardíamente desde el punto de vista evolutivo. Sin embargo, la distinción entre percepción y cognición no se realiza de la manera directa observada en el paradigma simbólico. A nivel subconceptual de análisis, tanto las representaciones conceptuales como las representaciones perceptuales son subsímbolos: patrones de activación en la red. En cambio, a nivel conceptual, se pueden entender aproximadamente en términos de estructuras mentales, como percepciones, conceptos, creencias, inferencias, entre otros. En consecuencia, la diferencia se trazaría en este segundo nivel.

En segundo lugar, el TAC no se refiere mayormente a la modularidad planteada por Fodor (1983, 1985), por lo que no establece diferencias en el procesamiento de tipos de representación. En las respuestas a las críticas realizadas a su artículo *On the Proper Treatment of Connectionism* (1988), Smolensky acepta que las comparaciones que establece entre las arquitecturas neurales y subsimbólicas tienen en consideración “los típicos modelos conexionistas actuales que constan de un solo módulo” (1988, p. 65). El autor reconoce la importancia de los módulos, y no niega la posibilidad de que los modelos conexionistas posteriores puedan exhibir una estructura modular. No obstante, plantea que

la verdadera cuestión es si la estructura modular de los futuros modelos conexionistas se relacionará con la estructura modular del cerebro, o si la arquitectura de los modelos será impulsada por consideraciones computacionales que resulten no estar profundamente ligadas a la arquitectura neural (Smolensky, 1988, p. 65).

En suma, la distinción que puede trazarse entre representaciones conceptuales y perceptuales no pertenece al nivel subconceptual, que es el nivel preferente del paradigma subsimbólico. Sumado a ello, no se postula una diferencia de procesamiento como en la modularidad fodoriana. En el próximo subapartado abordaré el segundo rasgo de las representaciones mentales en consideración, a saber, su vínculo con las palabras del lenguaje natural.

Relación entre representaciones mentales y palabras del lenguaje natural

El conexionismo problematiza el empleo del lenguaje como modelo teórico del paradigma simbólico, tanto a nivel sintáctico como a nivel semántico. Para Smolensky,

desde el nacimiento de la ciencia cognitiva, el lenguaje ha sido el modelo teórico dominante. Los modelos cognitivos formales han tomado su estructura de la sintaxis de los lenguajes formales, y su contenido de la semántica del lenguaje natural. La mente se ha concebido como una máquina de manipulación formal de símbolos, y los símbolos manipulados han asumido esencialmente la misma semántica que las palabras (...) (1988, p. 4).

El nivel de análisis preferentemente subconceptual del conexionismo concibe un sistema complejo gobernado por ecuaciones diferenciales que tiene como *inputs* y *outputs* vectores numéricos en lugar de representaciones similares a palabras. Si bien es posible

realizar interpretaciones a nivel conceptual con respecto a estos elementos, existe un cambio semántico.

Una de las críticas más importantes realizadas al paradigma subsimbólico se encuentra en una extensa discusión cuyo artículo principal es *Connectionism and Cognitive Architecture: A Critical Analysis*, de Fodor y Pylyshyn (1988). En él, los autores identifican como focos de desacuerdo entre el enfoque simbólico y subsimbólico dos puntos: la postulación de estructura para las representaciones y la sensibilidad de los procesos a dicha estructura, que constituyen los dos núcleos del enfoque simbólico. Este enfoque sostiene que las representaciones poseen sintaxis y semántica combinatoria, y las operaciones mentales aplican a ellas en virtud de su forma. De acuerdo con Fodor y Pylyshyn, el conexionismo no se compromete con estas dos afirmaciones, sino que limita las relaciones primitivas entre nodos a relaciones causales, no estructurales. Por ejemplo, en una máquina conexionista, el nodo etiquetado como *A&B* se conecta causalmente con el nodo etiquetado como *A* y el nodo etiquetado como *B*, pero sin una relación estructural. A diferencia de ello, en una máquina convencional, las instanciaciones de la expresión *A&B* tienen como constituyentes instanciaciones de la expresión *A* y de la expresión *B*, y su semántica se determina por la semántica de tales constituyentes.

En primer lugar, con respecto a las representaciones mentales, los sistemas conexionistas no operarían sobre símbolos con estructura, es decir, el lenguaje no es un medio de computación. Los nodos pueden tener interpretación semántica (una etiqueta con expresiones lingüísticas que indiquen su contenido), pero el nodo mismo no tiene estructura y las relaciones entre etiquetas no cumplen un rol causal en las operaciones del sistema. La distribución neuronal de las representaciones no implica que tengan constituyentes: la distribución ocurre entre entidades, mientras que las relaciones de constitución operan dentro de las entidades — sosteniendo que las entidades que son semánticamente evaluables tienen partes que también lo son—. Además, el conexionismo reemplaza la distinción entre símbolos atómicos y moleculares (véase pp. 98-99) por la distribución de representaciones en microcaracterísticas, pero ello tampoco exhibe constitución. En este marco, un concepto se distribuye en subconceptos con contenido representacional que se superponen. Por ejemplo, + *tiene-cordones* genera la impresión de que es constituyente del concepto de ZAPATILLA: se mantienen en un vínculo de conjunto-subconjunto. No obstante, la superposición no implica relaciones de parte-todo: a nivel sintáctico, *tiene-cordones* no pertenece a ZAPATILLA.

El conexionismo también puede representar las relaciones como microcaracterísticas: la creencia de que *Luis quiere a Mancha* es representada por un vector que incluye a Luis, a Mancha y a la relación *querer a*. Sin embargo, no todos los subconjuntos de características son constituyentes. (+*Luis-sujeto*; + *Mancha-objeto*) no es constituyente del símbolo *Luis quiere a Mancha*. Además, para distinguir una creencia como *Luis quiere a Mancha* y *Gladys odia a Diana* de otra distinta como *Luis quiere a Diana* y *Gladys odia a Mancha*, se requeriría emplear una cantidad de microcaracterísticas similar a la cantidad de oraciones del lenguaje natural, porque se representan por el mismo conjunto (+*Luis-sujeto*; +*quiere a*; *Mancha-objeto*; +*Gladys-sujeto*; +*odia a*; +*Diana-objeto*). Por otro lado, identificar un pensamiento con un vector solo funciona bajo la suposición de que sus nodos asociados son los únicos activos. No obstante, las situaciones suelen ser complejas y, por tanto, existirían otros nodos activos además de los que pertenecen a la proposición considerada. En esos casos, no se puede distinguir los subvectores que representan una proposición con respecto a otra: en una situación donde una persona está tomando café y pensando que *Luis quiere a Mancha*, los nodos que representan cada uno de estos componentes (LUIS; QUIERE A; MANCHA; PERRA; CAFÉ...) se activarían todos en un tiempo determinado, pero no todos se corresponden con los conceptos que constituyen la misma proposición. Por ejemplo, los nodos PERRA, QUIERE A y CAFÉ pueden expresar una proposición de tipo *La perra quiere al café*, y ello no pertenece a la situación original. Por ende, no basta con el criterio de activación/inactivación para especificar los pensamientos y distinguirlos de otros. Se requiere, además, apelar a la estructura constituyente: las representaciones que expresan los conceptos pertenecientes a una proposición se encuentran en una misma construcción, mientras que otras, aunque activas, se consideran como externas.

En segundo lugar, con respecto al tratamiento de los procesos mentales por parte del conexionismo, Fodor y Pylyshyn (1988) afirman que estos no operarían atendiendo a la estructura de las representaciones, puesto que no se postula tal estructura. Así, el paradigma subsimbólico se limitaría a los procesos de aprendizaje y razonamiento, que son entendidos mediante consideraciones propias del asociacionismo clásico (según el cual la probabilidad de que una idea provoque otra depende de la fuerza de asociación entre ambas, que a su vez depende de la correlación previa entre ellas). En el aprendizaje, los pesos de las conexiones se modifican en función del entrenamiento hasta que el comportamiento del sistema modele las propiedades estadísticas de sus *inputs* (1988, p. 30). Para ello, las unidades ocultas detectan patrones estadísticos en la actividad de las

otras unidades. En el razonamiento, se determina el *output* indicado para un *input* determinado en dependencia de propiedades de las conexiones tales como el peso, que a su vez dependen de propiedades estadísticas de los eventos del ambiente. La estructura sintáctica o semántica de las representaciones involucradas no se considera.

La productividad, sistematicidad y composicionalidad son, como mencioné anteriormente, rasgos de la cognición utilizados como argumentos para defender la estructura combinatoria de las representaciones mentales. De acuerdo con Fodor y Pylyshyn (1988), dado que estos rasgos solo se explican bajo el supuesto de la existencia de estructura, las teorías cognitivas adecuadas deberían considerarla. Los autores afirman que los sistemas conexionistas no postulan estructura. Ello conllevaría a que las representaciones se traten como una mera lista, por lo que “cualquier colección de estados representacionales (causalmente conectados) es una posible mente” (1988, p. 49). Por tanto, la arquitectura de la mente/cerebro no sería conexionista. Si el paradigma subsimbólico mantiene la falta de reconocimiento de estructura, no puede hacerse cargo de la sistematicidad y productividad; si abandona la arquitectura de redes para reconocer la estructura, pero manteniendo su asociacionismo en los procesos, no puede dar cuenta de la sensibilidad de los procesos a la estructura (porque las asociaciones no son sensibles a ella).

Las falencias anteriormente mencionadas conducen a Fodor y Pylyshyn a sostener que el conexionismo solo puede aspirar a limitarse a procesos como el análisis de inferencias estadísticas o a convertirse en una teoría de la implementación. En otras palabras, se convertiría en “una teoría sobre cómo los sistemas cognitivos (clásicos) podrían estar implementados en cerebros reales o en alguna ‘neurología abstracta’” (1988, p. 54). Podría complementar al enfoque simbólico por medio de la explicación de cómo una red cerebral o cerebralmente inspirada realiza procesos simbólicos. Sin embargo, dado que los detalles de la implementación son distintos al nivel cognitivo, ello implicaría que el conexionismo debe mantenerse neutral acerca de la naturaleza de los procesos cognitivos, sin plantear modelos como teorías de la cognición.

Frente a las críticas, Smolensky (1987, 1988) afirma que es un error sostener una carencia de estructura constituyente en el caso de los sistemas conexionistas distribuidos. Los estados mentales son patrones de actividad con estructura constituyente tanto a nivel conceptual como a nivel subconceptual, donde “los constituyentes a nivel conceptual de los estados mentales son vectores de actividad, que a su vez tienen una estructura

constituyente a nivel subconceptual: las actividades de las unidades individuales” (Smolensky, 1988, p. 17).

Afirmar que las representaciones conexionistas son atómicas se funda en una falacia de muñeco de paja: considerar las representaciones ultra locales de ciertos modelos (en los que las proposiciones se representan por nodos individuales) como si fueran los modelos conexionistas estándares. En las representaciones distribuidas, “las representaciones de entidades conceptuales de nivel superior tales como las proposiciones están distribuidas en muchos nodos, y los mismos nodos participan simultáneamente en la representación de muchas entidades” (1987, p. 144). Es un error sostener que las relaciones que se establecen entre entidades distribuidas no poseen consecuencias para la estructura dentro de ellas. En efecto, existen implicaciones con respecto a la composicionalidad, que se pueden extraer del análisis sugerido por el mismo Pylyshyn (1984). El análisis es el siguiente: si se toma la representación distribuida de *taza de café* y se sustrae la representación distribuida de *taza sin café*, el resultado sería la representación de *café*. Según el conexionismo, estas representaciones consisten en patrones de activación donde las unidades activas mantienen una correspondencia con las microcaracterísticas propias de la descripción del objeto. La representación de *café* resultante es un vector considerado como un símbolo en un contexto determinado, y cuyas unidades subconceptuales (su estructura interna) cambian de acuerdo con dicho contexto. Por ejemplo, se trataría de un líquido marrón, caliente y oloroso, pero en el contexto de la taza: sus bordes se amoldan a la forma de la taza y su superficie de porcelana. Si la representación inicial fuera *lata de café*, la representación resultante sería un polvo marrón oloroso cuyos bordes se amoldan a la forma de la lata cilíndrica. Así, “la representación compuesta se construye a partir del *café extraído* de la situación *taza de café*, junto con *taza* extraída de la situación *taza de café*, junto con su interacción” (1987, p. 147). Ello conforma una familia de patrones de actividad que solo mantienen un parecido de familia entre sí. Por tanto, las representaciones distribuidas complejas poseen estructura constituyente a nivel interior, aunque de manera aproximada, y son susceptibles a mecanismos inferenciales que extraen una parte del todo. La estructura constituyente se encuentra dentro, y no entre, las entidades.

No obstante, a pesar de que esta relación de constitución sea importante para analizar el comportamiento del modelo, no forma parte de su mecanismo causal: el procesamiento del vector no incluye descomponerlo en sus elementos constituyentes, sino que se debe considerar las relaciones entre entidades (actividades numéricas

individuales)⁶¹. Además, la descomposición de una representación compleja en sus constituyentes no es única, porque la noción de constitución pertenece a la computación simbólica que, aunque útil, es solo aproximada. Si bien las representaciones distribuidas poseen estructura, no implementan exactamente un lenguaje del pensamiento simbólico. La dependencia del contexto, la consideración de las interacciones entre entidades, la falta de identificación unívoca de los constituyentes y la noción de parecido de familia entre las representaciones de un mismo concepto como CAFÉ indican que la estructura constituyente conexionista no implementa literalmente una estructura constituyente sintáctica.

Por otra parte, Smolensky sostiene que se escapa de los objetivos del conexionismo de la época el tratamiento de los procesos mentales como sensibles a la estructura de las representaciones. A pesar de ello, el autor afirma que la resolución de la paradoja de la cognición (véase pp. 111-112) puede incluir que los procesos analizan las regularidades estadísticas de manera sensible a la estructura de la información representada en patrones numéricos distribuidos.

En suma, el conexionismo es capaz de dar cuenta de las dos afirmaciones centrales del enfoque simbólico. La composicionalidad conexionista permite instanciar ambas sin adoptar un lenguaje del pensamiento simbólico, y este lenguaje no provee una base exacta formal de las estructuras y procesos. Ambos paradigmas no difieren en cuanto a su aceptación de las dos afirmaciones, sino en cómo las instancian formalmente: el paradigma simbólico las formaliza a través de estructuras sintácticas y manipulación simbólica, y el paradigma subsimbólico por representaciones vectoriales distribuidas, composicionalidad y procesos basados en asociación y sensibles a la estructura.

Un segundo punto en la respuesta de Smolensky señala que la utilidad del conexionismo va más allá de la modelación de procesos de nivel inferior o la implementación de programas convencionales. Es posible, contrario a lo supuesto por Fodor y Pylyshyn (1988), aceptar las afirmaciones de estructura constituyente de los estados y sensibilidad a la estructura por parte de los procesos sin convertirse en una teoría de la implementación. Según Smolensky, el rol del conexionismo radica más bien en el desarrollo de modelos formales basados en matemáticas de sistemas dinámicos, y la producción de conceptos y técnicas para estudiar el desempeño a nivel inferior y superior.

⁶¹ Ello evita que la arbitrariedad de la definición de los constituyentes introduzca ambigüedad en el procesamiento de la red. La ambigüedad existe para quienes estudian el comportamiento del modelo.

Los modelos conexionistas y simbólicos se relacionan por medio del *principio de correspondencia cognitiva*, según el cual

la relación general entre los principios formales conexionistas y los principios formales simbólicos —dado que ambos son instanciaciones de nociones comunes no formales— radica en que, si tomas un análisis de nivel superior de lo que sucede en los sistemas conexionistas, encontrarás que coinciden, hasta cierto punto, con lo que ocurre en el formalismo simbólico (1987, p. 153).

A nivel superior, los sistemas conexionistas muestran propiedades simbólicas emergentes. Así, nociones no formales como memoria, inferencia o constitución son instanciadas por el conexionismo en principios formales basados en matemáticas de sistemas dinámicos. Por ejemplo, la inferencia lógica se instancia como inferencia estadística⁶².

La investigación conexionista presenta diversos límites y desafíos, sobre todo en lo que concierne a los constreñimientos fuertes de la cognición o la representación de estructuras complejas. Sin embargo, a pesar de no tener el poder lógico sintáctico de la computación simbólica, mantiene ciertas ventajas. Smolensky afirma que la opción de modelar los constreñimientos fuertes a partir de los constreñimientos blandos se justifica porque “si logramos construir símbolos y manipulación de símbolos a partir del ‘conectoplasma’, entonces tendremos una explicación de *dónde vienen los símbolos y la manipulación de símbolos*” (1987, p. 141). La teoría, de ser exitosa, explicaría cómo surgen los símbolos a partir de estructuras neuralmente inspiradas, lo que ayudaría a entender el funcionamiento del cerebro, los alcances y límites de los constreñimientos fuertes y de la aproximación simbólica. En consonancia con ello, este modelo muestra una mayor facilidad de implementación en el sistema nervioso y emplea matemáticas cercanas a los sistemas dinámicos físicos.

Las disputas entre Fodor y Smolensky no se detuvieron con los textos revisados. Smolensky (1990, 1995a, 1995b) insistió en la capacidad de las representaciones

⁶² El principio de la correspondencia cognitiva es formulado por Smolensky en analogía con el principio de correspondencia cuántica, porque cumplen un rol similar en la microteoría de la cognición y en la microteoría en física respectivamente. Existen principios físicos compartidos por el formalismo clásico y el cuántico, como el principio de conservación de la energía. Estos se instancian de una manera determinada en el formalismo clásico (que sería considerado de nivel superior) y, a nivel inferior, son reinstanciados en el formalismo cuántico. Sin embargo, el formalismo cuántico no implementa el formalismo clásico, sino que ambos se relacionan, y la mecánica clásica provee un marco aproximado de nivel superior. A pesar de ser aproximada, es importante para la explicación de fenómenos clásicos y para guiar el descubrimiento de principios cuánticos. Todo lo anterior aplicaría también al caso de los modelos cognitivos simbólicos y subsimbólicos.

distribuidas de exhibir estructura, mientras que Fodor (1997b) y Fodor y McLaughlin (1990) insistieron en su incapacidad de exhibir los constituyentes de las representaciones mentales complejas que se requieren para postular procesos sensibles a la estructura. Se escapa de los objetivos de la presente tesis la reconstrucción íntegra de esta disputa. Me limité a la caracterización de la primera parte, puesto que contiene los puntos principales sobre los cuales se articularon las discusiones. Sintetizando la respuesta de Smolensky al respecto, es posible sostener que las representaciones conexionistas poseen estructura a nivel conceptual y subconceptual, y los procesos podrían ser sensibles a ella. Sin embargo, la composicionalidad conexionista no instancia un lenguaje del pensamiento simbólico.

A continuación, me centraré en el último rasgo relacionado con la noción de REPRESENTACIÓN involucrada, a saber, la relación entre representaciones mentales y representaciones neurales.

Relación entre representaciones mentales y representaciones neurales

El nivel subsimbólico formaliza de manera abstracta el procesamiento que ocurre en el sistema nervioso (Smolensky, 1989, p. 237). No obstante, los modelos subsimbólicos no son modelos neurales, y la relación entre ambos puede ser abordada desde un punto de vista sintáctico y semántico. El primero de ellos, referido a la relación entre la arquitectura subsimbólica y la arquitectura cerebral, justifica la visión extendida según la cual el paradigma subsimbólico tiene una directa inspiración neural. En efecto, la operación de sus mecanismos de procesamiento es cercana al nivel neural, y existe una serie de similitudes con las características más generales de la arquitectura del cerebro: las unidades y sus conexiones son similares a las neuronas y sus sinapsis. Los estados de ambas redes se definen por variables numéricas continuas (en el caso del conexionismo, activaciones y fuerzas de conexión, y en la corteza cerebral, potenciales de acción y sinapsis). La interacción entre unidades o entre neuronas es de alta complejidad, y sus parámetros son modificables. Las variables de estado son numerosas y cambian en el tiempo de manera continua. Sin embargo, otras características de la corteza cerebral no son tomadas en consideración, puesto que “el nivel subsimbólico de análisis es mayor que el nivel neural” (1988, p. 10). Las neuronas y sus sinapsis tienen localización espacial, lo que influye en las interacciones entre neuronas lejanas, mientras que las unidades y sus conexiones no la poseen. Existe una topología compleja en las proyecciones distales entre

áreas de la corteza, mientras que, en las redes conexionistas, la topología de las proyecciones distales entre nodos es simple. Por último, la corteza cerebral posee numerosos tipos de señales (las interacciones entre neuronas distales se median por señales discretas, y existe integración compleja en una neurona), mientras que las redes conexionistas cuentan con un único tipo de señal con integración lineal e interacciones no discretas. Cabe destacar que la información disponible en la época sobre el sistema nervioso era mayormente sobre su estructura y no sobre sus procesos dinámicos, por lo que no ayudaba en el desarrollo de modelos neurales de funciones cognitivas como la resolución de problemas o el procesamiento de lenguaje.

Con respecto al ámbito semántico, la relación entre las representaciones entendidas como patrones de actividad y las representaciones neurales no es clara. Dado el desconocimiento que se poseía sobre las representaciones neurales de los dominios cognitivos de nivel superior, la semántica de las unidades de red se encuentra más próxima al nivel conceptual que al nivel neural. En conclusión, dada la cercanía entre el ámbito sintáctico y el nivel neural por un lado, y el ámbito semántico y el nivel conceptual por el otro, el autor concluye que el nivel subconceptual figura entre los niveles neural y conceptual (1988, p. 9). Por ende, el conexionismo no debe ser considerado como un modelo neural —y, de hecho, la arquitectura se ha determinado cada vez más por consideraciones matemáticas en lugar de neurales, según Smolensky.

En el presente subapartado caractericé el TAC propuesto por Paul Smolensky, con un énfasis nuevamente en la relación entre representaciones conceptuales y representaciones perceptuales, entre representaciones mentales y palabras del lenguaje natural, y entre representaciones mentales y representaciones neurales. En relación con la ecología conceptual, las influencias teóricas destacadas incluyen la consideración de la arquitectura conexionista, los sistemas dinámicos, la inferencia estadística y la distinción entre niveles de análisis. Todo ello repercute en la noción de REPRESENTACIÓN resultante, que se entiende como patrones de activación entre unidades o nodos en una red. A continuación, me detendré en el segundo debate donde se puede localizar un cambio en REPRESENTACIÓN en ciencias cognitivas: el sistema de símbolos perceptuales de Lawrence Barsalou.

2.3. Sistema de Símbolos Perceptuales, de Lawrence Barsalou

Las perspectivas de cognición corporizada y *grounded cognition*⁶³ realizan una crítica a las teorías cognitivas clásicas (sobre todo al paradigma simbólico), por la inadecuación de la manipulación de símbolos amodales similares a palabras. Una de las propuestas fundacionales en este marco es el SSP de Lawrence Barsalou (1999).

Barsalou es crítico con respecto a la separación entre cognición y percepción sostenida por la modularidad de la mente de Fodor (1983, 1985) (véase pp. 101-104) y mantenida por el paradigma subsimbólico (véase pp. 115-116). De acuerdo con Barsalou, a la base de esta separación se encuentra la concepción de que las representaciones cognitivas son amodales: sus estructuras internas “no mantienen correspondencia con los estados perceptuales que las produjeron” (Barsalou, 1999, p. 578). Se implementan en sistemas físicos distintos, usan esquemas representacionales diferentes y son gobernados por principios diferentes. Ello conduce a que las representaciones amodales y los estados perceptuales mantengan una relación arbitraria o convencional, lo que se vincula con la influencia del desarrollo de formalismos y lenguajes de programación en lógica y ciencias de la computación.

Los sistemas de símbolos amodales han sido ampliamente empleados dadas sus funciones simbólicas como la representación de conceptos y proposiciones, así como su capacidad de combinar símbolos productivamente y producir inferencias. Sin embargo, Barsalou (1999) indica un número importante de problemas con esta aproximación. En primer lugar, señala la falta de evidencia empírica en psicología y neurociencias que apoye la existencia de símbolos amodales, y la amplia evidencia que apoya el carácter perceptual. Se ha determinado, por ejemplo, que los símbolos conceptuales tienen carácter perceptual (Glaser, 1992) o que el conocimiento categorial está enraizado en regiones sensoriomotoras del cerebro (Damasio, 1989)⁶⁴. En segundo lugar, la capacidad

⁶³ Barsalou (2008) identifica su teoría dentro del *grounded cognition* y la diferencia del enfoque general de la cognición corporizada. Para él, la primera postura sostiene que “la cognición se encuentra típicamente fundamentada de múltiples modos, incluyendo simulaciones, acción situada y, en ocasiones, estados corporales” (2008, p. 619). Por su parte, la segunda sugiere, incorrectamente, un foco exclusivo en los estados corporales como fundamento (*ground*) de la cognición. No obstante, el *grounded cognition* a menudo se cuenta dentro del enfoque de la cognición corporizada.

⁶⁴ Las zonas de convergencia de Damasio (1989) son empleadas por Barsalou para ejemplificar el proceso de simulación (véase pp. 128-129) por medio del cual se explica el procesamiento conceptual. Las zonas de convergencia son un conjunto de áreas de asociación que se encargan de integrar información de distintas modalidades. Las áreas asociativas locales de las regiones sensoriomotoras “capturan patrones de

representacional de los símbolos amodales en la implementación de sistemas conceptuales es desafiada por las dificultades que tiene para representar algunas funciones computacionales o conocimientos, como es el caso del conocimiento espaciotemporal (Clark, 1997b). En tercer lugar, los sistemas de símbolos amodales poseen al menos dos problemas conceptuales. El proceso de transducción entre los estados perceptuales y los símbolos amodales carece de una explicación satisfactoria y, más aún, no se encuentra respaldado por evidencia neural. De manera similar, tampoco existe una explicación satisfactoria de cómo los símbolos amodales se vinculan a estados perceptuales y entidades del mundo, es decir, no se da cuenta de su fundamento (sin embargo, véase pp. 146). En cuarto lugar, no queda claro cómo el sistema puede tener una comprensión en los casos de entidades o eventos ausentes o futuros, puesto que, en ausencia de referentes físicos, el entorno no otorga fundamento a los símbolos. En quinto lugar, cuando son postuladas para explicar estas entidades o eventos, las representaciones amodales son redundantes, por lo que las teorías amodales no son parsimoniosas. Por ejemplo, una solución al problema del comportamiento anticipativo consiste en postular representaciones perceptuales que medien entre los símbolos amodales y expliquen el fundamento y la relación con eventos u objetos ausentes. No obstante, la opción no es satisfactoria porque las representaciones perceptuales hacen todo el trabajo: las representaciones perceptuales guardadas en la memoria de largo plazo son activadas durante la transducción, lo que activa a su vez el símbolo amodal asociado. Luego, la activación del símbolo amodal activa las memorias perceptuales asociadas, lo que da fundamento a la comprensión. Por último, los sistemas de símbolos amodales pueden proveer explicaciones *a posteriori* o *post hoc* (Anderson, 1978), pero no son capaces de una predicción *a priori* de los fenómenos (sobre todo perceptuales) tales como la rotación en la imagería mental (Kosslyn, 1980).

A raíz de las limitaciones anteriores, Barsalou opta por una teoría perceptual del conocimiento. En la literatura, se ha considerado que las teorías perceptuales no poseen el poder expresivo suficiente para implementar un sistema conceptual íntegro en comparación con las teorías amodales, dado que son concebidas como sistemas de grabación. Los sistemas de grabación son sistemas que capturan información física por medio de la creación de copias atenuadas de ella, sin interpretar cada una de sus partes o

representación perceptual” y luego “reactivan estas representaciones sensoriomotoras para simular la experiencia y apoyar así el procesamiento cognitivo” (Barsalou, 1999, p. 637).

elementos, como en el caso de los videos o las fotografías. Un sistema conceptual, en cambio, interpreta las entidades en la grabación, asociando instanciaciones específicas en la percepción “al conocimiento de tipos generales de cosas en la memoria (es decir, a los conceptos)” (1999, p. 581). Con ello permite ir más allá del *input* perceptual, por lo que es inferencial. Además, es productivo: es capaz de construir conceptos complejos desde conceptos simples. También permite la formulación de proposiciones mediante el enlace de un concepto a una instanciación de manera que sea susceptible de valor de verdad. Barsalou sostiene que es posible realizar una teoría perceptual del conocimiento que exhiba estas propiedades, y su línea argumental puede resumirse en los siguientes puntos:

- a. La cognición y la percepción comparten sistemas neurales.
- b. Los sistemas sensoriomotores representan tanto las entidades percibidas en su presencia como las conceptualizaciones de estas en su ausencia.
- c. Por tanto, es posible implementar un sistema conceptual empleando mecanismos sensoriomotores.

De acuerdo con el SSP, la cognición y la percepción comparten sistemas tanto cognitivos como neurales. En particular, los estados perceptuales nacen en sistemas sensoriomotores y subconjuntos suyos son extraídos por medio de atención selectiva y almacenados en la memoria de largo plazo. Allí, funcionan como representaciones que subyacen a la cognición: son símbolos que pueden ser manipulados y representar referentes del mundo. Un símbolo perceptual corresponde al registro de un aspecto limitado de los estados neurales que subyacen a la percepción, por lo que se entiende como una representación neural esquemática que opera primariamente a nivel inconsciente. El proceso de formación implica las siguientes etapas⁶⁵:

⁶⁵ Ello contrasta con el proceso de creación de símbolos en el enfoque simbólico. En él, los estados perceptuales surgen en sistemas sensoriomotores, pero un subconjunto suyo es transducido a un sistema representacional que no es perceptual. Los símbolos amodales resultantes son incluidos en estructuras representacionales mayores tales como esquemas, listas y redes semánticas. Estas estructuras constituyen un sistema simbólico funcional con sintaxis y semántica combinatoria, que lleva a cabo las funciones cognitivas de nivel superior (Barsalou, 1999, p. 578). Los símbolos amodales mantienen una relación estrecha con el lenguaje (véase pp. 104-108). En el caso del enfoque subsimbólico, es posible mantener la diferenciación entre la primera capa (entendida como un sistema perceptual) y las capas de unidades ocultas (entendidas como sistema conceptual), así como la relación de arbitrariedad entre los *inputs* perceptuales y las representaciones conceptuales. Los pesos iniciales de los enlaces entre las unidades de *inputs* y las unidades ocultas tienen valores aleatorios por lo que, luego del aprendizaje de la red, las representaciones conceptuales resultantes se relacionan arbitrariamente con los estados perceptuales que las originaron. Sin embargo, los esquemas conexionistas pueden ser distintos (no amodales) si se representara tanto la información cognitiva como la información perceptual en la misma red asociativa.

1. Los sistemas de neuronas de las regiones sensoriomotoras capturan información del cuerpo y el ambiente durante la experiencia perceptual.
2. El sistema cognitivo focaliza la atención selectiva en un aspecto limitado de la percepción.
3. El sistema cognitivo almacena la información seleccionada en la memoria de largo plazo, creando símbolos perceptuales esquemáticos.
4. Los símbolos perceptuales relacionados se organizan en simuladores que permiten construir simulaciones ilimitadas de un objeto o evento en su ausencia.
5. Los simuladores controlan la atención durante la simulación, lo que tiene consecuencias semánticas —por ejemplo, enfocarse en distintos puntos de una misma representación espacial origina la distinción entre *arriba*, *abajo*, *adelante* y *atrás*.
6. Las representaciones almacenadas son recuperadas en un proceso que también es entendido como simulación perceptual, puesto que es “un intento por simular la entidad o evento original” (1999, p. 605).
7. Luego de su almacenamiento, los símbolos funcionan de manera dinámica: las conexiones y activaciones del patrón original se alteran por cambio en el contexto o almacenamiento posterior de otros símbolos en la misma área asociativa.

Los simuladores son entendidos como colecciones de memorias de modalidad específica. Se componen de (i) un marco que integra los símbolos, definido como “un sistema integrado de símbolos perceptuales que es empleado para construir simulaciones específicas de una categoría” (1999, p. 590), y (ii) un conjunto de simulaciones potencialmente infinitas que pueden construirse a partir del marco. Con respecto a (i), los marcos contienen símbolos previos de objetos u eventos categorialmente relacionados, es decir, condensan información multimodal de la experiencia perceptual vinculada con ellos. Incluyen, de manera separada, información espacial y de contenido de las subregiones representadas, donde los contenidos son representados como especializaciones. Cada subregión posee diversas especializaciones y, en el caso de los eventos, se considera el tiempo. Cuando se percibe un segundo objeto o evento similar, se produce un recuerdo y una recuperación del conjunto de símbolos del objeto o evento previo. Este conjunto de símbolos recuperado guía el procesamiento del objeto o evento presente, lo que puede conducir a la extracción de información de contenido de otras partes para relacionarlas con las mismas subregiones que el contenido anteriormente extraído, o bien al procesamiento de otras subregiones para establecer nuevos símbolos

perceptuales. Las conexiones que se presentan en el marco son dinámicas: cambian frente a nuevas instancias, fortaleciéndose ante el procesamiento repetido, debilitándose con el tiempo, etcétera. Los simuladores son entendidos como los conceptos, esto es, el conocimiento y procesos que permiten la representación adecuada de entidades y eventos (1999, p. 587). Por otro lado, una conceptualización es una simulación específica de la entidad o el evento en cuestión. Los simuladores pueden implementar las funciones de un sistema conceptual una vez establecidos. Me detendré en ello en el subapartado *Relación entre representaciones mentales y palabras del lenguaje natural* (pp. 131-135).

En relación con (ii), un subconjunto determinado de la información contenida por el marco se activa para construir simulaciones específicas en la memoria de trabajo. La simulación es “la reenacción de los estados perceptuales, motores e introspectivos adquiridos durante la experiencia con el mundo, el cuerpo y la mente” (2008, p. 618). Se activa la representación volumétrica del objeto, donde la inclusión o exclusión de subregiones depende de la superficialidad o complejidad de la simulación. En cada subregión se activan las especializaciones que estén más asociadas a la subregión, a otras regiones activas y a otras especializaciones activas. En el caso de los eventos, las subregiones y sus especializaciones pueden cambiar en el tiempo. Durante la simulación, además de la recuperación de la información contenida en el marco, es posible transformar dicha información: fragmentarla, trasladarla en espacio o tiempo, etcétera.

Dada la operación de la atención selectiva, tanto los marcos como las simulaciones son más pobres que las percepciones que los originaron. Las simulaciones son parciales, sesgadas y distorsionadas. Por ejemplo, ante una serie de puntos organizados en fila, se percibe una línea. El sistema cognitivo desarrolla, desde la infancia, simuladores para eventos y objetos que son relevantes en el entorno. En su establecimiento y ejecución influyen factores tales como las predisposiciones genéticas para el procesamiento del espacio o el movimiento, y el ambiente determinado en el que se encuentra el individuo⁶⁶.

El surgimiento inicial en los sistemas sensoriomotores confiere a los símbolos perceptuales una naturaleza modal y analógica. Es modal porque se representan en los mismos sistemas que los estados perceptuales que los produjeron, y es analógica porque

⁶⁶ Los simuladores mantienen una relación con los modelos mentales propuestos por Craik y mencionados anteriormente (pp. 87-88). La diferencia radica en que, mientras que los simuladores tienen dos niveles de estructura (es decir, el marco y el conjunto de simulaciones potencialmente ilimitadas), los modelos mentales serían equivalentes a las simulaciones de objetos o eventos específicos.

su estructura mantiene una correspondencia con dichos estados⁶⁷. Más precisamente, estos son multimodales (puesto que se originan en todos los módulos de experiencia percibida), y neurales (distribuidos en áreas cerebrales de modalidad específica). Junto con operar para toda modalidad, aplica a la experiencia introspectiva, que incluye estados emocionales, representacionales y operaciones cognitivas como la transformación y la comparación. Su dependencia de los datos sensoriomotores resalta la corporización de los símbolos como característica importante. Esta corporización introduce diversas variaciones. Por ejemplo, un mismo concepto puede representarse de manera distinta pues sus símbolos perceptuales se derivan de entornos distintos, y las diferencias fenotípicas llevan a realizar discriminaciones diversas para conformar los símbolos. Todo ello posee ventajas adaptativas, porque permite adaptar el sistema conceptual humano a los ambientes específicos. Para Barsalou, la importancia de la corporización problematiza la perspectiva funcionalista:

El funcionalismo implica que el sistema computacional que subyace a la inteligencia humana puede entenderse de manera independiente del cuerpo humano. Además, implica que el mismo sistema básico de símbolos funciona en todos los seres humanos normales, al margen de su idiosincrasia biológica. Al igual que los *software* computacionales pueden caracterizarse independientemente del *hardware* concreto que los implementa, el sistema simbólico humano puede caracterizarse independientemente del sistema biológico que lo encarna (Barsalou, 1999, p. 598).

Luego de la presentación general del SSP, me enfocaré a continuación en la relación que se establece entre (i) representaciones conceptuales y representaciones perceptuales (pp. 130-131); (ii) representaciones mentales y palabras del lenguaje natural (pp. 131-135), y (iii) representaciones mentales y representaciones neurales (pp. 135-136).

Relación entre representaciones conceptuales y representaciones perceptuales

De acuerdo con Barsalou (1999), la percepción y la cognición comparten sistemas neurales y cognitivos, por lo que la demarcación entre ambas es inválida. Ello conduce a la problematización de la impenetrabilidad cognitiva de la percepción sostenida por Fodor (1983, 1985). Barsalou afirma que los sistemas sensoriomotores son penetrables por la

⁶⁷ Sin embargo, el autor aclara que la correspondencia no es entre los símbolos perceptuales y el mundo físico, sino entre los símbolos perceptuales y los estados perceptuales que les dieron origen. Si bien puede existir una correspondencia con el mundo físico en ocasiones, existen excepciones (por ejemplo, con cualidades secundarias como el color).

cognición, esto es, la cognición puede afectar su contenido. Un ejemplo de ello se halla en la literatura de la imaginación mental, donde la imaginación visual activa la corteza visual primaria sin que se presente un *input* físico. Por tanto, la ejecución de un simulador involucra la ejecución parcial *top-down* de estos sistemas sensoriomotores. No obstante, la penetrabilidad de la información *top-down* ocurre cuando la información *bottom-up* está ausente. Cuando se encuentra presente, pueden ocurrir dos situaciones: si ambas son compatibles, el procesamiento *top-down* penetra y complementa la información *bottom-up*. Si ambas entran en conflicto, la información *bottom-up* suele dominar (excepto, por ejemplo, en estados alucinatorios o similares).

Estas representaciones sensoriomotoras son empleadas por un conjunto de procesos cognitivos básicos —percepción, imaginación y concepción— y otros procesos no básicos —anticipación perceptual, interpretación y otras—. En la percepción, los mecanismos representacionales representan objetos físicos por un procesamiento *bottom-up*. En la imaginación y la concepción, en cambio, tales mecanismos representan objetos en su ausencia por medio de procesamiento *top-down*. Por su parte, en procesos no básicos, se fusionan ambos procesamientos. Por ejemplo, en la anticipación perceptual, se emplea experiencia de entidades percibidas en el pasado para simular una actividad futura, mientras que, en la interpretación, la representación conceptual de un estímulo sesga el procesamiento sensorial. El procesamiento *top-down* y el *bottom-up* se coordinan para establecer percepciones relevantes.

Dado que los sistemas sensoriomotores representan tanto las entidades percibidas en su presencia como las conceptualizaciones de estas en su ausencia, es posible sostener que las representaciones perceptuales y las representaciones conceptuales mantienen un vínculo neural y cognitivo. No existe una diferencia de procesamiento entre ellas como en la modularidad fodoriana. En el próximo subapartado, me centraré en el segundo rasgo relacionado con la noción de REPRESENTACIÓN, a saber, la relación entre representaciones mentales y las palabras del lenguaje natural.

Relación entre representaciones mentales y palabras del lenguaje natural

Con respecto a la relación entre representaciones mentales y palabras, Barsalou sostiene que

Los símbolos amodales guardan una importante relación con las palabras y el lenguaje. (...) Se asume que el pensamiento simbólico es análogo al lenguaje en

muchos aspectos importantes. Al igual que el procesamiento del lenguaje implica el procesamiento secuencial de las palabras en una oración, se asume que el procesamiento conceptual implica el procesamiento secuencial de los símbolos amodales en estructuras similares a las de una lista o una oración (1999, p. 579).

De acuerdo con Barsalou, el enfoque simbólico suele emplear formas lingüísticas para expresar símbolos amodales y establece una analogía entre el pensamiento simbólico y el lenguaje. Ello justifica la relación arbitraria que existe entre los símbolos amodales y los estados perceptuales que los originaron: “al igual que las palabras suelen tener relaciones arbitrarias con las entidades del mundo, los símbolos amodales tienen relaciones arbitrarias con los estados perceptuales” (Barsalou, 1999, p. 578). El SSP, en cambio, postula que los símbolos son multimodales, y que los símbolos lingüísticos se desarrollan junto con los símbolos perceptuales que se encuentran asociados con ellos. Según esta perspectiva, desde la infancia se desarrollan simuladores que representan entidades y eventos, lo que es anterior a la adquisición de lenguaje. Luego de su adquisición, se asocian las palabras aprendidas con los simuladores relevantes, se construyen simulaciones nuevas a partir de las preferencias (*utterances*) de otros y se construyen preferencias para transmitir simulaciones a otros.

Un símbolo lingüístico no es un símbolo amodal ni se acompaña de un símbolo amodal. Se desarrolla de manera análoga a los símbolos perceptuales, siendo “una memoria esquemática de un evento percibido, donde el evento percibido es una palabra hablada o escrita” (1999, p. 592). Nuevamente, en este caso opera la atención selectiva, y se extraen memorias esquemáticas de estados perceptuales que luego son integradas en simuladores que producen simulaciones de estas palabras. Tras el desarrollo de los simuladores para las palabras, estos se asocian con simuladores para las entidades y eventos a los que se refieren (entidades o eventos completos, o bien subregiones y especializaciones). Muchas simulaciones de palabras se asocian en el simulador de un concepto para “producir un campo semántico que refleje el campo conceptual subyacente” (1999, p. 592). Al reconocer una palabra, el sistema conceptual activa el simulador del concepto para simular un referente. Así, la comprensión del lenguaje es “la construcción de una simulación perceptual para representar el significado de una preferencia o texto” (1999, p. 605).

Los simuladores pueden implementar las funciones de un sistema conceptual una vez establecidos. Entre estas funciones se cuenta la realización de categorizaciones e inferencias categoriales, y funciones simbólicas tales como la productividad, el

establecimiento de proposiciones y la representación de conceptos abstractos. Con respecto a la realización de categorizaciones e inferencias categoriales, la pertenencia de una entidad a una categoría específica tiene lugar si el simulador para la categoría es capaz de producir una simulación exitosa de esa entidad. A diferencia de las teorías amodales, “el conocimiento que determina la categorización está representado aproximadamente de la misma manera que las entidades percibidas que deben ser categorizadas” (1999, p. 587). Las categorizaciones almacenan una simulación de la entidad categorizada y, ante la percepción de una entidad similar, se asigna a la categoría dada su concordancia con la simulación existente. Si se percibe una entidad que no concuerda, se construye una simulación nueva. Asimismo, las inferencias categoriales surgen por simulación. El conocimiento asociado con una categoría permite realizar predicciones acerca de las entidades categorizadas. Un simulador puede simular información que sobrepasa lo percibido en una entidad categorizada, u operar en ausencia de ella, dado que posee un amplio conocimiento multimodal sobre la categoría. Así, las simulaciones pueden anticipar rasgos de las entidades como su estructura o comportamiento.

Por otro lado, en relación con la implementación de funciones simbólicas, los simuladores poseen productividad: son capaces de formar un número ilimitado de simulaciones novedosas perceptualmente complejas a partir de recursos finitos por medio de la combinación y recursividad. Por ejemplo, la simulación de una relación espacial (arriba, abajo, a la izquierda, etcétera) entre dos objetos puede combinarse con una cantidad potencialmente ilimitada de objetos distintos entre los cuales se establece dicha relación. La recursividad viene dada por la capacidad de especializar una especialización de una región, lo que puede replicarse indefinidamente produciendo, en principio, un número ilimitado de simulaciones.

En segundo lugar, los simuladores pueden formular proposiciones con respecto a situaciones percibidas. Los sistemas de símbolos perceptuales establecen relaciones tipo-instanciación entre conceptos e individuos percibidos al vincular un simulador (tipo) con el individuo (instanciación) a través de la simulación. Además, este mapeo tipo-instanciación implícitamente constituye una proposición del tipo *Es cierto que el individuo percibido es X*⁶⁸. Una misma situación es plasmada de diferentes maneras en

⁶⁸ En esta perspectiva, una proposición verdadera corresponde a un enlace exitoso entre los simuladores y los individuos percibidos, mientras que una proposición falsa es un vínculo no exitoso entre ambos. Una proposición negativa ocurre “cuando la ausencia explícita de algo en una simulación se corresponde con una ausencia análoga en una situación percibida” (Barsalou, 1999, p. 596).

dependencia de la selección de aspectos específicos de la misma, o bien de los mismos aspectos contruidos de distinta forma. Es posible seleccionar y vincular una cantidad ilimitada de aspectos, por lo que se genera un número potencialmente ilimitado de proposiciones que describan la situación.

En tercer y último lugar, los simuladores son capaces de representar conceptos abstractos. Para su representación, se emplean tres mecanismos: *framing*, selectividad y símbolos introspectivos. Primero, se identifican las secuencias de eventos en los que se enmarca el concepto abstracto. Luego, la atención selectiva se enfoca en el contenido central del concepto. Finalmente, se especifican los símbolos introspectivos asociados. Por ejemplo, la representación del concepto de VERDAD se puede ejemplificar de la siguiente manera: el agente construye una simulación perceptual tras escuchar *Hay un globo sobre una nube*. Tras ello, percibe una situación física relacionada con este enunciado y mapea la simulación perceptual en ella. Al final, evalúa si la simulación representa adecuadamente la situación física. De ser el caso, el agente puede sostener *Es verdad que hay un globo sobre una nube*. Así, el concepto VERDAD es enmarcado por una secuencia de eventos; la atención selectiva se enfoca la parte central de la situación relacionada con el concepto, a saber, que la simulación mapea de forma correcta la situación, y los eventos introspectivos tales como el proceso de comparación entre la simulación y la situación son relevantes. Por tanto, en esta perspectiva, los conceptos abstractos son perceptuales. Su diferencia con los conceptos concretos radica en que se apoyan en mayor grado en configuraciones complejas de información multimodal distribuida en el tiempo.

Cabe destacar que incluso la manipulación formal de símbolos puede explicarse apelando a las simulaciones. En relación con la lógica y la matemática, se simulan símbolos arbitrarios. Por ejemplo, la experiencia perceptual con operaciones tales como el *modus ponens* y los símbolos involucrados permite desarrollar la capacidad para simular el procedimiento: simular las premisas seguidas de la conclusión, y simular la sustitución de las variables por constantes (1999, p. 606).

En conclusión, el SSP postula símbolos multimodales que problematizan la relación arbitraria entre los símbolos amodales y las palabras del lenguaje natural. Además, conforma un sistema conceptual que es capaz de implementar funciones simbólicas como la productividad, el establecimiento de proposiciones y la representación de conceptos abstractos. En el próximo subapartado me enfocaré en el

último rasgo relacionado con la noción de REPRESENTACIÓN propuesta por Barsalou, a saber, la relación entre representaciones mentales y representaciones neurales.

Relación entre representaciones mentales y representaciones neurales

La propuesta de Barsalou (1999) apela a la implementación neural como parte de su teoría. Afirma que la percepción y la cognición poseen bases neuroanatómicas que, aunque no son idénticas, comparten mecanismos representacionales. Ello lo conduce a sostener que “la cognición es inherentemente perceptual, compartiendo sistemas con la percepción tanto a nivel cognitivo como neural” (1999, p. 577). Para el autor, esta afirmación tiene implicancias en neurociencias. Dado que existe un conocimiento mayor sobre cómo se implementa la percepción en el cerebro en comparación con el conocimiento sobre la cognición, sostener que la percepción subyace a la cognición permitiría que descripciones neurales de los colores, las formas, entre otros, proporcionen información sobre la representación conceptual de esta información.

En la teoría perceptual del conocimiento sostenida por Barsalou, los símbolos perceptuales son entendidos como representaciones neurales. Por tanto, a diferencia de las representaciones del enfoque simbólico, puede sostenerse una identidad entre representación mental y representación neural. Esta identidad sobrepasa la inspiración neural presente en las representaciones distribuidas propias del conexionismo. Por ejemplo, como mencioné anteriormente, Smolensky (1988) sostiene que el nivel de análisis de su aproximación es un punto intermedio entre los modelos simbólicos y los neurales. No obstante, a pesar de que Barsalou apela a mecanismos neurales en común en su abordaje de la percepción y la cognición, afirma que

Las propiedades de esta teoría (...) no se basarán en mecanismos neurales específicos. En su lugar, esta formulación de la teoría debe considerarse como una explicación funcional de alto nivel de cómo el cerebro podría implementar un sistema conceptual utilizando mecanismos sensoriomotores (1999, p. 582).

Por ende, se enfoca en la plausibilidad de formular una teoría del conocimiento perceptualmente basada que sea capaz de implementar funciones simbólicas, y no en el fundamento de esta teoría en sistemas neurales (que correspondería a una tarea posterior).

En el presente subapartado caractericé el SSP de Lawrence Basalou, con el hincapié en los tres rasgos relevantes para la noción de REPRESENTACIÓN: la relación entre

representaciones conceptuales y perceptuales, entre representaciones mentales y palabras del lenguaje natural y entre representaciones mentales y neurales. Para el autor, las representaciones (en particular, los símbolos perceptuales) son neurales y multimodales. La ecología conceptual en este caso incluye la problematización del funcionalismo y de la tesis de la impenetrabilidad cognitiva de la percepción, junto con la apelación a los aportes de la neurociencia.

Con la revisión del SSP concluye la realización del primer paso definido en el primer capítulo para el análisis de las discontinuidades conceptuales, esto es, la caracterización de los sistemas conceptuales sucesivos. En el próximo apartado me concentraré en el segundo paso: la demostración de existencia de cambio y la determinación del tipo de cambio.

3. Demostración de existencia de cambio conceptual y determinación del tipo

La demostración de la existencia de cambio entre sistemas sucesivos se realiza por medio de la constatación de que el concepto es inexpresable en el vocabulario del sistema anterior, o bien que aparece con modificaciones. Así, el estudio de las elaboraciones teóricas entre la TRM, el TAC y el SSP arroja una línea de CC en la noción de REPRESENTACIÓN.

En primer lugar, la noción de representación ligada al enfoque simbólico sostenido por Fodor (1975, 1987, 1997a) y Fodor y Pylyshyn (1988) corresponde a REPRESENTACIÓN AMODAL. Una representación es un símbolo (un objeto físico) con contenido, poderes causales y estructura constituyente. La amodalidad apunta a la no perceptualidad: un símbolo amodal es un símbolo que no se corresponde con la modalidad perceptual que lo origina. Ello se vincula con la tesis de la impenetrabilidad cognitiva de la percepción, que sostiene que los módulos no son penetrables por la información del sistema de creencias del procesador central. A su vez, esta tesis se enmarca en la modularidad fodoriana, que distingue entre representaciones procesadas por dispositivos de *input* modulares y representaciones propias de los procesadores centrales no modulares. Por otro lado, la estructura constituyente es el núcleo de su similitud con los lenguajes naturales: las representaciones poseen una sintaxis y semántica combinatoria,

y las operaciones mentales aplican a ellas en virtud de su forma. En relación con las representaciones neurales, el fisicalismo no reduccionista de Fodor y su distinción entre un nivel cognitivo y un nivel de implementación justificarían una diferencia con las representaciones mentales. Por ende, las representaciones mentales no son representaciones neurales porque pertenecen al nivel cognitivo, no al nivel de implementación, y una no se reduce a la otra.

El paradigma subsimbólico mantiene los supuestos de representacionalidad y computacionalidad sostenidos por el paradigma simbólico, por lo que los CC entre ambos no son radicales. No obstante, se puede localizar un primer foco de transformación en la incorporación de la noción de REPRESENTACIÓN DISTRIBUIDA. Para el conexionismo, las representaciones son subsímbolos, patrones de activación de las unidades o nodos en una red. Las representaciones de entidades conceptuales de nivel superior se distribuyen en diversos nodos que participan en la representación de muchas entidades. Mientras que el enfoque simbólico utiliza símbolos o estructuras de datos que representan conceptos completos u otras entidades con contenido semántico, el nivel significativo de análisis conexionista es el patrón de activación. Su postura es cercana al modelo de representación distribuida o código de poblaciones en neurociencias, según el cual existe un patrón distribuido de actividad neuronal en distintas estructuras del sistema nervioso que se posiciona como correlato neuronal de conductas o estados cognitivos. Las distinciones entre representaciones conceptuales y representaciones perceptuales se realizan como descripciones aproximadas de nivel superior, y no pertenecen al nivel subconceptual, que es el nivel preferente de análisis del paradigma subsimbólico. Además, se problematiza el empleo del lenguaje como modelo teórico. Las representaciones conexionistas poseen estructura tanto a nivel conceptual como subconceptual, y los procesos podrían ser sensibles a ella. A pesar de ello, la composicionalidad conexionista no instancia un lenguaje del pensamiento simbólico. Por último, la relación entre representaciones mentales y representaciones neurales no es clara, dado el desconocimiento de la época acerca de las representaciones neurales de los dominios cognitivos de nivel superior. Cabe destacar, sin embargo, que los modelos subsimbólicos no son modelos neurales, y sus unidades pueden ser interpretadas como entidades neurales, pero también como entidades matemáticas.

No obstante, para Fodor y Pylyshyn (1988), la apelación a la distribución no es exclusiva del conexionismo. La arquitectura simbólica es neutral con respecto a la localización o distribución de los elementos, y no es incompatible con una distribución

neural siempre que se mantenga la idea de que la computación relevante para la cognición consiste en manipulación de símbolos. Este punto negaría la existencia de un CC incluso moderado en la noción conexionista de REPRESENTACIÓN, porque la distribución no sería una adición ni un cambio al concepto original. Esta afirmación surge a raíz de la consideración de la localización y la distribución como elementos pertenecientes a la teoría de la implementación, que no poseen repercusión directa en el nivel cognitivo. A propósito de esto, cabe plantearse la pregunta sobre la real compatibilidad entre las representaciones simbólicas y las representaciones distribuidas conexionistas. La representación en el enfoque simbólico es un símbolo que tiene contenido (o significado), y la distinción entre símbolo y contenido parece ser relevante para el punto presente. Por ejemplo, de acuerdo con van Gelder (1990), las condiciones requeridas para que una representación sea simbólica son incompatibles con que sea distribuida: es posible representar una estructura simbólica en forma distribuida, pero el contenido se mantiene simbólico. Ello se encuentra en consonancia con la diferencia central que establece Smolensky (1988) entre ambos paradigmas: a nivel sintáctico, los dos se pueden simular mutuamente, pero existe un cambio semántico en el conexionismo con respecto al paradigma simbólico. Para Smolensky, la representación (unidad de interpretación) es el patrón de activación de una red, y el nodo es la unidad de procesamiento. Por tanto, la noción de REPRESENTACIÓN DISTRIBUIDA del conexionismo parece constituir efectivamente un cambio.

Un segundo foco de transformación se encuentra en la noción de REPRESENTACIÓN MULTIMODAL introducida por el SSP de Barsalou (1999). Una representación multimodal es un símbolo que mantiene las características de la modalidad perceptual desde la cual se origina. Estos símbolos son analógicos, porque su estructura mantiene una correspondencia con los estados perceptuales que los produjeron, y neurales, distribuyéndose en áreas cerebrales de modalidad específica. Por ende, las representaciones conceptuales y las representaciones perceptuales comparten un vínculo cognitivo y neural. No existe una diferencia de procesamiento entre ellas, como la postulada por la modularidad fodoriana, o una distinción a nivel conceptual, como la aceptada por el conexionismo. Por otra parte, se problematiza nuevamente el lenguaje como modelo teórico y, en particular, la relación arbitraria que existe entre los símbolos amodales y las palabras del lenguaje natural. Los simuladores pueden implementar las funciones de un sistema conceptual, como la productividad, el establecimiento de proposiciones y la representación de conceptos abstractos. Por último, existe una relación

de identidad entre las representaciones mentales y las representaciones neurales: un símbolo perceptual es una representación neural, lo que es distinto a la inspiración neural conexionista.

Las diferencias expuestas anteriormente justifican la identificación de un CC primero en la incorporación de un carácter distribucional, y posteriormente en la incorporación del formato multimodal. El paso siguiente es identificar el tipo de cambio. Como mencioné en el primer capítulo, es posible condensar las clases de transformaciones siguiendo parcialmente la propuesta de Thagard (1992) de la siguiente manera: creación o adición de conceptos, eliminación de conceptos, reorganización básica, reorganización compleja y metaclasificación. Entre la noción de REPRESENTACIÓN involucrada en la TRM y la noción propia del TAC, el cambio parece corresponder a una reorganización básica, principalmente una descomposición. Las descomposiciones son adiciones o supresiones de relaciones de partes. En el primer capítulo (p. 54) mencioné el ejemplo planteado por Thagard (1992) de la descomposición de ÁTOMO tras el descubrimiento de la existencia de electrones. Ello puede aplicarse a los cambios vistos entre el paradigma simbólico y el paradigma subsimbólico: REPRESENTACIÓN es descompuesta en subsímbolos distribuidos en patrones de actividad. A nivel conceptual de análisis, puede ser interpretada como estructuras mentales propias del paradigma simbólico, pero ello solo ocurre de manera aproximada. Este análisis se encuentra en sintonía con las palabras de Smolensky (1987, 1988, 1989), para quien el paradigma subsimbólico busca, entre otras cosas, refinar al enfoque simbólico por medio del estudio de niveles inferiores de análisis.

Por otra parte, en el SSP, la distinción tajante entre representaciones conceptuales y representaciones perceptuales es rechazada: ambas comparten una base cognitiva y neural. Además, las representaciones mentales y las representaciones neurales mantienen una relación de identidad. Por tanto, nuevamente se puede observar una reorganización básica en la noción de REPRESENTACIÓN y, en particular, una coalescencia. Las coalescencias tienen lugar cuando las entidades que son consideradas como distintas en C1 son subsumidas en un concepto en C2. En el primer capítulo (p. 53) coloqué como ejemplo la coalescencia que identifica Carey (1985) de ANIMAL y PLANTA en el concepto SER VIVO. En el caso presente, REPRESENTACIÓN CONCEPTUAL y REPRESENTACIÓN PERCEPTUAL se subsumen en REPRESENTACIÓN NEURAL.

Luego de establecido el tipo de cambio, el último rasgo relevante en relación con este segundo paso del análisis consiste en la evaluación de la plausibilidad de describir el

resultado del proceso de CC en términos de inconmensurabilidad. Como mencioné en el primer capítulo, la comparación entre el estado inicial y el estado final de las discontinuidades fue clásicamente descrita por medio de la noción de inconmensurabilidad kuhniana. Los cambios estudiados en el caso de REPRESENTACIÓN son principalmente reorganizaciones básicas. Como tal, el nivel de radicalidad no es elevado. Entre REPRESENTACIÓN AMODAL, REPRESENTACIÓN DISTRIBUIDA y REPRESENTACIÓN MULTIMODAL existe una continuidad que niega que un concepto sea inexpresable en términos del sistema conceptual anterior.

El análisis realizado en este capítulo constituye una de las líneas de CC que se pueden estudiar con respecto a una noción central como REPRESENTACIÓN, lo que sugiere que existen otras líneas. Tras concluir el segundo paso estipulado para la realización del análisis, resta determinar el mecanismo central que posibilita el cambio. Abordaré este último tema en el apartado siguiente.

4. Mecanismos que posibilitan el cambio conceptual

En el primer capítulo, sostuve que los procesos de *bootstrapping* son considerados como principales en el impulso de discontinuidades conceptuales (véase pp. 32-39). En este marco, el razonamiento analógico constituye un mecanismo central a la base de las variaciones en la noción de REPRESENTACIÓN. Entre las analogías relevantes se encuentra la denominada *metáfora del computador*, la analogía con el funcionamiento neuronal y la problematización de la analogía con los sistemas de grabación.

La analogía que posibilita la noción de REPRESENTACIÓN AMODAL corresponde a la conocida metáfora del computador, entendiendo *computador* como un dispositivo de manipulación de símbolos. En la TRM, la analogía posee varias implicancias para la hipótesis del lenguaje del pensamiento y el compromiso ontológico con las actitudes proposicionales. Con respecto al lenguaje del pensamiento, se recurre a la analogía (i) para ilustrar la existencia de dispositivos que poseen un lenguaje de computación diferente al lenguaje de comunicación con el medio y (ii) como modelo de la relación entre lenguaje natural y código interno (Fodor, 1975). Para entender la aplicación de la analogía, es preciso identificar dos lenguajes distintos en los computadores: un lenguaje de *input/output* que es empleado para establecer comunicación con el entorno, y un lenguaje de máquina empleado para implementar las computaciones. Entre ambos median compiladores que especifican “bicondicionales cuyo lado izquierdo es una fórmula en el

código de *input/output* y cuyo lado derecho es una fórmula en el código de máquina” (Fodor, 1975, p. 66). Así, el compilador permite determinar las condiciones de verdad de las fórmulas del código de *input/output*. Sin embargo, no se requieren compiladores para el lenguaje de máquina, porque sus fórmulas mantienen correspondencia con los estados físicos y procesos relevantes computacionalmente, es decir, el computador está construido para usar ese lenguaje. Una situación análoga se puede establecer entre el lenguaje natural aprendido y el código interno: para emplear el lenguaje adquirido, se requiere traducir sus fórmulas a fórmulas que se correspondan con los estados físicos relevantes computacionalmente, esto es, a fórmulas del código interno.

En segundo lugar, con respecto al compromiso ontológico con las actitudes proposicionales, los computadores son descritos por Fodor como “una solución al problema de la mediación entre las propiedades causales de los símbolos y sus propiedades semánticas” (Fodor, 1987, p. 19). Los estados mentales tienen propiedades semánticas —se pueden individualizar por su contenido y evaluar por su relación con el mundo— y poderes causales —son causantes de la conducta y son causados por eventos del entorno o por otros estados mentales—. Las relaciones causales entre actitudes proposicionales respetan sus relaciones de contenido. Por ejemplo, inferencias tales como *Si Luis cree que Gladys aprobó el curso, entonces cree que alguien aprobó el curso* (si cree que $F(a)$, entonces cree que $\exists(x) F(x)$) pueden realizarse precisamente porque los contenidos se vinculan con estados mentales causalmente relacionados. Los computadores permiten entender la conexión entre propiedades semánticas y propiedades causales de los símbolos por medio de la sintaxis. Los símbolos interactúan causalmente en virtud de sus estructuras sintácticas, y el computador ofrece un entorno en el que la sintaxis determina el rol causal respetando la semántica. Dado que tener actitudes proposicionales implica mantener una relación con las representaciones, la metáfora permite el desarrollo de una teoría que explique las relaciones de contenido entre pensamientos causalmente relacionados. Para ello, se requiere adoptar un compromiso con la existencia de símbolos mentales con propiedades semánticas y sintácticas. En suma, la analogía con los computadores es central para el enfoque simbólico. Para Fodor y Pylyshyn (1988), va más allá de una mera metáfora: es una hipótesis empírica literal.

Con respecto a la REPRESENTACIÓN DISTRIBUIDA, la analogía central radica en ciertas propiedades del funcionamiento neuronal. El modelo de redes se conforma de unidades de procesamiento con valores numéricos de activación y patrones de conectividad entre ellas que poseen pesos. La relación entre la arquitectura subsimbólica

y la arquitectura cerebral justifica la visión según la cual el conexionismo tiene una inspiración neural. Se consideran algunas propiedades de las neuronas, sus sinapsis y los estados de las redes. Existen también diferencias entre ambas: no se consideran numerosos rasgos de la arquitectura cerebral y, a nivel semántico, representaciones distribuidas conexionistas y representaciones neurales no poseen una relación clara. El nivel subsimbólico de análisis es mayor que el nivel neural: considera detalles generales de la corteza cerebral en el ámbito sintáctico, y es cercana al nivel conceptual en el ámbito semántico. No obstante, el establecimiento de una analogía no toma en consideración todos los elementos de los dos dominios: se emplean algunas similitudes relacionales para realizar las inferencias. En última instancia, si bien los modelos subsimbólicos no son modelos neurales, formalizan de manera abstracta el procesamiento ocurrido en el sistema nervioso y aspiran a la explicación del surgimiento de los símbolos a partir de estas estructuras con inspiración neural. Ello ayuda, a su vez, a un mayor entendimiento del funcionamiento cerebral.

Por último, en el caso de la REPRESENTACIÓN MULTIMODAL, el rol de las analogías como mecanismos que impulsan CC consiste en su problematización en lugar de ser una fuente de caracterizaciones. La analogía relevante comprende los sistemas de grabación, como las fotografías o los videos. Como mencioné anteriormente, consisten en sistemas que crean copias débiles de la información física que capturan. Entender las teorías perceptuales del conocimiento como sistemas de grabación conlleva a considerar que carecen del poder expresivo suficiente para implementar un sistema conceptual. Los sistemas conceptuales poseen poder inferencial, productividad y pueden formular proposiciones. La postura de Barsalou sostiene que es posible una teoría perceptual del conocimiento que exhiba estas propiedades, por lo que cuestiona la analogía con los sistemas de grabación.

En el presente apartado llevé a cabo el último paso del análisis propuesto por la tesis, esto es, la determinación de un mecanismo importante para las transformaciones conceptuales en la noción de REPRESENTACIÓN. La visión general que se deriva apunta a la existencia de reorganizaciones básicas no radicales en este concepto, plasmadas en la incorporación de distribución y de multimodalidad. Además, un mecanismo importante que influye directamente en las reorganizaciones es el empleo de analogías, como la metáfora del computador, la analogía con el funcionamiento neuronal y la problematización de la

analogía con los sistemas de grabación. Antes de establecer las conclusiones, en el próximo apartado abordaré dos puntos que subyacen a las distinciones observadas.

5. Discusión

Sintetizando las definiciones de REPRESENTACIÓN que se desprenden de los trabajos anteriormente mencionados, se puede sostener que:

- a. En el enfoque simbólico representado por la TRM, las representaciones se entienden como símbolos amodales con contenido, poderes causales y estructura.
- b. En el enfoque subsimbólico representado por el TAC, las representaciones corresponden a patrones de activación entre unidades o nodos en una red.
- c. En el SSP, las representaciones son consideradas como símbolos neurales multimodales.

Las distinciones en las definiciones mencionadas dependen de la adopción de una serie de supuestos que pueden derivarse de la consideración de elementos relevantes de la ecología conceptual. En la articulación de los sistemas conceptuales sucesivos que propuse en el segundo apartado, diferencié las posturas de cada sistema con respecto a la relación entre representaciones conceptuales y perceptuales, representaciones mentales y palabras del lenguaje natural, y representaciones mentales y neurales. En particular, la amplitud o restricción en la noción de cognición define en buena medida las distinciones entre representaciones conceptuales y perceptuales, mientras que la adopción o rechazo de la realizabilidad múltiple determina la relación con el nivel de implementación (por ejemplo, neural) y con las palabras.

En primer lugar, como mencioné en el segundo capítulo, las teorías alternativas o no clásicas en ciencia cognitiva impulsaron una serie de discusiones sobre los límites de la cognición. Estas discusiones se observan, en parte, en el análisis que realiza Akagi (2007) acerca de la transformación progresiva que presenta esta noción. El autor observa una reorganización básica vista en la diferenciación entre una cognición superior o *highfalutin* y una cognición *inclusiva* que incorpora mecanismos sensoriomotores, afectivos e incluso artefactos u otros individuos. El enfoque simbólico de Fodor se vincula con la primera noción de cognición, pues recurre a la inferencia lógica como modelo. El enfoque subsimbólico sostenido por Smolensky transita hacia la consideración de la percepción (por ejemplo, procesos de categorización u otros) como fundamento de la inteligencia, pero mantiene la separación, a nivel conceptual, entre percepción y

cognición. El enfoque del *grounded cognition* de Barsalou se enmarca en la noción inclusiva de cognición: percepción y cognición no son separados, sino que mantienen sistemas neurales y cognitivos comunes.

En segundo lugar, la realizabilidad múltiple sostiene que un mismo estado mental puede ser realizado por estados físicos distintos. Esta es una consecuencia del funcionalismo, donde la identificación de los estados mentales se realiza por medio de su función en lugar de su constitución. En el caso de los sistemas de símbolos que tienen relación con la inteligencia humana, ello implica que la implementación física no es relevante para el modelo cognitivo: se puede implementar satisfactoriamente en sistemas tales como las computadoras, entendidas como dispositivos físicos. Tanto en el enfoque simbólico como en el subsimbólico, la teoría de la implementación no es constituyente de la teoría psicológica. Más aún, a nivel sintáctico, un computador von Neumann puede simular una red conexionista y viceversa. No obstante, en el SSP de Barsalou, el funcionalismo (y, con ello, la realizabilidad múltiple) se problematiza. Para el autor, una característica central de los símbolos radica en su corporización. El símbolo se relaciona de manera estrecha con el sistema físico que representa dado su proceso de formación (véase pp. 127-128). Los computadores digitales modernos no pueden implementar un sistema conceptual humano completo porque mantienen diferencias marcadas con el sistema sensoriomotor que representa los conceptos. Además, la naturaleza analógica de los símbolos perceptuales garantiza las relaciones estructurales que mantienen con sus referentes: cambios estructurales en un símbolo conducen a cambios en su referente. Ello problematiza la relación entre representaciones y palabras del lenguaje natural con respecto a la naturaleza referencial discreta: los cambios en la forma de las palabras no implican cambios analógicos en el referente, y entre ambos se establece un enlace convencional (Barsalou, 1999).

La definición de cognición adoptada y la aceptación o rechazo de la realizabilidad múltiple son, por tanto, dos puntos a la base de los compromisos epistemológicos y los problemas asociados con la ecología conceptual de la cual surge la noción de REPRESENTACIÓN elaborada por cada postura examinada. Junto con la operación de las analogías específicas, permiten entender las reorganizaciones básicas por las cuales ha transitado el concepto en esta línea de CC.

6. Conclusiones

El objetivo del presente capítulo consistió en concretar el análisis del CC en la noción de REPRESENTACIÓN en ciencia cognitiva. En particular, me enfoqué en una línea de discontinuidad conceptual que puede observarse en los debates entre el enfoque simbólico propuesto por Fodor (1975, 1983, 1987, 1997a; Fodor y Pylyshyn, 1988), el enfoque subsimbólico de Smolensky (1987, 1988, 1989) y el SSP de Barsalou (1999). A modo de conclusión, revisaré las implicancias de este análisis a raíz de la consideración de los tres rasgos de la ciencia cognitiva vistos en el segundo capítulo: los supuestos sustantivos, la interdisciplinariedad y la inmadurez, que imponen ciertos constreñimientos al estudio de las discontinuidades conceptuales en él.

En primer lugar, con respecto a los supuestos sustantivos, en las discontinuidades vistas se modifican algunos elementos relacionados con la computacionalidad y la representacionalidad. Como mencioné en el segundo capítulo, el supuesto computacional sostiene que la mente/cerebro corresponde, en buena medida, a un dispositivo computacional (es decir, un dispositivo de manipulación simbólica); mientras que el supuesto representacional afirma que también es un dispositivo representacional, porque contiene entidades con contenido semántico que son manipuladas por los procesos computacionales atendiendo a su estructura. El supuesto computacional se mantiene en el paradigma simbólico y el paradigma subsimbólico: solo varía la máquina de preferencia en la cual se implementa la arquitectura cognitiva (convencional o conexionista). En el SSP, en cambio, se cuestiona el supuesto en alguna medida. Los computadores digitales modernos plantean una noción de cognición que corresponde a la computación de símbolos amodales separados de los sistemas de percepción. Si bien pueden implementar un SSP, mantienen diferencias con el sistema sensoriomotor que da origen a los símbolos, por lo que se diferenciaría de la inteligencia humana. Ello no niega la posibilidad de implementación computacional de la teoría (Barsalou, 1999, p. 582). Dado que la simulación es “un principio computacional unificador de diversos procesos en el cerebro” (Barsalou, 2008, p. 622), las redes neurales son la alternativa más natural, pero tienen límites. Se requiere mayor desarrollo de mecanismos que permitan implementar las simulaciones y los circuitos de modalidad específica (Barsalou *et al.*, 2003, p. 90). Estos avances se encuentran en progreso (Barsalou, 2020, p. 7).

Por otra parte, dado que el CC incluye reorganizaciones básicas y no eliminaciones en la noción de REPRESENTACIÓN, se mantiene el supuesto representacional. Sin embargo, sus componentes fundamentales cambian. El portador de la representación corresponde a estructuras de datos en la TRM, patrones de activación de los nodos junto con estados que consisten en pesos específicos de los enlaces en el TAC, y registros de estados neurales en el SSP. En la TRM, el fundamento o *grounding* de las representaciones complejas depende de sus constituyentes. El contenido de estos constituyentes mantiene una relación causal con los objetos o estados de cosas del mundo que producen sus instanciaciones. Para Fodor, un concepto *C* significa la propiedad *P* porque una instanciación de *P* causa una instanciación de *C*. El concepto PERRO significa la propiedad *perro* porque la propiedad *perro* causa instanciaciones de PERRO. La existencia de instanciaciones de *Q* que causan instanciaciones de *C* es asimétricamente dependiente de la covariación causal entre instanciaciones de *P* y *C*. En otras palabras, PERRO significa la propiedad *perro* y no la propiedad *lobo* porque la existencia de instanciaciones de *lobo* que causen PERRO depende de que la propiedad *perro* cause PERRO, pero no viceversa. En el TAC, las representaciones son patrones de activación cuyas unidades activas mantienen correspondencia con las microcaracterísticas de la descripción del objeto. Su contenido depende del contexto en el que se encuentra el símbolo, y del modelo conexionista adoptado (que relaciona los patrones con las descripciones a nivel conceptual de los *inputs* y *outputs*). Por otro lado, en el SSP, el fundamento de las representaciones radica en los sistemas sensoriomotores del cerebro. La simulación es la “reenacción de los estados perceptuales, motores e introspectivos adquiridos durante la experiencia con el mundo, el cuerpo y la mente” (2008, p. 618). Las zonas de convergencia de Damasio (1989) constituyen un ejemplo del proceso: las áreas asociativas locales de las regiones sensoriomotoras capturan patrones de representación perceptual y, posteriormente, reactivan estas representaciones para simular la experiencia (1999, p. 637). Por último, el interpretante en un computador convencional refiere a los comportamientos de ejecución posibles de los procesos determinables definidos para la representación, junto con representaciones suplementarias relevantes⁶⁹ (von Eckardt, 1995, pp. 292-293); en el TAC, corresponde a las consecuencias computacionales de la

⁶⁹ Un proceso determinable es un tipo de operación computacional especificada por un algoritmo, que incluye rasgos determinables como *inputs* variables (von Eckardt, 1995, p. 291). El comportamiento de ejecución de un proceso, por su parte, son los distintos modos de ejecución, que varían en dependencia de los rasgos determinables.

activación de los nodos asociados a la representación⁷⁰, y en el SSP se mantiene indeterminado por el estado aún en desarrollo de la implementación de la teoría.

En lo tocante a la interdisciplinariedad, en el segundo capítulo mencioné la posibilidad de que las variaciones en la noción de REPRESENTACIÓN dependan de las subdisciplinas que estén en interacción en el enfoque en cuestión. En el análisis anteriormente realizado, algunas fuentes que influyen en la discontinuidad son la apelación a la lingüística en la TRM, las diferencias entre la TRM y el TAC con respecto a los avances en las ciencias de la computación que emplean, y el apoyo en la neurociencia en el SSP. No obstante, las tres posturas revisadas adoptan un representacionalismo por lo que, a pesar de las diferencias, preservan el rol explicativo de REPRESENTACIÓN, aunque caracterizado de formas distintas.

Por último, en relación con la inmadurez de la ciencia cognitiva, el análisis presente retoma una consideración ya adelantada con anterioridad (p. 61 y p. 94): es posible el CC sin cambio de teoría. El CC fue tradicionalmente investigado en las revoluciones científicas, lo que contrasta con la inmadurez que caracteriza a la ciencia cognitiva. En ella, las teorías desarrolladas son todavía tentativas (véase p. 5). El enfoque simbólico no ha sido reemplazado por el enfoque subsimbólico —y, más aún, tanto Fodor y Pylyshyn (1988) como Smolensky (1987, 1988, 1989) aceptan la complementación entre ambos, aunque en el primer caso implica reducir el conexionismo a una teoría de la implementación—. Además, el SSP no ha reemplazado a ninguno de los dos enfoques clásicos. Estos y otros enfoques mantienen su influencia, pero la noción de REPRESENTACIÓN efectivamente presenta reorganizaciones básicas. En conclusión, existen discontinuidades conceptuales entre posturas con vigencia simultánea que no constituyen un reemplazo de la otra. De ello se puede extraer un panorama de continuidad entre C1 y C2 cercano al que sostiene Thagard (véase pp. 45-46).

⁷⁰ Estas consecuencias computacionales usualmente son la activación de las unidades y las señales enviadas por medio de los enlaces.

CONCLUSIONES GENERALES

El objetivo general de la presente tesis consistió en la realización de un análisis del CC presente en el concepto de REPRESENTACIÓN en ciencia cognitiva. Por medio de una investigación documental, sostuve que el proceso principal de transformación de esta noción apunta a la incorporación de distribución y la incorporación de multimodalidad. Para ello, en el primer capítulo determiné los pasos y características importantes del estudio de las discontinuidades conceptuales mediante un análisis de obras relevantes en las disciplinas que conforman la aproximación clásica del CC. La metodología resultante consiste en la caracterización de los sistemas conceptuales sucesivos, la individuación de los conceptos a investigar de acuerdo con el rol que cumplen en la teoría, la demostración de existencia de un cambio entre los sistemas conceptuales sucesivos junto con la determinación del tipo de cambio, y el estudio de los mecanismos que posibilitaron el CC.

En el segundo capítulo puntalicé tres características de la ciencia cognitiva que constriñen el análisis de sus conceptos: sus supuestos sustantivos —computacionalidad y representacionalidad—, su interdisciplinariedad y su inmadurez. Además, abordé algunos antecedentes de los CC encontrados en la literatura especializada, que comprenden nociones tales como PROCESAMIENTO DE INFORMACIÓN, CONCEPTO, COGNICIÓN y REPRESENTACIÓN. En conjunto, sugieren la existencia de tres momentos de transformación conceptual: la revolución cognitiva, las discusiones de los límites de la cognición y la revolución neurocognitiva.

Por último, en el tercer capítulo realicé el análisis en la noción de REPRESENTACIÓN siguiendo los pasos planteados en el primer capítulo. Señalé las discontinuidades conceptuales vistas en los debates entre el enfoque simbólico, el enfoque subsimbólico y el SSP. En el primero, las representaciones se entienden como símbolos amodales con contenido, poderes causales y estructura. En el segundo, las representaciones son patrones de activación entre unidades o nodos en una red. En el tercero, las representaciones son símbolos neurales multimodales. El CC observado corresponde a reorganizaciones básicas plasmadas en la incorporación de distribución y de multimodalidad, e impulsadas en gran medida por el empleo de analogías como la metáfora del computador, la analogía con el funcionamiento neuronal y la problematización de la analogía con los sistemas de

grabación. Por último, relacioné el resultado del análisis con las tres características de la ciencia cognitiva mencionadas anteriormente.

Como punto final, cabe establecer dos limitaciones del análisis presentado. La primera limitación se vincula con la noción de CONCEPTO que fue estudiada. En la introducción sostuve que los conceptos son entendidos como representaciones mentales, y sus características definitorias dependen del enfoque cognitivo adoptado. Por ejemplo, Thagard (1992) sostiene que los conceptos son estructuras mentales similares a las palabras; Carey (2009) provee una definición según la cual son estados del sistema nervioso que tienen contenido, y Nersessian afirma que son estructuras neurales relacionadas con el sistema sensoriomotor. Sumado a ello, en el segundo capítulo mostré antecedentes de CC en CONCEPTO (véase p. 79), y en el tercer capítulo estudié precisamente algunas discontinuidades en la noción de REPRESENTACIÓN. Bajo el supuesto de que la definición de CC es influida por la manera en la que CONCEPTO es definido, las distintas definiciones provistas y la existencia de las discontinuidades nombradas anteriormente podrían influir en el establecimiento de qué se entiende por CC. Las implicancias de estas variaciones no fueron abordadas en la presente tesis. En segundo lugar, otra limitación que me interesa resaltar refiere a la elección de autores y discusiones para describir el proceso de CC. La referencia a los trabajos de Fodor, Smolensky y Barsalou se originan en la importancia de sus aportes en el área. No obstante, existen diversos autores —en otras vertientes de la cognición 4EA, por ejemplo— cuyos aportes no fueron considerados. En particular, el antirrepresentacionalismo fue abordado de manera somera en uno de los antecedentes de CC mostrados en el segundo capítulo (pp. 86-92). En relación con estas limitaciones, algunas futuras directrices del estudio de los CC en la ciencia cognitiva incluyen los siguientes puntos:

1. Evaluación de las implicancias de la adopción de distintas definiciones de CONCEPTO para el estudio de los CC.
2. Ampliación de la línea de discontinuidades conceptuales por medio de la incorporación de otros debates establecidos en el área, junto con su ecología conceptual y mecanismos de cambio.

Los conceptos poseen una centralidad notable tanto a nivel individual como a nivel colectivo, dado que permiten la representación del mundo y la comprensión del lenguaje, por lo que el estudio de su proceso de creación, alteración y comunicación es siempre un área de investigación provechosa.

REFERENCIAS BIBLIOGRÁFICAS

- Adams, F., y Aizawa, K. (2001). The Bounds of Cognition. *Philosophical Psychology* 4, 43–64.
- Adams, F., y Aizawa, K. (2008). *The Bounds of Cognition*. Malden, MA: Blackwell.
- Akagi, M. (2017). Rethinking the Problem of Cognition. *Synthese* 195(8), 3547-3570.
- Anderson, J.R. (1978). Arguments Concerning Representations for Mental Imagery. *Psychological Review* 85, 249–77.
- Arabatzi, T., y Kindi, V. (2008). The Problem of Conceptual Change in the Philosophy and History of Science. En S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change*. Nueva York: Routledge, 345-373.
- Barsalou, L.W. (1999). Perceptual Symbol Systems. *Behavioral and Brain Sciences* 22(4), 577-660.
- Barsalou, L.W., Simmons, W.K., Barbey, A.K., y Wilson, C.D. (2003). Grounding Conceptual Knowledge in Modality-Specific Systems. *Trends in Cognitive Sciences* 7(2), 84-91.
- Barsalou, L.W. (2008). Grounded Cognition. *Annual Review of Psychology* 59, 617-645.
- Barsalou, L.W. (2020). Challenges and Opportunities for Grounding Cognition. *Journal of Cognition* 3(1), 1–24.
- Bechtel, W. (1998). Representations and Cognitive Explanations: Assessing the Dynamicist's Challenge in Cognitive Science. *Cognitive Science* 22, 295–318.
- Brooks, R. (1999). *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press.
- Carey, S. (1985). *Conceptual Change in Childhood*. Cambridge, MA: MIT Press.
- Carey, S. (1999a). Knowledge Acquisition: Enrichment or Conceptual Change? En E. Margolis, y S. Laurence (Eds.), *Concepts: Core Readings*. Cambridge, MA: MIT Press, 459-487.
- Carey, S. (1999b). Sources of Conceptual Change. En E.K. Scholnick, K. Nelson, y P. Miller (Eds.), *Conceptual Development: Piaget's Legacy*. Mahwah, NJ: Lawrence Erlbaum Associates, 293-326.
- Carey, S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press.

- Carnap, R. (1981). Logical Foundations of the Unity of Science. En O. Hanfling (Ed.), *Essential Readings in Logical Positivism*. Oxford: Blackwell, 112-129.
- Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.
- Clark, A. (1997a). The Dynamical Challenge. *Cognitive Science* 21, 461–481.
- Clark, A. (1997b) Being There: Putting Brain, Body, and World Together Again. Cambridge, MA: MIT Press.
- Clark, A., y Chalmers, D. (1998). The Extended Mind. *Analysis* 58(1), 7-19.
- Craik, K. (1943). *The Nature of Explanation*. Cambridge: Cambridge University Press.
- Damasio, A.R. (1989). Time-Locked Multiregional Retroactivation: A Systems-Level Proposal for the Neural Substrates of Recall and Recognition. *Cognition* 33, 25-62.
- Dennett, D.C. (1986). *Content and Consciousness*. Londres: Routledge and Kegan Paul.
- diSessa, A.A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction* 10(2/3): 105-225.
- Fodor, J. (1975). *The Language of Thought*. Nueva York: Thomas Y. Crowell Company.
- Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Fodor, J. (1985). Précis of The Modularity of Mind. *The Behavioral and Brain Sciences* 8, 1-42.
- Fodor, J. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.
- Fodor, J. (1997a). The Representational Theory of Mind. *American Behavioral Scientist* 40(6), 829-841.
- Fodor, J. (1997b). Connectionism and the Problem of Systematicity (Continued): Why Smolensky's Solution Still Doesn't Work. *Cognition* 62, 109-119.
- Fodor, J., y Pylyshyn, Z. (1988). Connectionism and Cognitive Architecture: A Critical Analysis. *Cognition* 28, 2-71.
- Fodor, J., y McLaughlin, B. (1990). Connectionism and the Problem of Systematicity: Why Smolensky's Solution Doesn't Work. *Cognition* 35, 183–204.
- Fodor, J. (2008). *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.
- Franklin, S. (2014). History, Motivations and Core Themes. En K. Frankish, y W.M. Ramsey (Eds.), *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press, 15-33.

- Gallagher, S., y Zahavi, D. (2008). *The Phenomenological Mind. An Introduction to Philosophy of Mind and Cognitive Science*. Londres: Routledge.
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton, Mifflin and Company.
- Glaser, W.R. (1992). Picture Naming. *Cognition* 42, 61–105.
- Greiffenhagen, C., y Sherman, W. (2008). Kuhn and Conceptual Change: On the Analogy between Conceptual Changes in Science and Children. *Science and Education* 17(1), 1-26.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Keijzer, F. (2002). Representation in Dynamical and Embodied Cognition. *Cognitive Systems Research* 3(3), 275-288.
- Kosslyn, S.M. (1980). *Image and Mind*. Harvard: Harvard University Press.
- Kuhn, T. (1970). *The Structure of Scientific Revolutions*. Segunda edición. Chicago: University of Chicago Press.
- Kuhn, T. (1977). *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press.
- Kuhn, T. (1982). Commensurability, Comparability. Communicability. En P.D. Asquith, y T. Nickles (Eds.), *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 2. East Lansing, Michigan: Philosophy of Science Association, 669-688.
- Kuhn, T. (2000a). Possible Worlds in History of Science. En J. Conant, y J. Haugeland (Eds.), *The Road Since Structure: Thomas S. Kuhn*. Chicago: University of Chicago Press, 58-89.
- Kuhn, T. (2000b). Reflections on my Critics. En J. Conant, y J. Haugeland (Eds.), *The Road Since Structure: Thomas S. Kuhn*. Chicago: University of Chicago Press, 123-175.
- Kuhn, T. (2000c). Afterwords. En J. Conant, y J. Haugeland (Eds.), *The Road Since Structure: Thomas S. Kuhn*. Chicago: University of Chicago Press, 224-252.
- Loving, C.C., y Cobern, W.W. (2000). Invoking Thomas Kuhn: What Citation Analysis Reveals about Science Education. *Science and Education* 9, 187-206.
- Maldonado. P. (2007). What We See is How We Are: New Paradigms in Visual Research. *Biological Research* 40(4), 439-450.

- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W. H. Freeman and Company.
- McCloskey, M. (1983). Intuitive Physics. *Scientific American* 248(4), 122–130.
- Menary, R. (2010). Introduction to the Special Issue on 4E Cognition. *Phenomenology and the Cognitive Sciences* 9, 459-463.
- Miller, G. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review* 63, 81-97.
- Neisser, U. (1976). *Cognition and Reality*. San Francisco: W. H. Freeman.
- Nersessian, N. (1992). How do Scientists Think? Capturing the Dynamics of Conceptual Change in Science. *Cognitive Models of Science* 15, 3-44.
- Nersessian, N. (1995). Opening the Black Box: Cognitive Science and History of Science. *Osiris* 10, 194-211.
- Nersessian, N. (2008). *Creating Scientific Concepts*. Cambridge, MA: MIT Press.
- Newell, A., Shaw, C., y Simon, H. (1958). Elements of a Theory of Human Problem Solving. *Psychological Review* 65, 151-166.
- Newen, A., de Bruin, L., y Gallagher, S. (Eds.) (2018). *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press.
- Özdemir, G. y Clark, D.B. (2007). An Overview of Conceptual Change Theories. *Eurasia Journal of Mathematics, Science & Technology Education* 3(2), 351-361.
- Piaget, J. (1950). *The Psychology of Intelligence*. Londres: Routledge.
- Piaget, J. (1976). Piaget's Theory. En B. Inhelder, H. Chipman, y Z. Zwingmann (Eds.), *Piaget and His School: A Reader in Developmental Psychology*. Nueva York: Springer-Verlag, 11-23.
- Posner, G.J., Strike, K.A., Hewson, P.W., y Gertzog, W.A. (1982). Accommodation of a Scientific Conception: Towards a Theory of Conceptual Change. *Science Education* 66, 211–227.
- Pylyshyn, Z. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, MA: MIT Press.
- Rescorla, R. (1988). Pavlovian Conditioning: It's Not What You Think. *American Psychologist* 43, 151-160.
- Saal, F. (1975). Conductismo, neoconductismo y Gestalt. En N.A. Braunstein, M. Pasternac, G. Benedito, y F. Saal (Eds.), *Psicología: ideología y ciencia*. México: Siglo XXI, 261-278.

- Saavedra, M.A. (2018). Historia de la neuropsicología y neuropsiquiatría. En E. Labos, A. Slachevsky, T. Torralba, P. Fuentes, y F. Manes (Eds.), *Tratado de neuropsicología clínica*. Segunda edición. Buenos Aires: Akadia Editorial, 3-8.
- Schacter, D., y Tulving, E. (2005). *The Oxford Handbook of Memory*. Oxford: Oxford University Press.
- Shannon, C., y Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.
- Sinatra, G.M., y Pintrich, P.R. (2003). The Role of Intentions in Conceptual Change Learning. En G.N. Sinatra, y P.R. Pintrich (Eds.), *Intentional Conceptual Change*. Nueva Jersey: Lawrence Erlbaum Associates, 1-18.
- Skinner, B.F. (1953). *Science and Human Behavior*. Nueva York: Free Press.
- Smolensky, P. (1987). The Constituent Structure of Connectionist Mental States: A Reply to Fodor and Pylyshyn. *Southern Journal of Philosophy* 26(Supplement), 137-161.
- Smolensky, P. (1988). On the Proper Treatment of Connectionism. *Behavioral and Brain Sciences* 11(1), 1-23.
- Smolensky, P. (1989). Connectionist Modeling: Neural Computation/ Mental Connections. En J. Haugeland (Ed.) (1997), *Mind Design II: Philosophy, Psychology, Artificial Intelligence*. Cambridge, MA: MIT Press, 233-250.
- Smolensky, P. (1990). Representation in Connectionist Networks. En *Intellectica. Revue de l'Association pour la Recherche Cognitive* 9/10(1), 127-165.
- Smolensky, P. (1995a). Connectionism, Constituency, and the Language of Thought. En C. Macdonald, y G. Macdonald (Eds.), *Connectionism: Debates on Psychological Explanation*. Cambridge, MA: Blackwell, 164-198.
- Smolensky, P. (1995b). Reply: Constituent Structure and Explanation in an Integrated Connectionist/ Symbolic Cognitive Architecture. En C. Macdonald, y G. Macdonald (Eds.), *Connectionism: Debates on Psychological Explanation*. Cambridge, MA: Blackwell, 223-290.
- Stillings, N.A., Weisler, S.E., Chase, C.H., Feinstein, M.H., Garfield, J.L., y Rissland, E.L. (1995). *Cognitive Science: An Introduction*. Segunda edición. Cambridge, MA: MIT Press.
- Strike, K., y Posner, G. (1992). A Revisionist Theory of Conceptual Change. En R.A. Duschl, y R.J. Hamilton (Eds.), *Philosophy of Science, Cognitive Psychology, and*

- Educational Theory and Practice*. Nueva York: State University of New York Press, 147-176.
- Thagard, P. (1988). *Computational Philosophy of Science*. Cambridge, MA: MIT Press.
- Thagard, P. (1992). *Conceptual Revolutions*. Princeton, NJ: Princeton University Press.
- Thagard, P. (2011). Conceptual Change in Cognitive Science: The Brain Revolution. En W.J. Gonzalez (Ed.), *Conceptual Revolutions: From Cognitive Science to Medicine*. La Coruña: Netbiblo, 65-78.
- Thagard, P. (2012). Cognitive Architectures. En K. Frankish, y W. Ramsey (Eds.), *The Cambridge Handbook of Cognitive Science*. Cambridge: Cambridge University Press, 50-70.
- Thagard, P. (2014). Explanatory Identities and Conceptual Change. *Science & Education* 23(7), 1531-1548.
- Van Gelder, T. (1990). Why Distributed Representation is Inherently Non-Symbolic. En G. Dorffner (Ed.), *Konnektionismus in Artificial Intelligence und Kognitionsforschung. Informatik-Fachberichte 252*. Berlin, Heidelberg: Springer, 58-66.
- Van Gelder, T. (1995). What Might Cognition Be, if Not Computation? *The Journal of Philosophy xci*, 345-381.
- Varela, F. J., Thompson, E., y Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Vilarroya, O. (2017). Neural Representation. A Survey-Based Analysis of the Notion. *Frontiers in Psychology* 8, 1458.
- Von Eckardt, B. (1995). *What is Cognitive Science?* Cambridge, MA: MIT Press.
- Von Eckardt, B. (2006). Cognitive Science: Philosophical Issues. En L. Nadel (Ed.), *Encyclopedia of Cognitive Science*, 1-8.
- Vosniadou, S. (Ed.) (2008). *International Handbook of Research on Conceptual Change*. Nueva York: Routledge.
- Watson, J. (1959). *Behaviorism*. Chicago: University of Chicago Press.