

# La naturalización enactiva de la normatividad *agencia y criterios normativos*

Laura Mojica

2022

## Abstract

En esta tesis defiendo una concepción de normatividad fundamentada en la acción y sus consecuencias prácticas. Con ello busco contribuir a una explicación naturalista de los fenómenos normativos en general y de la cognición en particular. Empiezo sentando las bases conceptuales del fenómeno de la normatividad y defiendo que todo fenómeno normativo tiene dos elementos: por un lado, un agente individual capaz de seleccionar las interacciones que establece con el ambiente en el presente, y por otro, criterios normativos que se extienden en una escala temporal más amplia que tales interacciones y permiten determinar su corrección más allá de las impresiones del individuo que las establece. Con esta base conceptual, tomo como punto de partida para la naturalización de la normatividad la aproximación enactiva a la cognición que ofrece una explicación naturalista de agencia y normatividad en términos de autopreservación. Sin embargo, muestro que el enactivismo no distingue explícitamente entre agencia y criterios normativos, y esto limita su poder explicativo respecto a fenómenos como la cognición ultrarrápida y el surgimiento de normas irrelevantes o nocivas para la preservación de la agencia individual como las normas sociales o las funciones biológicas establecidas evolutivamente. Para encarar estas limitaciones, propongo refinar la concepción enactiva de normatividad como autopreservación, y defiendo que las interacciones normativas, primero, son establecidas por un sistema autónomo y adaptivo, y segundo, hacen parte de algún sistema autónomo que se extiende en el tiempo y establece los criterios normativos de tal interacción. Al contrario de lo que se asume actualmente en el enactivismo, en la definición que propongo el sistema autónomo que establece los criterios normativos no es necesariamente idéntico al agente que establece la interacción, aunque puede serlo, y esto permite explicar no solo cómo es posible que los agentes erren sin darse cuenta sino también por qué es posible que agentes individuales establezcan interacciones correctas pero irrelevantes o nocivas para su automantenimiento.

## Agradecimientos

Para empezar, quiero agradecer a mis directores. A Tom Froese por creer en mis ideas incluso cuando son poco ortodoxas, por entusiasmarse con ellas y por inspirarme a desarrollarlas de la mejor manera posible; y a Jorge Galindo por su apoyo empeinado para ayudarme a navegar las burocracias y lograr sacar adelante proyectos que me llevaron a crecer inmensamente. A ambos les agradezco por haberme acogido en los momentos difíciles que pasé durante el doctorado. También agradezco a Max Martínez por apoyarme desde el principio, por abrirme espacios para desarrollar mi investigación, en especial el seminario PhiBio, y por recordarme constantemente que al hacer filosofía hay que tener siempre una posición crítica. Ninguna investigación es posible sin el diálogo con otros, por eso quiero agradecer al grupo 4E Cognition en México y a The Embodied Cognitive Science Unit en Okinawa por las estimulantes discusiones de las que aprendí enormemente y por ser espacios en donde todos encontramos una voz, donde podemos esculpir nuestras ideas y donde encontramos aliados para ello. Una investigación tampoco es posible sin financiamiento, por eso agradezco al CONACyT en México y al Okinawa Institute of Science and Technology en Japón por financiar mi investigación y por dedicar fondos públicos a la construcción de conocimiento.

Quiero agradecer especialmente a Juanca por acompañarme y apoyarme durante el doctorado, por construir conmigo un hogar de cuidado mutuo y generosidad, un hogar que nos permite y nos alienta a ambos a perseguir nuestras ambiciones y a encontrar nuestras voces creativas, y por hacerme de comer en los momentos de trabajo más intenso. Agradezco a mi hermana Dani por llevarme a ser más determinada y a ser una mejor persona, a mi hermano Simi por escucharme y por su sentido del humor, a mi mamá por su apoyo incondicional y a mi abuela por su amor incondicional. Y finalmente, agradezco a mis amigos cuyo apoyo fue fundamental en estos años. A Manuel por ser un excelente interlocutor en temas filosóficos y no filosóficos, por su empatía y por apoyar decididamente el desarrollo de mi carrera y de las ideas que me apasionan. Agradezco a Manuel, Lorena y Stephen por las aventuras en la isla paradisíaca, por sus excepcionales habilidades de cocina y por enseñarme a disfrutar de la vida durante los meses que compartimos en Okinawa. A Mark por su genuinidad y su generosidad, y por convertirse muy rápidamente en un excelente amigo. Y finalmente a Clau por su sabiduría y por darme empujones motivacionales decisivos para terminar la tesis. ¡A todos muchas gracias!

# Tabla de contenidos

Abstract.....	2
Introducción .....	6
<b>Capítulo 1 Agencia mínima y criterios de corrección.....</b>	<b>11</b>
1. 1. <i>La aproximación representacional: sentido y verdad</i> .....	11
1. 2. <i>Dos condiciones necesarias para la normatividad</i> .....	16
1.2.1. Agencia y propósitos intrínsecos.....	16
1.2.2. Criterios normativos.....	19
1. 3. <i>Normatividad como consecuencias prácticas</i> .....	23
1.3.1. Primer paso: Sentido del lenguaje como uso dentro de prácticas compartidas.....	24
1.3.2. Segundo paso: La agencia cognitiva como pertenencia a prácticas mentales .....	31
1.3.3. Tercer paso: Regularidades del sentido, regularidades nomológicas y regularidades biológicas 34	
1.3.4. Cuarto paso: Criaturas con comportamiento normativo .....	36
1.3.5. ¿Condiciones de verdad y representaciones?.....	41
1. 4. <i>Conclusión</i> .....	44
<b>Capítulo 2 La naturalización de la agencia mínima.....</b>	<b>46</b>
2. 1. <i>Teoría autopoietica: el proyecto mecanicista</i> .....	47
2.1.1. Autonomía .....	48
2.1.2. Sistemas dinámicos y determinismo estructural .....	50
2. 2. <i>Recuperar la teleología: antecedentes</i> .....	54
2.2.1. El problema de la teleología en Kant .....	55
2.2.2. Jonas y el salto enactivo a la normatividad .....	59
2. 3. <i>Recuperar la teleología: el fenómeno de la normatividad del enactivismo</i> .....	61
2.3.1. Adaptividad y automantenimiento .....	65
2.3.2. Precariedad.....	71
2.3.3. Temporalidad .....	73
2.3.4. Grados y espacio normativos .....	74
2. 4. <i>Críticas a la teleología enactiva</i> .....	78
2.4.1. La amenaza del psicologismo y de la circularidad.....	78
2.4.2. Teoría autopoietica clásica y la [ir]respetabilidad científica.....	79
2.4.3. El problema de la motivación .....	87
2.4.4. El problema de la temporalidad y los criterios de corrección.....	89
2. 5. <i>Conclusión</i> .....	91

<b>Capítulo 3</b>	<b>Extensión temporal y criterios de corrección .....</b>	<b>92</b>
3. 1.	<i>Adaptividad, temporalidad y cognición ultrarrápida.....</i>	92
3.1.1.	Esquemas y redes sensoriomotores .....	94
3.1.2.	Cognición ultrarrápida .....	96
3.1.3.	Determinismo estructural y anticipación: una tensión hasta ahora no reconocida ....	97
3.1.4.	Posibles soluciones .....	99
3.1.5.	Expansión temporal y subdeterminismo .....	104
3. 2.	<i>Teleosemántica.....</i>	105
3.2.1.	Dos críticas: intencionalidad e intensionalidad.....	107
3.2.2.	Giro hacia la acción: criterios normativos extendidos en el tiempo .....	111
3. 3.	<i>Conclusión .....</i>	115
<b>Capítulo 4</b>	<b>Normatividad relacional y situada: una propuesta integracionista ...</b>	<b>117</b>
4. 1.	<i>Variedades de normatividad: la continuidad vida-mente-agencia social .....</i>	118
4.1.1.	Normatividad biológica: enactivismo e interactivismo .....	119
4.1.2.	Normatividad sensoriomotora .....	123
4.1.3.	Interacción y normas sociales .....	126
4. 2.	<i>Una brecha en el continuo.....</i>	129
4. 3.	<i>Cerrar la brecha: normatividad relacional y situada .....</i>	138
4. 4.	<i>Conclusión .....</i>	149
<b>Capítulo 5</b>	<b>Normatividad y psicología ecológica .....</b>	<b>150</b>
5. 1.	<i>Affordances y su carácter normativo: la visión recibida.....</i>	150
5. 2.	<i>Reivindicando la normatividad de las affordances .....</i>	152
5.2.1.	El carácter público y social de la normatividad .....	152
5.2.2.	Propósitos y la distinción nomológica-normativa .....	154
5.2.3.	Objeciones y respuestas.....	156
5. 3.	<i>Criterios de normatividad y formas de vida.....</i>	161
5.3.1.	Ambiente, umwelt y hábitat .....	161
5.3.2.	Formas de vida como sistemas autónomos.....	164
5.3.3.	Individuo y umwelt.....	171
5.3.4.	Formas de vida compartidas y hábitat .....	173
5. 4.	<i>Conclusión .....</i>	176
<b>Recapitulación y conclusiones.....</b>		<b>177</b>
<b>Bibliografía .....</b>		<b>182</b>

## Introducción

Las aproximaciones naturalistas conciben la cognición como un fenómeno natural que puede ser estudiado por la ciencia. El tipo de ciencia particular que se considere adecuado para esto dependerá de la forma en que se conciba su objeto de estudio, pero, en cualquier caso, tal objeto de estudio es un fenómeno con condiciones de corrección. Esto significa que, a diferencia de otros fenómenos de la naturaleza como la trayectoria de una piedra que cae parabólicamente, los comportamientos cognitivos tienen la característica de que pueden ser correctos o incorrectos y de esta característica deben dar cuenta las diferentes aproximaciones a la cognición.

Hay al menos dos aproximaciones generales a la cognición y su carácter normativo. Por un lado, está la aproximación más popular que considera que el agente establece una relación cognitiva con el ambiente para representárselo de alguna forma y, por otro lado, están las aproximaciones alternativas que consideran que los agentes establecen relaciones cognitivas con el ambiente fundamentalmente para actuar. A estas dos aproximaciones generales subyacen dos nociones de sentido distintas: en el primer caso, el sentido se entiende paradigmáticamente como proposiciones con condiciones de verdad, y en el segundo caso, el sentido se entiende como las posibilidades de acción que el ambiente ofrece al agente. Pero más allá de las diferencias, ambas concepciones implican que la relación cognitiva que el agente establece con el ambiente es normativa en algún sentido: o bien debe preservar la verdad o bien debe preservar las posibilidades de acción; e implica también que existe la posibilidad del error para el agente: o bien puede hacerse una representación falsa o irrelevante del ambiente, o bien puede seleccionar un curso de interacción con el ambiente que no le permita conseguir sus propósitos o que limite sus posibilidades de llevar a cabo una acción completa o acciones en el futuro. El carácter normativo de la cognición es, en cualquier caso, una de sus características centrales.

Una estrategia para aproximarse de forma naturalista a la cognición es empezar por una categoría más general, a saber, los fenómenos normativos y explicar cómo estos pueden surgir en la naturaleza. Esta estrategia tiene la ventaja de que no es necesario explicar de entrada cómo surge la capacidad de hacer diferenciaciones tan sutiles como las que hacen los sistemas cognitivos, sino que es suficiente, para empezar, con explicar formas de diferenciación más básicas como las que se encuentran, por ejemplo, en otras formas de vida. Empezar abordando formas básicas de normatividad nos lleva a una segunda ventaja de esta estrategia y es que explicar la normatividad como un fenómeno natural permite situar a la cognición como parte de una familia de fenómenos que ocurren en la naturaleza, en vez de aproximarnos a esta como un fenómeno extraordinario y exclusivo de los seres humanos. En este trabajo contribuyo a esta estrategia y ofrezco una definición de normatividad que sea útil para las aproximaciones alternativas orientadas a la acción, en particular, para la aproximación enactiva a la cognición.

En el Capítulo 1 muestro que la cognición es normativa y argumento que hay dos elementos necesarios para que existan los fenómenos normativos: se necesita un agente cuya

representación, juicio o selección de curso de acción sea errada o acertada, correcta o incorrecta, y se necesitan además criterios normativos que permitan determinar la corrección más allá de las impresiones presentes del agente. Estas dos dimensiones, sin embargo, no son exclusivas de la cognición, sino que son propias de cualquier fenómeno normativo. Por ejemplo, los fenómenos biológicos, como el metabolismo y la forma en que los organismos obtienen la energía que debe ser metabolizada, son fenómenos normativos. La razón es que estos procesos pueden fallar u ocurrir correctamente y en esa medida están sujetos a criterios normativos; y son además procesos que hacen la diferencia para el organismo, es decir, su existencia depende de estos procesos y puede seleccionar diferentes maneras para mantenerlos ocurriendo, como respirar o comer. Los fenómenos biológicos son fundamentales para las explicaciones naturalistas de la cognición pues de su carácter normativo básico pueden emerger formas complejas de normatividad como la cognitiva.

El enactivismo autopoietico propone una explicación naturalista en este sentido y da cuenta de cómo surgen estas dos dimensiones, agencia mínima y criterios normativos, simultáneamente en la naturaleza. Su propuesta se cimienta en la identificación y explicación de la forma más básica de comportamiento normativo en la naturaleza: la vida. La aproximación enactiva defiende que la forma más básica de normatividad es el automantenimiento, más específicamente, defienden que la naturaleza autónoma, precaria y adaptiva de los seres vivos instaura el propósito básico de constituirse y la posibilidad mantenerse a sí mismos como seres vivos. Esto significa, primero, que el mantenerse con vida es el criterio normativo para evaluar tanto las interacciones de los seres vivos con su ambiente como sus respuestas metabólicas-organísmicas, y segundo que estos tienen la posibilidad de seleccionar trayectorias de interacción que hacen la diferencia respecto a su propia existencia y las seleccionan con el propósito intrínseco de mantenerse con vida. En el Capítulo 2 expongo la explicación del enactivismo y muestro que esta sienta las bases para desarrollar una explicación del surgimiento de la normatividad en la naturaleza, en particular, de cómo surge la agencia mínima y los criterios normativos en la manera más básica en la naturaleza.

La aproximación enactiva a la normatividad deja, sin embargo, una pregunta importante abierta, a saber, cuáles son las condiciones de posibilidad conceptuales para que surjan formas de normatividad que sean irrelevantes o amenacen el mantenimiento de la vida de un organismo. Aclarar cómo la noción de normatividad del enactivismo permite el surgimiento de este tipo de normatividades es crucial porque comportamientos y procesos fundamentales de los seres vivos como la reproducción son precisamente este tipo de fenómeno normativo, y porque la gran mayoría de fenómenos cognitivos, como el uso del lenguaje, son irrelevantes para mantenernos con vida.

Defiendo que un paso en esta dirección es desarrollar una concepción de normatividad para el enactivismo en la que se hagan explícitas dos distinciones. Por un lado, la distinción temporal entre la perspectiva presente del agente y su desarrollo como agente en el tiempo a lo largo de su vida. Por otro lado, la distinción entre criterios normativos y agencia normativa. Esta distinción es crucial en el nivel básico de normatividad como automantenimiento de los seres vivos porque hace posible reconocer que los criterios normativos de una interacción están en

la escala temporal de su vida y son irreducibles a su perspectiva presente. En esa medida, esa distinción permite aclarar por qué no es necesario que los agentes calculen las trayectorias y consecuencias a futuro de sus respuestas e interacciones presentes para contar como una respuesta o interacción normativa. Este punto se pone de manifiesto con los resultados experimentales de la cognición ultrarrápida que examino en el Capítulo 3 y motivan la naturaleza temporalmente extendida de las respuestas cognitivas.

Los casos de cognición ultrarrápida se refieren a resultados experimentales en los que los agentes exhiben respuestas cognitivas en tiempos tan cortos que parecen respuestas automáticas y entonces no es claro en qué sentido pueden ser respuestas normativas. Muestro que este punto puede aclararse reconociendo que los criterios normativos esenciales para hacer de una interacción un comportamiento normativo están extendidos en el tiempo y están constituidos por la historia de interacciones del agente y de la forma de vida a la que pertenece; es decir, estos criterios se constituyen y existen más allá de la interacción particular presente.

Reconocer que los criterios normativos están extendidos temporalmente permite reconocer también que el sentido de una interacción presente está determinado por sus consecuencias prácticas a futuro y por su situación histórica, y por ende, sobrepasa el presente inmediato del agente y su perspectiva presente en la interacción. Que la historia sea determinante para el sentido de las acciones presentes es una idea familiar entre las propuestas teleosemánticas que defienden que la normatividad se establece en la historia de selección de una especie. Así que propongo una reinterpretación de la teleosemántica, en particular, la defendida por Millikan y la defendida por Shea, que dé pie a una reconciliación con propuestas de naturalización de la normatividad centradas en el organismo individual como la propuesta enactiva.

Tal reconciliación no solo sería fructífera para el enactivismo sino también para las propuestas teleosemánticas. Si bien estas últimas dan cuenta particularmente bien de la emergencia de criterios normativos y de su naturaleza extendida en el tiempo a partir de procesos evolutivos de selección, una de sus limitaciones es que pierden de vista la constitución de la agencia mínima, pero explicar este punto es esencial porque solo del comportamiento de agentes recaen los criterios normativos. Así que en el Capítulo 4 regreso a la propuesta enactiva y propongo una definición de normatividad que busca reconciliar las explicaciones naturalistas de la normatividad centradas en la agencia, como la enactiva, y las centradas en los criterios normativos, como la teleosemántica.

Mi propuesta de normatividad busca contribuir a esta reconciliación desde la aproximación enactiva y la fundamento en la idea enactiva de que la agencia mínima emerge de un sistema autónomo y adaptivo. Para motivar mi propuesta, comienzo discutiendo una pregunta que queda abierta en esta aproximación y que surge de su definición actual de normatividad, la pregunta por cómo surgen los fenómenos normativos no centrados en el individuo. El problema es el siguiente: el enactivismo autopoiético defiende la continuidad entre las normas biológicas, cognitivas y sociales fundamentada en su concepción de normatividad básica como automantenimiento. De acuerdo con el enactivismo, el automantenimiento de la identidad sensoriomotora de un agente establece las normas cognitivas que regulan su



comportamiento, y el automantenimiento de su identidad social determina las normas sociales. Sin embargo, no hay una explicación clara de cómo los individuos, que por su propia constitución están movidos a interactuar con el mundo bajo la norma del automantenimiento, podrían interactuar con el mundo guiados por normas no centradas ni constituidas por ellos mismos. Además, entender toda la normatividad como automantenimiento hace que no sea claro cómo los agentes establecen interacciones sociales y adquieren hábitos que no tienen ninguna implicación en su constitución individual. La falta de claridad en estos puntos proviene de la definición enactiva de normatividad como automantenimiento, y pueden evitarse refinando esta definición para que incorpore la distinción entre agencia mínima y criterios normativos sin dejar de estar orientada a la acción y a las consecuencias prácticas.

Defiendo que un fenómeno normativo es una interacción que se establece por un sistema con agencia, como es definido por el enactivismo; a su vez, tal interacción forma parte de un sistema que se mantiene a sí mismo, i.e., un sistema autónomo, que establece los criterios normativos para evaluar la interacción y estos criterios son su propio mantenimiento. El sistema autónomo que establece los criterios normativos puede coincidir o no con la identidad del agente que interactúa; de hecho, este es el caso para la forma más básica de normatividad, esto es, para formas de vida mínimas en donde el criterio normativo de sus interacciones y respuestas particulares es si permiten o no el mantenimiento del sistema en la escala temporal de su vida. Sin embargo, en una gran cantidad de fenómenos normativos, el sistema autónomo que establece los criterios normativos no es el mismo sistema que regula la interacción, es en cambio un sistema más amplio del que la interacción particular hace parte. Por ejemplo, una instancia de uso del lenguaje iniciada por un agente está sujeta a las prácticas lingüísticas de su comunidad que forman un sistema autónomo que sobrepasa a, aunque esté constituida por, cada uno de los agentes particulares y sus acciones. Es importante anotar que un sistema autónomo no es necesariamente un sistema con agencia. Para ello deben cumplirse otras dos condiciones, la adaptividad y la precariedad, de manera que el sistema tenga la capacidad de hacer la diferencia sobre su propia existencia. Sistemas autónomos como comunidades de animales o tornados no cumplen con estas características y por ende no son agentes aunque potencialmente sí puedan instaurar criterios normativos.

Distinguir entre estas dos dimensiones de normatividad permite hacer diferenciaciones un poco más finas de los fenómenos normativos. Primero, se pueden diferenciar como ya lo ha hecho el enactivismo según si el comportamiento evaluado se trata de una respuesta interna del organismo, una interacción sensoriomotora o una interacción social. Se puede hacer además una segunda diferenciación de estos fenómenos según los criterios normativos que los rigen; estos pueden ser el mantenimiento de su propia vida, el mantenimiento de la forma de vida compartida por la especie o el mantenimiento de las prácticas sociales compartidas por la comunidad.

Esta definición de normatividad implica que los patrones de comportamiento y respuestas orgánicas establecidos en la evolución son sistemas autónomos que establecen los criterios normativos de los comportamientos biológicos y cognitivos. Sin embargo, esta idea está lejos de ser evidente por sí misma, así que dedico el capítulo final, el Capítulo 5, a defender que las

prácticas sociales constituyen sistemas autónomos. Me centro en las prácticas sociales porque normalmente estas establecen los criterios normativos relevantes para la cognición, y aunque este tipo de criterios incidentalmente también puedan ser establecidos por la historia evolutiva, los criterios normativos evolutivos suelen concernir el mantenimiento de la vida biológica y por lo tanto suelen constituir fenómenos normativos biológicos.

La psicología ecológica ha construido herramientas conceptuales cruciales para defender el carácter autónomo de las prácticas sociales, en particular, los conceptos de affordance y de forma de vida. El concepto de affordance se refiere a las posibilidades de acción que los agentes perciben en el ambiente dadas sus habilidades. Empiezo discutiendo en detalle un par de posibles objeciones a la atribución de normatividad a las affordances y muestro que, a pesar de que las razones que estas objeciones ofrecen apuntan a distinciones importantes entre diferentes fenómenos de la naturaleza, no son suficientes para rechazar su carácter normativo. Luego discuto dos conceptos clave para entender la relación cognitiva entre agente y ambiente y su carácter situado: forma de vida y prácticas sociales. Argumento que las interacciones cognitivas agente-ambiente están siempre situadas dentro de formas de vida compartidas que constituyen sistemas autónomos porque en estas se constituyen las habilidades de involucramiento con affordances y las consecuencias prácticas de su ejecución. Esto significa que tales formas de vida establecen los criterios de corrección para tales interacciones y establecen sus condiciones de posibilidad. Por otro lado, argumento que las prácticas sociales entendidas como formas de vida compartidas consisten en patrones de comportamiento de un grupo de personas en un entorno material que se sostienen en el tiempo y que establece las condiciones que le permiten mantener su propia existencia. Puesto que las formas de vida entendidas como prácticas sociales hacen posibles interacciones sociales particulares y mantienen el ambiente material que permite a los agentes actuar, y a su vez tales acciones y entornos materiales mantienen la organización que constituyen la práctica, las prácticas sociales son sistemas autónomos. Si este es el caso, las prácticas sociales en tanto sistemas autónomos establecen además los criterios normativos que permiten evaluar y dar sentido al comportamiento de los agentes sociales.

# Capítulo 1

## Agencia mínima y criterios de corrección

En este capítulo defiendo que hay dos condiciones necesarias para que haya fenómenos normativos: primero, debe haber un sistema que sea agente en un sentido mínimo, eso es, que sea un sistema que seleccione un curso de interacción, un comportamiento o se forme un estado mental y no otro de acuerdo con sus propósitos propios, y segundo, debe haber criterios de corrección que determinen la corrección de tal interacción, comportamiento o estado mental. Cualquier intento de explicación de los fenómenos normativos, y en particular, del fenómeno normativo de la cognición, debe dar cuenta de ambas condiciones y parte de ello es definir más precisamente en qué consisten agencia mínima y criterios de corrección, y dar cuenta de condiciones necesarias para que surjan en la naturaleza. Tanto agencia mínima como criterios de corrección han sido entendidos de formas distintas en la literatura. Por ejemplo, la filosofía analítica de la mente de finales del siglo XX entendió estos conceptos en términos de verdad y de contenido proposicional, y se centró en explicar la cognición humana. Esta, por supuesto, no es la única manera en que tales conceptos pueden interpretarse, así como tampoco la pregunta por la cognición humana es la única en la que tales conceptos son útiles. Termino el capítulo defendiendo una interpretación alternativa de los conceptos de agencia mínima y criterios normativos, esta es una interpretación orientada a la acción que permitirá poner ambos conceptos en el contexto más amplio de la pregunta por la normatividad en general y no solo la pregunta por la cognición humana. Tal interpretación pragmática será el primer paso conceptual hacia la definición de normatividad que propongo para un proyecto de naturalización enactiva y ecológica.

### 1. 1. La aproximación representacional: sentido y verdad

En esta sección discuto la aproximación representacional a la mente y reviso varios de sus supuestos fundamentales. Me concentro en las propuestas de finales del siglo XX defendidas por autores como Fodor, Dennett y Davidson, quienes defienden posturas muy diferentes pero cuyos puntos en común me permiten caracterizar de manera general los elementos básicos de las aproximaciones representacionales. A pesar de que se ha avanzado mucho en estas aproximaciones desde entonces, me concentro en ellas por dos razones. Primero, porque son el antecedente histórico inmediato de las aproximaciones a la cognición más populares actualmente, a saber, las aproximaciones representacionales, y es de donde vienen muchos de los conceptos y suposiciones en las que estas aproximaciones todavía se fundamentan. Estas aproximaciones son además opuestas a la aproximación a la cognición, al sentido y a la normatividad que adoptaré a lo largo de la tesis. Así que identificar puntos comunes en ambas aproximaciones me permitirá en este capítulo destilar elementos mínimos necesarios que caracterizan el fenómeno que ambas aproximaciones buscan explicar, esto es, la cognición en tanto fenómeno normativo.

## 1ra suposición: hacer sentido del mundo es representárselo

Las aproximaciones representacionales parten de la idea de que las instancias en que puede observarse que un sistema hace sentido del mundo son los estados mentales, y conciben tales estados mentales como representaciones. Los estados mentales son, por ejemplo, creencias, deseos o incluso emociones, se supone que los atribuimos cotidianamente a otros y a nosotros mismos, y que nos permiten explicar el comportamiento de otros y dar cuenta del nuestro propio como un comportamiento guiado por razones para actuar. A esta atribución cotidiana de estados mentales se le conoce como *folk psychology* y está estrechamente relacionada con nuestra capacidad de usar y entender el lenguaje. De acuerdo con la aproximación representacional, la estructura básica de un estado mental (ej., creer que P) consiste en tener una actitud (creer) hacia un contenido (P), que el agente puede expresar lingüísticamente para dar razón de su comportamiento; por ejemplo, «creí que era el botón de la luz, no de la alarma».

A partir de esta suposición inicial, surge una primera pregunta, a saber, si las representaciones existen en tanto entidades o no. Como es de esperarse, las aproximaciones representacionales a la cognición asumen que las representaciones mentales sí existen y son aquello a lo que nos referimos cuando atribuimos estados mentales a los agentes. Fodor (1990) denomina esta posición realismo respecto a las actitudes proposicionales y ha sido uno de sus defensores más prolíficos y mejor conocidos. Autores como Dennett (1981), por otro lado, rechazan tal posición realista y defienden que las adscripciones de estados mentales que hacemos a otros son instrumentos para entender y predecir su comportamiento, y no afirmaciones sobre la existencia de entidades mentales representacionales. Para Dennett (1981), las representaciones capturan patrones de procesamiento de información, pero no hay razones suficientes para asumir que la forma en que de hecho ocurren en la mente coincide con cómo suponemos cotidianamente que sucede. La diferencia entre ambas posiciones, realistas como la de Fodor e instrumentalistas como la de Dennett, es el compromiso ontológico con cómo se constituye la estructura subyacente en la mente que permite el comportamiento cognitivo.

Sin embargo, en ambos casos se asume que en un nivel intuitivo y cotidiano los agentes hacen sentido de su ambiente en términos de actitudes proposicionales; esta es, además, la concepción de la mente más comúnmente aceptada en filosofía y ciencias cognitivas. Sea solo en el nivel personal, como metáfora guía o como verdad literal de los mecanismos cognitivos, el hacer sentido del ambiente se entiende en términos de entidades mentales que presentan el mundo de cierta manera al sujeto; tales entidades tienen propiedades semánticas, es decir, tienen sentido, se refieren a algo, son acerca de algo, pueden ser verdaderas o falsas, correctas o incorrectas, etc. y tienen también propiedades sintácticas, es decir, una estructura en virtud de las cuales tiene relaciones de inferencia lógica con otras representaciones. Las representaciones son los elementos básicos y esenciales de la mente, es decir, las unidades psicológicas del sentido, y en esa medida, la cognición consiste en el procesamiento de tales representaciones.

Una representación tiene dos elementos: (i) una actitud proposicional que se dirige a o que es acerca de (ii) un contenido. Mientras el contenido se refiere a un aspecto del mundo, objeto o estado de cosas que es susceptible de ser caracterizado en términos de verdad<sup>1</sup>, la actitud proposicional le da la estructura a la representación que permite que se conecte inferencialmente con otras representaciones de una manera específica (Pitt, 2020). Por ejemplo, creer que el gato tiene hambre implica creer, aunque sea implícitamente, que el gato está lo suficientemente sano como para sentirla, mientras que desear que el gato tenga hambre está implicado por la creencia de que el gato probablemente no está lo suficientemente sano como para sentir hambre junto con el deseo que el gato esté sano. En el ejemplo, dos representaciones distintas –un deseo y una creencia– con el mismo contenido proposicional –el gato tiene hambre– tiene diferentes relaciones de inferencia con otras representaciones en virtud de que se trata de diferentes actitudes proposicionales.

Esta caracterización de las representaciones como entidades con propiedades semánticas, i.e., son acerca de algo e implican y son implicadas por otras representaciones, permite diferenciarlas de otras entidades o eventos no-cognitivos. Por ejemplo, la respuesta de las células ciliadas del oído, a diferencia de las representaciones mentales, no tienen propiedades semánticas ni sintácticas, solo responden a un estímulo sonoro del ambiente, pero no son acerca de nada ni son implicadas ni implican ninguna afirmación; en suma, la actividad de estas células es puramente causal y no-cognitiva. La aproximación representacional distingue relaciones causales de relaciones de sentido, y ubica la marca de lo mental en representaciones internas en la mente de los agentes. Esta aproximación surge como contraposición al conductismo que explica el comportamiento cognitivo en términos de estímulo-respuesta pero que hace imposible distinguir comportamientos determinados exclusivamente por leyes mecánicas y automáticas de comportamientos determinados también por la atribución de sentido al mundo y el razonamiento del agente (Fodor, 1981). El representacionalismo entonces busca reivindicar como parte de la realidad y como explicación psicológica respetable lo que ocurre en nuestra mente y el sentido que le damos al mundo.

## 2da suposición: La relación entre agente-ambiente es de input-output

La forma en que se entiende la relación entre agente y ambiente nos lleva a la segunda suposición que caracteriza la aproximación representacional a la cognición. Un agente recibe

---

<sup>1</sup> Hay autores que defienden, por diferentes motivos, la existencia de contenidos no proposicionales que no pueden formularse en términos de verdad (ej. Evans 1982, Cussins 1990, Bermúdez 1998, Grzankowski 2015). Sin embargo, la caracterización paradigmática de los contenidos mentales es en términos de verdad, es decir, como proposiciones susceptibles de ser verdaderas o falsas, o como el conjunto de los mundos posibles en donde el contenido es verdadero, etc., y la razón es que se suele tomar como característica definitoria de los contenidos el que tengan propiedades semánticas (Pitt, 2020). Puesto que el propósito de este trabajo es ofrecer una definición de normatividad para las aproximaciones enactivas y corporizadas a la cognición, no discutiré las complejidades de los contenidos no proposicionales en esta brevísima caracterización de las aproximaciones representacionales a la cognición.

los estímulos del ambiente, esto es, el input, y produce un output que puede ser una acción o una declaración lingüística. La aproximación representacional busca explicar lo que sucede en medio de input y output. En palabras de Shea, la aproximación representacional mínimamente requiere postular “una explicación putativa de [la operación o el comportamiento del sistema respecto de su ambiente] que dependa en parte en la atribución de propiedades representacionales al sistema (ej., hacer seguimiento de p, buscar que q, etc.)” que si es exitosa, ofrezca “una caracterización del tipo de propiedades que tendrían las propiedades representacionales del sistema tendría para que sea exitoso al explicar [la operación y comportamiento del sistema en relación con su ambiente]” (2013b, p. 503). Esto significa que la explicación representacional es esquemáticamente la siguiente: un agente percibe el ambiente como significativo porque se forma una representación que transforma el input sensorial que recibe en una imagen significativa (i.e., una representación), esta representación es luego procesada en virtud de sus propiedades semánticas y sintácticas<sup>2</sup>, y produce como output una acción o locución, es decir, un comportamiento o una operación en la que el agente se relaciona con su ambiente (Shea, 2013b, p. 503). Tal output es significativo en virtud de que hay una o más representaciones que constituyen la razón por la que el agente ejecuta tal acción o locución. La explicación representacional es exitosa cuando la atribución de ciertas propiedades representacionales logra predecir el output de comportamiento dado un input sensorial.

### 3ra suposición: un error cognitivo es una representación inadecuada

Finalmente, las aproximaciones representacionales entienden la posibilidad del error de los sistemas cognitivos como una representación inadecuada (*misrepresentation*). Dentro de los estados mentales y las formas de hacer sentido es posible que los agentes tengan representaciones falsas o inadecuadas, esto es, descripciones de objetos que no existen en el mundo o representaciones de estados de cosas que no son el caso. Para que el agente pueda formar creencias verdaderas sobre el mundo, las representaciones deben estar conectadas de una forma adecuada con este, por ejemplo, mediante conexiones causales con los objetos que representan. Encontrar cuáles son las formas de conexión entre representaciones y referentes que garantizan razonablemente la verdad de las creencias que forma el agente es la pregunta de la epistemología, en otras palabras, la epistemología busca criterios normativos para diferenciar creencias justificadas de creencias no justificadas.

Pero hay una pregunta anterior, a saber, la pregunta por cómo ocurren estas representaciones y cómo se determina su contenido para que un estado mental pueda ocurrir en primer lugar independientemente de su verdad o falsedad. El punto más general de esta pregunta es diferenciar estados mentales de eventos naturales que por sí mismos no tienen sentido y de los que no se puede decir que son cognitivamente correctos o incorrectos. Esta pregunta es, en

---

<sup>2</sup> Autores como Stephen Stich (1983) argumentan que los procesos cognitivos ocurren en función de las propiedades puramente sintácticas de las representaciones: solo las propiedades físicas y sintácticas de los estados mentales tienen poder causal, y por ello, a pesar de que los aspectos semánticos de las representaciones hagan parte de la psicología folk, estos deben permanecer fuera de la psicología científica.

otras palabras, la pregunta por los criterios de diferenciación entre sistemas capaces de hacer sentido de los demás sistemas de la naturaleza.

Una característica crucial que permite diferenciar los sistemas capaces de hacer sentido de los demás es que los primeros son susceptibles de errar, mientras que el comportamiento de los sistemas no-cognitivos simplemente ocurre, no es correcto o incorrecto en sí mismo. Bajo la concepción representacional de la mente, la capacidad de errar se entiende como la posibilidad de representarse estados de cosas que no son el caso. Esta posibilidad es crucial para dar cuenta adecuadamente de la cognición representacionalmente por al menos dos razones.

En primer lugar, hay estados mentales que se caracterizan por representar estados de cosas que suelen no corresponden con el mundo actual, aunque no sean necesariamente un error. Por ejemplo, la capacidad cognitiva de desear requiere un objeto de deseo que es una representación de un estado de cosas que normalmente no es el caso; tales representaciones orientan el comportamiento del agente para cambiar el ambiente y conseguir que sean el caso. En la medida en que este tipo de estados mentales busca cambiar el mundo para que se ajuste a lo que el agente tiene en mente, se dice que estos estados mentales tienen dirección de ajuste (*direction of fit*) de mundo a mente; esto en oposición a estados mentales como las creencias cuya dirección de ajuste es de mente a mundo.

En segundo lugar, a pesar de que presumiblemente exista una relación causal entre objetos del mundo y representaciones, los agentes cognitivos pueden tener creencias erradas. Sin esta posibilidad, esto es, la posibilidad del error, tampoco existiría la posibilidad de estar en lo correcto, y entonces sería imposible distinguir entre creencias verdaderas y falsas, deseos apropiados e inapropiados, percepciones adecuadas e inadecuadas, etc. Bajo la aproximación representacional a la mente, el error cognitivo ocurre al menos de dos maneras: cuando hay un error de razonamiento, es decir, una inferencia o paso injustificado de una representación a otra, o bien, cuando el contenido de una creencia no coincide con el estado de cosas del mundo y entonces la creencia es falsa, o cuando el objeto de un deseo es imposible y entonces el deseo es inapropiado, etc. Por supuesto, un agente cognitivo puede tener una creencia falsa o hacer un razonamiento incorrecto a pesar de tener la impresión de estar en lo correcto, así que el que un agente crea que un contenido es verdad no permite determinar la corrección de tal creencia. Por tanto, para determinar la corrección es necesario apelar a criterios externos a las impresiones inmediatas del agente.

Como veremos más adelante, esta no es la única manera de concebir la cognición ni el error cognitivo; sin embargo, revisar las aproximaciones representacionales al hacer sentido, la relación entre agente y ambiente, y sobre todo, la posibilidad del error característica del comportamiento de los sistemas cognitivos permite destilar aspectos esenciales de los fenómenos normativos que deben mantenerse en cualquier aproximación a la cognición y a la normatividad en general.

## 1. 2. Dos condiciones necesarias para la normatividad

En esta sección defiendo dos aspectos esenciales: por un lado, un sistema de cuyo comportamiento esté producido por sus propios propósitos, y por otro lado, criterios normativos para evaluar tal comportamiento que vayan más allá de los estados del agente individual durante la ejecución de su comportamiento o la formación de sus creencias y deseos.

### 1.2.1. Agencia y propósitos intrínsecos

La discusión de si se deben considerar o no a los seres vivos como entidades con propósitos en sí mismos tiene raíces en la discusión sobre teleología de Kant. Esta discusión la trataré en las secciones 2. 2 y 2. 3; por ahora, en esta sección, quisiera mostrar que es necesario que un sistema tenga propósitos propios y la posibilidad de seleccionar entre diferentes cursos de acción para que su comportamiento sea genuinamente normativo y exista entonces la posibilidad del error. Por comportamiento genuinamente normativo me refiero al comportamiento que es correcto o incorrecto, mejor o peor, etc. en virtud de su propia existencia y no solo en función de la existencia de otro sistema. Para aclarar esta distinción y mostrar por qué los sistemas genuinamente normativos tienen propósitos propios y posibilidad de selección quisiera empezar considerando su contraparte: los sistemas no normativos.

Los sistemas que más claramente son no-normativos son objetos de la naturaleza que no tienen la capacidad de establecer por sí mismos interacciones con su ambiente, por ejemplo, las piedras. El comportamiento de una piedra no puede evaluarse normativamente respecto a la piedra misma porque esta no tiene ningún propósito propio. Si bien, es posible evaluar su corrección en función de, por ejemplo, si golpea a un gatito al caer o si es lo suficientemente pesada para mantener una puerta abierta, esta evaluación solo tiene sentido respecto a los propósitos propios de otro sistema: el gato que quiere mantenerse con vida y el humano que quiere mantener la puerta abierta. Los comportamientos no normativos simplemente ocurren, no son ni correctos ni incorrectos en sí mismos porque son efectos de causas que no tienen propósitos o preferencias; mientras que aquello que puede ser evaluado normativamente, es decir, aquello que puede acertar o cometer un error, es producido por sistemas que tienen un propósito propio.

Por otro lado, y aunque quizás menos evidente, los objetos diseñados por seres vivos, ej. chimpancés o humanos, tampoco son genuinamente normativos. Un artefacto, por ejemplo, un termostato, puede ser descrito en términos de propósitos, por ejemplo, mantener la temperatura del cuarto estable. De esta manera, podría pensarse que el artefacto es capaz de diferenciar, primero, los estímulos del ambiente relevantes para ese propósito de los que no son relevantes, y segundo, entre los que sí son relevantes, los que llaman a subir la temperatura, los que llaman a bajarla y los que llaman a mantenerla igual. Sin embargo, las respuestas del termostato, como las de cualquier artefacto, carecen de la flexibilidad que tienen las respuestas de los agentes cognitivos, la razón es que los termostatos así como las



piedras no pueden seleccionar entre formas distintas de responder a los cambios del ambiente, así que una explicación causal en términos de las propiedades físicas del metal del que está hecho el termostato basta para entender su comportamiento. El termostato no tiene la posibilidad de seleccionar entre distintas posibilidades de interacción y respuesta a los cambios del ambiente porque su comportamiento está determinado: no existe la posibilidad de que el metal no se contraiga con el frío o que no se expanda cuando aumenta la temperatura. Esto implica además que el termostato tampoco tiene la posibilidad de actuar de acuerdo con sus propios propósitos aún si los tuviera. Por esta razón y puesto que el termostato, así como ningún artefacto, tiene la posibilidad de seleccionar, tampoco tiene la posibilidad de errar y entonces su comportamiento no es genuinamente normativo.

La posibilidad del error requiere que haya un comportamiento que sea correcto o incorrecto. Respecto a tal comportamiento o evento hay dos posibilidades: puede ser causado por el mismo sistema para quien hace la diferencia o puede no serlo. Si no es causado por el sistema para quien hace la diferencia, aquello que es valorado como correcto o incorrecto respecto a los propósitos del sistema no es correcto o incorrecto en sí mismo porque el hecho de que ocurran estos eventos es independiente del sistema que valora. Por ejemplo, un sistema puede valorar la disminución de glucosa disponible en el ambiente como inconveniente respecto a su propósito de alimentarse, pero este no es en sí mismo un evento correcto o incorrecto, solo lo es relativo a los propósitos del sistema. Por otro lado, cuando el evento es causado por el sistema mismo, es decir, cuando es su propio comportamiento y apunta a conseguir sus propios propósitos, este es un comportamiento genuinamente normativo, i.e., correcto o incorrecto. Por ejemplo, si el sistema se mueve hacia un espacio en el que la concentración de glucosa es mayor, su comportamiento es correcto con respecto al propósito que motiva el comportamiento mismo. Este tipo de eventos, i.e., los comportamientos seleccionados y relevantes para los propósitos propios del sistema, son comportamientos normativos; y cuando tales propósitos son cognitivos y los criterios de corrección también lo son, el comportamiento es un comportamiento con sentido, y por ende constituyen el objeto de estudio de las ciencias cognitivas y la filosofía de la mente.

Quisiera hacer dos aclaraciones respecto a la idea de propósitos propios como elemento constitutivo del tipo de sistemas que nos interesan. Primero, por propósito me refiero tanto a fenómenos psicológicos que implican planeación a futuro, ej., tener el propósito de aprender a hacer bonsáis, como a fenómenos como las preferencias que implican una diferenciación básica de situaciones mejores y peores, pero que no implican habilidades psicológicas complejas, ej., el propósito de las plantas de orientarse al sol. Por supuesto, el concepto de propósito es un concepto normativo en sí mismo, lo que me lleva al segundo punto. Una buena estrategia de naturalización de la normatividad debe ser capaz de explicar el surgimiento de sistemas con propósitos o preferencias propias a partir de conceptos o fenómenos que no sean normativos en sí mismos. En el siguiente capítulo veremos que esta es precisamente la estrategia de la aproximación enactiva a la cognición.

En suma, la posibilidad del error requiere mínimamente que sea producido por un sistema que tenga la posibilidad de seleccionar su comportamiento y sus respuestas a los cambios en el

ambiente. Es necesario además que el sistema tenga forma de diferenciar entre sus posibilidades de selección y, para ello, el sistema debe tener propósitos propios, por ejemplo, alguna característica del sistema que deba mantenerse o una función que el sistema deba cumplir. A los sistemas que cumplen con estas dos condiciones llamo agencia mínima.

### Agencia y propósitos intrínsecos en la aproximación representacional

Bajo la aproximación representacional, los propósitos que permiten tanto a agentes cognitivos hacer sentido de su ambiente como a observadores externos reconocer su comportamiento como significativo y evaluar su corrección son entendidos de una forma que se supone abstracta, universal y en gran medida independiente de propósitos instrumentales y propios de circunstancias particulares: el propósito de un agente cognitivo es representarse el mundo de forma verdadera y conservar la verdad.

Las aproximaciones representacionales entienden la agencia cognitiva como un sistema que procesa representaciones que son entidades semánticas y cuyo propósito es representarse el mundo de forma verdadera; por ello, suele afirmarse que los sistemas cognitivos son capaces de diferenciar entre aspectos del mundo de manera *intensional*. Esto significa que los sistemas cognitivos clasifican los objetos y aspectos del mundo en conjuntos definidos mediante descripciones que los presentan de alguna manera, por ejemplo, el conjunto de los amigos de Carlos, mientras que los sistemas no-cognitivos responden al mundo y a sus categorías solo extensionalmente, por ejemplo, el conjunto cuyos elementos son Jaime, José, Min (Fodor, 1982).

Vale la pena notar que esta imagen de la mente sostiene y es sostenida por la idea subyacente de que la forma paradigmática de hacer sentido del ambiente y de actuar con sentido es mediante el lenguaje. Esto es evidente en la naturaleza proposicional del contenido de las representaciones que se asume por defecto y se acepta ubicuamente –por ejemplo, uno de sus defensores más conocidos, Fodor, propone explícitamente un lenguaje del pensamiento (1975). Bajo esta concepción de mente y significado, una acción (comportamiento con sentido) se diferencia de otro tipo de eventos y comportamientos porque es consecuencia de representaciones, un deseo y una creencia paradigmáticamente, cuyo contenido puede ser formulado en términos lingüísticos con condiciones de verdad y falsedad. Las acciones son el output de las representaciones y pueden diferenciarse del comportamiento no-cognitivo en tanto que el agente puede ofrecer razones de sus acciones, pero no del comportamiento no-cognitivo (Davidson, 1982). Consecuentemente, no solo se asume que la capacidad de interactuar con sentido con el mundo es un fenómeno fundamentalmente lingüístico, sino que el sentido mismo se define como las condiciones en las que una afirmación es verdadera, esto es, el sentido se define como significado del lenguaje y el significado del lenguaje se define como condiciones de verdad y falsedad.

Bajo la aproximación representacional de la mente, las interacciones que no están guiadas por estados mentales con condiciones de verdad y falsedad no constituyen un comportamiento cognitivo. Entonces, la caracterización del ejemplo anterior, el ejemplo del sistema que responde a la disponibilidad de glucosa, no es suficiente para determinar si se trata de un

comportamiento cognitivo o no porque no es claro si el sistema se representa internamente tales estados de cosas o si su comportamiento es una respuesta automática. De manera que solo si el sistema tiene la creencia de que la concentración de glucosa es mayor en determinada dirección, y esta creencia motiva su comportamiento se trata de un comportamiento cognitivo y no meramente de un reflejo.

### 1.2.2. Criterios normativos

En la sección anterior vimos que un fenómeno normativo requiere de un sistema con agencia mínima, y definí agencia mínima como un sistema capaz de seleccionar algunos de sus cursos de interacción con el ambiente y con propósitos propios que le permitan diferenciar entre los cursos de acción a seleccionar. Sin embargo, estas condiciones, si bien necesarias, no son suficientes para dar cuenta de los fenómenos normativos en muchos de los casos que nos interesan. Es necesario además que haya criterios que determinen la corrección de tales interacciones, esto es, criterios normativos. En esta sección discuto brevemente en qué consisten los criterios normativos y muestro por qué es necesario diferenciarlos de la agencia mínima y no equiparlos a los propósitos propios del agente.

Una de las características centrales de muchos fenómenos normativos es que el carácter correcto o incorrecto de un comportamiento no depende de lo que el agente pueda creer, sentir o de sus propios procesos internos. Por ejemplo, para un gato es incorrecto comerse las hojas de un lirio o frotarse contra este, a pesar de que sus hojas le puedan parecer sabrosas o sienta placer al frotarse contra este; así mismo, afirmar que  $(a + b)^2 = a^2 + b^2$  es incorrecto así le parezca claro y evidentemente correcto a quien hace la afirmación. En ambos casos hay criterios normativos que sobrepasan la interacción gato-lirio y el acto de afirmar, y que son además independientes de las impresiones de placer o corrección que puedan tener los agentes. Por supuesto, hay diferencias importantes entre los dos ejemplos: mientras uno es incorrecto en virtud de criterios biológicos, el otro es incorrecto en virtud de criterios matemáticos que se usan en la práctica propia de las matemáticas. En el Capítulo 4 discuto la diferencia entre estas formas de normatividad, y uso la concepción de normatividad que propongo en ese mismo capítulo para explicar la diferencia en la constitución de tales criterios normativos.

Para preservar la independencia entre la corrección del comportamiento y las impresiones del agente es crucial no equiparar los criterios normativos que determinan la corrección de los comportamientos con los propósitos propios del agente, y menos aún con estados o impresiones internas que el agente pueda tener mientras selecciona y lleva a cabo sus comportamientos. La razón para evitar la segunda equiparación es relativamente clara: presumiblemente, cuando un agente selecciona un curso de interacción suele hacerlo bajo la impresión de que es una buena selección, correcta, benéfica de alguna forma, etc. o al menos bajo la impresión de que no es incorrecta. Sin embargo, tal comportamiento puede ser incorrecto, y de no separar impresiones del agente de criterios de corrección sería imposible dar cuenta de la posibilidad del error en estos casos.

La razón para mantener la distinción entre propósitos propios del agente y criterios normativos es menos evidente. A diferencia de las impresiones, los propósitos propios del agente suelen extenderse en el tiempo como, por ejemplo, mantener un sistema de creencias consistente, evitar creencias falsas, mantenerse con vida, ser parte de una comunidad, etc. Por ello, tales propósitos propios pueden servir como criterios de corrección. Este es el caso en el ejemplo anterior del gato y los lirios porque un propósito propio del gato es mantenerse con vida. Sin embargo, en el caso de la afirmación algebraica, puede ser fácilmente el caso que el agente no tenga como propósito propio hacer afirmaciones algebraicamente correctas, sino quizás su propósito es ganar dinero en un programa de concursos o aprobar una materia para no tener que lidiar con esos temas nunca más en su vida, etc. Estos propósitos propios del agente no son en efecto los criterios que determinan la incorrección de afirmar que  $(a + b)^2$  es igual  $a^2 + b^2$  porque tal afirmación es falsa aún si el agente logra ganar dinero en el concurso o pasar la materia. Entonces, para poder dar cuenta de qué tan correcto, apropiado o pertinente es un comportamiento es necesario distinguir propósitos propios de criterios normativos y así acomodar en la explicación del fenómeno normativo el que el comportamiento pueda sea correcto o incorrecto en función de criterios ajenos al agente.

En suma, la corrección del comportamiento de un agente no depende de las impresiones o estados internos del agente durante la interacción ni exclusivamente de sus propósitos propios, sino de las propiedades de la interacción misma respecto a criterios de corrección. Para terminar esta sección, mostraré cómo se han entendido los criterios externos relevantes para la cognición desde una aproximación representacional. En las próximas secciones defenderé una aproximación alternativa: una aproximación centrada en la acción y no en la verdad o la representación.

### Criterios normativos en la aproximación representacional

A partir del giro lingüístico, la posibilidad de reconocer los sistemas cognitivos como sistemas que hacen sentido y que pueden hacerlo correcta o incorrectamente ha sido un punto crucial en filosofía analítica. Bajo las aproximaciones analíticas y representacionales a la mente, la posibilidad de reconocer un sistema cognitivo se entiende como el reconocer que el comportamiento está guiado por actitudes proposicionales y que este puede ser juzgado como correcto o incorrecto según el contenido proposicional que lo motive (Davidson, 1982). Mantener cierta independencia entre el significado y aquello que el agente cree que está significando permite establecer criterios para determinar si algo hace sentido o no, y si sí lo hace, para determinar su corrección. Por esa razón, la pregunta por los criterios normativos del comportamiento cognitivo, es decir, criterios de corrección determinados por algo que va más allá del comportamiento específico, tomo forma en las tradiciones analíticas en filosofía de la segunda mitad del siglo XX como la pregunta por cómo determinar el significado del lenguaje más allá de lo que un hablante particular pueda creer que significa su uso del lenguaje. Por lo tanto, mantener la distinción entre significado (i.e., sentido del lenguaje) e impresión de significado es crucial para reconocer la posibilidad del error propia de los sistemas cognitivos, y por ello es una de las intenciones explícitas de la concepción representacional funcionalista de la mente propuesta por autores como Fodor (1990, p. xi).

Incidentalmente, representarse el mundo de forma verdadera y mantener esa verdad en los razonamientos puede ser útil porque tales representaciones pueden proveer la información necesaria al sistema para navegar el ambiente según sus propósitos prácticos particulares y contingentes. En ese sentido, las representaciones tienen contingentemente la virtud de poder guiar el comportamiento con sentido de los agentes, es decir, sus acciones, en cualquier contexto y para cualquier propósito. Sin embargo, esta utilidad contingente de las representaciones no constituye sus criterios de corrección ni determinan su sentido, ni es tampoco el propósito intrínseco de los sistemas cognitivos: la cognición se trata fundamentalmente de representarse el mundo de forma verdadera. Esta concepción de cognición permite diferenciar comportamientos cognitivos, racionales o causados por una mente de comportamientos que solo parecen cognitivos. Por ejemplo, Davidson defendió en los años 80 que los animales, a pesar de su capacidad de reaccionar de formas complejas a estímulos diferenciados, no son racionales, no tienen mente y no son agentes cognitivos porque no son capaces de formarse creencias con condiciones de verdad y falsedad (1982). La razón es que ser agente cognitivo o ser racional requiere que el sujeto sea capaz de diferenciar entre lo objetivo y lo subjetivo, y que domine esta diferencia; esto es en otras palabras, el sujeto debe tener la idea de que sus propias creencias no necesariamente coinciden con las creencias de otros y, crucialmente, no necesariamente coinciden con la verdad objetiva. De acuerdo con Davidson, la capacidad de reconocer esta diferencia depende de que el agente domine un lenguaje natural. Por supuesto, esta idea ha sido desafiada en las últimas décadas y autores como Beck (2012) han rechazado la necesidad de contenidos proposicionales y han defendido la posibilidad de cognición en animales no humanos dentro de la aproximación representacional.

Esta interpretación de los propósitos propios de los sistemas cognitivos y de los criterios de corrección como preservación de la verdad es una aproximación posible a la cognición, pero no la única. Explicar los sistemas cognitivos como sistemas con un propósito abstracto, universal e independiente de los contextos particulares no se sigue necesariamente del fenómeno de la cognición en la naturaleza, y, como veremos más adelante, existe la posibilidad de aproximarnos de manera distinta.

La figura a continuación resume el argumento hasta ahora:

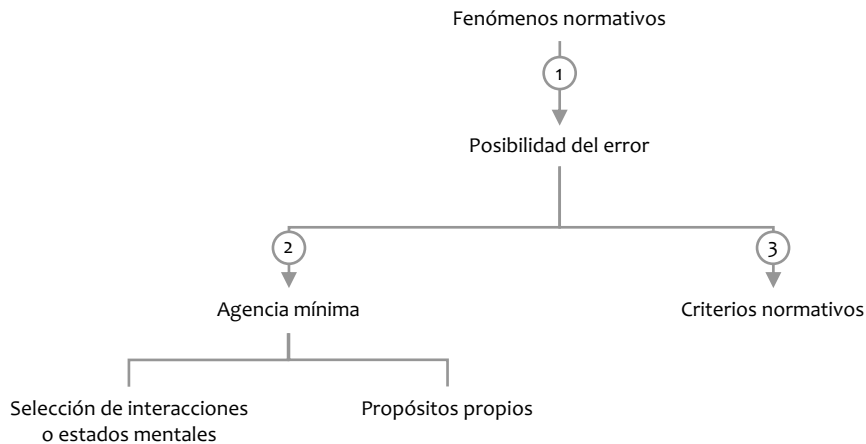


Figura 1

Las flechas simbolizan una relación de implicación, de manera que  $A \downarrow B$  significa que A implica B, o, lo que es lógicamente lo mismo, que B es condición necesaria de A. ① Mostré que para que un fenómeno sea normativo es necesario que exista la posibilidad del error y esta posibilidad hace que tenga sentido diferenciar entre fenómenos correctos e incorrectos, adecuados o inadecuados, etc. Los fenómenos normativos pueden ser comportamientos, interacciones entre un agente y su ambiente o estados mentales de un agente. Argumenté que hay dos condiciones necesarias para que exista la posibilidad del error, una agencia mínima y criterios normativos. ② Definí la agencia mínima como constituida por dos características. En la gráfica, la línea sólida simboliza la relación de constitución, de manera que  $A \downarrow B \downarrow C$  significa que A está constituido por B y C conjuntamente. La definición de agencia mínima es la siguiente: un sistema es agente mínimamente cuando tiene la posibilidad de seleccionar las interacciones que establece con otros y con el ambiente material; y cuando además tiene la posibilidad de distinguir entre los cursos de interacción que selecciona y para ello argumenté que el sistema debe tener propósitos propios. ③ Finalmente argumenté que es necesario que haya criterios normativos para determinar la corrección del comportamiento o el estado mental seleccionado por el agente. Mostré que tales criterios normativos son, en muchos casos, distintos a los propósitos propios del agente y por ello existe la posibilidad del error incluso cuando los agentes tienen la impresión de estar en lo correcto o cuando actúan de acuerdo con sus propósitos propios.

## 1. 3. Normatividad como consecuencias prácticas

En esta sección defiendo una concepción de normatividad, agencia y criterios normativos en términos de consecuencias prácticas. Esta concepción de normatividad es una interpretación alternativa a la interpretación representacional de normatividad y de las dos condiciones necesarias para que haya fenómenos normativos que identifique, y tiene dos funciones en la tesis: primero, sustancia el fenómeno que busco capturar con la definición naturalizable de normatividad que propongo en el Capítulo 4, y segundo, sirve como guía para entender y evaluar las aproximaciones a la cognición que examinaré en el resto de la tesis.

Para construir esta noción pragmática de normatividad partiré de las observaciones de Wittgenstein en las *Investigaciones filosóficas* (2009 [1953]) y las *Observaciones sobre la filosofía de la psicología* (1980). En particular, tomaré el compromiso radical de Wittgenstein en su filosofía tardía con la concepción de significado como uso (IF, §§ 7, 38, 43, 390) que resulta de llevar hasta sus últimas consecuencias dos convicciones de la filosofía analítica del siglo XX: primero, el sentido, que en la modernidad se supuso privado, debe ser público y, segundo, la mente puede explicarse solo a través de una explicación del lenguaje (Cely Ávila & Mojica, 2019, p. 917). Esta última idea es altamente problemática porque elimina la posibilidad de que existan mentes sin lenguaje. Esta posibilidad, sin embargo, no puede ser desechada en principio en una definición, sino que debe ser una pregunta abierta, dada la noción de mente que adoptemos, y debe responderse empíricamente. A pesar de ello, muestro en esta sección, las concepciones de lenguaje y sentido como uso de Wittgenstein permiten trazar un continuo de sentido entre formas de vida humanas y no humanas, y deja abierta la posibilidad de formas de hacer sentido no lingüísticas.

Propongo cuatro pasos para llegar a la idea de normatividad como consecuencias prácticas: en el primer paso, continúo con la discusión sobre cómo concebir el sentido centrada en el sentido lenguaje y defiendo una concepción wittgensteiniana de sentido como uso. Veremos que esta aproximación da pie a una concepción de sentido y mente que no se limita a los seres humanos. En el segundo paso, a partir de esa concepción de sentido, defiendo una concepción de agencia cognitiva orientada al uso en donde los criterios normativos propios de la cognición se establecen en las regularidades que forman las prácticas sociales compartidas. El tercer paso está dedicado a discutir estas regularidades, en particular, a distinguirlas de otras regularidades, las nomológicas, que no constituyen criterios normativos. Esta discusión me permitirá identificar un tercer tipo de regularidades, las biológicas, que aunque no se establecen en prácticas sociales, son normativas. Finalmente, en el cuarto paso, desarrollo la idea de que ser agente cognitivo es pertenecer a prácticas sociales y muestro además que hay criaturas que participan de regularidades normativas, pero no sociales, como las regularidades biológicas, y en esa medida, tienen comportamiento normativo. En suma, defiendo que los criterios normativos propios de un fenómeno normativo están constituidos por regularidades subdeterministas que se forman por los comportamientos, interacciones y la materialidad específica que habitan los individuos que conforman tal regularidad, mientras que la agencia mínima consiste en participar de tales regularidades y tener la posibilidad de seleccionar comportamientos o cursos de acción individuales que se alineen o no con tales regularidades.

### 1.3.1. Primer paso: Sentido del lenguaje como uso dentro de prácticas compartidas

Vimos que las aproximaciones representacionales conciben la normatividad en términos de verdad, más concretamente, en términos de representar adecuadamente los estados de cosas del mundo y monitorear de manera confiable sus cambios. Formarse una representación adecuada del mundo puede considerarse un propósito en sí mismo respecto del cual pueden diferenciarse los juicios cognitivos entre mejores o peores (Pitt, 2020). En un giro pragmático, autores como Cussins (2002) y Clark (2015) han propuesto que el propósito de las representaciones es fundamentalmente orientar al agente para que actúe en el ambiente. Sea cual sea el propósito, las aproximaciones representacionales definen los criterios de normatividad como la correspondencia entre contenidos representacionales y propiedades del mundo, en donde las propiedades relevantes son o bien las que permiten hacer inferencias válidas o bien las que permiten al agente actuar. La aproximación al sentido que quiero defender es más radical que este giro pragmático de las últimas décadas. Defiendo una aproximación al sentido, incluido el sentido del lenguaje, que prescindiera de las representaciones como explanans básico y que se base en cambio en el uso, la acción, la interacción y la práctica.

Para motivar la concepción de sentido como uso, quisiera comenzar mostrando los problemas de la noción de sentido como representación con condiciones de verdad. Bajo la concepción representacional, el caso paradigmático de significado es una afirmación. Una afirmación es significativa cuando es causada por una representación de un estado de cosas que se asume verdadero y que puede ser de hecho verdadero o falso. A partir del significado de las afirmaciones puede derivarse el significado de los demás tipos de oraciones: una pregunta corresponde a un conjunto de representaciones de estados de cosas de las que el agente busca saber cuál es verdadera y cuál es falsa, una orden es una representación de un estado de cosas que no es el caso y que se le pide a quien va dirigida la orden que lo haga verdadero, etc. Esta idea de significado viene de la impresión de que el significado del lenguaje es una colección de nombres que designan proposiciones o representaciones en la mente o en el cerebro de un agente en virtud de la cual el lenguaje es significativo. Bajo esta misma imagen, una acción es significativa en tanto que es causada por una creencia y un deseo que son entidades mentales representacionales que pueden ser nombradas con una proposición —ej. creo que si  $x$ , entonces  $y$ ; deseo que  $y$ . En suma, el lenguaje es significativo porque nombra proposiciones o representaciones.

Sin embargo, “esta concepción de lenguaje como colección de nombres tiene varios problemas. Para comenzar, hay expresiones lingüísticas problemáticas como «¡ayuda!», «hola», «¡no!» para las que es difícil ver cómo pueden nombrar algo, pero que son perfectamente significativas (IF, § 27). Segundo, asumir de entrada que las expresiones lingüísticas significativas deben nombrar algo crea problemas metafísicos: ¿Qué tipo de entidades son los números, el amor, la justicia, etc.?” (Mojica, 2014, p. 15). Y si no son entidades metafísicas y se asume que aquello que nombran son estados o eventos en el cerebro de quien usa tales expresiones, ¿cómo podría esa persona asegurarse de que de hecho



está ocurriendo lo que debe ocurrir en su cerebro cuando habla de números o de amor o del color azul?. Quizás, lo que es azul para ella es rojo para otra persona y lo que es amor para ella es lástima para otra. Esta pregunta nos lleva a un tercer problema, un problema epistemológico: se vuelve enigmático cómo podemos aprender un lenguaje, en especial, este tipo de palabras, cuando aquello que constituye su significado, o bien es escurridizo y abstracto como en el caso de las ideas metafísicas, o bien está oculto en el cerebro de las personas. Aún más sorprendente, ¿cómo es que podemos usar esas palabras tan frecuentemente y entendernos los unos a los otros sin si quiera saber con certeza su significado real? (Mojica, 2014, p. 15). Concebir el significado del lenguaje como uso evita estos problemas metafísicos y epistemológicos.

De acuerdo con Wittgenstein, “tanto lenguaje como acción son significativos sólo en la medida en que las acciones concretas y los usos situados del lenguaje tienen sentido para otros dentro de la comunidad de la que el agente hace parte. El significado no pertenece a una esfera interna y privada, sino a nuestras actividades públicas y compartidas. La razón es que el significado necesita criterios de corrección disponibles para todos los hablantes, es decir, criterios que nos permitan a todos actuar y hablar significativamente, y reconocer lo que tiene sentido y cuál es su sentido en contraposición de lo que no lo tiene” (Cely Avila & Mojica, 2019, p. 917). Si este no fuera el caso, es decir, si los criterios de corrección no estuvieran disponibles para todos los hablantes, ellos no podrían distinguir lo que es de hecho correcto de lo que solo parece correcto tanto para sí mismos como para otros. Pero sabemos que de hecho todos los hablantes son capaces de reconocer los usos correctos del lenguaje, es decir, los usos que tienen sentido, y son capaces de reconocer cuál es su sentido —esto lo sabemos porque los hablantes son capaces de hacer actividades que requieren el uso correcto del lenguaje, como interactuar con otros y vivir en sociedad. El único terreno común que permite distinguir lo que tiene sentido de lo que no disponible para todos los hablantes es en el patrón de hacer, esto es, acciones y usos del lenguaje, que constituyen la forma vida cotidiana de una comunidad de hablantes y que estos comparten en el contexto material en el que viven. En otras palabras, el criterio último de sentido son las prácticas compartidas.<sup>3</sup>

De acuerdo con la concepción de sentido como uso, sonidos y marcas en el papel tienen sentido en la medida en que hacemos algo con ellas. Para ilustrar este punto, vale la pena considerar un ejemplo muy simple: imaginemos cómo el término lingüístico ‘manzana’ puede ser aprendido por alguien que carece de lenguaje. “Somos entrenados<sup>4</sup> para responder a

---

<sup>3</sup> Es posible que haya fenómenos normativos distintos al lenguaje y cuyos criterios de corrección no estén establecidos por prácticas sociales. Como argumento en la primera mitad del capítulo 5, este es precisamente el caso para la normatividad biológica cuyo criterio de normatividad más básico es la vida y la muerte que son eventos o estados públicos y que hacen una diferencia práctica para los individuos.

<sup>4</sup> Wittgenstein usa el término entrenamiento en vez de enseñanza, aprendizaje o socialización para enfatizar que se trata de un proceso en donde fundamentalmente se moldea el comportamiento de los agentes y se abren posibilidades de acción que adquieren sentido en la acción misma y la interacción y reconocimiento de otros. Este énfasis se contraponen a la imagen del proceso de socialización o aprendizaje como una implantación de ideas en la mente interna que parece ser implicado por la concepción representacional del sentido en el que las representaciones son entidades internas. Este término además insinúa una continuidad entre nuestras formas de

‘manzana’ tomándola de la mesa y dándosela a alguien para que se la pueda comer, dibujándola y señalando otros dibujos de manzana, tomándola de entre otras varias frutas, etc. La similitud de estas actividades con juegos —el niño sigue las reglas del juego cuando su cuidador dice ‘manzana’— revela una característica esencial del lenguaje, a saber, que está constituido por juegos del lenguaje de nuestras expresiones lingüísticas en las que hacemos algo con ellas y seguimos reglas. Estas reglas no son completamente deterministas y, la gran mayoría de las veces, no son formuladas explícitamente: así como un pianista puede tocar una misma pieza de formas distintas, el niño puede tomar la manzana de distintas formas y seguir jugando el juego de la ‘manzana’” (Mojica, 2014, p. 16). Así que aprender un lenguaje, en este caso el término ‘manzana’, es participar en estas actividades, en estas formas de vida (IF, § 23); y en la medida en que aprendemos más expresiones lingüísticas, adquirimos más herramientas para involucrarnos en actividades más elaboradas alrededor de las manzanas y nuestra vida se vuelve más sofisticada; por ejemplo, podemos hablar de las dinámicas mundiales de comercio de las manzanas, podemos investigar su origen evolutivo, podemos acumular, compartir y crear recetas con manzanas, etc. Es por ello que “entender un lenguaje significa dominar una técnica” (IF, § 199). El significado de nuestras expresiones es entonces una familia de haceres en una familia de contextos similares que no implica formarse representaciones internas y bautizarlas.

“En el famoso argumento del lenguaje privado (IF, §§ 243-272), Wittgenstein se enfrenta a un caso en el que aparentemente el significado no proviene de nuestras prácticas compartidas sino de una ocurrencia mental indudablemente privada: las sensaciones, en particular, el dolor. Si aceptáramos que nuestras expresiones de dolor son significativas en virtud de que se refieren a un estado mental interno y privado, admitiríamos que el significado de tales expresiones es sólo accesible para un único individuo: quien padece el dolor. Pero si este fuera el caso, no sólo los otros sino el individuo mismo no tendría forma de corroborar si se está refiriendo o no a la misma sensación que supuestamente dio significado originalmente a la expresión” (Cely Avila & Mojica, 2019, pp. 917-918). Cualquier sensación que al individuo le parezca que corresponde con la representación de la instancia ejemplar de dolor, contaría como dolor. Pero si este fuera el caso, el individuo mismo no tendría manera de corroborar dentro de su mente privada si en verdad se trata de lo mismo, “entonces no existirían criterios del uso correcto de tales expresiones, y en esa medida estaría embarcado en un pseudo-juego sin reglas y, por ello, sin sentido (IF, §§ 244, 258, 261, 268). Apelar a las circunstancias o a las expresiones corporales, como regularmente hacemos, es apelar a criterios públicos, y es esto lo que nos permite dar sentido a nuestras propias sensaciones y a las de los otros. De manera que apelar a un fenómeno interno y privado como las representaciones o a un significado metafísico no juega ningún papel en la configuración de sentido ni para otros ni para nosotros mismos.” (Cely Avila & Mojica, 2019, pp. 917-918). En la siguiente sección discutiré con detalle las consecuencias que esta concepción de significado como uso compartido en prácticas sociales tiene para nuestra concepción de fenómenos mentales y de la mente en general.

---

vida humanas y formas de vida no humanas que parece apropiado dado el propósito de naturalizar la cognición y la normatividad.

Si el sentido de las expresiones lingüísticas es que estas permiten actuar a los hablantes, es crucial poder diferenciar acciones de comportamientos sin sentido. El caso del lenguaje, en particular, los criterios de corrección que establecen qué cuenta como uso del lenguaje, si es un uso correcto y cómo entenderlo muestran cómo hacemos estas distinciones. Los criterios para distinguir usos del lenguaje de, ej., balbuceos, deben estar disponibles para todos los hablantes del lenguaje pues todos somos capaces de hacer tales distinciones y de responder adecuadamente a los usos del lenguaje de los otros. De nuevo, la respuesta está en los patrones de comportamientos compartidos. Cuando nos enseñan un lenguaje por primera vez, nos entrenan para hacer lo que hacen otros, i.e., usar el lenguaje como los otros, responder a sus acciones e interactuar con ellos como ellos lo hacen, etc. Los criterios para diferenciar entre lo que tiene sentido y lo que no están constituidas precisamente por el curso de comportamientos e interacciones<sup>5</sup> que forman un patrón en comunidad. Tal patrón está disponible públicamente en tanto que otros pueden ser entrenados para hacer parte de ese patrón, y en tanto que todos quienes han sido entrenados para seguir este patrón tienen la capacidad de reconocer si alguien usa el lenguaje con significado o no, esto es, pueden reaccionar a este para interactuar con él porque hace sentido, y si no lo hace, pueden corregirlo o excluirlo. En suma, las regularidades de nuestra vida compartida y nuestro uso del lenguaje son regularidades que constituyen el criterio del significado del lenguaje, y en ese sentido, el piso de inteligibilidad es el patrón de haceres compartidos que constituyen nuestra vida cotidiana.

Hay dos puntos que quisiera resaltar respecto a la constitución de los patrones de haceres compartidos como criterios normativos.

1. Primero, los haceres y reacciones aceptables no están exhaustivamente definidos en cada caso posible por nuestro patrón actual de haceres compartidos. Estos patrones dejan espacio para casos sub-determinados en los que seleccionamos un hacer o una reacción para los que no existe aún un patrón de reconocimiento y respuesta, pero para los que se puede construir tal patrón con otros. En la analogía con el juego este caso sería como hacer una interpretación novedosa de las reglas del juego, expandirlas o incluso inventar un nuevo juego a partir de otro. Por ello, las reglas que le dan sentido a nuestras acciones y usos del lenguaje no son normatividades omniabarcantes, son en cambio patrones que se mantienen y se interpretan constantemente en nuestras acciones y reacciones concretas.
2. Segundo, el contexto material en el que la vida humana ocurre constriñe las posibilidades de acción que tenemos en nuestras prácticas sociales, y en ese sentido, constriñe también el sentido y los criterios de corrección de lo que hacemos y lo que creemos. Algunas de estas posibilidades son incorporadas en nuestras prácticas

---

<sup>5</sup> El concepto de interacción en este contexto y en el contexto de los debates de ciencias cognitivas enactivas y ecológicas que veremos más adelante se refiere a las interacciones agente-agente y agente-ambiente en tiempo real. Este uso difiere del uso estándar que se le da a este concepto en la tradición de las ciencias sociales en el que se limita a las interacciones agente-agente. En el Capítulo 5 discutiré conceptos sociológicos y allí será importante tener en mente el uso amplio de las ciencias cognitivas del concepto de interacción.

sociales como parte del piso sobre el que estas se sostienen. Por ejemplo, el hecho de que no podamos volar sin ayuda de artefactos o que no podamos ir de un lugar a otro instantáneamente están incorporados en nuestras prácticas de transporte (caminar, ir en bus, carro, metro, bicicleta, avión, etc.) y de citarnos los unos con los otros (hacer reuniones con antelación suficiente para llegar a tiempo, evitar caminos para evitar acumulaciones de personas, etc.). Estos hechos son rara vez parte de lo que discutimos explícitamente, pero hacen parte de las certezas prácticas sobre nuestras posibilidades e imposibilidades de acción. En ocasiones, las certezas prácticas que sostienen una forma de vida pueden volverse creencias explícitas y ser refutadas, como sucedió, por ejemplo, con la creencia de que la tierra es plana. Esta creencia hizo parte de nuestras certezas prácticas hasta que la tecnología nos permitió incorporar en nuestra cotidianidad las posibilidades que su forma esférica ofrece, por ejemplo, poder viajar literalmente alrededor del mundo.

Hasta ahora hemos visto que los criterios que determinan el sentido y la corrección deben estar disponibles públicamente para todos los individuos que hacen parte de una forma de vida. El sentido de las expresiones lingüísticas está determinado por lo que hacemos con ellas dentro de actividades que compartimos con otros y tales actividades tienen sentido en tanto que hacen parte de hacer y de reaccionar a otros que compartimos con la comunidad a la que pertenecemos.

Por supuesto, para comportarse con sentido no es necesario siempre interactuar con otro o ser de hecho evaluado por otro en el momento en que el comportamiento es llevado a cabo. Es posible, por ejemplo, tener un diario del que nadie jamás sepa y que nadie jamás lea o hacer un ejercicio matemático que no sea nunca revisado por nadie, y hacer sentido al llevar a cabo estas actividades. La razón es que tanto llevar un diario como hacer ejercicios matemáticos son actividades que hacen parte de prácticas sociales más amplias para las que somos entrenados en la escuela y en casa durante varios años de nuestra vida. Así, hacemos el ejercicio matemático para, ej., practicar una habilidad nueva que estamos aprendiendo o verificar si una estimación que hicimos unas horas antes fue acertada; llevamos un diario para, ej., organizar nuestras emociones o resaltar los eventos del día que con los que aún no sabemos qué hacer o por los que nos sentimos agradecidos. Por lo tanto, estas actividades específicas son significativas no solo porque fuimos entrenados por otros para llevarlas a cabo, sino porque hacen parte del patrón amplio de nuestra vida en sociedad en donde de hecho somos reconocidos e interactuamos con otros. Esto implica que el ejercicio de evaluación normativa es parte constitutiva de nuestras prácticas. Tal ejercicio de evaluación siempre ocurre con otros miembros de la comunidad y en esa medida ocurre desde un agente interno a la práctica.

Esta concepción de criterios de normatividad tiene dos consecuencias que quisiera resaltar. Primero, el sentido está extendido temporalmente. El sentido de un evento o comportamiento está determinado por el patrón de actividades en que se inserta. Esto implica que el sentido es un fenómeno que está extendido temporalmente porque su significado depende tanto de la historia de entrenamiento del individuo por parte de otros para pertenecer a prácticas sociales

como por las consecuencias prácticas respecto a las posibilidades de interacción con otros y de llevar a cabo actividades exitosamente. Es importante notar que ni el entrenamiento pasado de los individuos ni las posibilidades que un evento abre determinan completamente el sentido de un comportamiento: puesto que el sentido de un comportamiento depende de su lugar dentro del patrón compartido de haceres de una comunidad, y los patrones existen solo extendidos en el tiempo desde el pasado y hasta el futuro.

La segunda consecuencia de esta concepción de sentido que quiero resaltar permite mostrar por qué la concepción de sentido como representación es problemática: ni las reglas lógicas ni definiciones abstractas ni los fenómenos internos pueden servir como criterios públicos de significado porque no están disponibles para todos los hablantes. Así mismo, representaciones internas no pueden ser significativas en sí mismas porque, dado su carácter privado, estas no están disponibles públicamente para ser juzgadas o para servir como criterios para juzgar la corrección y el significado del lenguaje. Por lo tanto, naturalizar la normatividad requiere una concepción pública de normatividad. Vale la pena revisar un segundo argumento para justificar por qué el sentido del lenguaje es un fenómeno público.

Wittgenstein nos invita a considerar cómo determinamos si el uso de una palabra como «cubo» es correcto o no (IF, §138-150). Uno podría imaginarse que es correcto cuando coincide con una imagen mental o una representación interna de un cubo, o incluso, cuando se subsume a una definición general de cubo. Pero en cualquiera de esos casos, podemos preguntar: ¿Cómo puedo estar segura de que hice coincidir correctamente el objeto que llamo «cubo» en esta instancia de uso particular con la representación interna que tengo? ¿Cómo puedo estar segura de que apliqué la definición de forma correcta? Podríamos intentar dar una regla general más o un modelo interno que indique cómo aplicar la definición, pero surge el mismo problema: ¿cómo puedo estar segura de que estoy interpretando esa regla o modelo general de aplicación correctamente? La única salida a este potencial regreso al infinito, un regreso al infinito que llevó a Kripke (1982) a su interpretación escéptica de la filosofía de Wittgenstein y así dar a luz a Kripkenstein, es apelar a nuestras prácticas compartidas del uso del lenguaje. El que una expresión lingüística sea usada correctamente o no está determinado por si otros hablantes del lenguaje pueden entenderla, es decir, por si pueden o no actuar con el otro o en respuesta al otro —v.g., «corta la berenjena en cubos», «un cubo tiene exactamente ocho caras, mientras que una pirámide tiene al menos cuatro», «para dibujar una casa, empieza por dibujar un cubo», «determine el área de un cubo cuyo lado mide 0.3 metros», etc. Tales reacciones a las acciones de otro y a sus usos del lenguaje son una manifestación de que lo entendemos, y esto es posible cuando usamos un lenguaje en la medida en que existe una costumbre, una práctica compartida que los conecta (IF, §198, 202).

Estos problemas de interpretación o correspondencia entre representaciones mentales y usos del lenguaje o aspectos en el mundo “sugieren que no hay un conjunto de reglas abstractas que sostengan y definan universalmente cómo usar e interpretar el lenguaje correctamente, ni en tanto una característica metafísica constitutiva ni como una representación en nuestras mentes” (Mojica, 2014, p. 17). Si fueran aquello que determina de hecho la corrección del lenguaje, “el uso del lenguaje sería una interpretación de tales reglas; pero, en palabras de

Wittgenstein, toda interpretación pende en el aire con aquello que interpreta, y no puede darle ningún apoyo. Las interpretaciones por sí mismas no determinan el significado. (IF, §198). Las reglas abstractas son irrelevantes para nuestro curso de acción, y son insuficientes para explicar que hagamos esto en vez de aquello al usar el lenguaje porque estas necesitan ser interpretadas. En cambio, el significado está determinado por el uso, y por tanto, reglas no pueden ser el criterio para determinar qué tiene sentido y qué no” (Mojica, 2014, p. 17).

Podemos volver ahora a la afirmación que hice arriba: las reglas de la lógica no sirven como criterios públicos de normatividad para determinar el sentido del lenguaje. La primera razón es que el sentido de muchas de nuestras expresiones lingüísticas no está determinado por sus condiciones de verdad porque no las tienen; por ejemplo, los saludos, las expresiones onomatopéyicas, etc. La segunda razón y la más importante es que aun cuando sea posible analizar lógicamente otras muchas clases de las expresiones lingüísticas que usamos como las afirmaciones, las preguntas, las indicaciones, etc., estas formas de análisis hacen parte de una disciplina especializada y las formas de usar y responder al lenguaje cuando se hace lógica no están disponibles para todos: solo una minoría de los hablantes podría reconocer los conectores lógicos relevantes para la lógica clásica o la lógica modal, y aun menos hablantes podrían usar con éxito las reglas de derivación lógica clásica o modal o cualquier otra una vez se tengan las traducciones lógicas de sus expresiones en lenguaje natural. Sin dominar estas reglas lógicas es imposible entender la forma lógica de nuestros enunciados y la validez de nuestros argumentos en términos lógicos, y puesto que solo una muy pequeña minoría domina tales reglas, pero todos hablamos el mismo lenguaje, estas reglas lógicas no pueden ser el criterio público de normatividad que determina la corrección y el sentido del lenguaje.

La Figura 2 resume un aspecto importante de la concepción de sentido como uso que he defendido en esta sección: el sentido de lo que hace un individuo se fundamenta en última instancia en si este le permite continuar interactuando con otros y llevando a cabo nuestras actividades de forma fluida. Para que esto sea el caso, es decir, para que interacciones y actividades continúen de forma fluida, deben

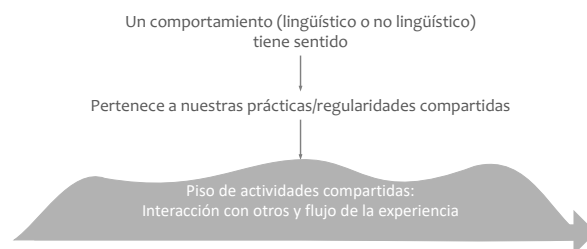


Figura 2

Las flechas simbolizan una relación de implicación, de manera que  $\downarrow$  significa que A implica B, o, lo que es lógicamente lo mismo, que B es condición necesaria de A.

continuar los patrones del pasado y permitir que se sigan construyendo hacia el futuro. Vale la pena notar que continuar construyendo patrones de haceres compartidos hacia el futuro no implica conservar las formas en las que se han llevado a cabo hasta ahora; es posible innovar en nuestros haceres y continuar con patrones que constituyen el sentido sin conservar las formas y valores, por ejemplo, la forma en que hoy continuamos con la práctica del matrimonio y permitimos matrimonios entre parejas del mismo sexo continúa e innova la práctica del matrimonio del pasado.

En la sección anterior vimos que las aproximaciones representacionales a la mente consideran que el propósito propio de la cognición es formarse representaciones verdaderas, adecuadas, etc. y conservar esa verdad. Mostré que esa concepción de cognición se sigue de una concepción de sentido en términos de verdad. Sin embargo, si el sentido está determinado por los usos que constituyen un patrón compartido en una comunidad, es decir, nuestras actividades compartidas, es dudoso que la cognición pueda definirse a partir de una normatividad veritativa; después de todo, las menos de nuestras actividades tienen como propósito encontrar y preservar la verdad. ¿Qué concepción de mente se sigue entonces de la concepción de sentido como uso que he defendido hasta ahora? Una primera pista es que ser un agente cognitivo depende de la capacidad de incorporarse a los patrones de hacer y usos del lenguaje que constituyen nuestras formas de vida. En la siguiente sección desarrollo esta idea y muestro que el tener mente depende crucialmente de ser un cuerpo vivo, i.e., un cuerpo cuyas acciones expresan propósitos intrínsecos y una vida afectiva.

### 1.3.2. Segundo paso: La agencia cognitiva como pertenencia a prácticas mentales

En esta sección discuto la concepción de agente cognitivo que se sigue de la concepción de sentido como uso y reconocimiento en prácticas compartidas que defendí en la sección anterior. Estas actividades incluyen no solo lo que decimos, sino lo que hacemos con nuestros cuerpos, cómo interactuamos con otros, las circunstancias en las que nos encontramos, etc. Así que, puesto que el sentido se constituye en nuestras prácticas compartidas, la estrategia más adecuada para comenzar es examinar las prácticas en las que atribuimos mente a otros. De esa manera, podemos examinar los rasgos esenciales del fenómeno de ser un agente o del tener mente de la forma en que lo entendemos cotidianamente y de la forma en que hace parte constitutiva de nuestras vidas. No todas las prácticas de nuestra vida cotidiana son relevantes para este propósito en tanto que no es necesario ser un agente cognitivo para pertenecer a ciertas prácticas, por ejemplo, para mover objetos mecánicamente o para seguir algoritmos. En oposición, para atribuir sensaciones, emociones o creencias a otros es crucial atribuir vida mental a otros y ese es el tipo de prácticas en las que me centraré para abordar la pregunta por en qué consiste ser un agente cognitivo.

#### Prácticas mentales

Para entender qué significa para nosotros cualquier fenómeno de nuestra vida mental y en particular el dolor, Wittgenstein nos invita a considerar nuestro uso de expresiones lingüísticas relacionadas con estos fenómenos, esto es, hacer una investigación conceptual. En general, las investigaciones conceptuales que Wittgenstein lleva a cabo en las *Investigaciones Filosóficas* buscan revelar los rasgos constitutivos de las prácticas en las que usamos el lenguaje. A diferencia de las investigaciones empíricas, las investigaciones conceptuales no buscan descubrir las causas específicas de la formación de conceptos, sino hechos generales de la naturaleza que fundamentan la forma de vida dentro de la que el uso del lenguaje tiene sentido (IF, xii). Como Dromm (2003) ha señalado, estas investigaciones conceptuales se hacen a través de historias naturales ficticias en las que imaginamos “cómo aprenden los

niños a usar ciertas palabras y los hechos generales de la naturaleza que podrían pertenecer a las actividades implicadas en nuestro uso de dichas palabras... para revelar cómo sonidos y marcas en un papel son expresiones con significado” (Mojica, 2014, p. 19). Así que una investigación conceptual sobre nuestras expresiones lingüísticas de dolor muestra los hechos generales de nuestra forma de vida en las que el fenómeno del dolor tiene algún sentido.

Un bebé aprende a usar las expresiones lingüísticas de dolor fundamentalmente como reemplazo de su llanto. Cuando el bebé llora, sus cuidadores hablan con ella, reconocen sus expresiones y las circunstancias en las que se encuentra y atienden lo que identifican como sus necesidades: «¿ya tienes hambre? Sí, ¡es hora de comer!» o «ay, ¿te pegaste en la cabeza? Ya, ya, ya pasará» o «¿qué pasa? ¿será que tienes fiebre? No, ¿será que estás cansada? Seguro, estuvimos paseando toda la tarde. Vamos a tomar la siesta». Con el paso de los años, los niños aprenden a usar ese lenguaje ellos mismos. Esto implica que aprenden a reconocer las circunstancias en las que, en vez de llorar, es apropiado decir, por ejemplo, «me caí y me duele la pierna» como expresión de dolor; implica entender que aquello que siente en su cuerpo es dolor, que tiene una causa y que puede ser atendido con estiramientos, descansando, limpiando la herida si la hay, etc.

Aprender un lenguaje es tan natural para nosotros como aprender otras actividades que sostienen la vida humana: [o]rdenar, preguntar, relatar, charlar pertenecen a nuestra historia natural tanto como andar, comer, beber, jugar. (IF, §25). El ejemplo del dolor del párrafo anterior ilustra un fenómeno más general respecto a nuestra capacidad de sentir: aprender a usar el lenguaje significa aprender a darle forma a nuestras sensaciones, emociones y estados de ánimo, expresarlos usando el lenguaje o embarcándonos en actividades socialmente construidas como ir a una cantina a celebrar o comprar rosas y enviarlas para mostrar gratitud, aprendemos a entender lo que sentimos como eventos que nos ocurren en circunstancias específicas que podemos explicar a otros, que se sienten en nuestro cuerpo de maneras específicas y que nos hace acercarnos y entender el mundo de maneras específicas. En términos aún más generales, nuestros términos lingüísticos que se refieren a habilidades y actividades mentales como dudar, creer, recordar, calcular, especular, imaginar, interpretar, etc. tienen sentido dentro de prácticas compartidas con otros en las que le damos forma y sentido a nuestros comportamientos, a nuestra historia de vida y a nuestra experiencia. Tales actividades, las actividades que expresan nuestra vida mental pertenecen a prácticas sociales y constituyen “partes de nuestra vida que se influyen mutuamente y juntas le dan forma a nuestra cotidianidad” (Mojica, 2014, p. 18). Estos conceptos son precisamente los que nos interesa estudiar en ciencias cognitivas y psicología. Al respecto, Wittgenstein afirma:

Los conceptos psicológicos son conceptos cotidianos, nada más. No están recientemente fabricados por la ciencia para su propio propósito, como sí lo están los conceptos de la física y la química. Los conceptos psicológicos están relacionados con aquellos de las ciencias exactas como los conceptos de la ciencia médica lo están a aquellos de las ancianas que pasaban su tiempo cuidando a los. (1980, § 62)

Por lo tanto, para aproximarnos de manera naturalista a la mente, es indispensable no perder de vista que nuestra mente se constituye en nuestras prácticas y en situaciones materiales



específicas. Esto no significa que en principio no exista otra posibilidad para que se constituyan criterios normativos para hacer sentido. Es posible que haya sistemas, naturales o no, que sean capaces de seleccionar y establecer interacciones con el ambiente que calificaríamos como cognitivas, pero que estén sujetos a criterios normativos que no estén constituidos en prácticas sociales. Para saber si esta posibilidad es de hecho el caso es necesario, primero, ofrecer una definición operacionalizable de cognición que capture la constitución mayormente social de los criterios normativos que rigen el comportamiento cognitivo humano, y segundo, mostrar empíricamente que existen sistemas que tienen este comportamiento. Pero ofrecer tal definición está fuera del alcance de esta tesis en la que me concentro en ofrecer una definición de normatividad que contribuya a la naturalización enactiva y ecológica de la mente.

Antes de continuar, quisiera abordar brevemente una pregunta sobre la compatibilidad de la filosofía de Wittgenstein con el proyecto que propongo en esta tesis, no solo para justificarlo mejor, sino para aclarar la manera en que busco que sean tomadas las caracterizaciones del sentido, significado, normatividad, etc. que hago en estas secciones a partir de la filosofía del Wittgenstein tardío.

¿Es realmente compatible la filosofía de Wittgenstein con un proyecto de naturalización?

Algunos autores han interpretado las observaciones sobre psicología y lenguaje de Wittgenstein en las *Investigaciones* como hipótesis empíricas sobre la adquisición del lenguaje; por ejemplo, David Bloor ha afirmado que, si las ideas de Wittgenstein hubieran recibido un tratamiento completo, serían como el conductismo de Skinner (1983, p. 63 en Dromm 2003, p. 676). Si este fuera el caso, se podría interpretar la concepción de normatividad, agencia y criterios normativos que ofrecí en la sección anterior como un conjunto de hipótesis que deben ser probadas por las ciencias cognitivas.

Esta lectura, sin embargo, es altamente problemática por dos razones. La primera es porque va en contra de la forma en que Wittgenstein mismo caracterizó sus observaciones. No solo él no ofrece nunca soporte ni motivación empírica para sus observaciones, sino que él mismo afirma:

Si la formación de conceptos puede ser explicada por hechos de la naturaleza, ¿no deberíamos estar interesados, no en la gramática [i.e., una investigación conceptual], sino en su fundamento en la naturaleza? ---Estamos de hecho interesados también en la correspondencia entre conceptos y hechos muy generales de la naturaleza. (Estos hechos generalmente no nos llaman la atención por su generalidad.) Pero nuestro interés no es por esa razón arrojado a las posibles causas de la formación de conceptos; no estamos haciendo ciencias naturales; ni tampoco historia natural --- porque podríamos también inventarnos historias naturales ficticias para nuestros propósitos. (PI, PPF, § 365)

La segunda razón y la más importante por la que no es adecuado tomar las concepciones de normatividad, agencia y criterios normativos que ofrecí en la sección anterior como un conjunto de hipótesis que deben ser probadas por las ciencias cognitivas es que ese no es el papel que tienen en el proyecto que propongo en esta tesis. No solo la manera en que caractericé estos conceptos apunta a hechos generales de la naturaleza carece de la especificidad para ser operacionalizados como hipótesis falseables, sino que su función es apuntar a hechos generales de la naturaleza que sirven como piso para que la diferenciación entre eventos, cosas o escenarios mejores o peores, correctos o incorrectos, etc. tenga sentido como un evento de la naturaleza.

Esta caracterización general sirve como primer paso en un proyecto de naturalización porque establece el sentido del fenómeno que vamos a investigar y eventualmente naturalizar. Caracterizaciones como esta son esenciales porque la ciencia, sino que se apoya en conceptos fundamentales que le dan sentido a los fenómenos que investiga, y esto incluye, por supuesto, a una ciencia preocupada por el sentido como fenómeno de la naturaleza como lo es la ciencia cognitiva.

### 1.3.3. Tercer paso: Regularidades del sentido, regularidades nomológicas y regularidades biológicas

En la sección 1.3.1 mostré que el criterio que nos permite diferenciar lo que tiene sentido de lo que no y cuál es su sentido es si pertenece o no al patrón de comportamientos compartidos por la comunidad a la que pertenece el agente. Mostré que tales patrones son públicos en tanto que, primero, otros pueden ser entrenados para hacer parte, y segundo, todos pueden reaccionar a otros hablantes para interactuar con ellos porque su comportamiento hace sentido, y si no lo hace, pueden corregirlos o excluirlos. Los patrones de actividades que constituyen el sentido forman regularidades, pero estas regularidades no son equiparables a las regularidades nomológicas, es decir, aquellas formadas por los eventos naturales.

Una diferencia crucial es respecto a la normatividad. Las regularidades que constituyen el sentido son regularidades respecto a las que podemos errar o estar en lo correcto, y en ese sentido, son normativas. En cambio, las regularidades nomológicas simplemente son y esto implica que, si un evento continúa o no tal regularidad, el evento no es correcto o incorrecto, simplemente es. El siguiente ejemplo ilustra esta diferencia normativa: llamar para preguntar los horarios en que abre un restaurante, contestar el teléfono y responder las inquietudes de los clientes pertenecen a la regularidad de hacer y responder preguntas y a la práctica más general de ir a comer a restaurantes. Los individuos pueden acertar o errar a la hora de seguir estas regularidades, ej., si el individuo marca el número de teléfono de su mamá (que no trabaja en un restaurante) o si usa un plátano para llamar, su comportamiento es incorrecto en el primer caso y en el segundo no solo es incorrecto sino que roza el sinsentido. En cambio, el que teléfonos y plátanos no desaparezcan de repente para aparecer en otro lugar del mundo es parte de las regularidades nomológicas y esperaríamos que las sigan el teléfono o el plátano particular que usa el individuo; pero si alguno de tales objetos no siguiera la regularidad, no se trataría de un evento incorrecto, sería quizás un evento extraordinario desde nuestra

perspectiva pero no uno evaluable normativamente. El que una regularidad sea normativa y la otra no implica que en un caso hay agentes, individuos o entidades que seleccionan un curso de acción y en el otro caso no hay selección de por medio. Implica además que en un caso hay consecuencias prácticas que hacen la diferencia para quien selecciona y en el otro caso, en tanto que no hay una entidad que selecciona, no es posible hablar de consecuencias prácticas.

Hay dos puntos respecto a la posibilidad de selección que vale la pena distinguir. En primer lugar, la especie humana se caracteriza porque su comportamiento es extremadamente moldeable. Nuestro repertorio de acción es tan pobre al nacer que somos incapaces de siquiera mantenernos con vida a nosotros mismos, pero esto permite que seamos entrenados para comportarnos de una u otra manera y que lleguemos a adquirir un amplísimo repertorio de posibles cursos de acción entre los que podemos seleccionar en circunstancias específicas. Eventualmente podemos escoger las formas concretas y las actividades específicas de entre ese repertorio no solo para suplir nuestras necesidades básicas vitales, sino también las actividades que constituyen nuestra vida cotidiana y cuyo valor no depende de su impacto en nuestro organismo.

El segundo punto respecto a la selección que quiero resaltar es que, si bien la maleabilidad y la amplitud del repertorio de acción humana es excepcional en la naturaleza, la posibilidad de selección no es exclusiva de la especie humana. Por el contrario, todas las formas de vida seleccionan cursos de acción para mantenerse con vida y su comportamiento puede ser correcto o incorrecto, exitoso o no con respecto a ello. El comportamiento de las criaturas vivas no está completamente determinado por su historia evolutiva ni por su genética, aunque ese sea el piso que lo hace posible, es en cambio un comportamiento flexible que responde a los cambios sutiles del ambiente y que busca seleccionar, la mayoría de las veces, lo más conveniente para mantener la vida del organismo. En ese sentido, la posibilidad de selección y el carácter normativo de tal selección se extiende por todo el reino de lo vivo. Sin embargo, esto no implica, pero tampoco excluye la posibilidad de que las regularidades y la normatividad implicada en este tipo de selección sean constitutivas de sentido. Esta observación sirve como motivación para distinguir un tercer tipo de regularidad: la regularidad biológica que, si bien es normativa con respecto a mantener la forma de vida de las especies, no surge del entrenamiento para pertenecer a prácticas compartidas. En la próxima sección regresaré a este punto y discutiré en qué medida podemos considerar que hay sentido y cognición en las regularidades biológicas.

Bajo la concepción de sentido como uso dentro de prácticas compartidas, el que un comportamiento que se selecciona esté sujeto a criterios de corrección o de éxito significa que tal comportamiento tiene consecuencias prácticas para el agente. Por un lado, en el caso particular de los seres humanos, la consecuencia relevante es si el agente logra llevar a cabo la actividad o no. Lo que esto implica concretamente depende de la práctica particular que consideremos. En el ejemplo anterior, la consecuencia práctica que se busca al llamar es lograr obtener la información de los horarios del restaurante y eventualmente ir a comer en el momento adecuado. En cualquier caso, la consecuencia última está determinada por si el patrón de comportamiento del agente pertenece o no a la práctica o la costumbre, y esto

significa en última instancia la posibilidad de interactuar con otros y, en esa medida, el ser reconocido por ellos como un agente que hace sentido. Por otro lado, las consecuencias prácticas de los comportamientos seleccionados biológicamente no radican en el reconocimiento de sentido del otro dentro de una práctica específica, sino en la posibilidad misma de mantenerse con vida. Esta posibilidad, sin embargo, implica también el mantener la interacción y las posibilidades de interacción con aspectos materiales del ambiente y con otras criaturas vivas que también lo habitan. Así que, en ambos casos, un buen indicador de que se está actuando conforme a los patrones normativos es que los agentes mantengan sus posibilidades de interacción; esto es sin embargo un indicador simplificado y general, y la evaluación normativa siempre dependerá del evento concreto en las circunstancias específicas.

Hasta ahora he mostrado que las regularidades que constituyen el sentido entre humanos son regularidades constituidas por el patrón de lo que hacen agentes que se comportan de formas parecidas, que interactúan entre ellos y que se entienden los unos a los otros. Mostré que estas son regularidades normativas en tanto que permiten determinar si la selección de comportamiento de un agente es correcta o incorrecta, y argumenté que el que tal selección de comportamiento sea correcta o incorrecta consiste en las consecuencias prácticas que tiene para el agente en términos de sus posibilidades de interacción. En suma, el fundamento y criterio último del sentido es que permita el flujo de actividades compartidas en un contexto material específico. Mostré además que hay una segunda regularidad normativa: las regularidades biológicas. Estas constituyen la forma de vida de cada una de las especies seres vivos, y son regularidades que surgen de los comportamientos y situaciones seleccionadas por las criaturas vivas. Los criterios de corrección en este caso no dependen exclusivamente del reconocimiento de otros como miembro de la práctica, sino de que se mantenga la vida individual y a lo largo de generaciones.

### 1.3.4. Cuarto paso: Criaturas con comportamiento normativo

#### Comportamiento cognitivo

Ahora que la naturaleza de las prácticas en las que se constituye el sentido es más clara, podemos responder una primera pregunta crucial: ¿Qué significa ser un agente cognitivo bajo la concepción de sentido como uso? Mostré que una condición necesaria es hacer parte de regularidades normativas, así que para responder esa pregunta es necesario aclarar en qué consiste pertenecer a regularidades normativas. Además, quedó abierta la pregunta por el tipo de regularidades normativas a las que se debe pertenecer para ser un agente que hace sentido, en particular, si las regularidades normativas biológicas pueden constituir sentido o no.

El criterio de corrección para saber si un comportamiento pertenece o no a una práctica es si el agente lleva a cabo una actividad exitosamente, y llevar a cabo una actividad exitosamente implica muchas veces que los otros puedan interactuar en respuesta a nuestros comportamientos: ej., nuestras preguntas reciben respuesta o hacen que alguien nos dirija a alguien más que quizás tiene la respuesta; al extender la mano, el otro nos la agarra y nos pregunta cómo hemos estado; insertamos un billete en una máquina y tomamos la bebida que

sale de una de sus puertas, etc. Situar el criterio normativo en la posibilidad de interacción dentro de las actividades mismas implica que para hacer sentido no es necesario y ni siquiera relevante que los agentes internalicen normas; pero entonces surge la pregunta ¿en qué consiste hacer sentido para el agente mismo, es decir, desde su punto de vista?

Un agente que hacen sentido tiene la intención de llevar a cabo una actividad, ej., tiene la intención de preguntar para saber las horas de atención del restaurante porque a su vez tiene la intención de ir esta noche o tal vez quiere engañar a otro y hacer parecer que tiene la intención de ir a cenar con él. Tener tales intenciones y propósitos estructuran el comportamiento del agente y le dan sentido. Si bien esta descripción es correcta, sería un error asumir que tales intenciones y propósitos son un algo interno en el agente, por ejemplo, una representación, una experiencia específica, un patrón de activación neuronal específico, una entidad mental simple o compuesta por la combinación de otros dos estados, un deseo y una creencia. Como argumenté arriba, explicar los comportamientos que tienen sentido apelando a entidades o fenómenos internos es errado porque no están disponibles para todos quienes son capaces de entender al agente. Pero ¿en qué consiste entonces tener una intención o propósito? Así como las sensaciones y el dolor, los propósitos no son un algo, pero tampoco una nada:

“¿Pero seguramente admitirás que hay una diferencia entre el comportamiento de dolor acompañado de dolor y el comportamiento de dolor sin dolor?”—¿Admitirlo? ¡Qué diferencia más grande podría haber!— “y aún así llegas una y otra vez a la conclusión de que la sensación en sí misma es una nada”—¡No, en absoluto. No es un algo, pero no es una nada tampoco! La conclusión fue solo que una nada puede servir tan bien como un algo de lo que nada se puede decir. Solo rechazamos la gramática que trata de imponerse sobre nosotros aquí.

La paradoja desaparece solo si rompemos radicalmente con la idea de que el lenguaje siempre funciona de una manera, siempre tiene el mismo propósito: comunicar pensamientos—que pueden ser sobre casa, dolores, el bien y el mal o lo que se te antoje. (IF, §304)

Los propósitos existen en tanto que son realizados o son hechos reales en el mundo. Esto significa que son expresados en alguna medida, y no que existan solo en tanto que sean realizados exitosamente; pero el tipo de actividades que significan la expresión de un propósito varían entre prácticas y actividades. Por ejemplo, el propósito de viajar a Islandia puede expresarse en la entrada de un diario, preguntar a otros por sus experiencias en ese lugar, buscar información en internet, darle ‘me gusta’ a las fotos de ese lugar, etc. En cambio, el propósito de saludar a un conocido en la calle puede realizarse en el saludo mismo o al tropezar con una viga al intentar hacerlo, etc.

Hay dos implicaciones que quisiera resaltar. Primero, actuar de acuerdo con propósitos no siempre requiere planes concretos ni decisiones deliberadas. En muchas de nuestras prácticas, actuar con propósitos consiste en interactuar con el mundo y con los otros, es decir, orientarnos corporalmente para interactuar de ciertas formas estructuradas. Segundo, aquello que estructura el comportamiento y lo hace una acción con sentido es que hace parte del

patrón compartido de actividades. De manera que, en ningún caso los propósitos se definen ni las acciones se estructuran por alguna experiencia o entidad interna característica, en cambio, lo que estructura el comportamiento y lo hace un comportamiento cognitivo con propósito es el pertenecer a prácticas compartidas.

Hay tres condiciones para ser un agente que entra a patrones compartidos y actúa con propósitos: es necesario (i) ser una entidad individual, (ii) interactuar constantemente con el ambiente y (iii) tener algo que perder y poder actuar para evitarlo. Las primeras dos condiciones pueden ser cumplidas, aunque de maneras distintas, por humanos y artefactos<sup>6</sup>. Por el contrario, la última condición es uno de los hechos generales de la naturaleza que permiten que nuestras formas particulares de hacer sentido surjan. Este hecho general es que somos cuerpos vivientes, sintientes y expresivos, capaces de actuar por motivación propia, capaces de percibir el ambiente y capaces de moldear los patrones de interacción propia. El ser un cuerpo vivo es crucial para que podamos decir con sentido que el otro tiene mente, en ese sentido, a la hora de discutir de qué tiene sentido decir que tiene vida mental Wittgenstein afirma:

Solo de un ser humano vivo y de lo que se parece a (lo que se comporta como) un ser humano viviente puede uno decir: tiene sensaciones, ve, está ciego, oye, está sordo, está consciente o inconsciente. (IF, § 281).

### Comportamiento normativo

Los argumentos wittgensteinianos considerados hasta ahora fueron formulados originalmente para ser aplicados al uso del lenguaje y su significado cuyos cimientos de observabilidad y de criterios públicos son las prácticas sociales compartidas por humanos en el contexto material en el que viven. Sin embargo, no solo los rasgos que nos permiten atribuir mente y sensaciones no se limitan a lo humano, sino que tampoco las condiciones que nos permiten hablar de fenómenos normativos se limitan a formas de vida humanas. Lo que nos permite reconocer comportamientos normativos no humanos son los comportamientos de un cuerpo vivo que es expresivo, que puede interactuar flexible y espontáneamente con su ambiente a lo largo del tiempo y que lo hace de acuerdo con sus propios propósitos y dentro de su propia forma de vida. Consideremos de nuevo rasgos de nuestro uso del concepto de dolor.

Imaginemos que una piedra siente dolor. ¿Cómo podría suceder? ¿Cómo trataríamos a la piedra en ese caso? (IF, § 284). Para darle sentido a ese escenario, quizás tengamos que imaginar una cara interna de la piedra que hace una mueca de dolor o un alma que grita por dentro, quizás un alma humana que quedó atrapada en una piedra, como en los cuentos de hadas. Es imposible darle sentido a la idea de una piedra que siente dolor si no la conectamos

---

<sup>6</sup> De acuerdo con el enactivismo, y con varios de sus antecesores intelectuales como Kant, Jonas y Simondon, los artefactos no son individuos *per se*, solo lo son en virtud de que sus partes son ensambladas como un sistema individual por su diseñador. En cambio, humanos y todas las especies vivas se constituyen a sí mismos como individuos. Esta distinción será crucial en los siguientes capítulos porque la naturaleza autonconstituída de los seres vivos permitirá explicar el surgimiento de los fenómenos normativos en la naturaleza.

de alguna manera con aquello a lo que le atribuimos dolor con sentido, pero aun así es confuso, no sabríamos cómo tratar su dolor o en general qué hacer respecto a su dolor. Para salir de esta confusión, Wittgenstein escribe:

Ahora mira a la mosca retorciéndose y enseguida estas dificultades desaparecen y el dolor parece poder agarrarse de aquí, sonde antes todo era, por así decirlo, demasiado liso para ello (IF, § 284).

Entendemos de inmediato que la mosca que se retuerce tiene dolor, y que quizás que está muriendo. Vemos de inmediato que la mosca es un cuerpo vivo que siente y que hace lo posible por mantenerse con vida. Lo vemos no solo en el retorcimiento de dolor de la mosca, sino en sus exploraciones para buscar comida, en sus intentos por escapar de nuestras torpes manotadas, etc. Aunque su forma de vida sea radicalmente distinta a la nuestra, es, como nosotros, un cuerpo vivo que busca mantenerse con vida. Esto resuena con la aproximación de Hans Jonas (2001 [1966]) a la vida, para quien solo la vida puede ser conocida por la vida: no dudamos de la vida del otro, lo entendemos de inmediato como un ser vivo porque nosotros mismos somos criaturas vivas.

Aunque la forma de vida y los comportamientos de animales no-humanos no son moldeables ni pueden ser entrenados en la misma medida en que los humanos sí lo son, los animales no-humanos también son cuerpos vivos, que sienten, expresan y actúan por motivación propia, también son capaces de percibir el ambiente y responder a sus cambios para evitar perder aquello que valoran. Esto significa que los animales no-humanos tienen algo que perder y pueden actuar para evitarlo, y como señalé arriba esta es una condición esencial para ser un agente normativo. Por ello, tiene sentido atribuirles sensaciones, propósitos en su comportamiento, y entendimiento o falta de este. Esto es evidente en que tiene sentido y de hecho nos preguntamos si, ej., los cerdos pueden reconocer su imagen en un espejo y buscamos resolver esa pregunta empíricamente, pero en cambio no tiene sentido hacernos la misma pregunta de una piedra.

Preservar una forma de vida no solo hace una diferencia práctica para la criatura en sí misma y perfectamente pública para observadores potenciales. Sus comportamientos son evaluables de acuerdo con criterios públicos, por ejemplo, el poder formar vínculos con otros individuos o terminar aislado, poder reproducirse o no, lograr nutrirse o caer en la desnutrición, etc. En suma, salud y enfermedad, vivir con otros miembros de la especie o no lograr hacerlo, y en últimas la vida y la muerte son los criterios públicos y observables que dan significado al comportamiento de las criaturas vivas. En ese sentido, hay una continuidad entre el significado del uso del lenguaje y otros comportamientos humanos y no-humanos. Ambos son normativos en tanto que hacen parte un patrón de comportamientos compartido con otros en una forma de vida y en tanto que tienen consecuencias prácticas para la vida de un individuo al que le importa continuar existiendo<sup>7</sup>.

---

<sup>7</sup> En el Capítulo 4 discuto casos como el suicidio en donde esto no parece ser el caso, y en cierto sentido, no lo es.

Los ejemplos de formas de vida no humanas que he usado hasta ahora han sido de animales cuyos cuerpos y formas habitar el mundo son cercanas a las nuestras y por eso hacen más fácil que los veamos como agentes mínimos que les importa mantenerse con vida. Pero formas de vida más lejanas a las nuestras, como las de plantas y hongos son también normativas, lo que significa que estos también son agentes mínimos. Uno de los obstáculos para reconocer la agencia mínima de estas formas de vida es que nuestras intuiciones sobre, ej., las sensaciones de dolor de las plantas, son más cercanas a nuestras intuiciones sobre las piedras. Aunque sin duda las plantas no experimentan dolor, o al menos no como los animales lo hacemos, estas sí cumplen con las dos características esenciales de la agencia mínima que identifiqué arriba: primero, tienen la capacidad de responder de diferentes maneras a los cambios en el ambiente, esto es, tienen la capacidad de seleccionar sus cursos de interacción con el ambiente; y, segundo, tienen propósitos propios que permiten que sus selecciones hagan la diferencia para ellas mismas, a saber, buscan crecer y reproducirse.

Con estas consideraciones en mente, ya estamos en posición de abordar la siguiente pregunta: ¿en qué medida podemos considerar que hay sentido y cognición en las regularidades biológicas?

Vimos que una diferencia importante entre las regularidades constitutivas del sentido y las regularidades biológicas es que las primeras están constituidas por los comportamientos de agentes que son entrenados para ello, mientras que las segundas no requieren entrenamiento porque surgen de la constitución biológica del organismo, de las posibilidades del ambiente y de los patrones de interacción organismo-ambiente que se forman a lo largo de generaciones. Esta diferencia respecto a la necesidad de entrenamiento es una indicación importante de lo que diferencia el hacer sentido de otras actividades normativas: el propósito del comportamiento cognitivo no es satisfacer necesidades vitales inmediatamente, ej., respirar o metabolizar alimentos. En cambio, la cognición se trata de actividades estructuradas en el tiempo, que atienden propósitos indeterminados por las necesidades orgánicas pero determinados por criterios de corrección compartidos con otros, de ahí la impresión de que la cognición trasciende de alguna manera nuestra naturaleza de organismos vivos. Esto significa que la cognición implica actividades cuyos criterios de corrección se construyen por patrones intersubjetivos y que a través del entrenamiento dan forma a los propósitos y al comportamiento de los agentes individuales.

La diferencia entre criterios de corrección biológicos y del sentido implica también una diferencia en la manera en que se es agente biológico y en la manera en que se es agente cognitivo. Como señalé arriba, en ambos casos, parte crucial de ser agente es ser un individuo con algo que perder y tener la capacidad práctica de responder a ello para contrarrestarlo; la forma en que los agentes responden está sujeta a los criterios de normatividad impuestos por los patrones de acción de la comunidad o especie a la que pertenecen. Los agentes que están sujetos normativamente y constituyen las regularidades biológicas tienen evidentemente algo que perder: su propia existencia, su vida. Lo que está en juego en el caso de la agencia cognitiva, no es la vida del organismo, es su pertenencia a la comunidad. Pertenecer a la comunidad, como he insistido hasta ahora, implica hacer parte de las regularidades de hacer



que comparten los miembros de tal comunidad. Estas regularidades del hacer construidas socialmente son cruciales, en primera instancia, porque sin ellas no podríamos adquirir las capacidades para mantenernos a nosotros mismos con vida. Esto no solo significa que debemos aprender cómo cuidar de nosotros mismos y procurarnos, ej., descanso y alimento, sino más generalmente, debemos aprender a interactuar con el mundo para construir y manifestar nuestra existencia como individuos, ej., aprender a construir y manifestar preferencias, opiniones, habilidades, etc. En esa medida, la agencia cognitiva es construida intersubjetivamente sobre la base biológica, moldeable y subdeterminada del tipo de cuerpo vivo que somos los humanos.

### 1.3.5. ¿Condiciones de verdad y representaciones?

Para terminar esta subsección, quisiera discutir cómo se conciben y qué papel tienen las condiciones de verdad y las representaciones dentro de la concepción de sentido como uso que he defendido hasta ahora. Muestro que ambas tienen significado solo dentro de ciertas familias de prácticas y, dada la importancia que le damos a tales prácticas dentro de nuestras formas de vida, tendemos a usarlas como imagen para entender las múltiples formas de hacer sentido que tenemos.

Consideremos, en primer lugar, la concepción de significado como condiciones de verdad. De la concepción de sentido como uso que defendí arriba, se sigue que el significado del lenguaje no está determinado en todos los casos por sus condiciones de verdad; esto no significa, sin embargo, que en ningún caso las condiciones de verdad de un enunciado sean determinantes para su significado. De hecho, este es el caso en algunas de nuestras prácticas, en particular, en nuestras prácticas científicas, por ejemplo, a la hora de formular hipótesis. Si bien el significado de una hipótesis depende del contexto teórico en el que es formulada, sus condiciones de verdad son cruciales para determinar su significado. Por un lado, el contexto teórico y técnico al que pertenece una hipótesis determina, ej., el significado de los términos especializados que hacen parte de esta, sus implicaciones respecto a otras teorías dentro del mismo campo científico, las formas en que es conveniente medir las predicciones que hace la hipótesis, etc. En ese contexto, las condiciones que harían que las afirmaciones de una hipótesis fueran verdaderas determinan en gran medida qué hacemos con tal hipótesis, es decir, el tipo de experimentos que haríamos para ponerla a prueba y los resultados que la apoyarían o que contribuirían a refutarla. Lo mismo ocurre con muchas de las teorías y leyes de la ciencia: su significado está determinado tanto por su conexión con otras hipótesis y teorías como por las condiciones de verdad de lo que afirman. Es importante notar que el significado específico de hipótesis y teorías científicas está determinado tanto por la práctica social de la ciencia como por las interacciones directas y mediante instrumentos con el mundo material. En esa medida, las prácticas científicas no son proyectos completamente relativos al contexto social en el que son llevados a cabo, pero tampoco son acercamientos a la realidad o abstracciones de esta completamente independientes del contexto social en el que son llevados a cabo.

Consideremos, en segundo lugar, la concepción de representación dentro de la concepción de lenguaje como uso que he defendido. El sentido usual de las representaciones surge de prácticas en las que representamos, por ejemplo, cuando en clase queremos representar una idea para explicarla más claramente y dibujamos una imagen en el tablero, o cuando estipulamos que 'x' representa cualquier número real al plantear una ecuación. En ambos casos, las representaciones juegan un papel específico dentro de la práctica que permite llevar a cabo actividades propias de esta: en el primer caso puede facilitar el diálogo entre maestra y estudiantes, y en el segundo caso, puede servir como elemento variable con el que operar y resolver ecuaciones. El significado de tales representaciones depende del uso concreto que se le dé dentro de las prácticas, y son representaciones en virtud de sustituye y simplifica otro elemento o acción más compleja dentro de la práctica. Este patrón de sustituir para simplificar o para resaltar o para caricaturizar, etc. tienen sentido en tanto que pertenecen a una práctica en la que llevamos a cabo actividades compartidas; estas son en los ejemplos, la práctica de la enseñanza en la escuela y la práctica del cálculo y las matemáticas en general o la práctica de la física. En esa medida, representar es una especie de práctica cuyas formas se adaptan a las actividades de otras prácticas y cuyo significado depende de que pertenezcan a tales actividades. Por ello, representar puede ser visto como una actividad que llevan a cabo personas que pertenecen a prácticas y cuyo sentido depende de la actividad más amplia que estemos llevando a cabo.

¿Qué podría significar entonces afirmar que la mente es representacional en un contexto científico? Si queremos adoptar las representaciones dentro de prácticas científicas, debemos precisar condiciones de verdad para que estas sean entidades o procesos que expliquen procesos mentales. Análogamente, si queremos usar las representaciones como entidades mentales a las que puedan reducirse los procesos cognitivos, es necesario poder caracterizarlas e identificarlas como entidades.

Bennett y Hacker (2003) discuten exhaustivamente el papel de las adscripciones de predicados psicológicos al cerebro como frecuentemente se hace en ciencias cognitivas y filosofía de la mente; una de estas adscripciones es la capacidad de representarse el mundo y de inferir información de estas representaciones. Bennett y Hacker revisan ejemplos explícitos de este tipo de adscripciones y muestran que se trata de una forma confusa de hablar o, más frecuentemente, se trata de una atribución de capacidades psicológicas al cerebro sin significado claro y de la que se derivan conclusiones ilegítimamente como si el cerebro fuera una persona capaz realmente de entender, imaginar, representar.

Estas atribuciones constituyen una falacia tan común que Bennett y Hacker le dan un nombre: la falacia mereológica. Esta falacia consiste en atribuir un predicado a un sujeto cuya categoría no es aplicable para ese predicado, en particular, atribuir predicados psicológicos como representar, entender o imaginar a un órgano del cuerpo y no a un individuo; y derivar de esta atribución afirmaciones o explicaciones sobre nuestra psicología o sobre el órgano en cuestión. Bennett y Hacker discuten extensa y convincentemente las diferentes respuestas que se han dado en la literatura a esta acusación, y muestran que el problema central es que la atribución de predicados psicológicos al cerebro no tiene sentido. La razón esencialmente es

la siguiente: cuando decimos de alguien que representa algo tenemos una idea clara de lo que esto significa, es decir, sabemos reaccionar de maneras adecuadas a la representación del otro, así, si la representación es con fines explicativos, podemos pedirle a quien representa que haga aclaraciones respecto a algún punto; si la representación es una estrategia para recordar el camino, esperamos que quien representa pueda llegar a su destino, etc. Sin embargo, cuando decimos que el cerebro puede representar, es un sinsentido esperar que el cerebro nos aclare algo o que el cerebro pueda llegar a algún destino. Si bien el cerebro en tanto órgano del cuerpo es necesario y central como condición material para la ejecución de actividades humanas, en particular, actividades psicológicas, no tiene sentido decir que el cerebro es quien representa y no el individuo. Así como la actividad de los pulmones es esencial para poder fumar, es el individuo quien fuma y no sus pulmones. En esa medida, atribuciones estados, actividades y capacidades mentales al cerebro difícilmente se trata de una explicación de nuestros estados psicológicos; es, en el mejor de los casos, una metáfora y en el peor, una falacia.

Más allá del sinsentido de la idea literal de que el cerebro representa, la imagen de la mente representacional tiene dos consecuencias que deben evitarse. Como mostré arriba, la imagen representacional de la mente está estrechamente ligada a la concepción de sentido como condiciones de verdad porque asume que la función fundamental de la mente es representar el mundo de manera verdadera o adecuada y que los procesos mentales son entonces operaciones que la mente hace con tales representaciones. La primera consecuencia de esta imagen de la mente que debe evitarse es que caracteriza a los agentes con mente como sujetos cuyo propósito intrínseco es formarse una representación verdadera del mundo. Bajo la concepción representacional de la mente, se asume que actuar en el mundo o interactuar con otros es una actividad secundaria y, en esa medida, estas actividades son explicadas en términos de formación adecuada de representaciones —de ahí que la interacción con otros se entienda en términos de la formación de una teoría sobre su mente, esto es, la popular ToM (Theory of Mind), que se forma o bien a partir de una simulación del otro o usando herramientas conceptuales para inferir qué puede estar pasando por su mente. Sin embargo, como mostré arriba, ser agente cognitivo humano implica tener múltiples capacidades de interactuar con el mundo que consideramos psicológicas o mentales y que son constituidas, cultivadas y compartidas por la comunidad; así que es problemático asumir una concepción de mente en donde solo una capacidad, la capacidad de representar, se asume como fundamental ignorando capacidades más básicas y generalizadas como la capacidad de actuar con un propósito.

Hay dos opciones para enfrentar estas dificultades: se puede mantener la concepción representacional de la mente e intentar hacer los ajustes conceptuales necesarios para evitar o solucionar estos problemas como efectivamente se ha hecho en las últimas décadas con autores como Shea (2018), Cussins (2002) y Clark (2015), o bien, se puede adoptar una noción alternativa de sentido y normatividad. Esta última es la opción que adopto en esta tesis. Si aceptamos la noción de sentido como uso que defendí arriba, los fenómenos y procesos mentales son procesos que ocurren en la interacción, bien sea con el mundo o con otros. Sin embargo, concebir la mente bajo la imagen representacional, oscurece este carácter público e

interactivo de la mente y esa es la segunda consecuencia de la concepción representacional de la mente que debe evitarse. El problema de la mente representacional y de la idea de que su propósito intrínseco es la representación del mundo con condiciones de verdad adecuadas es que centra la explicación de la mente bajo la imagen de un individuo que observa el mundo para representárselo. Esto oscurece la multiplicidad de interacciones y actividades que manifiestan nuestra vida mental y que no implican una actitud o punto de vista que busca representar sino uno que busca actuar.

Un último problema de considerar las representaciones como unidades explicativas internas de la mente es que, como señalé arriba, una entidad que no está disponible públicamente no puede ser el criterio o el lugar de dónde emerge el sentido. Esto significa que el comportamiento no es significativo fundamentalmente en virtud de nada interno y no disponible públicamente; en cambio, nuestras actividades son significativas y, por así decirlo, manifestaciones de que somos agentes con mente, en tanto que están sujetas a criterios públicos de corrección. La mente no es, por ende, una esfera interna y por definición escondida de la mirada de otros, y los estados mentales no son entidades que habitan esa esfera; la mente es un patrón de actividades que construyen los individuos públicamente y de manera compartida.

## 1. 4. Conclusión

En este capítulo mostré que hay dos condiciones necesarias para que existan fenómenos normativos: una agencia mínima y criterios normativos; y defendí además una interpretación wittgensteiniana y pragmática de estas dos condiciones.

1. Empecé discutiendo las aproximaciones representacionales a la mente y mostré que dentro de estas aproximaciones es necesario que haya una agencia mínima y criterios normativos para que haya cognición; y, bajo estas aproximaciones, estas son entendidas en términos de verdad: la agencia mínima se entiende como sistemas que buscan formarse representaciones y creencias verdaderas y, por otro lado, los criterios normativos se entienden como criterios de verdad objetiva o correspondencia entre representaciones y realidad.
2. Más allá de esta y de cualquier aproximación a la cognición y a los fenómenos normativos en general, mostré que agencia mínima y criterios normativos son necesarios para que haya fenómenos normativos. La agencia mínima consiste en un sistema que puede seleccionar el curso de sus interacciones, sus comportamientos o la formación de sus estados mentales, y para el que la selección de uno u otro de estas interacciones, comportamientos o estados mentales hace la diferencia. Por su parte, los criterios normativos son aquello que permite determinar la corrección de aquello que selecciona el agente más allá de sus propias impresiones.
3. En la última sección defendí una interpretación de agencia mínima y criterios normativos orientada a la acción, para ello me basé en una concepción wittgensteiniana del lenguaje como uso que extrapolé a fenómenos normativos en general y a formas de vida no humanas. Primero, defendí que los criterios normativos propios de un fenómeno

normativo están constituidos por regularidades que no determinan completamente cómo deben ser continuadas y que se forman por los comportamientos, interacciones y la materialidad específica que habitan los individuos que conforman tal regularidad, mientras que la agencia mínima consiste en participar de tales regularidades y tener la posibilidad de seleccionar comportamientos o cursos de acción individuales que se alineen o no con tales regularidades. Extrapolé esta concepción de sentido como uso dentro de prácticas sociales a otras formas de normatividad apelando a las regularidades biológicas, que a diferencia de las regularidades nomológicas no son completamente deterministas y se constituyen por los patrones de aquello que es seleccionado por los agentes, y por ello pueden establecer criterios normativos para el comportamiento de las criaturas biológicas.

## Capítulo 2

### La naturalización de la agencia mínima

Una estrategia posible para naturalizar la normatividad es apelar a las instancias de la naturaleza que exhiben la capacidad de comportarse normativamente: los seres vivos. Dentro de estas aproximaciones se pueden encontrar dos estrategias generales: apelar, o bien, a procesos evolutivos de selección, o bien, a la constitución misma de los organismos en tanto seres vivos. Esta última es la estrategia del enactivismo que ofrece una explicación naturalista de la agencia y es por la que quisiera empezar. Veremos más adelante (sección 3. 2) que los procesos evolutivos de selección. La aproximación enactiva defiende que la naturaleza autónoma y precaria de los seres vivos instaaura el propósito básico de mantenerse a sí mismos con vida, y esto implica una norma y una teleología inmanente en su comportamiento. A pesar de que las normas y propósitos inmanentes son centrales para el enactivismo actual, defenderlos contradice sus orígenes históricos: La aproximación enactiva es heredera de la teoría autopoiética en biología propuesta por Maturana y Varela en la década de los 70 cuyo propósito era eliminar cualquier imputación de teleología en las explicaciones biológicas.

En este capítulo muestro cómo ocurre la transición conceptual de una propuesta a otra; en particular, me centro en las motivaciones del enactivismo para defender que la normatividad y los propósitos son propiedades intrínsecas del comportamiento de los agentes vivos. Esta revisión me permitirá mostrar que la propuesta enactiva tiene al menos dos ventajas. Primero, a diferencia de la teoría autopoiética original, el enactivismo permite dar cuenta de la naturaleza corporizada y situada de los agentes. Este es el caso gracias a que la concepción básica de normatividad que defiende el enactivismo se basa en la diferencia práctica que un agente puede hacer sobre su propia existencia. Esta concepción de normatividad coincide con la concepción de sentido que defendí en la sección 1. 3, en tanto que ambas se fundamentan en las consecuencias prácticas del comportamiento del agente en su ambiente, aunque, como discuto en el Capítulo 4, estas no son formas de normatividad equiparables. La segunda ventaja de la propuesta enactiva es que sienta las bases para desarrollar una explicación del surgimiento de la normatividad en la naturaleza que permite trazar un continuo entre formas básicas de normatividad como la biológica y formas más complejas. A pesar de estas ventajas, quedan varias preguntas abiertas y posibles objeciones a la propuesta enactiva que discuto al final del capítulo. Por un lado, discuto las acusaciones de circularidad, psicologismo e irrespetabilidad científica y muestro que el enactivismo cuenta ya con respuestas al respecto. Por otro lado, señalo dos puntos que permanecen oscuros en la propuesta enactiva actualmente, a saber, la extensión temporal de los fenómenos normativos y la posibilidad de separar los criterios de corrección de las necesidades orgánicas de los agentes. La propuesta de normatividad que defenderé en los capítulos finales apunta a aclarar estos dos puntos.

## 2. 1. Teoría autopoietica: el proyecto mecanicista

En las décadas de los 70 y 80, Maturana y Varela proponen la teoría autopoietica para explicar qué hace que un ser vivo sea una unidad autónoma, capaz de reproducirse y que activamente mantiene su identidad ante daños o amenazas. Una de sus motivaciones es llenar el hueco explicativo que surge con el énfasis en la especie por sobre el individuo de la teoría de la evolución, y lo hacen explicando el origen de la individualidad biológica (Maturana & Varela, 1980, p. 73). Más específicamente, el propósito de la teoría autopoietica es explicar en términos científicos la individualidad autónoma que caracteriza a los seres vivos y que es condición necesaria para la teoría de la evolución misma: no hay supervivencia evolutiva de la especie si no hay seres vivos individuales de dicha especie.

Maturana y Varela parten explícitamente de dos supuestos, primero, que esta explicación debe darse en términos de las relaciones entre los procesos que componen al ser vivo, y segundo, que esta explicación debe darse en términos mecanicistas:

Nuestra aproximación será mecanicista: no se apelará a ninguna fuerza o principio que no se encuentre en el universo físico. Pero nuestro problema es la organización de lo viviente y, por lo tanto, nuestro interés no será en las propiedades de los componentes, sino en procesos y relaciones entre procesos realizados a través de componentes (Maturana & Varela, 1980, p. 75)

Este compromiso mecanicista implica que la explicación científica de la autonomía debe prescindir de cualquier componente teleológico. En cambio, debe describirse a los seres vivos desde el mismo punto de vista y de la misma manera que se hace con el comportamiento de sistemas no vivos, como el sistema solar (Maturana & Varela, 1980, pp. 85-87): así como la rotación de la tierra al rededor del sol no se explica suponiendo que la tierra tiene ese propósito, se busca explicar el comportamiento de las células y demás seres vivos sin asumir que estos tienen propósitos como reproducirse, alimentarse o mantenerse con vida.

Aunque Maturana continua con este objetivo mecanicista de investigación y con la idea de que una explicación naturalista de la agencia biológica debe eliminar explicaciones teleológicas (Maturana, 2002), Varela empieza a dudar de esta aproximación en la década de los 90, y en el 2002 se publica póstumamente en coautoría con Weber el artículo que inauguraría el giro normativo dentro del enactivismo (Barrett, 2017; Weber & Varela, 2002) y en el que derivan explícitamente teleología intrínseca de autonomía. Este giro implica tanto el reconocimiento de las limitaciones explicativas del concepto de autonomía como de sus implicaciones metafísicas. Merece la pena repasar este desarrollo conceptual desde la propuesta original tanto para caracterizar la noción actual de normatividad en el enactivismo como para hacer explícitas tensiones conceptuales que aún no se han resuelto.

### 2.1.1. Autonomía

Para explicar la individualidad biológica, la teoría autopoietica clásica operacionaliza el concepto de autonomía, esto es, propone una definición de este concepto lo suficientemente precisa para hacer ciencia, y a partir de allí estudia la forma mínima de vida: la célula. El concepto de autopoiesis se refiere a la conceptualización de la forma mínima de vida como la forma mínima de autonomía. La virtud de esta aproximación es que permite ver las condiciones mínimas de posibilidad de la vida, y en esa medida, sienta las bases sobre las que otras formas de vida mucho más complejas pueden ocurrir en la naturaleza. ¿Cuál es la definición de autonomía que propone la teoría autopoietica clásica?

Los seres vivos son autónomos porque mantienen activamente su identidad a través del tiempo y tienen la peculiaridad de que constantemente renuevan la materialidad de los procesos y los elementos que los componen. La identidad de los seres vivos entonces no depende de estos últimos sino de la forma en que se relacionan sus componentes y procesos internos entre sí, y es justamente esa relación lo que permanece constante a través del tiempo. Por ello, la teoría autopoietica clásica define el concepto de autonomía como una propiedad formal de la organización de los seres vivos. En palabras de Maturana y Varela, un sistema autónomo se define como:

“una red de procesos de producción (transformación y destrucción) de componentes que produce los componentes que: (i) a través de sus interacciones y transformaciones continuamente regeneran y realizan la red de procesos (relaciones) que los producen a ellos mismos; y (ii) constituyen el sistema como una unidad concreta en el espacio en el que sus componentes existen al especificar el dominio topológico de su realización como tal red.” (Maturana y Varela 1980, p. 79)

La instanciación más simple de este tipo de organización es la célula y permite ver más claramente en qué consiste esta definición. Por un lado, una célula se distingue física y topológicamente del ambiente (ii) gracias a su membrana. Esta sirve como barrera material que separa del ambiente los procesos que componen a la célula y los agrupa en una unidad. La membrana celular es a su vez producto de los procesos intracelulares que suplen constantemente las moléculas necesarias para que se renueve. Pero estos procesos internos a la célula dependen también del entorno favorable que crea la membrana dentro de la célula. En suma, la organización interna de procesos produce la membrana celular constantemente y la membrana celular posibilita tales procesos. Esta relación de posibilitación mutua ocurre entre todos los procesos que componen a la célula y, por definición, en cualquier otro sistema autónomo: cada uno de sus procesos posibilitan a algún otro y la relación entre todos constituye al sistema completo.

La célula además interactúa con el ambiente mediante su membrana de manera que, idealmente, entra solo lo que puede ser transformado en un insumo para la renovación del sistema y salen los residuos de tales transformaciones. Estas transformaciones que convierten lo que entra desde afuera en partes de los procesos propios de la organización de la célula son



los procesos metabólicos. Los procesos metabólicos permiten mantener estable el medio interno de la célula y producen los insumos para otros componentes y otros procesos dentro de esta, como la membrana celular; pero estos procesos a su vez solo pueden suceder dentro de tal medio intracelular estable. De manera que la célula cumple también con la primera característica organizacional de los sistemas autónomos que vimos arriba (i): es una red de procesos que continuamente se produce a sí misma en tanto red.

Esta organización de procesos significa que la célula, así como cualquier otro sistema autónomo, es cerrada organizacionalmente, es decir, está organizada de tal manera que sus elementos o procesos internos resultan en el mantenimiento del sistema mismo y no en un producto externo. Por ejemplo, una célula se produce a sí misma, pero un sistema de ensamblaje produce algo ajeno, como una computadora. La capacidad de autoproducción y automantenimiento de estos sistemas tienen una consecuencia crucial para la pregunta original por la constitución del individuo biológico y el agente cognitivo, a saber, su propia organización implica que su identidad no depende de los límites o propósitos impuestos por un observador externo sino de su organización misma. La razón es que el sistema mismo es capaz de producir tanto los límites físicos que lo diferencian del ambiente como la organización de procesos y elementos que le permiten continuar diferenciándose del ambiente activamente a través del tiempo.

Los sistemas autónomos en general pueden mantener su identidad de dos maneras, o bien manteniendo su organización y sus componentes estáticos, o bien manteniendo las relaciones entre componentes, pero renovando constantemente tales componentes (Maturana & Varela, 1980, p. 81). Como hemos visto hasta ahora, los sistemas vivos se caracterizan porque mantienen su identidad de la segunda manera. La razón es que los procesos que componen al sistema son precarios, es decir, decaen y eventualmente se extinguen si no se renuevan constantemente; y si los procesos son precarios también lo es la organización autónoma entre ellos. Esto implica que los sistemas vivos necesitan estar abiertos a interactuar con el ambiente para conseguir los recursos materiales y energéticos necesarios para renovar sus componentes y así mantener la relación de clausura organizacional entre ellos que define su identidad. Los seres vivos son entonces sistemas precarios cerrados organizacionalmente y abiertos termodinámicamente.

Esta apertura termodinámica, es decir, la interacción del sistema con el ambiente es selectiva. Como vimos arriba en el ejemplo de la célula, idealmente la membrana celular está abierta solo a los elementos externos que puede metabolizar y se cierra a lo que podría ser una amenaza para el sistema. Este es el caso para todos los sistemas autónomos y precarios: todos establecen su relación con el ambiente en términos del impacto que este tiene sobre su identidad, es decir, sobre la organización de sus procesos, y no como una relación input-output en la que las perturbaciones del ambiente son tomadas como eventos independientes (Maturana & Varela, 1980, pp. 81, 89). Esto implica que el sistema establece tanto el dominio de interacción como las condiciones de interacción con el ambiente de manera que busca lo que requiere para renovar los procesos que pertenecen a su organización y compensa las perturbaciones del ambiente para no perder su autonomía.

La capacidad de los seres vivos de establecer su propio dominio de interacciones es consecuencia de su organización circular y precaria. Puesto que el sistema debe renovar constantemente los procesos que lo componen, estos procesos de renovación deben ser cíclicos y repetirse a través del tiempo para mantener la misma organización que identifica al sistema como sistema autónomo. Las interacciones con el ambiente son parte esencial de estos procesos de renovación porque proveen los recursos energéticos y materiales necesarios para que ocurran. Así que si en algún momento de los procesos de renovación internos del sistema ocurre una interacción de la que el sistema obtiene los recursos necesarios para mantener su autonomía, el sistema puede predecir que esta interacción ocurrirá de nuevo. En palabras de Maturana:

“Cada interacción es una interacción particular, pero cada predicción es la predicción de una clase de interacciones que se definen por los rasgos de sus elementos que permitirán al sistema vivo mantener su organización circular después de la interacción, y por ende, interactuar de nuevo. Esto hace de los sistemas vivos sistemas inferenciales y su dominio de interacciones un dominio cognitivo.” (Maturana & Varela, 1980, p. 10)

De manera que la auto-determinación y el auto-mantenimiento que se siguen de la organización autónoma y precaria de los seres vivos implica que sus interacciones son procesos cognitivos. En otras palabras, la mente surge de la interacción de un ser vivo con el ambiente para suplir las necesidades orgánsmicas-estructurales para seguir existiendo.

En suma, de acuerdo con la teoría autopoietica clásica, los seres vivos son sistemas autónomos precarios y esto, hemos visto hasta ahora, significa que (i) se auto-determinan porque generan sus propios límites que separan su interior del ambiente y (ii) se auto-mantienen porque los procesos que los constituyen están organizados de forma que producen al sistema mismo. Vimos que la autonomía es una propiedad de la organización del sistema que se mantiene a través del tiempo, mientras que los procesos concretos que constituyen esta organización son precarios y por ende el sistema debe interactuar con el ambiente para obtener los recursos para renovarlos y no perecer. En esa medida, el sistema establece el dominio de posibles interacciones: aquellas que contribuyen al mantenimiento de su autonomía y las perturbaciones que puede compensar sin perder su autonomía. Precisamente en estas interacciones surge la cognición.

### 2.1.2. Sistemas dinámicos y determinismo estructural

Bajo esta concepción de seres vivos como sistemas autónomos es posible aclarar qué significa la aproximación mecanicista con la que se comprometen Maturana y Varela. Como mencioné arriba, la teoría autopoietica clásica busca explicar la constitución del individuo biológico apelando solo al universo físico. Esto, sin embargo, no significa que para explicar el comportamiento de los seres vivos se tenga que apelar exactamente a las mismas leyes que explican el movimiento de los sistemas inertes. De hecho, parte crucial de la teoría autopoietica y, más adelante, del enactivismo y de las aproximaciones a la mente que toman el cuerpo como elemento crucial, es descubrir las leyes y regularidades propias del

comportamiento de los seres vivos (Port & Van Gelder, 1995). El aparato conceptual de la teoría de sistemas dinámicos en matemáticas ha sido la herramienta clave para este propósito, y ha permitido a las ciencias cognitivas corporizadas acercarse a la formalización de aspectos importantes del comportamiento de los seres vivos (Richardson & Chemero, 2014).

Desde la perspectiva de teoría de sistemas dinámicos, la mente es un sistema complejo que comprende el sistema nervioso, el resto del cuerpo y el ambiente; y los procesos cognitivos son modelados como una clase de cambios dentro de ese sistema que ocurren en el tiempo. De acuerdo con la teoría de sistemas dinámicos, un sistema dinámico es un conjunto autocontenido de aspectos cambiantes del ambiente. Esto significa que el cambio en cualquier elemento del sistema depende del estado de los demás elementos, y si hay algo que induzca cambios dentro del sistema, esto debe pertenecer también al sistema. Un aspecto que influye pero no es influenciado por el sistema es operacionalizado normalmente como un parámetro y no como parte del sistema *per se*, por ejemplo, la gravedad es un parámetro porque influye el comportamiento de los seres vivos pero no es influenciada por este. Así que definir la mente como un sistema dinámico compuesto por cerebro, cuerpo y ambiente implica que ninguno de los tres puede considerarse como un sistema cognitivo por sí mismo en sentido estricto; sino que conforman juntos un mismo sistema de cuyo comportamiento emergen los procesos cognitivos.

Tanto en la teoría autopoietica clásica como en el enactivismo contemporáneo, el objeto de estudio relevante para entender la cognición es la interacción de un sistema autónomo con su ambiente. Para ambos, sin embargo, un sistema autónomo y el ambiente con el que se acopla o interactúa también conforman un sistema. De acuerdo con el enactivismo contemporáneo la mente surge en este sistema amplio organismo-ambiente; sin embargo, es crucial que el organismo que hace parte de tal sistema amplio conforme en sí mismo un sistema autónomo. Así que, cuerpo y cerebro forman un sistema autónomo cuyo carácter autónomo permite que la mente surge en el sistema más amplio organismo-ambiente. En el resto del capítulo usaré la palabra 'sistema' en estos dos sentidos y será claro por el contexto a cuál de las dos acepciones me refiero: sistema amplio organismo-ambiente u organismo como sistema autónomo.

La mente es también un sistema complejo lo que significa que está compuesto por elementos que interactúan entre sí y cuyo patrón de comportamiento como conjunto no puede ser predicho por el comportamiento de sus componentes individuales. De acuerdo con Richardson y Chemero (2014), los sistemas complejos tienen al menos tres características que los diferencian de otros sistemas y que son cruciales para caracterizar la mente.

(i) Los sistemas complejos son autoorganizados, lo que significa que su organización y patrones de comportamiento no resultan de un elemento que los controle de manera centralizada, sino que resultan de las interacciones mismas que unen a los componentes como parte de un mismo sistema (p. 40). Por ejemplo, la organización de una estrella, que es un sistema complejo, no resulta de un componente central sino de la unión paulatina de partículas de hidrogeno, helio y algunos metales en menor medida que eventualmente forman una masa gracias a la gravedad. Los procesos de combustión y transformación de elementos

característicos de las estrellas resultan de las condiciones de presión y temperatura que la organización misma de la estrella produce, y no de la instrucción de un componente central.

(ii) Los sistemas complejos son de ensamblaje suave o *soft-assembly*. Esto significa que el comportamiento de los componentes individuales no está determinado rígidamente, sino que depende del contexto, en particular, de si hacen parte de una unidad coherente dentro del sistema (p. 40). Un elemento en una estrella se comporta de manera distinta según si hace parte del núcleo en donde el elemento mismo será transformado, ej. hidrógeno en helio, o si hace parte de la zona de radiación en donde transportará la energía residual de las reacciones del núcleo hacia fuera de la estrella. Esta flexibilidad del comportamiento de los elementos individuales contrasta con sistemas no complejos como una computadora cuyos componentes están rígidamente designados para una tarea.

(iii) Los sistemas complejos son sistemas no lineales. De acuerdo con Richardson y Chemero, esto significa lo siguiente: “el output del sistema no es directamente proporcional al input, esto en oposición a un sistema lineal en el que el output puede ser representado simplemente como la suma ponderada de sus componentes (2014, p. 41). Esta característica es evidente en las estrellas en tanto que las velocidades y ritmos en que los componentes de la estrella fluyen de dentro hacia afuera y viceversa dependen de forma no lineal de la evolución de la estrella (Nowakowski & Dziembowski, 2003).

El sistema que forman cerebro-cuerpo y organismo es también un sistema autoorganizado, de ensamblaje suave y no lineal. La idea detrás de la teoría autopoiética y ahora también de los enfoques radicalmente corporizados a la cognición es que estudiar empíricamente los detalles del comportamiento de este sistema revela los detalles particulares de lo que caracteriza a un ser vivo y un sistema cognitivo.

Adicionalmente, un sistema dinámico es determinista y esta característica será crucial para la discusión al final de este capítulo. En palabras de Port y van Gelder, el determinismo de un sistema significa que los estados futuros del sistema siempre están determinados, de acuerdo con alguna regla, por su estado presente exclusivamente (Port & Van Gelder, 1995, p. 6). Así que cualquier estado pasado del sistema cerebro-cuerpo-ambiente determina completamente los estados futuros del sistema. Esto implica que los procesos cognitivos pueden ser modelados como cambios en el sistema que pueden ser descritos por algún conjunto de reglas; normalmente, estas son ecuaciones no lineales, pero pueden ser también operaciones geométricas u otras herramientas matemáticas. En particular, como vimos en la sección anterior, los sistemas vivos establecen interacciones según sus necesidades estructurales; entonces, la organización o estructura del organismo juega un papel determinante en los cambios del sistema. Los agentes cognitivos son, por ende, sistemas deterministas estructuralmente.

El determinismo estructural captura una convicción ontológica de la teoría autopoiética, de la teoría de sistemas dinámicos y presumiblemente de la mayor parte de explicaciones científicas y naturalistas hasta ahora. Esta convicción es que cualquier cambio en un sistema está completamente determinado por su estructura y su estado anterior, y no por ningún factor

externo al sistema. Una de las motivaciones detrás de este axioma del determinismo estructural es presumiblemente que una explicación no debe apelar a entidades cuyo estatus ontológico sea más dudoso u oscuro que el de lo que se quiere explicar, por ejemplo, explicar el comportamiento de los seres vivos apelando a entidades metafísicas como el alma o como un propósito natural último. La otra motivación es que una explicación debe mostrar por qué ocurre un fenómeno eliminando la posibilidad de que hubiera sido de otra manera, y esto supone la convicción ontológica de que la naturaleza se presta para ser explicada de esa manera. En tanto convicción ontológica, el determinismo estructural es una propiedad común a todos los sistemas, vivos o inertes, concretos o abstractos. Esto significa que los seres vivos no tienen un estatus ontológico distinto a los sistemas no vivos, y por ende, de acuerdo con la formulación original de la teoría autopoietica, una explicación de su constitución y comportamiento debe eliminar los propósitos y valores que como observadores externos tendemos a adscribirles (Maturana & Varela, 1980, pp. 85-86).

Concebir a seres vivos y agentes cognitivos como sistemas deterministas tiene al menos dos implicaciones importantes y problemáticas para naturalizar la normatividad. Primero, cualquier cambio o comportamiento de un sistema se sigue de sus condiciones anteriores y solamente de estas. Esto implica que el comportamiento de un sistema determinista no podría ser de otra manera y, por ende, el sistema no puede seleccionar su curso de acción: no hay selección genuina si no hay más de una posibilidad. Segundo, una explicación completa de un fenómeno idealmente debe incluir todos los elementos y parámetros que lo determinan como parte del sistema. Dada la complejidad de los sistemas naturales, en particular, de los sistemas vivos, en los modelos reales suelen incluirse solo ciertos elementos como parte del sistema y los demás se modelan como parámetros o como ruido. La impredecibilidad del comportamiento de los seres vivos y la impresión de que siguen propósitos y valores suele modelarse como ruido, pero, de acuerdo con el axioma del determinismo estructural, ontológicamente se siguen de procesos anteriores dentro del sistema de forma determinista. Por ende, apelar a propósitos y a decisiones de acuerdo con valores en la explicación del comportamiento cognitivo es parte de nuestras limitaciones epistemológicas y no de una característica real de los sistemas cognitivos bajo una descripción científica (Maturana, 2016; Villalobos, 2013).

Hasta ahora hemos visto que, para la teoría autopoietica, la forma de vida mínima es un sistema autopoietico, esto es, un sistema mínimamente autónomo y precario cuya identidad se constituye y mantiene por la relación entre los procesos que lo componen. Vimos tres características importantes de este tipo de sistemas. Primero, en virtud de su organización interna, los resultados de los procesos que constituyen al sistema contribuyen al mantenimiento de su organización, es decir, su función es mantener la identidad del sistema. Por eso los sistemas autónomos son sistemas que se autodeterminan y automantienen. Segundo, los sistemas autónomos crean su propio dominio de interacciones y lo hacen en términos de perturbaciones a su identidad y compensaciones, y no en términos de input-output. Crucialmente, los sistemas vivos son sistemas cognitivos, puesto que de la regularidad de las necesidades estructurales del organismo y su satisfacción mediante interacciones con el ambiente surgen predicciones sobre las interacciones y necesidades futuras. Tercero, los

sistemas autopoieticos son estructuralmente deterministas, es decir, su comportamiento está completamente determinado por su estructura y su estado anterior, de manera que no podría haber ocurrido de otra forma. Esta definición de ser vivo es la base para proponer una explicación del comportamiento cognitivo sin apelar a nada distinto que a los mecanismos que gobiernan sus procesos constitutivos y a su organización cerrada en tanto sistema. Esto significa que la definición original del ser vivo como sistema autopoietico busca eliminar la apelación a propósitos o preferencias para explicar sus interacciones con el mundo.

## 2. 2. Recuperar la teleología: antecedentes

Vimos en la sección anterior que uno de los objetivos de la teoría autopoietica era eliminar la atribución de propósitos y preferencias de las explicaciones científicas de la agencia biológica y cognitiva. Sin embargo, la caracterización de los seres vivos como seres autónomos es compatible también con el objetivo opuesto, el objetivo de darle a los propósitos y preferencias un lugar en la naturaleza. Este giro en los objetivos teóricos fue inaugurado por Varela mismo y es ahora central en una de las tres variantes del enactivismo: la variante que se conoce como la clásica, autopoietica, varelina o como ‘enactivismo’ a secas, que fundamenta el surgimiento de la agencia en la organización autónoma del organismo. Las otras dos variantes del enactivismo, la radical y la sensoriomotora, se centran respectivamente en la concepción anti-representacional de la cognición mínima (Hutto, 2011; Hutto & Myin, 2013, 2017) y en la explicación de la percepción y la acción con sus características intencionales y experienciales apelando a las posibilidades del cuerpo y su interacción con el ambiente (O'Regan & Noë, 2001b). El enactivismo radical será discutido en el siguiente capítulo porque ofrece una explicación naturalista del surgimiento de los criterios de normatividad en la historia evolutiva. Por su parte, la variedad sensoriomotora no juega un papel protagónico en este debate de la naturalización de la normatividad, puesto que se ha centrado en la experiencia perceptual y no en una explicación general de la mente.

Varela empieza a separarse de la posición anti-teleológica de la teoría autopoietica que, por el contrario, Maturana sí mantiene, ya desde los 80 (Varela, 1981; Villalobos, 2013). De acuerdo con Froese (2011), la motivación de Varela para dar este giro fue el cambio de enfoque de principios universales y teoremas abstractos al cuerpo vivo concreto que incluye no solo sus propiedades materiales y termodinámicas, sino también sus propiedades fenomenológicas y experienciales (p. 633). En los años 90, Varela defiende explícitamente el componente fenomenológico de los organismos autónomos y les atribuye la capacidad de hacer sentido en *The Embodied Mind* (Varela et al., 1991) –libro del que surgen las tres variedades del enactivismo–, y crucialmente, en el 2002 con ‘Life after Kant’ (Weber & Varela, 2002), defiende que los organismos vivos en tanto sistemas autónomos son intrínsecamente teleológicos. Este último argumento junto con el trabajo de, por ejemplo, Thompson y Di Paolo, de principios de la década de los 2000 ha sido el punto de partida del desarrollo de la noción de normatividad y sentido dentro del enactivismo clásico en las últimas décadas. Este desarrollo ofrece una concepción naturalista de la agencia que promete explicar cómo surge la teleología intrínseca de la organización autónoma de los seres vivos y con ello un agente mínimo normativo.

Defender sólidamente esta propuesta requiere ofrecer dos tipos de argumentos. Por un lado, es necesario ofrecer razones independientes para abandonar la idea de que la teleología no debe hacer parte de las explicaciones científicas de la vida, es decir, se debe argumentar por qué quisiéramos mantener la teleología dentro de una ontología naturalista. Este tipo de razones motivan filosóficamente el proyecto, porque apuntan tanto a la pregunta ontológica de si la teleología hace parte de la naturaleza o no, como a la pregunta epistemológica de cuál es la forma más apropiada de explicar los fenómenos naturales, en particular, la vida. Por otro lado, defender enactivamente la naturalización de la normatividad requiere mostrar cómo es de hecho posible derivar una teleología intrínseca de la constitución de los seres vivos. Para ello es necesario examinar cada caso particular y mostrar los procesos específicos que explican el surgimiento de la agencia mínima para un tipo de organismo particular en su nicho. Puesto que el propósito de esta tesis es examinar las aproximaciones naturalistas a la normatividad de los enfoques corporizados a la cognición, en este capítulo me dedico al primer tipo de argumentos. En el último capítulo considero los casos de estudio específicos que se han defendido bajo este enfoque, (por ejemplo, Moreno & Mossio, 2015), a la luz de los requisitos de naturalización de la normatividad que defiendo a lo largo de toda la tesis.

### 2.2.1. El problema de la teleología en Kant

La postura de la eliminación de la teleología de la teoría autopoiética en su formulación original se sitúa en un debate más amplio: el debate sobre el lugar que tiene la vida dentro de una ontología naturalista. A grandes rasgos, el debate es el siguiente: idealmente, una explicación científica del comportamiento de los seres vivos debería prescindir de elementos teleológicos como explanandum; en cambio, los rasgos teleológicos del comportamiento y la constitución de los seres vivos deberían entenderse como atribuciones arbitrarias del observador que se pueden y deben eliminar de la ciencia una vez se encuentren los mecanismos subyacentes que producen su comportamiento. Sin embargo, atribuir propósitos y funciones hasta ahora ha probado ser ineludible en biología. ¿Qué lugar tiene entonces la teleología en la naturaleza?

La concepción de teleología que está en juego en este debate es la teleología como rasgo intrínseco en oposición a una teleología relativa. Por un lado, la teleología relativa se refiere a los propósitos característicos de un artefacto cuyas partes y cuya organización como un todo se explican por el propósito que el diseñador tenía al concebirlo o a la forma en que una parte de la naturaleza es útil para los humanos. Este tipo de teleología ha sido descartada ampliamente como explicación del comportamiento de los seres vivos, es decir, ni la voluntad de dios ni el diseño de un ingeniero ni ningún otro factor externo explica el comportamiento de los seres vivos (Weber & Varela, 2002, p. 106). Hay al menos dos razones para esto. Primero, asumir un diseñador de los seres vivos no es explicativo porque tal diseñador es improbable e imposible de estudiar. Segundo, como señala Kant, la relación de las partes de un organismo con la totalidad del organismo es distinta a la relación entre las partes de un artefacto y el artefacto como totalidad: mientras las partes de un artefacto son material y lógicamente independientes del artefacto mismo y además lo anteceden temporalmente, las partes de un organismo dependen de la totalidad a la que pertenecen para existir y surgen

simultáneamente con el organismo en sí o posteriormente a este. Por eso los artefactos son analizables en partes, pero los organismos no (Thompson, 2007, pp. 135-136). De manera que los organismos, así como los artefactos, dependen de sus partes y la organización entre ellas para existir como totalidad, pero a diferencia de los artefactos, cada una de las partes del organismo dependen de la totalidad para existir. Esto sugiere que el propósito de un organismo es mantenerse a sí mismo a través del mantenimiento de sus partes, y este no un propósito externo a sí mismo. Esto nos lleva al segundo tipo de teleología, la teleología intrínseca.

La teleología intrínseca se refiere a propósitos que surgen del organismo mismo y que guían su comportamiento; por ejemplo, la función de un órgano determina cómo debería comportarse, y el propósito de mantenerse con vida de un organismo determina si su comportamiento es apropiado o no. Los propósitos y funciones de este tipo son precisamente el elemento teleológico que no se ha podido eliminar de las explicaciones biológicas, aunque se incluyan frecuentemente solo como metáforas inescapables y no como enunciados ontológicos. De esta imposibilidad, surge la tensión conceptual no solo para la biología sino para la naturalización del comportamiento cognitivo, la tensión entre el ideal de una ciencia puramente descriptiva, por un lado, y la imposibilidad de eliminar la teleología en las explicaciones de los seres vivos, por el otro (Allen & Neal, 2020).

La formulación de este problema se la debemos a Kant. En la *Crítica de la razón pura* (2010 [1787]), Kant sostiene que nuestro conocimiento científico de la naturaleza se sigue de las formas y principios a priori de nuestras facultades cognitivas. Las formas del espacio y tiempo y las relaciones causales son los principios a priori que estructuran lo que percibimos por los sentidos, y de los que se siguen por fuerza de necesidad las leyes de la naturaleza. De ahí que Kant conciba la naturaleza como un sistema objetivo de relaciones físico-matemáticas, necesarias y a priori; y que sostenga que conocer la naturaleza científicamente consiste en explicar sus fenómenos en términos de causalidad mecánica (2010 [1787]). Sin embargo, la existencia misma de los seres vivos no se sigue de tales categorías a priori, porque su existencia misma no puede derivarse de las formas y relaciones causales de la materia en el espacio-tiempo. Kant reconoce en la *Crítica del juicio* (1992 [1790]) que, a diferencia de otros cuerpos de la naturaleza, entendemos a los seres vivos como organismos que tienen propósitos en sí mismos; es decir, los entendemos como cuerpos cuyo comportamiento no solo sigue leyes mecánicas, sino que también sigue causas finales. Esta organización es completamente contingente porque no está implicada en los principios a priori y solo podemos conocerla mediante la experiencia. Incluso, como el mismo Kant reconoce en la *Crítica del juicio*, nos es imposible si quiera familiarizarnos con la organización y el comportamiento de los seres vivos sin apelar a propósitos y causas finales:

Es, en efecto, completamente cierto que ni siquiera podemos con suficiencia tomar conocimiento de los seres organizados y de su posibilidad interna con arreglo a principios meramente mecánicos de la naturaleza, y mucho menos explicárnoslos; y ello con tal certeza que se podría decir atrevidamente que para los hombres es absurdo el solo trazar un plan semejante, o esperar que acaso pudiera nacer alguna vez un



Newton que hiciese concebible no más que la generación de una brizna de hierba con arreglo a leyes naturales a las que no ha ordenado ninguna intención, sino que debe negársele absolutamente esta inteligencia a los hombres. (Kant, 1992 [1790], §75)

Es imposible deducir la existencia de seres vivos, con su comportamiento y organización particulares, de las leyes mecánicas newtonianas, y esto implica su vez que el principio teleológico y el principio mecánico de investigación a la naturaleza no pueden subsumirse el uno al otro –son entonces principios inconmensurables. Kant captura esta tensión en la antinomia del juicio teleológico en la que tesis, el principio de investigación mecanicista, y antítesis, la particularidad de los seres vivos, se contraponen:

La primera máxima de ésta es la tesis: toda generación de cosas materiales y de sus formas tiene que ser juzgada como posible según leyes meramente mecánicas.

La segunda máxima es la antítesis: algunos productos de la naturaleza material no pueden ser juzgados como posibles según leyes meramente mecánicas (su enjuiciamiento requiere una ley de causalidad enteramente distinta, a saber, la de las causas finales). (1992 [1790], §70, 312-308)

La inconmensurabilidad de estos dos principios de investigación implica que hacemos dos clases de juicios distintos de una misma cosa inaprehensible: la naturaleza. Kant, sin embargo, reconoce una jerarquía entre estas dos clases de juicios cuando se trata de entender a los seres vivos (Kant, 1992 [1790], §78): los seres vivos deben ser explicados en términos mecánicos tanto como sea posible, pero estas explicaciones causales siempre estarán subordinadas a su teleología (Weber & Varela, 2002, p. 108).

La inevitabilidad de la teleología en el entendimiento de los seres vivos es aún más problemática si, como Kant, se considera que es imposible una aproximación científica a la organización particular de los seres vivos. Kant, anticipando la teoría de sistemas dinámicos y la teoría autopoiética, reconoce que los seres vivos tienen una organización particular en tanto que son causa y efecto de su propia existencia, y eso implica que su propia existencia no solo es un fin en sí mismo sino su propio medio (Kant, 1992 [1790], §77). Pero para Kant, en su época, era imposible concebir relaciones causales diferentes a las leyes mecánicas, y por ello, era imposible imaginar una aproximación científica a esta forma de organización.

Kant deja abierta la pregunta por el lugar de la teleología, si es parte de la naturaleza o de nuestras limitaciones, y por cómo podría la naturaleza acomodar simultáneamente una ontología mecanicista y una que admita elementos teleológicos.

Respecto a este problema hay dos alternativas. Por un lado, se puede asumir que el componente teleológico de nuestras explicaciones de los seres vivos refleja nuestras limitaciones epistemológicas, o bien porque hasta ahora no hemos desarrollado el marco conceptual o las herramientas adecuadas, o bien porque nuestras capacidades cognitivas mismas no nos permiten entender la vida sin apelar a la teleología. La teoría autopoiética originalmente asumió la primera alternativa y propuso una concepción de seres vivos que

buscaba superar tal limitación epistemológica. Actualmente esta posición es defendida no solo por Maturana mismo, sino por autores como Villalobos y Ward. Por otro lado, se puede asumir que la persistencia de la teleología en biología es característica de los seres vivos y parte de los fenómenos naturales que deben ser explicados. Así, en lugar de buscar eliminar las explicaciones teleológicas de la biología, se puede buscar explicar cómo surgen en la naturaleza sistemas capaces de comportamiento que describimos como teleológico. Esta es la alternativa que asume el enactivismo en donde se adopta la teoría autopoietica como herramienta para fundamentar la teleología y no como herramienta para eliminarla.

La teoría autopoietica y el enactivismo hacen parte de las varias respuestas que se han dado en filosofía de la biología a la pregunta de cómo explicar la teleología de forma naturalista. Esta pregunta se ha convertido en la discusión más específica de cómo determinar la función biológica, en la que hay dos tipos de respuesta, una con un enfoque evolutivo y una con enfoque organizacional. Defensores del enfoque evolutivo sostienen que la función de las características biológicas está determinada por la selección natural. Una de las defensoras más importantes de esta posición es Ruth Millikan (2017) quien propone que la función propia de un ítem –por ejemplo, el corazón– está determinada por la función que cumplieron los ítems que lo originaron –en el ejemplo, la función de los protocorazones y los corazones de los ancestros del organismo–, siempre y cuando tal función haya llevado causal e históricamente a que el ítem –el corazón– se reprodujera (Millikan, 1989, p. 288). Crucialmente, lo que determina la función de un órgano, comportamiento o característica es su historia causal evolutiva y no sus propiedades y circunstancias actuales. Si bien este tipo de respuestas evolutivas son fundamentales para responder preguntas como el origen de los criterios normativos, no permiten responder cómo es posible en primer lugar que una historia evolutiva comience y que se configure un ser vivo cuando antes no había ninguno; en esa medida, caen en explicaciones *ad infinitum*: todo lo que ahora es, ej., un corazón, un organismo, es por que viene de un sistema u organismo que llegó a ser, pero este sistema u organismo llegó a ser porque otro sistema u organismo anterior llegó a ser, etc. En el capítulo III, discutiré en detalle los méritos y limitaciones de la propuesta de Millikan y cómo partes de esta pueden complementar una concepción naturalista de normatividad basada en el enactivismo (ver sección 3. 2).

Por otro lado, el enfoque organizacional, del que hace parte el enfoque enactivo, defiende que la función de un órgano, comportamiento o característica está determinada por su contribución al mantenimiento de la vida del organismo al que pertenecen. La función está entonces determinada por la organización y las circunstancias actuales del ser vivo, y no por su historia evolutiva. Esta forma de entender las funciones está estrechamente ligada con la forma en que el enactivismo deriva la teleología y normatividad intrínsecas de la organización del organismo: función, teleología y normatividad están determinadas por lo que permite al organismo mantenerse con vida o lo que lo amenaza; en esa medida, la organización del ser vivo sustenta una respuesta naturalista a la pregunta de cómo surge la agencia mínima. Este último argumento es el tema de la siguiente sección. En el capítulo 3 volveré a los enfoques históricos/evolutivos y defenderé que ofrecen una respuesta naturalista al origen de los criterios de normatividad.

### 2.2.2. Jonas y el salto enactivo a la normatividad

El paso decisivo de la teoría autopoiética al enactivismo contemporáneo fue la publicación de *The Embodied Mind* (TEM) de Varela, Thompson y Rosch (1991). TEM fue crucial por al menos tres razones. La primera es que esboza cómo los conceptos de autonomía y autopoiesis pueden usarse para defender la incorporación de la teleología y la normatividad a las ciencias cognitivas; en particular, defiende que la autonomía de los organismos es central y es lo que fundamenta su experiencia (1991, p. 55). Segundo, defiende que la percepción consiste en acción perceptualmente guiada y que los patrones recurrentes que esta forma dan lugar a que emerjan estructuras cognitivas (1991, p. 173). Tercero, defiende ampliamente la incorporación de la experiencia de primera persona a las ciencias cognitivas, no solo como un fenómeno que debe ser examinado y explicado, sino como una característica ineludible de los investigadores mismos, del punto de vista científico (1991, pp. 10-12). El compromiso con reincorporar la experiencia a la ciencia cognitiva es tan fuerte en TEM que proponen usar el budismo y la meditación como herramientas para examinar la perspectiva de primera persona. Aunque esta propuesta fue mayormente abandonada, TEM es el libro que inaugura el proyecto enactivo porque perfila tanto la conexión conceptual entre experiencia y autopoiesis que guía al proyecto enactivo, como la concepción de percepción en particular y cognición en general como fundamentada en patrones sensoriomotores. TEM y la aproximación a las ciencias cognitivas centrada en la experiencia que inaugura inauguran el enactivismo como proyecto de naturalización de la teleología y la normatividad.

Weber y Varela (2002) motivan la introducción de la teleología dentro del enactivismo a partir de la tensión identificada por Kant: el carácter teleológico del comportamiento de los seres vivos debería ser reducible a explicaciones mecánicas, sin embargo, eliminarlo de nuestras explicaciones sigue siendo imposible. De acuerdo con Weber y Varela, Kant mismo no solo esbozó una solución a esta tensión al proponer concebir los seres vivos como seres que son causas y efectos de sí mismos, y por ende medios y fines de sí mismos; sino que en sus últimos escritos sugiere que las categorías a priori no son el fundamento último de la experiencia sino que lo a priori está fundamentado en el cuerpo (Weber & Varela, 2002, p. 110). Por ejemplo, Weber y Varela (2002, p. 109) citan los escritos póstumos de Kant en donde sostiene:

Las fuerzas móviles de la materia son lo que el sujeto móvil mismo hace con su cuerpo en otros cuerpos. – Los efectos que corresponden a estas fuerzas están contenidos en los actos simples mediante los que percibimos los cuerpos mismos (Akad.-A. XXII, p. 326).

En la segunda mitad del siglo XX, Hans Jonas desarrolla la fenomenología organicista en la que critica la fundamentación abstracta de la experiencia que defiende Kant en su periodo crítico y propone al cuerpo vivo como fundamento de la experiencia. Particularmente importante para el enactivismo, Jonas soluciona la tensión entre teleología y explicación mecanicista de los seres vivos argumentando que la teleología ocupa un lugar en la naturaleza, que esta se deriva del metabolismo y que es central en nuestra experiencia del mundo.

Veamos cómo cada uno de estos dos elementos, experiencia y teleología, de la filosofía organicista de Jonas es adoptado por Weber y Varela y posteriormente por el enactivismo.

Jonas defiende que la teleología es parte de nuestra propia experiencia y no un principio intelectual oscuro e imposible de fundamentar. Su estrategia, típicamente fenomenológica, es señalar que nuestra experiencia es en primer lugar la experiencia de seres vivos y no de científicos o filósofos indiferentes y desconectados del ambiente. Esto significa, de acuerdo con Jonas, que somos primariamente centros de experiencia y sujetos activos en el mundo, y nuestra idea más básica de causalidad viene precisamente de nuestra experiencia de actuar y de sentir los efectos de la materia que reacciona a nuestras acciones. Después de discutir las concepciones de causalidad de Hume por asociación y la de Kant por razón, Jonas afirma:

El aspecto primario de la causalidad no es la conexión regular, ni siquiera una conexión necesaria, sino fuerza e influencia; que estos son ellos mismos los contenidos de la experiencia y no una interpolación entre contenidos de la experiencia (= perceptos) por una función sintética, sea por asociación o por razón; ...la fuente de esta experiencia es, de hecho, no percepción sensorial, sino nuestro cuerpo esforzándose en la acción. (Jonas, 2001 [1966], p. 33)

La experiencia de la causalidad se subsume a un impulso más fundamental, el deseo primordial de mantenernos con vida: estamos movidos a afirmar nuestra propia existencia y este propósito que nos mueve a actuar es precisamente la experiencia de la teleología que nos es propia por ser cuerpos vivos (Weber & Varela, 2002, p. 110). Nuestra propia teleología nos lleva a establecer un punto de vista desde donde damos sentido y categorizamos nuestra interacción con el mundo: el ambiente se vuelve un mundo de valor en tanto nos permite o no afirmar nuestra existencia. Además, nuestro propio punto de vista teleológico nos permite entender a otros seres vivos como centros de experiencia y acción que se comportan teleológicamente para afirmar su propia existencia.

Puesto que el cuerpo vivo tiene un papel central en la experiencia, Jonas propone analizar las formas mínimas de vida para comprender con más profundidad cómo se configura la experiencia. Jonas parte del hecho biológico de que todos los seres vivos y sólo los seres vivos son seres con metabolismo, es decir, seres que constantemente transforman elementos del ambiente en energía y materia para mantenerse a sí mismos con vida, renovarse, crecer, etc. En esa medida, la existencia de los organismos es resultado constante de su propia actividad metabólica, y si tal actividad se detiene, el organismo deja de existir. Jonas destila dos consecuencias fenomenológicas de este hecho biológico. Por un lado, puesto que el organismo mantiene su identidad mientras la materia de su cuerpo fluye y cambia constantemente, el sujeto tiene cierta independencia de las particularidades de la materia que pasan a través de su metabolismo. Pero, por otro lado, la identidad y la existencia del sujeto dependen precisamente de que la materia siga fluyendo y transformándose en su cuerpo, y en ese sentido su existencia es precaria y depende de la materialidad. Por ello, al organismo le concierne afectivamente la materia que le permite evitar perecer. Esta doble relación del organismo con la materia que lo constituye es una relación de libertad necesitada (*needful*

*freedom*): su forma y organización lo hacen parcialmente libre de la materia, pero necesita de la materia para continuar existiendo.

El papel que Jonas da al metabolismo coincide con el concepto de autonomía no solo en que ambos surgen de la insatisfacción con las concepciones fragmentadas de vida como procesos moleculares e información genética, y de mente como procesamiento de información; sino también en que capturan concepciones de vida y cognición como actividades del agente, y no como eventos o propiedades que le ocurren al sujeto pasivamente. Más precisamente, el metabolismo es una clase de sistema autónomo, y el concepto de autonomía captura de forma generalizada la organización circular de los sistemas que se autoproducen y automantienen activa y constantemente. La filosofía de Jonas ha sido acogida y actualizada por el enactivismo para derivar la agencia, y el carácter teleológico y normativo de la autonomía de los seres vivos.

## 2. 3. Recuperar la teleología: el fenómeno de la normatividad del enactivismo

Para reconstruir la propuesta enactiva de naturalización de la normatividad, empezaré discutiendo dos implicaciones del concepto de autonomía: (i) el sistema se diferencia de su medio en virtud de su propia actividad, es decir, se autodiferencia y (ii) en virtud de esa diferenciación, establece un dominio de interacciones con el ambiente que le permite automantenerse.

- (i) *Autodiferenciación física*: Un sistema autónomo es un sistema constituido por procesos que se relacionan entre sí de manera que cada uno de ellos es posibilitado por algún otro dentro del sistema, es decir, otro lo causa o crea las condiciones que necesita para ocurrir. En virtud de su organización, estos procesos crean un espacio en el que son posibles, –incluso, un espacio delimitado físicamente como es el caso de la célula–, y que se diferencia del espacio exterior en el que no podrían persistir –por ejemplo, cuando un órgano es extraído del cuerpo vivo decae rápidamente y se descompone.
- (ii) *Automantenimiento y dominio de interacciones*: Aquello que está fuera del agente es, sin embargo, esencial para que el agente pueda obtener los recursos materiales y energéticos para renovar constantemente sus procesos. El sistema debe interactuar constantemente con el ambiente para mantenerse como la unidad autodiferenciada que es. Estas interacciones ocurren en los términos del sistema, es decir, el impacto de los eventos externos sobre el agente es asimilado de manera que se incorpora a su organización y por ende le permite mantener su identidad. Si un evento no puede ser asimilado en esos términos, el sistema pierde su organización, esto es, su autonomía y, por ende, pierde su identidad. Este último tipo de eventos está fuera del dominio posible de interacciones del agente. Por lo tanto, la organización del sistema mismo, es decir, su autonomía, determina su

propio dominio de interacciones y este es el dominio que permite su automantenimiento.

Moreno y Barandiaran (2004) caracterizan estas dos implicaciones como la explicación naturalista de la dicotomía dentro-fuera que involucra dos procesos complementarios pero asimétricos. El primero es un proceso constructivo en el que el sistema se autoconstituye recursivamente, es decir, genera las condiciones físicas para mantener próximos sus componentes materiales y así constituir un cuerpo discretamente diferenciable (Villalobos & Razeto-Barry, 2020); la forma más común en que los seres vivos logran tal proximidad es generando y manteniendo una barrera física que los distingue de su medio y que mantiene sus flujos internos de materia y energía dentro de tal barrera. Este proceso de autoconstrucción ocurre en todos los seres vivos y en algunos sistemas inertes, por ejemplo, las estrellas (ver p. 51), y no son suficientes para crear una dicotomía real entre dentro y fuera:

una dicotomía real dentro-fuera solo surge cuando el sistema es capaz de crear su propio mundo externo, no meramente como algo que no pertenece a la organización del sistema, sino en el sentido de un mundo de procesos interactivos motivados por el sistema mismo y que son ‘significativos’ para este en el sentido de que son funcionales o disfuncionales para el mantenimiento del sistema. En ese sentido, la dicotomía dentro-fuera sugiere que la organización del sistema, para ser entendida como “interna”, debe tener consecuencias operacionales más allá del sistema que constituye (Moreno & Barandiaran, 2004, pp. 13-14)

Estas consecuencias son precisamente el segundo proceso que proponen Moreno y Barandiaran, el proceso interactivo, que es exclusivo de los seres vivos. En estos procesos el sistema modula, genera o controla las condiciones externas necesarias y las interacciones con estas para su automantenimiento. Este tipo de interacción es crucial porque es donde ocurre la normatividad en la naturaleza. En lo que resta de esta sección explicaré cuáles son las condiciones necesarias para atribuir estos términos cargados de normatividad y valor como ‘comportamiento motivado por el propósito de mantener la identidad’ que identifica el enactivismo. Tales condiciones son al menos tres: que el sistema sea autónomo, precario y adaptivo; las tres son propiedades de sistemas dinámicos que no presuponen en sí mismas e individualmente comportamientos normativos ni atribuciones de valor al ambiente y a las interacciones.

En la teoría autopoietica, esta relación se concibe de manera abstracta como *acoplamiento estructural* que se refiere a la interacción dinámica y recurrente entre dos sistemas que intercambian y transforman materia y energía (Maturana, 2002, pp. 16-17). Esta interacción ocurre solo en un nivel material, y el término ‘estructura’ se refiere precisamente a esa materialidad. Por su parte, la organización del sistema, que se refiere a las relaciones entre sus procesos y componentes es una propiedad exclusivamente formal que, aunque se instancia materialmente, no depende de ella: la organización autónoma es una propiedad que un sistema material tiene o no tiene, y en esa medida puede preservarse o perderse en el acoplamiento estructural del sistema material que la instancia. De acuerdo con la teoría autopoietica clásica, la materialidad por un lado y la organización por el otro son dos dominios distintos que no

interactúan causalmente, y la relación generativa que un observador podría identificar entre estos dos dominios se debe a “una relación histórica que hace el observador al correlacionar la dinámica del dominio de composición del sistema [su organización] con lo que ocurre con este como totalidad resultante [su estructura] en el dominio en el que existe como totalidad” (Maturana, 2002, p. 12). Tal relación histórica e impresión subjetiva no implican que la organización resulte causalmente de su nivel material.

La separación entre los dominios organizacional y material que hace la teoría autopoietica clásica es, sin embargo, insostenible por tres razones. Primero, de acuerdo con Di Paolo (2018, p. 82), el acoplamiento estructural no permite diferenciar las interacciones organismo-ambiente más allá de si se mantiene la organización del sistema o no. Esto implica que no permite entender los matices de la respuesta del organismo a las perturbaciones del ambiente y mucho menos diferenciar una interacción cognitiva de una no cognitiva (profundizo en este punto en la sección de Adaptividad y automantenimiento). Segundo, también de acuerdo con Di Paolo, la separación entre los dominios estructural y organizacional son idealizaciones en las que se quisiera que la materia y energía que entran al sistema fueran asimiladas perfectamente en la organización sin cambiarla, pero este no es el caso (2018, p. 83). Esto implica que la organización es parte de la materialidad misma del sistema y no un dominio independiente. Tercero, siguiendo a Froese (2017), la materialidad no es una propiedad contingente de la realización de la organización de un organismo capaz de interactuar con su ambiente. La materialidad de un sistema implica que además de ser autónomo es también un sistema termodinámico, y que mantener su organización autónoma es costoso termodinámicamente. Crucialmente, esto implica que eventualmente desaparecerá. Precisamente ese riesgo que viene dado con su materialidad implica que el sistema dejará de existir a menos que obtenga de su ambiente la materia y energía suficientes para mantener su propia identidad. Así que el dominio material o estructural del sistema debe influir causalmente en la organización del agente y su interacción con el ambiente. Más adelante, en la sección Precariedad, profundizaré en este punto, pero por ahora merece la pena notar que la precariedad es una propiedad termodinámica común a todos los sistemas, vivos e inertes, cognitivos y no-cognitivos. Por ende, la precariedad en sí misma no implica la capacidad de interactuar con el mundo en términos de valor y sentido; aunque sí juega un papel importante cuando el sistema es además autónomo y adaptivo como veremos más adelante en la siguiente sección.

El enactivismo busca derivar del concepto de autonomía no solo la autodiferenciación del sistema, su diferenciación del ambiente y la definición de su dominio de interacciones; busca también derivar los términos mismos en los que ocurre su interacción con el ambiente. Siguiendo a Jonas y su concepto de libertad necesitada (*needful freedom*), el enactivismo busca caracterizar esta relación como una en la que el agente hace sentido de su ambiente en términos de valor, de lo que requiere para mantener su identidad. A esta relación se le conoce como *hacer-sentido* (*sense-making*).

De acuerdo con la historia recibida del enactivismo, la autonomía del agente, en particular, su autodiferenciación, establece un criterio básico de normatividad: un evento es bueno en la

medida en que permite al organismo mantener su identidad y malo si la amenaza. Esto significa que la organización autónoma no solo permite al organismo diferenciarse del medio en el que existe, sino también establece una norma básica que rige la manera en que interactúa con su medio y esa norma es preservar su organización, la organización que define su identidad. Así, de acuerdo con las primeras formulaciones del enactivismo, cuando el organismo se autoconstituye se establece también a sí mismo como una agencia mínima desde donde el impacto que tiene el ambiente sobre este adquiere una valencia básica, y esta agencia mínima está determinada por el propósito de mantener su propia identidad (Thompson, 2007; Weber & Varela, 2002). En otras palabras, la autonomía del sistema establece la normatividad básica de permanecer dentro de su dominio de interacciones.

Sin embargo, la instauración del propósito intrínseco del automantenimiento y de la norma básica que se sigue de la autonomía no son suficientes para explicar las interacciones normativas. De acuerdo con Di Paolo (2005), la autodiferenciación del organismo únicamente instaaura una norma de todo o nada, es decir, solo permite diferenciar entre estados y tendencias que están dentro de los límites de viabilidad y los que están fuera. Entonces esta norma no permite dar cuenta de los grados en que las tendencias internas del organismo y sus interacciones con el ambiente lo acercan o lo alejan de los límites en los que perdería su identidad. El problema de esta diferenciación tan básica es que realmente no le permite al sistema reconocer los estados o comportamientos que, si se mantienen como están, lo llevarían a desaparecer. Así que, sin manera de diferenciar grados, estar corriendo directamente hacia el precipicio o en dirección o puesta, o estar enfermo y estar recuperándose son equivalentes respecto al propósito de mantener la identidad porque en todos casos el agente mantiene su organización, es decir, está dentro de sus límites de viabilidad por un tiempo (Di Paolo, 2005).

Pero si la autonomía no permite dar cuenta de este tipo de diferenciación, tampoco permite dar cuenta de qué hace que un sistema tenga una de las propiedades que, argumenté en el capítulo anterior, es característica de la agencia: tener la capacidad de diferenciar entre cursos de acción distintos. Después de todo, bajo la diferenciación binaria de dentro o fuera de los límites de viabilidad, todos los estados y comportamientos son buenos excepto los que lo llevan más allá del límite de viabilidad; pero si esto ocurre y el sistema sobrepasa ese límite, entonces deja de existir y, por lo tanto, ya no hay sistema que pueda diferenciar entre una cosa u otra.

Quisiera llevar el argumento de Di Paolo más allá y derivar una implicación adicional de que la autonomía no permita hacer diferencias de grado. La idea misma de una norma básica instaurada solo por la autonomía de un sistema tampoco puede defenderse: si el organismo no tiene la capacidad de distinguir sus propias tendencias, tampoco puede modular, generar o controlar sus interacciones o sus condiciones externas en función de sus preferencias propias, esto es, en función de lo que necesita para mantenerse. Pero, como argumento en el Capítulo 1, una norma existe solo en la medida en que hace una diferencia práctica en el patrón de comportamiento de los agentes. Así que, si el agente no tiene forma de distinguir los grados en que se acerca o aleja de sus límites de viabilidad, es decir, las consecuencias prácticas de



continuar o cambiar un curso de acción respecto a la norma básica de la autopreservación, no hay realmente ni agencia ni una norma. Por ende, la autodiferenciación que permite la autonomía no puede siquiera establecer una normatividad ni una teleología básica.

### 2.3.1. Adaptividad y automantenimiento

La solución que propone Di Paolo es hacer explícita una segunda característica definitoria de la agencia, la adaptividad, que ya estaba implícita en las formulaciones anteriores de la autonomía como base naturalista de la cognición. La definición que ofrece es la siguiente:

La adaptividad es definida como la capacidad de un sistema, en algunas circunstancias, de regular sus estados y su relación con el ambiente con el resultado de que, si los estados están lo suficientemente cerca al límite de viabilidad,

1. las tendencias son distinguidas y se actúa sobre ellas dependiendo de si los estados se acercarán o alejarán del límite y, como consecuencia,
2. las tendencias del primer tipo son acercadas o transformadas en tendencias del segundo tipo y así se previene que los estados futuros alcancen el límite con una velocidad hacia afuera (Di Paolo, 2005, p. 438)

En palabras más simples, la adaptividad es la capacidad de un organismo de (1) monitorear si sus tendencias alejan al organismo de sus límites de viabilidad o lo acercan a estos y (2) regularlas cuando sea necesario. Además, los procesos adaptivos son normativos porque siguen la norma intrínseca del organismo, a saber, preservar su forma de vida; de manera que la forma de vida permite que la norma emerja, y la norma, a su vez, mantiene la forma de vida.

Los procesos adaptivos pueden ocurrir tanto como procesos internos dentro de los límites físicos del organismo como por fuera de tales límites en la interacción del agente con el ambiente. Un proceso adaptivo se diferencia de otros procesos internos e interacciones porque el agente impone las condiciones en que se da el proceso o interacción. Di Paolo, Buhrmann y Barandiaran llaman a esta propiedad *asimetría interaccional* y la identifican como una capacidad necesaria para que el organismo sea adaptivo (2017, p. 118). Esta capacidad permite diferenciar, por ejemplo, un organismo que cae por un precipicio de uno que lo hace como parte de su forma de vida, por ejemplo, una ardilla voladora que salta a un precipicio y abre sus brazos para planear y controlar su caída. Pero incluso, procesos adaptivos internos dependen de la capacidad del organismo de abrirse o no a la interacción con el ambiente, es decir, de su capacidad de imponer las condiciones de interacción o no interacción. Por ejemplo, en procesos como la fotosíntesis, el organismo permite que el dióxido de carbono y el agua entren para ser metabolizados porque la energía solar necesaria para ello está disponible. La obtención de recursos necesarios para estos procesos internos es una interacción adaptiva. La adaptividad, sin embargo, no se limita a cerrarse a la interacción o a metabolizar materia y energía dentro de los límites físicos del organismo. Un proceso adaptivo puede también implicar la transformación de la materialidad de ciertos aspectos del ambiente que habita el organismo para poder interactuar adaptivamente con estos. Los

procesos de cocción de alimentos hacen precisamente esto en tanto que, mediante, ej., procesos de molienda y cocción, transforman elementos imposibles de metabolizar por el cuerpo humano, como las semillas de linaza o trigo, en alimentos, es decir, en sustancias metabolizables<sup>8</sup>.

Como los ejemplos anteriores sugieren, la adaptividad no requiere de mecanismos estables y dedicados para monitorear y regular las tendencias del organismo, sino que los procesos adaptivos pueden emerger de procesos distribuidos espacialmente (Mojica & Froese, 2019, pp. 123-124). Entonces, la capacidad adaptiva puede observarse primariamente en el comportamiento del sistema, más específicamente, en su patrón de respuestas a las perturbaciones que amenazan su autonomía. Cuando la manera en que este patrón de comportamiento es implementado autónomamente por el organismo y responde a las necesidades intrínsecas de automantenimiento del sistema, surge una agencia, esto es, un sistema que en su interacción con el ambiente es capaz de imponer y seguir su normatividad intrínsecamente. Una de las maneras en la que la adaptividad puede ser instanciada es a través del automantenimiento recursivo, esto es, la capacidad de interactuar de maneras distintas con el ambiente para automantenerse cuando las condiciones del ambiente cambien. En la sección 3.2 desarrollo esta idea.

Como señalé al final de la sección anterior, la autonomía por sí misma es incapaz de establecer realmente una norma para el sistema porque no permite establecer diferencias prácticas en el comportamiento del agente. La adaptividad, es decir, la capacidad del sistema de monitorear y regular sus interacciones, crea un espacio de posibles interacciones en donde sí se pueden diferenciar ciertos cursos de interacción de otros en tanto sus consecuencias prácticas para el sistema. En esa medida, la adaptividad tiene dos funciones: no solo permite al organismo evaluar el ambiente y actuar de acuerdo con la norma, sino que establece la norma misma de manera situada, es decir, establece las consecuencias más allá de las impresiones o estados presentes del agente, consecuencias que permiten determinar el éxito o fracaso de sus interacciones. Esto significa que, en el nivel básico de las formas mínimas de vida, tanto la agencia como la norma objetiva con la que se evalúa un comportamiento son establecidas por la adaptividad.

Barandiaran y Egbert (2014) apuntan a una distinción similar, y muestran la forma en que la adaptividad puede establecer la norma de interacción de manera situada. Barandiaran y Egbert parten de la distinción introducida por Canguilhem (1966 [1978]) entre el establecimiento de una norma y el seguimiento de una norma para mostrar cómo podría precisarse qué significa que el comportamiento adaptivo siga la norma de automantenimiento establecida por la

---

<sup>8</sup> No bajo todas las concepciones de adaptividad en la literatura enactiva este tipo de procesos se consideran adaptivos. Tanto Moreno y Mossio (2015) desde la biología y la filosofía de la biología, como Villalobos (2020) desde la filosofía de la mente y de las ciencias cognitivas han defendido que los procesos adaptivos son exclusivamente los procesos cuyos componentes constitutivos hacen parte de la red metabólica del sistema. Puesto que los procesos de cocción no necesitan hacer parte de una red metabólica para existir, estos no cuentan como procesos adaptivos. Contrario a estas posiciones, la definición de adaptividad que adopto es más flexible y permite que cualquier proceso normativo y cuyos parámetros de interacción estén impuestos por el agente cuente como proceso adaptivo.

autonomía del sistema. Su estrategia es analizar un modelo de sistema precario mínimo cuyo automantenimiento depende de una sola variable, a saber, mantener la concentración de químicos estable dentro del sistema, y cuyo ambiente de interacción tiene diferentes concentraciones de recursos materiales y energéticos disponibles para el sistema. La viabilidad del sistema depende de la relación entre sus variables internas y los recursos disponibles en su ambiente. Esto les permite clasificar esta relación en cinco posibles escenarios (ver Figura 3). Primero, puede ser el caso que los recursos del ambiente sean suficientes y el estado interno del sistema se mantenga en equilibrio; en tal caso, el sistema está *adaptado* y por ahora no necesita hacer nada, ni compensar internamente ni buscar un ambiente distinto, para mantener su autonomía. Segundo, el sistema puede estar en equilibrio termodinámico porque perdió por completo su organización; en tal caso, el sistema está *muerto*.

Los otros tres escenarios son escenarios de precariedad y ocurren cuando el sistema debe moverse a un ambiente con más recursos porque la relación entre recursos actuales internos y externos, de mantenerse en la tendencia en la que están, llevarían inminentemente al sistema a la muerte. Dentro de este escenario, una primera posibilidad es que la relación entre el estado interno del sistema y los recursos externos sea suficiente para que el sistema alcance a moverse a una región donde haya más recursos externos y asimilarlos antes de perecer; este es el tercer estado del sistema, el estado *adaptivo*, y esta es la situación en donde surge realmente la normatividad. Cuando el sistema se encuentra en esta situación hay interacciones específicas que le permiten seguir viviendo y otras que lo llevan a la muerte, y en esa medida, hay interacciones situadas, concretas y específicas que son buenas para el sistema y debe seguirlas, y otras que debe evitar porque constituirían un error por ser fatales. En esta situación, la manera en que el sistema conduce sus interacciones hace una diferencia práctica, la diferencia más práctica posible, la diferencia entre la vida y la muerte.

En las otras dos situaciones, o bien los recursos externos son tan escasos que el sistema no puede hacer nada para evitar morir, independientemente de su estado, este es un estado *fútil*. Finalmente, el quinto estado es el estado *terminal* en que el estado interno del sistema está tan lejos de la adaptación que está en curso de morir independientemente de la cantidad de recursos externos disponibles. Si se acepta una noción de normatividad pragmática como la que definiendo en el capítulo 1, es decir, una normatividad en la que actuar de manera correcta e incorrecta hace una diferencia práctica, en estos dos últimos estados es imposible que el sistema interactúe con su ambiente de forma normativa: todos los caminos conducen a la muerte y ninguno a la adaptividad.

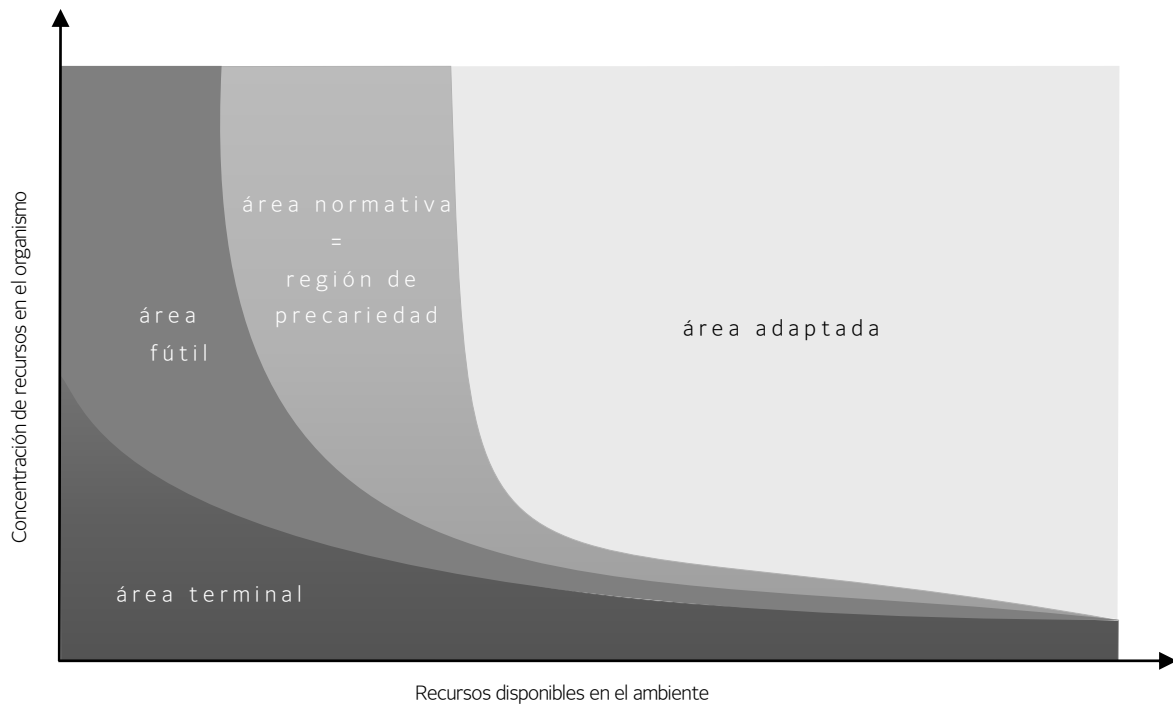


Figura 3

Adaptada de Barandiaran y Egbert (2014, Figure 8).

El área normativa o región de precariedad representa las circunstancias en que (i) el estado y recursos internos del organismo necesitan ser renovados para evitar la situación en la que la muerte es inminente, y (ii) los recursos disponibles en el ambiente pueden suplir lo que el organismo necesita. Esta es la única circunstancia en la que el curso de acción que seleccione el organismo implica una diferencia entre la muerte inminente y la adaptividad. Los valores por debajo de cero en la concentración de recursos en el organismo, es decir, la región bajo el eje x, representa la muerte del organismo. En la región arriba del eje x el organismo está vivo, incluso si este está a punto de morir. Las áreas gris oscuro y gris medio oscuro, áreas terminal y fútil respectivamente, son situaciones en las que el organismo no puede hacer nada para evitar morir. Finalmente, el área gris claro, el área adaptada, representa circunstancias en las que los recursos internos del organismo y de su ambiente son suficientes y el organismo no está en riesgo de perecer. En este caso y durante algún tiempo limitado, el organismo puede seleccionar cualquier curso de acción sin poner en riesgo su vida.

Hay dos conclusiones que Barandiaran y Egbert derivan del modelo mínimo que proponen y que permiten caracterizar de manera más precisa la normatividad. En primer lugar, su modelo permite mostrar que la norma del automantenimiento no existe en abstracto; en cambio, su sentido depende de que haya comportamientos adaptivos posibles para el sistema y necesarios para mantener su identidad del sistema en un momento y circunstancia dada con sus necesidades particulares. Esto implica que no hay un mecanismo ni un momento fundacional ni una variable del sistema dedicada a establecer la norma; en particular, la autonomía de un sistema por sí sola no establece la norma de interacción del sistema, sino que esta es establecida solo cuando el comportamiento adaptivo es necesario, posible, y hace una

diferencia práctica para el organismo respecto a requerimientos específicos del sistema. Crucialmente, esto implica que la normatividad, en tanto criterio normativo de evaluación de las interacciones, existe únicamente en circunstancias específicas, en las circunstancias en que hace una diferencia práctica respecto a requerimientos específicos para el mantenimiento de su propia forma de vida. Fuera de ahí, las interacciones no son normativas, al menos no respecto a esa dimensión particular.

Segundo, es posible que la respuesta del sistema sea errónea cuando una interacción adaptiva era necesaria. Puede ocurrir que eventos internos o externos al sistema eviten que el sistema transforme su tendencia y se encamine hacia la adaptación o que no lo haga con la eficiencia requerida. En esos casos, hay un error. Puesto que la normatividad se trata de procesos interactivos para regular tendencias, seguir una norma o errar en un momento y circunstancia dada es un asunto gradual: el sistema puede contrarrestar más o menos su tendencia hacia la muerte inminente. Por lo tanto, ni los criterios de normatividad ni los comportamientos del sistema siguen una norma de todo o nada, sino que se trata de una evaluación matizada y necesariamente atada a circunstancias específicas.

La idea de que el automantenimiento es la norma más básica que surge en la naturaleza como lo defiende el enactivismo puede ponerse en duda por las varias formas de vida o seres vivos que ponen en riesgo su propia vida. Consideremos un ejemplo en el que no solo un ser vivo pone en riesgo su vida, sino que biológicamente su propósito es perderla: los gametos, ej., los óvulos y espermatozoides humanos, cuya función es fusionarse con el gameto opuesto para dar lugar a un individuo distinto. En este proceso, cuando es exitoso, el espermatozoide empieza perdiendo sus partes, la cola y la vesícula acromosomal, o modificando su membrana en el recorrido dentro de útero; luego, cuando alcanza el óvulo, une sus membranas con este y finalmente, si la formación del cigoto es exitosa, los núcleos del óvulo y de lo que queda del espermatozoide se unen y, con ello, reorganizan y unen sus cromosomas y dan lugar a un nuevo individuo (Cornejo, 2009). La célula que resulta de este proceso no es idéntica ni al espermatozoide ni al óvulo, y por ello, su función biológica propia no es automantenerse.

Este caso, sin embargo, no constituye una refutación a la caracterización enactiva de los propósitos propios más básicos. En primer lugar, los gametos buscan activamente mantenerse con vida hasta el momento que empieza el proceso de fecundación por eso, metabolizan glucosa y generan energía para mantenerse con vida. Esto no significa que su función biológica no sea unirse al gameto opuesto y morir; significa, en cambio, que estas células pueden tener propósitos individuales y propósitos establecidos por la evolución que van en contra entre sí. En segundo lugar, estas células hacen parte de un sistema autónomo más amplio, el cuerpo humano y en condiciones normales no sobrevivirían fuera de este. Pertenecer a este sistema autónomo más amplio del que dependen hace que estén sujetas a criterios de normatividad adicionales a su propio automantenimiento. Sin embargo, la propuesta enactiva que he presentado hasta ahora no ofrece las herramientas necesarias para explicar este fenómeno. Más adelante, volveré a este problema general de la posibilidad de que los agentes estén sujetos a criterios normativos conflictivos y en el Capítulo 4 propondré una solución.

Merece la pena notar que las circunstancias en la que el sistema no está adaptado pueden transformarse a lo largo del tiempo; por ejemplo, gracias a la exposición progresiva a una sustancia nociva, o a daños por la precariedad misma del sistema, es posible que aumente o disminuya el rango de estados internos y externos que requieren una respuesta adaptiva. Regresaré a este punto más adelante.

La propuesta de Barandiaran y Egbert (2014) ofrece una articulación sólida aunque incipiente de cómo se constituyen las normas respecto a las cuales puede ser evaluado un comportamiento. Puesto que esta constitución se sigue de las necesidades adaptivas propias del agente, tales normas son normas intrínsecas a la forma de vida del sistema. Es decir, normas que el agente crea para sí mismo, y en ese sentido son autodeterminantes. Sin embargo, esta propuesta deja abierta la pregunta por cómo el agente puede actuar de acuerdo con estas normas o fallar al hacerlo, esto es, cómo de hecho el organismo es movido a seguir tales normas.

No solo los propios Barandiaran y Egbert (2014) reconocen que su propuesta no responde esta pregunta, sino que es un problema que ha sido señalado por otros autores. Barrett (2017), por ejemplo, articula este problema en términos de motivación. Para él, la pregunta que no se ha resuelto dentro de la propuesta enactiva es cuál es el mecanismo que motiva al sistema a actuar de acuerdo con la norma del automantenimiento. Barrett propone entender el criterio normativo no en términos de automantenimiento sino en términos de optimización termodinámica de alguna clase (Barrett, 2019, 2020). Si lográramos tener una respuesta precisa de cómo los agentes son movidos a actuar, podríamos explicar cómo surge una agencia mínima desde donde *importa* si la interacción entre agente y organismo resulta de una forma u otra. Esto permitiría responder cómo surge el error en tanto un camino que hace la diferencia para quien actúa y que hace cualitativamente distinto su comportamiento al de otros sistemas autónomos como el de las llamas de fuego o las estrellas. Esta es una pregunta abierta que requiere una investigación que sobrepasa los límites de esta tesis. Por ahora explicaré cuál es la perspectiva del agente hacia su ambiente que el concepto de adaptividad busca explicar.

## Mundo de valor

El concepto de adaptividad permite caracterizar el dominio de interacciones de una forma más rica y cercana a la experiencia de los sujetos en interacción con su ambiente. Arriba expliqué que el dominio de interacciones es el conjunto de intercambios de materia y energía en las que el agente no pierde su autonomía. En la medida en que un agente se autodetermina y necesita mantenerse a lo largo del tiempo, también determina qué intercambios de materia y energía con el ambiente entran a su dominio de interacciones y cuáles no (ver (ii) *Automantenimiento y dominio de interacciones*). Bajo esa caracterización basada únicamente en la autonomía, el dominio de interacciones es homogéneo porque hay diferencia entre las múltiples formas en que el agente puede mantener su organización. Sin embargo, la adaptividad, en tanto capacidad del organismo de monitorear y regular sus tendencias respecto a sus propios límites de viabilidad, permite hacer una distinción más matizada del dominio de interacciones:

permite al organismo reconocer si una interacción lo acerca o lo aleja de sus propios límites y entonces puede reconocer diferentes grados de valor en el ambiente, a saber, lo bueno y lo malo, lo más y lo menos conveniente para mantener su propia vida. De esa manera, el concepto de adaptividad busca explicar cómo el ambiente con el que interactúa el organismo es significativo para este como un mundo de valor.

El concepto de mundo se refiere entonces al ambiente percibido por el agente en términos de lo que permite mantener su forma de vida o la amenaza. Es importante notar que *mundo* se trata de la relación que el agente establece con su ambiente, no se refiere al ambiente físico que existe independientemente del agente ni tampoco a las necesidades del agente exclusivamente. El concepto enactivo de mundo es equivalente al concepto de *umwelt*. Baggs y Chemero (2018) toman este concepto de la propuesta original de von Uexküll (1957) y lo incorporan a la psicología ecológica: el *umwelt* es el ambiente percibido en términos de oportunidades de acción por un agente en particular. Baggs y Chemero proponen diferenciar el *umwelt*, por un lado, del hábitat, que se refiere al ambiente con el que pueden interactuar los miembros de una especie o una forma de vida, y por otro lado, del ambiente físico, que se refiere a la materialidad independiente de la forma en que cualquier ser vivo se relacione con este. En el capítulo 3 examino el concepto de hábitat y propongo entenderlo como correlato material, social y cultural de la normatividad externa de la cognición.

### 2.3.2. Precariedad

La organización de los seres vivos es precaria, y esta característica es crucial: la precariedad es lo que lleva al sistema a interactuar con el ambiente y es lo que hace que haya una diferencia práctica para el sistema entre una y otra tendencia, entre una y otra interacción o no interacción. De manera que, en virtud de la precariedad de la identidad, tanto las interacciones con el ambiente como la identidad misma se vuelven valiosas y significativas para el agente (Froese, 2017).

La precariedad no es, sin embargo, una propiedad única de los seres vivos, en cambio, es una propiedad determinista y universal porque todos los sistemas, vivos e inertes, tienden al equilibrio termodinámico y esto significa que eventualmente la organización de cada sistema que existe y existirá se disipará. La precariedad se vuelve una característica crucial para los sistemas vivos y normativamente determinante para sus interacciones cuando ocurre a una velocidad que coincide con las capacidades adaptivas del sistema, es decir, cuando ocurre a una velocidad que permite al sistema detectarla y contrarrestarla. Si un sistema autónomo no es precario, no necesita interactuar con su ambiente para mantener su autonomía y entonces no puede reconocer en su ambiente un mundo de valor ni puede instaurarse una norma intrínseca para su comportamiento. En esa medida, la precariedad no es una propiedad accidental de los sistemas vivos, sino que es esencial para su identidad y su capacidad de comportarse adaptivamente.

## Tensión: apertura del automantenimiento y clausura de la autodeterminación

La precariedad lleva al organismo a una tensión intrínseca que lo lleva a interactuar adaptivamente con el ambiente. La razón, como reconoce Di Paolo (2018), es la siguiente: si un sistema que se autodiferencia no es precario, no necesita interactuar con el ambiente y bastaría con que se cerrara a todo intercambio de materia y energía con el medio exterior para continuar existiendo. Así que, como señala Di Paolo (2018), la autodiferenciación de los sistemas autónomos en condiciones ideales llevaría al sistema a cerrarse. Los sistemas vivos, por el contrario, son sistemas precarios, es decir, sistemas que decaen rápidamente si no renuevan los procesos que lo componen. Por eso no basta con que un organismo se autodiferencie, sino que debe automantenerse a través del tiempo, y esto implica que necesita obtener materia y energía de su interacción con el ambiente. La capacidad de automantenimiento del agente, lo llevaría idealmente a abrirse a la interacción con el medio exterior, pero esta tendencia se opone a la tendencia a cerrarse a la interacción implicada por la autodiferenciación (Di Paolo, 2018). Así que hay una tensión en la relación entre agente y su ambiente que se sigue de la identidad misma del agente como sistema autónomo y precario.

Di Paolo propone que esta tensión se resuelve dialécticamente por la adaptividad que permite al sistema abrirse o cerrarse a interactuar con el ambiente en virtud de si sus propias tendencias internas lo requieren o no. En virtud de su precariedad, el organismo constantemente necesita obtener recursos del ambiente; por ello, mientras siga viviendo, nunca superará la tensión por completo y constantemente tendrá que interactuar adaptivamente con su ambiente. Esto implica que, si el sistema no fuera precario, no estaría movido a interactuar adaptivamente, y por esa razón no se puede considerar la materialidad del sistema como una característica accidental: la precariedad es esencial para que los sistemas puedan interactuar normativamente con su ambiente. A pesar de que esta propuesta de Di Paolo en el 2018 permite especificar con algo más de detalle cómo el organismo puede estar movido a actuar adaptivamente, a saber, para superar la tensión entre sus dos tendencias intrínsecas; su propuesta no permite resolver los dos problemas que identifiqué arriba que son el problema de la motivación (p. 70) y la necesidad de dar cuenta de la cognición como un proceso subdeterminista (p. 89).

## Identidad-mundo cambiantes

La materialidad del sistema no solo implica su precariedad, sino implica también que sus interacciones no están perfectamente limitadas a conseguir recursos para el automantenimiento: las interacciones, presumiblemente las que tienen cierta regularidad, pueden transforman tanto la identidad del organismo como a su mundo. La transformación del mundo del organismo ha sido tratada con profundidad por la psicología ecológica y la trataré en el capítulo 4. La transformación de la identidad del agente por su parte ha sido reconocida pero no tratada a profundidad por el enactivismo.

Di Paolo, por ejemplo, sustenta la transformación de la identidad de los organismos señalando que la separación abstracta entre relaciones estructurales –v.gr. las que proveen recursos–, y



relaciones organizacionales –v.gr. las que impactan la identidad organizacional del sistema– es una idealización útil para el diseño de artefactos, pero no como modelo para entender la relación entre organismo y su ambiente (Di Paolo, 2018, p. 84). Incluso en las aplicaciones tecnológicas esta separación tampoco es el caso; por ejemplo, aunque se supone que el combustible y la energía eléctrica provean energía pura al motor de los autos o a los electrodomésticos respectivamente, con el tiempo las impurezas del combustible y las fluctuaciones de energía deterioran la organización del motor y electrodomésticos, y hacen que su funcionamiento y eficiencia cambie. Así mismo, la organización de los seres vivos es parte de la materialidad misma del sistema y no un dominio independiente, de manera que su organización se transforma en los intercambios con el ambiente. Este tipo de transformaciones en la organización de los seres vivos son evidentes, por ejemplo, en la adquisición de una habilidad (Di Paolo, 2018) o en la exposición repetida a una sustancia (Barandiaran & Egbert, 2014).

Más allá de la materialidad de las interacciones, todos los seres vivos se caracterizan porque su organización cambia regularmente desde el inicio hasta el final de su vida. Esto ocurre tanto en formas mínimas y unicelulares de vida en sus procesos de crecimiento y reproducción, como en organismos multicelulares cuya organización cambia regularmente desde su concepción, pasando por el crecimiento y la adultez hasta la vejez. De manera que la identidad del organismo y, por ende, su forma de interactuar con eventos del ambiente se transforma no como una propiedad accidental sino como una propiedad propia de la vida. Sin embargo, esta transformación de la identidad del organismo a lo largo del tiempo, presumiblemente adaptativa, no ha sido tratada a profundidad dentro de la literatura enactiva a pesar de que ha sido indicada incipientemente, y a pesar de que es un punto importante para ofrecer una explicación enactiva completa de cómo se constituyen las capacidades de hacer sentido de agentes y especies. Quizás la razón no es únicamente el tiempo relativamente corto que lleva el enactivismo consolidándose, sino que no es evidente cómo puede explicarse este cambio regular y característico de la vida bajo la premisa autopoietica de que los seres vivos buscan intrínsecamente mantener su organización. Quizás el propósito intrínseco de mantener la propia forma de vida no es equivalente a mantener la organización, sino que implica en sí mismo un elemento de cambio. De manera que esta pregunta por el papel del cambio en la constitución de la vida está relacionada con el problema de la motivación que señalé más atrás y que trataré más adelante: probablemente los seres vivos no están movidos intrínsecamente a mantener su organización, sino que su motivación es, siguiendo a Spinoza, la reafirmación en el cambio.

### 2.3.3. Temporalidad

La precariedad es una característica inherentemente temporal: un sistema o proceso es precario si decae en una escala temporal compatible con la vida, es decir, ni demasiado rápido ni demasiado lento para poder ser compensado adaptivamente por el organismo. La idea misma de decaimiento es imposible de entender en términos estáticos porque un sistema estático, un estado congelado en el tiempo, no decae. Así que, puesto que los sistemas

cognitivos son precarios, son también intrínsecamente temporales. Esto tiene al menos dos implicaciones importantes.

Primero, las interacciones adaptivas son intrínsecamente temporales. La razón es que, para monitorear y contrarrestar las tendencias del organismo hacia sus límites de viabilidad, el organismo debe poder llevar a cabo tanto procesos que puedan detectar en el tiempo los cambios en sus tendencias, como procesos que desaceleren o reviertan la velocidad con la que este se acerca a sus límites de viabilidad. En particular, una respuesta adaptiva debe sincronizarse rítmicamente con los tiempos de la precariedad del organismo de manera que la respuesta esté en el rango temporal y el rango de velocidades adecuadas para contrarrestar las amenazas externas y la propia precariedad del organismo, y así permanecer con vida. Si los tiempos y ritmos de estos procesos no son tomados en cuenta —por ejemplo, las respuestas adaptivas se consideran estados del organismo detenidos en el tiempo como genes que se encienden o se apagan—, es imposible determinar si un sistema es precario o no y si es adaptivo o no. Esto a su vez haría imposible encontrar diferencias prácticas entre sistemas cerrados operacionalmente adaptivos y no adaptivos, por ende, haría imposible diferenciar entre sistemas capaces de hacer sentido de su ambiente de sistemas que no. Este papel fundamental del tiempo es precisamente lo que motiva el uso de sistemas dinámicos como el modelo más adecuado para entender los procesos cognitivos (Port & Van Gelder, 1995; Richardson & Chemero, 2014).

Segundo, la adaptividad no solo permite que el ambiente se presente como un espacio significativo para el organismo, sino también permite que este tenga una noción de temporalidad. Arriba mostré que, de acuerdo con el enactivismo, la adaptividad permite al organismo interactuar con su ambiente en términos de valor y sentido, es decir, en términos de lo bueno y lo malo para el mantenimiento de su identidad. Análogamente, el carácter temporal de la adaptividad en tanto proceso que se despliega en sincronía con la precariedad del organismo hace que el paso del tiempo se vuelva significativo para el organismo. Los cambios internos y externos que llevan al organismo de un estado adaptado a uno precario son cambios significativos para este y le permiten hacer la diferenciación básica entre pasado adaptado y presente precario que llama a la adaptividad. Presumiblemente, también los cambios en la identidad misma del organismo en el desarrollo, vejez o enfermedad (p. 72), dada su irreversibilidad, su carácter presumiblemente adaptivo y la transformación profunda que implican en la manera que el organismo se relaciona con el ambiente, permiten al organismo experimentar el paso del tiempo como una sucesión significativa de cambios irreversibles.

#### 2.3.4. Grados y espacio normativos

Hasta a ahora he mostrado la manera en que la aproximación enactiva da cuenta de cómo surge en la naturaleza una agencia mínima, es decir, un individuo que es capaz de diferenciar entre un curso de acción y otro de acuerdo con una norma o propósito, y cuyas acciones son también evaluables respecto a sus propios propósitos. Es un buen momento para recapitular el

argumento y las limitaciones que he identificado. El punto de partida conceptual es el surgimiento de individuos autónomos en la naturaleza.

La explicación que ofrece el enactivismo es que los seres vivos son sistemas que se autodiferencian, es decir, sistemas que mediante sus propios procesos realizan su propia identidad y se diferencian físicamente del medio en el que existen. Esta explicación permite situar en la naturaleza misma de los organismos y no en el espectador la constitución de su propia identidad. Sin embargo, la autonomía es una propiedad formal que no es suficiente para explicar características esenciales de los seres vivos, por ejemplo, su precariedad. La precariedad de los organismos hace que necesiten obtener materia y energía a los ritmos adecuados para seguir existiendo. Esto los lleva a abrirse a la interacción con el ambiente, pero esta tendencia, representada en la Figura 4 con la línea punteada, contradice la tendencia a cerrarse que implica su autodiferenciación, representada en la Figura 4 con la línea continua.

La tensión entre ambas tendencias se resuelve en tanto que el organismo establece una relación adaptiva con el ambiente. Las interacciones adaptivas son interacciones en las que el organismo monitorea y regula sus propias tendencias con respecto a si se acercan o alejan a sus límites de viabilidad. Así que el organismo se cierra o se abre a la interacción con el ambiente en la medida y en los ritmos que sus necesidades de automantenimiento imponen, en ese sentido la relación adaptiva es asimétrica y normativa. A esta relación adaptiva se le conoce también como *sense-making* porque implica que el organismo se relaciona con un ambiente en términos de lo que tiene valor para su vida. En otras palabras, aquello en el ambiente que percibe y con lo que interactúa el organismo es un mundo de valor al que no es indiferente.

El concepto de *sense-making* o ‘hacer-sentido’ es extremadamente popular en la literatura enactiva autopoiética, por eso vale la pena hacer una aclaración conceptual al respecto en este punto. ‘Hacer-sentido’ se refiere a las respuestas adaptivas del organismo en las que este interactúa con el ambiente, y se contrapone entonces a las actividades adaptivas no interactivas, por ejemplo, la regulación de los latidos del corazón. Esta definición, sin embargo, admite casos como la regulación de la temperatura del cuerpo en donde ocurre un intercambio entre el organismo y el ambiente regulado por el organismo. Si bien este tipo de interacciones son normativas en tanto que buscan mantener los procesos homeostáticos del cuerpo y pueden ser exitosas o fallidas, puede ser confuso usar la palabra ‘sentido’ para referirse a estas. Para evitar las presuposiciones asociadas con esta palabra y las posibles confusiones o implicaciones ilegítimas que estas puedan traer, me separaré de la tradición enactiva autopoiética y evitaré usar el término ‘hacer-sentido’ excepto cuando sea necesario para la exposición de conceptos bien conocidos en la literatura como el ‘hacer-sentido participativo’.

Como se muestra en la Figura 4, el mundo de valor, es decir, el ambiente que habita y que es relevante para el organismo, hace parte de un espacio relacional más amplio, el espacio relacional de precariedad que, como definí arriba, consiste en el conjunto de situaciones en las que el organismo podría actuar y en las que el ambiente podría ofrecer los recursos necesarios para que su acción haga la diferencia para su propio mantenimiento (ver Figura 3). El espacio de precariedad es entonces el espacio en el que el organismo podría establecer una relación adaptiva, y por ende significativa, con el ambiente, y esto incluye tanto el espacio que actualmente habita (i.e., el mundo de valor) como las situaciones potenciales, pero no actuales en las que el organismo podría establecer una relación adaptiva con el ambiente.

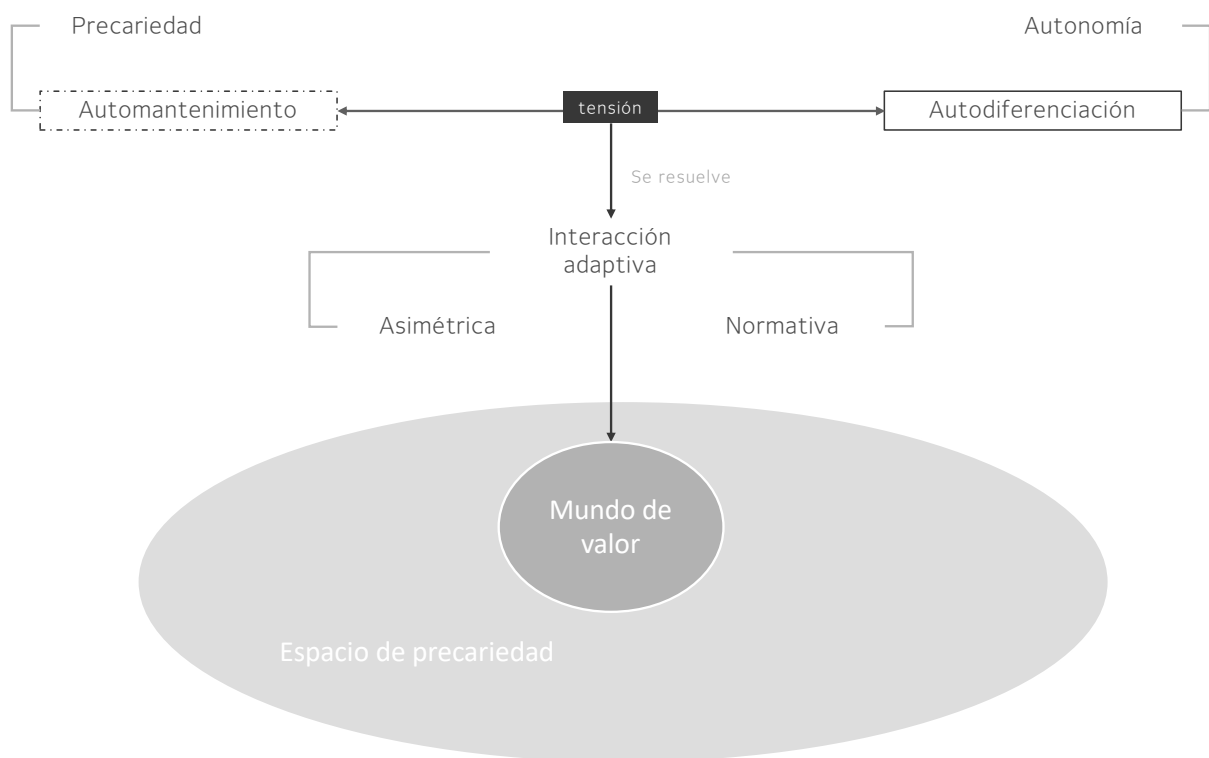


Figura 4

La concepción enactiva de agencia mínima y de la relación adaptiva, asimétrica y significativa que establece con el ambiente permite ofrecer una concepción de normatividad naturalista basada en el propósito de mantener la propia vida que esboza tanto el surgimiento de la agencia mínima, como el surgimiento de los criterios que determinan si un comportamiento o respuesta es normativamente adecuado o no.

Por un lado, los criterios de normatividad surgen en circunstancias que cumplen dos condiciones. La primera condición es que el organismo necesite renovar alguno de los procesos o elementos que pertenecen a la red que lo constituyen. En el modelo mínimo de Barandiaran y Egbert (2014) esta condición es definida en los siguientes términos: la concentración de un aspecto esencial para mantener la forma de vida del organismo está perdiendo concentración dentro de su organización, y de continuar esta tendencia el

organismo pronto empezaría a morir irreversiblemente. La segunda condición es que tanto las capacidades como los recursos externos disponibles para el organismo le permitirían revertir tal curso de acción y llevarlo fuera de la región de peligro. Esta región es un espacio de precariedad en el que el curso de acción que seleccione el organismo tiene implicaciones prácticas importantes para este y que pueden diferenciarse por grados entre dos extremos: las respuestas que sacan al organismo del espacio de precariedad y le permiten mantener su forma de vida de la manera más eficiente y las respuestas que lo llevan más rápida e irreversiblemente hacia la muerte. De esta manera los criterios de normatividad se definen de forma situada y en función de las capacidades y las circunstancias particulares del organismo. Es importante notar que esa es una definición del espacio de posibilidades normativas para un agente en una circunstancia en particular, pero que es independiente del curso de acción que de hecho seleccione el organismo en el sentido en que, si el organismo selecciona un curso de acción que lo perjudica, su comportamiento será inapropiado incluso cuando al organismo mismo no se lo parezca. Veremos en el Capítulo 4 que en la gran mayoría de los casos los comportamientos de un agente están sujetos a diferentes criterios normativos y por ello pueden ser inapropiados bajo cierto criterio, ej. para mantener la viabilidad del organismo, pero ser apropiados bajo otro criterio, ej. alcanzar los estándares de excelencia propios de la gimnasia olímpica.

Por otro lado, la agencia mínima surge de sistemas autónomos cuya capacidad adaptiva les permite monitorear y regular sus propias tendencias a un ritmo adecuado para contrarrestar su naturaleza precaria y continuar existiendo. La adaptividad es una capacidad de los organismos cuya implementación puede ser a través de mecanismos dedicados o emerger de procesos distribuidos espacialmente, pero que crucialmente surge autónomamente del organismo, es decir, surge de la red de procesos que constituyen al organismo mismo. Si la capacidad de preservar la autonomía precaria del sistema surge del sistema mismo, surge la agencia mínima en la naturaleza en tanto que el sistema constituye intrínsecamente la norma que su comportamiento sigue.

Hay, sin embargo, varias preguntas que quedan abiertas con la concepción enactiva de normatividad. Primero, algunos de los conceptos cruciales de la propuesta enactiva parecen implicar implícitamente aquello que quieren explicar: una agencia mínima. Segundo, se le ha acusado de ser poco respetable científicamente. Tercero, como anticipé arriba, se ha señalado que deja abierto el problema de la motivación. Y cuarto, en la propuesta enactiva, no queda claro cómo es posible que surja una forma de normatividad que no esté atada a las necesidades del individuo, pero esta forma de normatividad es crucial para que exista el comportamiento genuinamente cognitivo. A continuación, revisaré los méritos de estas críticas y de la propuesta enactiva.

## 2. 4. Críticas a la teleología enactiva

### 2.4.1. La amenaza del psicologismo y de la circularidad

Una de las primeras críticas a la propuesta enactiva es que el uso de conceptos como ‘propósito’, ‘precariedad’, ‘experiencia’, ‘motivación’, ‘sobrevivencia’, etc. no solamente son atribuciones injustificadas y psicologistas a formas de vidas no humanas, sino que asumen lo que se quiere explicar, esto es, una agencia mínima.

Primero, respecto al psicologismo, las atribuciones a formas de vida como el propósito de sobrevivir o la motivación para hacerlo parecen atribuciones exageradas y antropomorfistas. Esta objeción al enactivismo señala que su explicación atribuye rasgos del comportamiento humano a entidades no humanas (Villalobos & Ward, 2016). Hacer esta atribución es problemática porque, como señalan Villalobos y Ward (2016), ninguna explicación de un fenómeno de la naturaleza debe atribuir rasgos humanos a entidades no humanas para ser respetable científicamente. Aunque para algunos esta afirmación parece obvia y justificada en sí misma, es necesario mostrar por qué hacer atribuciones de rasgos humanos es injustificado en objetos de estudio concretos, en particular, cuando se estudian formas de vida no humanas, después de todo, humanos y formas de vida no humanas hacen parte del mundo de lo vivo y en esa medida comparten características y necesidades básicas: necesidad metabólicas y reproductivas. Más adelante me ocuparé de otras razones que soportan la acusación de Villalobos y Ward de irrespetabilidad científica, pero por ahora me ocuparé de una razón que soporta la acusación específica del antropomorfismo.

El significado de nuestros conceptos como ‘motivación’ y ‘propósitos’ está constituido por nuestras prácticas sociales, es decir, por los patrones de usos de tales expresiones lingüísticas y las actividades en las que los usamos. Nuestros propósitos y motivaciones suelen estar atados a situaciones en las que pedimos y damos razones de nuestro comportamiento y del de otros. Consideremos las siguientes expresiones lingüísticas y las circunstancias en las que las usamos—¿Cuál es tu propósito profesional y por qué aplicar a esta posición contribuye a ese propósito? ¿Qué te motiva para querer trabajar con nosotros? —¿Cuáles son tus verdaderas intenciones con él? ¿Qué buscas al llamarlo si ni siquiera te cae bien? —¿Para qué le cortas el pelo a tu muñeca? ¿Sabes que no le volverá a crecer? —¿Por qué lloras, bebé? Ya debes tener hambre. En todos los casos, las motivaciones, propósitos o razones del comportamiento se articulan en el lenguaje. En los dos primeros casos se trata de prácticas compartidas por adultos que ya saben cómo utilizar los términos, es decir, saben qué cuenta como un propósito o motivación y qué no, y qué es aceptable y qué no. En cambio, los últimos dos diálogos se insertan en actividades en las que enseñamos a niños e infantes a usar el lenguaje y, con ello, a dar razones aceptables para su comportamiento e incluso a comportarse de formas específicas, ej., en vez de llorar al tener hambre, aprendemos a hablar para pedir comida.

Evidentemente, es un sin sentido atribuir propósitos y motivaciones profesionales a formas de vida no humanas, o ‘verdaderas intenciones’ en sus relaciones o motivaciones incorrectas al jugar con sus muñecas, etc. Hay, sin embargo, un continuo entre el bebé que llora porque

tiene hambre, el gato doméstico que maúlla porque tiene hambre, la rana que caza la mosca porque tiene hambre, etc. En todos los casos, se trata de seres vivos movidos a expresar y actuar por sus necesidades vitales, es decir, están movidos por las necesidades que su propia constitución corporal impone para continuar existiendo. No es necesario postular estados mentales proposicionales ni mecanismos internos antropomórficos para explicar esta motivación. Precisamente este tipo de motivaciones corporizadas fundamentales es de donde se agarran nuestras prácticas sociales. En ellas, reemplazamos los comportamientos con los que expresamos nuestras necesidades más básicas y en esa medida refinamos nuestras motivaciones y expresiones; por ejemplo, aprendemos a reemplazar el llanto por expresiones lingüísticas para pedir alimento, y eventualmente aprendemos a llevar a cabo una red compleja de acciones para alimentarnos: anticipamos que nos va a dar hambre y nos movilizamos para comprar alimento en un lugar que nos gusta.

Segundo, respecto a la impresión de circularidad, el concepto de precariedad que es central en la propuesta enactiva parece implicar aquello que se busca mostrar: que un sistema sea precario significa que cualidades valiosas del sistema pueden decaer y hasta desaparecer si no son mantenidas activamente. Por ejemplo, en las respuestas a inquietudes sobre el libro *Sensorimotor life*, Di Paolo (2019) define exactamente de esa manera el concepto de precariedad. Pero el calificar tales propiedades como valiosas ya asume una valoración normativa de las propiedades y estados que deben ser preservados e importan en oposición a los que no, y por ende la normatividad no puede surgir del carácter precario de los sistemas autónomos, sino que esta caracterización ya la asume.

Sin embargo, no solo no es necesario definir la precariedad de esta forma circular, sino que en la literatura ya se han ofrecido definiciones no circulares. Como defendí en la sección 2.3.2 Precariedad, esta puede ser entendida como una relación entre el ritmo de decaimiento de propiedades del sistema y las capacidades de respuesta del organismo, de manera que el sistema es precario en la medida en que tiene las capacidades de contrarrestar su propio ritmo de decaimiento. Todos los sistemas que existen eventualmente decaerán y algunos de ellos tienen mecanismos dedicados o emergentes que pueden activarse en la temporalidad adecuada para revertir ese decaimiento. Ninguna de estas propiedades depende de valoraciones normativas y en esa medida, el concepto de precariedad no es circular.

#### 2.4.2. Teoría autopoietica clásica y la [ir]respetabilidad científica

Una de las críticas más minuciosas y tajantes a la naturalización de la normatividad enactiva viene de quienes han continuado el desarrollo de la teoría autopoietica original. Villalobos, en particular, construye de manera detallada y bien articulada una crítica al enactivismo, a saber, que no es un proyecto científicamente respetable porque mantiene la teleología y la normatividad en sus explicaciones de la cognición. Villalobos ofrece varias razones para tal acusación a lo largo de su carrera y revisirlas me permitirá motivar una de las preguntas abiertas más importantes de la concepción enactiva de cognición, a saber, la pregunta por su ontología. Argumentaré, sin embargo, que las críticas más provocadoras de Villalobos están insuficientemente fundamentadas.

La respetabilidad científica es central en las críticas de Villalobos, y de acuerdo con él y colegas, esta se caracteriza por al menos tres presupuestos ontológicos respecto a los fenómenos naturales:

[los fenómenos naturales] son lo que son y ocurren de la forma en que ocurren (1) sin tener ninguna meta, propósito o intención; (2) sin tener libertad alguna de elegir o seleccionar qué hacer y (3) sin ser gobernados por ninguna consideración normativa en términos de qué puede ser bueno o malo, benéfico o dañino. (Villalobos & Ward, 2015, p. 226)

La razonabilidad de estos supuestos se puede ver con un ejemplo: si quisiéramos explicar la actividad de un volcán y el movimiento de la lava, sería un error asumir que el volcán tiene el propósito de hacer erupción y que es bueno para él hacerlo, o que la lava tiene el propósito de destruir el pueblo vecino, que destruirlo es bueno para ella porque la ‘alimenta’ y que pudo haber elegido otro camino. En cambio, explicar científicamente estos fenómenos naturales requiere que apelemos a sus causas, que pueden ser eventos o propiedades materiales, no a supuestos deseos o propósitos. Por ejemplo, la composición química de los materiales de erupción y su temperatura permiten determinar la fluidez de la lava, la cantidad de ceniza, la intensidad de la explosión, si produce gases tóxicos o no, su velocidad y alcance territorial, etc. Así, una explicación científicamente respetable del comportamiento de los volcanes (1) no apela a elementos teleológicos o normativos, (2) es determinista, y además, (3) es indiferente a juicios normativos tanto respecto al sistema mismo –lava y volcán–, como respecto a otros sistemas, –el pueblo vecino–.

Villalobos y Ward argumentan que la vida es un fenómeno natural y por ende debe ser estudiado bajo los mismos presupuestos ontológicos bajo los que explicamos el comportamiento de volcanes, ríos, moléculas, etc. (Villalobos & Ward, 2015, p. 226). Sin embargo, no es claro por qué estos presupuestos, que funcionan para explicar el comportamiento de sistemas inertes, deben ser mantenidos también para sistemas vivos. En particular, no es claro por qué asumir que el comportamiento de un ser vivo debería ser indiferente a sus propios juicios normativos si precisamente esta es una característica que, de existir, podría explicar la diferencia cualitativa que observamos entre el comportamiento de los seres vivos y de los demás sistemas. En ese sentido, el tercer supuesto debe ser una pregunta abierta para la investigación de los seres vivos y no una suposición de respetabilidad para la biología. Asumirlo como definición es entonces una suposición *ad hoc*.

Por otro lado, no es tan evidente que el primer requerimiento sea apropiado para la biología; de hecho, en la práctica científica se siguen usando rutinariamente y a sabiendas conceptos normativos y teleológicos como ‘función’ y ‘propósito’ sin que esto haga que la biología sea menos respetable científicamente ni que deba deshacerse de los elementos teleológicos (Allen & Neal, 2020). Aun así, parece legítimo exigir, quizás de otras ciencias o de la filosofía, de una explicación de la cognición que no apele a elementos teleológicos y normativos como explicación última de un comportamiento para evitar la circularidad de la explicación. De manera que, si hay elementos teleológicos y normativos, estos deben ser explicados sin apelar a elementos de la misma naturaleza. Este es precisamente uno de los objetivos del



enactivismo: explicar cómo el comportamiento teleológico y normatividad surgen en la naturaleza sin apelar a otra cosa que su organización misma. Finalmente, respecto al segundo punto, gran parte del enactivismo asume el determinismo al usar sistemas dinámicos como modelo explicativo (esto es de hecho un problema en el enactivismo y lo trataré más adelante).

Las razones para argumentar que el enactivismo no es un proyecto científicamente bien fundamentado no terminan con esta estipulación de criterios de respetabilidad científica. También junto a Ward (2016), Villalobos argumenta que el proyecto enactivo hereda el antropomorfismo de la filosofía de Jonas y un proyecto antropomorfista no es un proyecto válido científicamente. Definen antropomorfismo de la siguiente manera:

En términos generales, el antropomorfismo es la práctica de atribuir rasgos humanos a entidades no humanas. Podemos, por ejemplo, explicar la forma en la que el río Nilo fluye hacia el mar diciendo que quiere alcanzar el mar o que tiene el propósito de hacerlo. (Villalobos & Ward, 2016, p. 803)

La atribución a otros seres vivos de categorías existenciales que son propias de la experiencia humana es, de acuerdo con Villalobos y Ward, una instancia de antropomorfismo. Estas categorías son, por ejemplo, libertad y necesidad, autonomía y dependencia, etc. En particular, como Villalobos y Ward señalan, la derivación de la teleología que defiende el enactivismo está basada en el argumento de Jonas quien toma como evidencia de la teleología en otros seres vivos nuestra propia experiencia teleológica, y por ende no está motivada por propiedades y procesos materiales independientes de nuestro propio punto de vista. En ese sentido, la estrategia de derivación de la teleología del enactivismo se basa en una atribución de características humanas a entidades no humanas y depende de nuestro punto de vista.

Sin embargo, la atribución de teleología a seres vivos no humanos no está solo motivada por la influencia de Jonas en el enactivismo y de nuestra aproximación cotidiana a los seres vivos; también es inevitable en la práctica biológica. No solo se asumen funciones y propósitos para explicar científicamente el comportamiento de organismos y especies, sino que la capacidad misma de delimitar el fenómeno a estudiar, es decir, los seres vivos, requiere que atribuyamos un punto de vista teleológico o una agencia mínima. Como señala Kant (1992 [1790], §75), nos es imposible si quiera familiarizarnos con la organización y el comportamiento de los seres vivos sin entenderlos como regidos por propósitos intrínsecos; esto en oposición a artefactos, cuyo comportamiento sigue propósitos extrínsecos, los de su diseñador, y en oposición a sistemas inertes cuyo comportamiento no sigue ningún propósito y puede ser explicado enteramente en términos de causas eficientes. Si bien es cierto que la atribución de teleología depende de nuestro punto de vista, como señala Villalobos, esta atribución no solo se basa en la filosofía antropomorfista de Jonas sino en la incapacidad práctica tanto en ciencia como fuera de ella de entender los seres vivos de otra manera.

Villalobos señala al menos otros dos problemas con la atribución de teleología a seres vivos no humanos. El primero es que la distinción misma entre organismo y ambiente es una atribución arbitraria del observador y no una distinción propia del sistema. Más específicamente, Villalobos y Ward sostienen:

Un observador puede, en aras de ciertos propósitos descriptivos, tratar la estimulación sensorial como punto de partida y la respuesta motriz como el punto final. Sin embargo, hacer esto sería reflejo de los intereses del observador más que de la dinámica del sistema (2015, p. 236).

Siguiendo la teoría autopoietica original, para Ward y Villalobos, la interacción entre organismo y ambiente, en particular la interacción a través de su sistema sensoriomotor es indistinguible de cualquier otro tipo de interacción entre los elementos de un sistema. Después de todo, tanto sistemas vivos como inertes, están sujetos a las mismas leyes causales. Si este es el caso, argumentan Villalobos y Ward, no hay razón para suponer que el organismo realmente se distingue de su ambiente y, por ende, tampoco hay razón para suponer que el organismo, a diferencia de su ambiente, puede comportarse de acuerdo con normas intrínsecas o establecer genuinamente una relación intencional con el ambiente de acuerdo con sus propósitos intrínsecos.

Este argumento niega directamente en sus premisas una de las características que Di Paolo, Buhrmann y Barandiaran (2017) identifican como condición necesaria de la agencia: la asimetría interaccional (ver p. 65). Arriba expliqué que el comportamiento adaptivo de los agentes es entendido como un proceso o interacción con el ambiente en el que el organismo impone los parámetros de interacción. De manera que, si bien este tipo de interacciones no violan las leyes causales que rigen a toda la materia, el curso específico que toman es impuesto por el organismo de acuerdo con sus propósitos propios. La asimetría interaccional es necesaria para que el organismo pueda regular efectivamente sus propias tendencias en función de su automantenimiento; así que, si un sistema es adaptivo es también capaz de imponer asimétricamente las condiciones de interacción de acuerdo con sus propios propósitos, preferencias o necesidades. Por ende, mostrar cómo puede darse en la naturaleza la asimetría interaccional y normativa daría razones a favor de que hay una distinción entre organismo y ambiente en la que el organismo puede establecer activamente la interacción con el ambiente.

Villalobos, sin embargo, ofrece también una crítica a la adaptividad y, si está en lo correcto, no habría razón para asumir que hay alguna diferencia genuina entre las interacciones adaptivas y las demás interacciones entre cualquier tipo de sistema. Villalobos (2013) argumenta que la adaptividad, en tanto capacidad de monitorear y regular las tendencias del organismo, implica que el organismo tiene la capacidad de anticipar lo que sucede con sus propias tendencias. El problema es que, para Villalobos, la capacidad de anticipar es inconcebible sin asumir implícita o explícitamente que el organismo forma representaciones internas que ocupan el lugar de los eventos futuros. Apelar a representaciones sería profundamente problemático para el enactivismo porque al hacerlo su propósito de ofrecer una propuesta radical en las ciencias cognitivas que prescindiera de las representaciones quedaría desvirtuado. Si Villalobos está en lo correcto, la adaptividad no sería un buen criterio de diferenciación entre organismos vivos por un lado, y su ambiente y otros sistemas por el otro; y entonces, no habría tampoco buenas razones para sostener que los seres vivos son y se comportan teleológicamente.

El problema con este último argumento de Villalobos es que hay de hecho formas de adaptividad que no requieren de representaciones. Por ejemplo, Agmon, Gates y Beer desarrollaron una protocélula que puede desarrollar respuestas adaptivas a perturbaciones en el ambiente sin necesidad de mecanismos dedicados ni de representaciones (2015). Estas protocélulas son modelos dinámicos que simulan células con sus respectivas membranas y procesos autocatalíticos internos con sus diferentes concentraciones moleculares, estas células decaen a cierta velocidad y están en un ambiente que provee agua, alimento y perturbaciones al sistema. El propósito de este modelo es simular la emergencia de la agencia y sus respuestas adaptivas. Sin entrar en detalles, encontraron lo siguiente:

No se encontraron mecanismos regulatorios explícitos durante esta transición adaptiva emergente. La adaptividad es el resultado de procesos perturbados, y se explica mejor por su dinámica espacio-temporal emergente. El análisis demuestra que cambios locales en las distribuciones químicas, como el engrosamiento de las membranas, tienen consecuencias en el comportamiento subsecuente. En términos de sistemas dinámicos, la adaptividad está determinada por la forma de la cuenta de atracción de la configuración, y cómo las interacciones con el ambiente mueven al sistema a través del espacio de fase. Una transición es adaptiva si lleva al individuo a una configuración con una cuenca más conveniente (Agmon et al., 2015, p. 223).

Este resultado implica que las protocélulas pueden desarrollar autónomamente respuestas adaptivas, es decir, respuestas para mantener su viabilidad en el ambiente que no requieren representaciones.

Sin embargo, vale la pena notar que la protocélula desarrollada por Agmon y colegas es un modelo computacional y esto implica que no es realmente un sistema adaptivo ni un agente mínimo a pesar de que logre simular rasgos importantes de los sistemas adaptivos y por lo tanto no es un sistema normativo. Hay dos razones para esto. Primero, los modelos computacionales no son sistemas precarios, aunque puedan simular condiciones precarias, la razón es que la precariedad es introducida como una variable del sistema y no es una propiedad de los procesos que constituyen al modelo. Tales procesos algorítmicos pueden continuar y repetirse indefinidamente, por lo tanto, no son precarios. El que estos procesos no sean precarios implica el sistema que constituyen no tiene la necesidad de mantener su propia existencia, por ende, que las respuestas del sistema no hagan ninguna diferencia práctica para este mismo y por ello, no tiene la capacidad de los sistemas mínimamente agentes de diferenciar entre cursos de acción.

Segundo, el comportamiento de los sistemas modelados computacionalmente es determinista, y esto implica que no tienen genuinamente la capacidad de seleccionar entre diferentes cursos de acción. Pero, como argumenté en el Capítulo 1, esta posibilidad es crucial para que el sistema se trate de un agente y para poder evaluar su comportamiento como correcto o incorrecto, mejor o peor, etc. Por ello, las respuestas aparentemente adaptivas de las protocélulas de Agmon y colegas, solo simulan los comportamientos exitosos, pero realmente no existía la posibilidad de fallar en esa simulación. El valor de esta simulación es que muestra que existe la posibilidad de que emerja una respuesta adaptivamente correcta porque,

si ocurre en la simulación, entonces es una posibilidad de selección para los sistemas realmente autónomos, precarios y adaptivos, aún si llegaran a no seleccionarla.

Vimos además que la adaptividad requiere de asimetría interaccional y que observarla socavaría el escepticismo de Villalobos de la incapacidad de distinguir entre organismo y ambiente. Di Paolo, Buhrmann y Barandiaran (2017) ofrecen una interpretación dinámica de la asimetría interaccional que permite identificarla en la interacción entre organismo y ambiente. A grandes rasgos la idea es la siguiente: cuando se describe la interacción entre organismo y ambiente como un sistema dinámico, es posible identificar cuándo el cambio de variables de un subsistema, es decir, el comportamiento del organismo o del ambiente, cambia los parámetros del comportamiento del sistema organismo-ambiente. Por ejemplo, cuando un oso se tira al agua para nadar, cambia los parámetros de interacción con el ambiente porque deja de estar regido por las leyes de movimiento en tierra firme y pasa a estar regido por leyes de movimiento en el agua. Si bien estas asimetrías interaccionales son observables y permiten distinguir organismo y ambiente, aún queda por resolver la pregunta de si la forma en que el organismo impone los parámetros de interacción puede ser normativa, es decir, seguir sus propios propósitos, o si esta interpretación teleológica no tiene más fundamento que nuestro punto de vista arbitrario.

Atribuir teleología y consecuentemente comportamiento normativo a los seres vivos implica también atribuirles una agencia mínima desde donde las posibilidades que ofrece el mundo pueden diferenciarse en términos de lo que les permite mantenerse con vida o lo que amenaza su existencia. Hay dos razones por las que esta atribución va más allá de una atribución arbitraria antropomorfista. La primera razón es que la atribución de teleología en la aproximación científica a la vida responde de manera concreta tanto a la convicción naturalista de que las capacidades cognitivas humanas en su forma más mínima son un fenómeno natural continuo con otras formas de vida, como al propósito de explicarlo sin eliminarlo de la naturaleza. La idea es tomar el fenómeno del comportamiento teleológico y normativo en los seres vivos y explicarlo en términos naturales, no asumirlo como un primitivo ontológico pero tampoco eliminarlo. Bajo esta aproximación, apelar al punto de vista propio como evidencia como lo hace Jonas, no es evidencia de que de hecho todos los seres vivos tengan el punto de vista que nosotros tenemos, sino que es evidencia última e irrefutable de que el fenómeno del punto de vista teleológico y normativo de hecho existe en la naturaleza. Así que cuando Villalobos y Ward se preguntan:

Si el lenguaje teleológico de propósito, agencia y valuación es normalmente inapropiado en descripciones científicas y explicaciones de fenómenos naturales, ¿deberíamos usarlo cuando estudiamos la subclase de fenómenos naturales compuesta por seres vivos? ¿Hay algo radicalmente distinto en los seres vivos que los separe del resto de fenómenos naturales y que necesite el uso de herramientas explicativas y descriptivas teleológicas? (Villalobos & Ward, 2015, p. 225)

Contrario a la intención retórica de estas preguntas, la respuesta es sí. Lo que es radicalmente distinto es tener algo que perder y la consistencia con propósitos del comportamiento de los seres vivos.

La segunda razón es que, si entendemos la agencia mínima, en particular sus propósitos propios, no como un estado interno que no está disponible a la constatación pública, sino como un patrón que expresa vitalidad y sensibilidad afectiva a lo que puede afectar la existencia propia, y que consistentemente contrarresta las amenazas a la existencia del organismo, no parece haber razón de entrada para negar que se trata de un organismo con propósitos propios. Como mostré en el capítulo 1, apelar al comportamiento y no a mecanismos o representaciones internas es una manera más pragmática de entender en qué consiste la normatividad; así que un comportamiento que es consistente con un propósito y que tiene la apariencia de la vida es una buena razón para hipotetizar que se trata del comportamiento de un agente teleológico a menos que se demuestre lo contrario. Como señala Taguchi (2016), la posición de Villalobos y Ward está basada en la suposición peligrosa de que la apariencia de la vida puede ser separada de lo que es *realmente* la vida. Esta suposición es peligrosa porque sin la apariencia de la vida, un organismo no es más que un conjunto de procesos inertes. Nuestra capacidad de reconocer la vida es tan fundamental y rudimentaria que se da por sentada en las ciencias naturales y por eso, de acuerdo con Taguchi, es necesario apelar a la fenomenología, o alternativamente, como mostré en el primer capítulo, es necesario apelar al piso compartido de sentido: nuestras prácticas de haceres compartidos.

En esta discusión sobre la respetabilidad científica he argumentado en contra de las exigencias de Villalobos y colegas para calificar una investigación como buena ciencia. Estos argumentos pueden aplicarse, *mutatis mutandis*, a concepciones de naturalismo en las que las ciencias naturales, en particular la física, se toman como estándar a emular o como ciencias fundamentales y a la que se reducen tanto ontológica como metodológicamente todos los demás fenómenos. Esta concepción de naturalismo no es a la que apunta el concepto de normatividad que propongo en esta tesis. Quisiera esbozar brevemente la noción de naturalización detrás del proyecto que defiende en esta tesis y la manera en que entiendo la tensión entre el determinismo de las ciencias, en particular de los sistemas dinámicos, y la noción de agencia que defendí en el Capítulo 1 y la que defiende el enactivismo.

Respecto al naturalismo, como la gran mayoría de filósofos contemporáneos, asumo que no existe nada más allá de la naturaleza, ninguna entidad supranatural (Papineau, 2021), y en ese sentido, asumo que los fenómenos normativos son fenómenos naturales. Sin embargo, no asumo que la naturaleza pueda ser descrita exhaustivamente por una ciencia ni por un conjunto selecto de ciencias, y esto implica que ni la metodología ni los primitivos explicativos de una ciencia como la física o las ciencias naturales son primarios ontológicamente ni más respetables que los de las demás ciencias. Las diferentes ciencias son en cambio formas distintas de describir, explicar e interactuar sistemáticamente con aspectos del mundo que, para usar la metáfora de Davidson (1985) conviven y dialogan entre sí, así como suburbios, colonias o barrios distintos conviven y dialogan entre sí —uno de estos suburbios, y uno relativamente nuevo, es el de las ciencias cognitivas. Bajo esta imagen de naturalismo, aunque haya diálogos y posibilidades de apropiación entre ciencias distintas, ninguna es reducible a la otra ni está en posición de prescribir los primitivos explicativos ni las metodologías correctas a otra. Así que la concepción de normatividad que propongo en

esta tesis será una concepción para las ciencias cognitivas, en particular, para las ciencias cognitivas corporizadas, que, aunque potencialmente pueda servir para responder preguntas en otras ciencias, no busca prescribir cómo debe ser entendida la normatividad en estas.

Hay sin embargo una tensión entre determinismo y selección que he señalado brevemente en varios pasajes hasta ahora. La tensión es la siguiente: por un lado, tenemos la suposición de la ciencia de que todos los eventos o estados de la naturaleza están unívocamente causados por un evento o estado anterior. Esto significa que un estado dado siempre causa una misma consecuencia y solo esa consecuencia, y tal consecuencia, a su vez, es causa de otro estado y solo de este estado, etc., de manera que la naturaleza se concibe como una cadena de estados que ocurren inevitablemente como consecuencia del estado anterior, y por lo tanto, todos los eventos están determinados por el estado inicial. A esta idea se le conoce como determinismo estructural y es uno de los axiomas de la teoría de sistemas dinámicos que es una de las herramientas más importantes de operacionalización en las aproximaciones corporizadas y enactivas.

Pero, por otro lado, argumenté en el Capítulo 1 que la idea de agencia mínima implica que el sistema sea capaz de seleccionar sus cursos de interacción, de lo contrario, no existiría la posibilidad del error y, por tanto, tampoco la posibilidad del comportamiento exitoso. Estas posibilidades son constitutivas de la cognición y de todos los fenómenos cognitivos en general, y específicamente, son parte del concepto de adaptividad en tanto que la regulación del sistema sería imposible si este no puede cambiar su trayectoria actual. Sin embargo, si el axioma del determinismo estructural es cierto, estas posibilidades realmente no existen porque no hay posibilidad de seleccionar entre más de un curso de acción. Esta tensión es aún más evidente en el uso de la teoría de sistemas dinámicos como herramienta explicativa de las aproximaciones corporizadas. Ante esta tensión parece que nos enfrentamos a un dilema en el que tenemos que desechar la idea de la posibilidad de selección y con ello cualquier forma de normatividad o bien tenemos que desechar el axioma del determinismo estructural y, no solo prescindir de los sistemas dinámicos como herramienta para las ciencias cognitivas, sino ir en contra de una de las convicciones más útiles y arraigadas en las ciencias naturales.

Aunque abordar este dilema a fondo y encontrar una solución está fuera del alcance de esta investigación, quisiera cerrar esta discusión sobre mi idea de naturalismo esbozando una posible solución. Puede argumentarse que es posible mantener el axioma del determinismo estructural en gran parte de la práctica científica como una herramienta metodológicamente conveniente y fructífera sin asumirlo como un principio que se sostiene para toda la naturaleza. En otras palabras, se pueden mantener los beneficios epistemológicos del determinismo sin asumir su universalidad. Esto es posible, quisiera proponer, si concebimos las leyes de las ciencias naturales como regularidades y no como principios universales. Dentro de las regularidades que identificamos en la naturaleza hay unas que se acercan asintóticamente a la determinación completa, por ejemplo, las leyes de la física clásica, mientras que otras son más irregulares o inciertas, como los mecanismos identificados en la biología. Las ciencias cognitivas describirían entonces un fenómeno regular que frecuentemente se distancia de la asíntota de la determinación. Bajo esa concepción de

naturaleza, podría argumentarse que los sistemas dinámicos tienen dos papeles en las ciencias cognitivas corporizadas: simulan una de las posibles trayectorias que pueden seleccionar los sistemas cognitivos o bien simulan las regularidades de los fenómenos cognitivos que más se acercan a un comportamiento determinista y que, en la práctica, son trayectorias sobre las que los agentes no tienen posibilidad de seleccionar una trayectoria distinta. Esto es consistente con la teoría de la irrupción en la que actualmente trabaja Tom Froese (2021) de acuerdo con la cual los momentos en los que los agentes seleccionan y su subjetividad hace una diferencia práctica en sus interacciones con el mundo pueden observarse como irrupciones, cortes o resets en las trayectorias deterministas de sus interacciones.

### 2.4.3. El problema de la motivación

Arriba señalé que en la propuesta enactiva no queda claro cómo un organismo autónomo y adaptivo está motivado para actuar de acuerdo con la normatividad situada que sus necesidades de automantenimiento imponen. Si bien hay una explicación clara de las circunstancias en que surge la normatividad y cómo se establece la diferencia normativa entre las posibilidades de comportamiento del agente, no queda claro cómo el organismo es movido a seleccionar una posibilidad sobre otra. Barrett (2019) muestra este problema a partir del concepto de ‘actividad motivada’. Él define actividad motivada como un comportamiento orientado a fines y en el que coinciden la fuerza sentida para actuar del agente y la causa del comportamiento. Esta definición es análoga a la definición de acción intencional en la que las razones para actuar del agente son también las causas de su acción. La actividad motivada puede entenderse como un comportamiento causado por causas finales en el que tales causas finales son también el propósito que experiencialmente mueve al agente a actuar. En particular, para el enactivismo, la causa final que mueve al agente a actuar es la autopreservación.

De acuerdo con Barrett, esta propuesta tiene dos problemas. Por un lado, no es claro cómo podría asumirse que la autopreservación es la causa final que motiva el comportamiento del agente sin asumir también que el organismo tiene una representación de sí mismo que busca preservar. Más concretamente, no es claro cómo la capacidad adaptiva de monitorear y regular las tendencias para mantener la identidad propia puede ocurrir sin una representación reflexiva de cuál es la identidad propia que hay que mantener. Por otro lado, buscar la autopreservación no permite explicar por qué los agentes están movidos a actuar y a explorar el ambiente cuando su identidad no está siendo amenazada, es decir, cuando no es necesario autopreservarse. Barbaras (2008) identifica el mismo problema desde una perspectiva fenomenológica: la vida no se caracteriza por la necesidad ni por el principio negativo de la autopreservación que implica quietud a menos que haya una necesidad que suplir o una amenaza que contrarrestar; la vida, en cambio, se caracteriza por el movimiento constante que manifiesta el deseo de lograrse, afirmarse y re-crearse constantemente a uno mismo. El deseo constante de autoafirmación se manifiesta en la interacción constante con el mundo porque para un organismo interactuar con el mundo es afirmarse ante lo que es distinto a sí mismo. Tanto Barbaras como Barrett señalan por vías distintas un mismo problema del enactivismo: asumir la autopreservación de una identidad estática como principio normativo no permite dar

cuenta del movimiento o interacción constantes, en particular, en ausencia de amenazas para la identidad del sistema.

Hay al menos dos posibilidades para responder a estas críticas, una, mantiene la autopreservación como propósito y causa final del comportamiento adaptivo de los organismos, y la otra propone modificar la concepción de causa final característica de los seres vivos. Para responder al problema de la motivación bajo la primera opción, un primer paso es concebir a los organismos que están constantemente en movimiento y exploran su ambiente como una red de diferentes sistemas autónomos y adaptivos –por ejemplo, el sistema inmunológico, el sistema nervioso o el sistema de hábitos sensoriomotores. Bajo esta concepción y siguiendo a Varela (1991), un organismo es un entramado de identidades sin un *yo* central que se afirman en diferentes circunstancias y respecto a diferentes aspectos del ambiente con los que potencialmente pueden interactuar. Cada uno de estos sistemas o identidades es precario y requiere que el organismo busque interactuar con el ambiente para automantenerse, así que el movimiento constante y la exploración están en efecto motivados por la necesidad de alguno de los diferentes sistemas autónomos que constituyen al organismo. En otras palabras, el movimiento y la exploración constante podrían explicarse en principio como interacciones motivadas a sacar a alguna de las identidades o subsistemas del organismo de su zona de precariedad. Esta posibilidad es consistente con la historia recibida del enactivismo, pero debe ser comprobada empíricamente no solo de la mano de las ciencias sociales, naturales y del comportamiento, sino también en la experiencia propia de los seres vivos. Sin embargo, es de entrada dudoso que todo nuestro comportamiento pueda estar motivado por la necesidad de evitar perecer y ser vivido fenoménicamente de esa manera y no como el deseo de autoafirmación. Esto nos lleva a la segunda opción.

Podríamos prescindir de la autopreservación como motivación y causa final que explica el comportamiento constante de los seres vivos, y buscar otro principio que logre explicar el problema de la motivación. Prescindir de la autopreservación podría dar un lugar central a la posibilidad del cambio en la concepción de seres vivos, y en ese sentido, concebirlos como sistemas cuya organización cambia en función de sus respuestas e interacciones adaptivas con el ambiente. Vimos más atrás, que la materialidad y su precariedad inherente son esenciales para diferenciar a los seres vivos de otros sistemas autónomos, y hacen necesarias las interacciones con el ambiente en la medida en que la autopreservación está amenazada. Si abandonamos la autopreservación como motivación y causa final, podríamos concebir la relación entre precariedad y autonomía de una manera más estrecha: los seres vivos son sistemas precarios autónomos cuya forma de autonomía se diferencia de los no vivos porque es una organización que cambia de manera situada sin dejar de existir. Mientras que otras organizaciones autónomas como los cristales no cambian constantemente, la organización de los seres vivos es inherentemente abierta al cambio.

Barrett (2019, 2020) argumenta que se debe abandonar la autopreservación como motivación más fundamental de la vida, y propone entender la adaptividad como actividad que busca maximizar propiedades termodinámicas, conocidas como propiedades termodinámicas *extremales*. Una propiedad extremal es la tendencia de un sistema a llevar al extremo una



propiedad, es decir, maximizar o minimizar tal propiedad tanto como sea posible en las circunstancias específicas en las que se encuentra. Barrett usa la maximización de entropía como ejemplo de una de estas propiedades que podría servir como principio motivador y causa final de los seres vivos. De acuerdo con él, la organización de los sistemas vivos es una estructura que permite disipar las tendencias entrópicas en su interior y a la vez maximizar la entropía de su entorno. Así que el sistema que forman organismo y ambiente en su interacción es un sistema que maximiza la entropía en virtud de la organización de uno de sus subsistemas, el organismo. Esto implica que el automantenimiento del organismo es una forma de mantener la tendencia a maximizar la entropía, en particular, es la forma característica en que todos los seres vivos lo hacen; pero implica también que el automantenimiento no es la motivación primaria ni la causa final del comportamiento del organismo, en cambio, es la maximización de la entropía. Esto permite explicar por qué un organismo podría estar movido a actuar sin que su identidad esté amenazada, y explicaría también por qué un organismo estaría movido a cambiar su organización misma: en ambos casos está movido a maximizar la entropía con las posibilidades que ofrece el ambiente. Para Barrett, una de las ventajas de esta propuesta es que no es necesario apelar a una representación o a un estado de cosas específico que anticipe y lo motive a actuar porque su motivación no es mantener su identidad específica sino mantener la tendencia a la entropía.

Sin embargo, la propuesta de Barrett tiene dos limitaciones fundamentales. La primera es que no es claro cómo podría entenderse el aspecto fenoménico de esta motivación, para usar las palabras de Barrett mismo, no es claro cómo la tendencia hacia la entropía podría *sentirse* como una fuerza motivacional. Esta limitación sin embargo no es directamente relevante para el proyecto de naturalización de la normatividad. La segunda limitación sí lo es y es la siguiente: no es claro cómo esta tendencia universal de todos los sistemas se convierte en una *motivación* en los seres vivos mientras que en los demás sistemas no lo es. En otras palabras, no es claro cómo una regularidad nomológica de la naturaleza, una tendencia inescapable de todos los sistemas que existen, puede convertirse en una tendencia normativa en el sentido en que motiva al sistema a actuar y su comportamiento podría lograr la maximización o errar al intentarlo. Como defendí en el capítulo 1, una motivación o un propósito que mueva al organismo a comportarse teleológicamente tiene un lugar real en la naturaleza si el comportamiento que el organismo selecciona de entre sus posibilidades hace la diferencia entre conseguir o no ese propósito. En la propuesta de Barrett, sin embargo, la motivación del organismo hacia la maximización de la entropía es completamente irrelevante porque no hay forma en que el organismo seleccione un comportamiento que no siga esta tendencia. Este problema, sin embargo, no es exclusivo de las propiedades termodinámicas extremas, sino que surgirá siempre que el comportamiento motivado se equipare a algún tipo de causa o tendencia determinista.

#### 2.4.4. El problema de la temporalidad y los criterios de corrección

Vimos arriba que los seres vivos son precarios porque, como cualquier sistema, inevitablemente tienden a desintegrarse. Mostré que la particularidad de los seres vivos es que el ritmo al que esto ocurre coincide con las capacidades de monitorear y regular sus propias

tendencias, es decir, con su capacidad adaptiva, de manera que logran mantener su organización en un equilibrio subóptimo en donde el óptimo es la disipación completa de su estructura, es decir, su muerte. En particular, para la aproximación enactiva, vimos que la sincronía temporal entre precariedad y procesos adaptivos es crucial para el automantenimiento del organismo (ver p. 71); por ejemplo, la protocélula de Agmon et al. (2015) es un modelo dinámico que es temporal en el sentido en que tiene un ritmo de decaimiento, y los procesos adaptivos que produce se sincronizan con tales ritmos y los contrarrestan. Los enfoques radicalmente corporizados, es decir, enfoques enactivos y ecológicos, coinciden en que la existencia misma de la cognición se desenvuelve temporalmente: como argumenta Kirchhoff (2015), los procesos cognitivos son constituidos dinámicamente y diacrónicamente. Di Paolo (2001) añade que los sistemas cognitivos son capaces de modular sus propias dinámicas a diferentes escalas temporales de manera que la novedad y las transiciones cualitativas de un estado a otro pueden ocurrir en el sistema, esto significa que la asimetría interaccional es una capacidad intrínsecamente temporal.

En el Capítulo 1 argumenté que, aunque la cognición y la constitución biológica u organizacional son ambos fenómenos normativos, tienen una diferencia fundamental: mientras que el propósito de la constitución biológica es mantener la identidad propia en tanto organismo individual, los propósitos propios de la cognición no dependen de su impacto en el organismo, sino de la posibilidad de interactuar con otros y de ser reconocido por otros. Mostré que esto se sigue de una concepción de sentido como actividad compartida cuyos criterios de corrección están disponibles para todos los que son capaces de llevar a cabo la actividad e interactuar con otros que también lo hacen. Bajo esta concepción, argumenté que los patrones normativos propios de la cognición están constituidos por haceres compartidos en una comunidad; y así mismo la identidad de los agentes cognitivos está parcialmente constituida por que son miembros de una comunidad de agentes que interactúan frecuentemente.

Sin embargo, la explicación enactiva de la co-emergencia de una agencia mínima solo da cuenta de una normatividad biológica mínima en la que agencia y criterios de normatividad coinciden en el propósito de mantener al organismo con vida, y en donde la escala temporal de las consecuencias prácticas evaluables y los criterios de normatividad se limitan al tiempo de vida del organismo y a sus consecuencias en el corto plazo. En contraste, los criterios de normatividad cognitiva y agencia no necesariamente coinciden, y los eventos y criterios normativos pueden estar extendidos en el tiempo. La separación temporal y de propósitos entre criterios de corrección y agencia de un agente es entonces una pregunta crucial. En el siguiente capítulo, examino casos de cognición ultrarrápida para mostrar que es necesario que los procesos cognitivos tengan una escala temporal extendida, y motivar precisamente la separación entre esas dos dimensiones de la normatividad, y en el capítulo 4 muestro con detalle por qué surge este problema en la propuesta enactiva y propongo una concepción de normatividad que permite resolverlo.

## 2. 5.Conclusión

En este capítulo mostré que el enactivismo puede entenderse como un proyecto prometedo de naturalización de la agencia y la normatividad. De acuerdo con esta aproximación, agencia y criterios de normatividad emergen simultáneamente. Por un lado, la agencia surge cuando un sistema se produce y mantiene a sí mismo y además puede seleccionar entre diferentes posibilidades de interacción con diferentes consecuencias para el mantenimiento de su propia identidad. Por otro lado, el éxito de sus interacciones y de sus respuestas a las perturbaciones está determinado por qué tan efectivamente logra contrarrestar lo que sea que amenace su identidad, de manera que la salud o la enfermedad y en últimas, la vida o la muerte que resultan de la selección adaptiva del organismo ofrecen un criterio de error que va más allá de las impresiones del organismo mismo. En un nivel fundamental de vitalidad básica, tener agencia es estar sujeto al tribunal de la experiencia, un tribunal que decide entre la salud y la enfermedad, entre la vida y la muerte del organismo.

Mostré que la definición enactiva de agente como sistema autónomo, precario y adaptivo permite establecer las circunstancias específicas en las que emerge el comportamiento normativo, esto es, las circunstancias en las que la trayectoria de comportamiento que el organismo selecciona hace una diferencia práctica para el organismo mismo respecto a otras trayectorias posibles. En la formulación más básica y generalizada, el comportamiento evaluable normativamente y los criterios para evaluarlo surgen cuando dos condiciones se cumplen simultáneamente: primero, la identidad del organismo está en riesgo y, segundo, los recursos internos y los recursos disponibles en el ambiente abren la posibilidad de que el organismo pueda seleccionar al menos una trayectoria de interacción que revierta tal riesgo. Esta explicación del surgimiento de la normatividad en circunstancias concretas muestra que la normatividad no se determina primariamente como una norma o principio universal que posteriormente permite evaluar cada interacción específica. Por el contrario, la normatividad surge en las condiciones específicas en las que seleccionar una u otra trayectoria de interacción hace la diferencia entre la vida y la muerte del organismo.

## Capítulo 3

### Extensión temporal y criterios de corrección

En los capítulos I y II mostré que los fenómenos normativos implican, por un lado, una agencia mínima que es capaz de provocar cambios en el mundo en una escala temporal y de manera que hagan una diferencia en sus propias trayectorias o condiciones individuales, y, por otro lado, criterios normativos que permitan evaluar objetivamente tales comportamientos individuales de acuerdo con una norma. En el Capítulo 2 mostré una explicación naturalista de la co-emergencia de una agencia mínima y criterios de corrección que coinciden en criaturas vivas básicas; en tales casos la consecuencia práctica que hace una diferencia para ellos y que es observable externamente es la vida o la muerte, la salud o la enfermedad del individuo. Al final del capítulo mostré que esta propuesta tiene limitaciones aún, una de ellas es que no da cuenta del desacoplamiento entre criterios de corrección y agencia.

En este capítulo examino casos de cognición ultrarrápida para mostrar que es necesario que los procesos cognitivos tengan una escala temporal extendida, y motivar la separación entre esas dos dimensiones de la normatividad. Casi toda la sección 3. 1 consiste en la traducción de este argumento que desarrollé junto con Tom Froese y fue publicado bajo el título *On the spatiotemporal extensiveness of sense-making: ultrafast cognition and the historicity of normativity* (Mojica & Froese, 2019). Para abordar este reto de la extensión temporal de la normatividad cognitiva, propongo apelar a la propuesta teleosemántica de Millikan que explicaré en la sección 3. 2. Defiendo una interpretación de la teleosemántica de Millikan en la que la historia de selección natural establece los criterios normativos para las interacciones cognitivas entre agente y ambiente. Sin embargo, a diferencia de lo que defendí en los primeros dos capítulos, Millikan se compromete con una concepción representacional de la mente. Discutiré dos críticas que se han hecho en ese sentido, a saber, la de Fodor, quien argumenta que la selección natural no establece criterios normativos con la especificidad y propiedades semánticas necesarias, y la de Hutto y Myin, que defienden que añadir propiedades representacionales no cumple realmente un papel explicativo. Concluyo en la sección 3.2.2, ofreciendo una motivación para llevar la propuesta de Millikan a dar un giro hacia la acción y argumento que esto es posible si (i) se abandona el requerimiento del isomorfismo y (ii) se incorpora al organismo individual como fuente de normatividad también.

#### 3. 1. Adaptividad, temporalidad y cognición ultrarrápida

Como mostré en el capítulo anterior, el enactivismo entiende a los agentes como sistemas autónomos y adaptivos. La autonomía implica que el sistema se produce continuamente a sí mismo, y su autoproducción establece el fundamento de la relación intencional que los agentes entablan con el mundo (Thompson, 2007): el mundo tiene significado en términos de lo que permite o lo que amenaza el propósito del agente de mantener su identidad (Weber & Varela, 2002). Mientras que la auto-producción explica la normatividad básica de la interacción agente-ambiente, la adaptividad explica cómo los agentes pueden llevar a la

práctica tal perspectiva de forma adecuada y diferenciada. Siguiendo la distinción que introduce en el capítulo 1, esto significa que la auto-producción establece los criterios normativos, mientras que la adaptividad captura la agencia cuyo comportamiento puede ser evaluado bajo esos criterios y que actúa de acuerdo con un propósito propio. De acuerdo con Di Paolo:

Un análisis cuidadoso del hacer-sentido muestra que las diferentes propiedades de la adaptividad (automonitoreo, control de la regulación interna y control de los cambios externos) están implicados al asumir que los organismos tienen una perspectiva con sentido sobre el mundo, de ahí que esta propiedad no sea simplemente una adición, sino que es esencial (junto con la autopoiesis) para naturalizar el hacer-sentido (2005, p. 430)

Más específicamente, la idea es que la adaptividad refleja la habilidad del organismo –necesaria para hacer-sentido– de evaluar las necesidades y los medios que se despliegan para tal propósito (Di Paolo, 2005, p. 445). Crucialmente, la adaptividad implica que monitorear y regular son momentos conceptualmente constitutivos del hacer-sentido.

Existe la preocupación de que apelar a procesos de evaluación o monitoreo y control o regulación de la actividad de un organismo establezca parámetros demasiado altos para formas básicas de hacer-sentido; o bien, que esta apelación atribuya erróneamente capacidades psicológicas para explicar el comportamiento de entidades que no tienen tales capacidades. Tras estas preocupaciones está la tendencia a interpretar los procesos adaptivos de manera centralizada y asumir que algo dentro del agente monitorea y controla el comportamiento exterior. Esto haría al enfoque enactivo vulnerable a los mismos cargos de homuncularismo que imputa a las teorías computacionales de la mente. En otras palabras, una cosa es afirmar incontrovertiblemente que en ocasiones monitoreamos y controlamos nuestro comportamiento, por ejemplo, mientras llevamos una taza llena de café caliente a la oficina tratando de no quemarnos y de no derramar nada en el suelo. Pero otra cosa es afirmar que procesos similares funcionalmente ocurren dentro de nosotros mientras llevamos a cabo tal acción, o incluso afirmar que estos procesos son condiciones esenciales para todas las instancias en que se hace sentido, incluida la quimiotaxis bacteriana para tomar un ejemplo popular.

De hecho, hay casos de comportamiento adaptivo que no involucran mecanismos dedicados a monitorear y regular. El modelo de dinámicas de concentración molecular en un espacio-tiempo continuo de Agmon, Gates y Beer (2015) ilustra este punto. En este modelo emerge una protocélula que se auto-produce y que es capaz de adaptarse a ciertas perturbaciones al reorganizar espontáneamente su estructura interna. Un análisis de estas transiciones adaptivas mínimas revela que emergen de procesos distribuidos en el espacio. Así que, puesto que monitorear y regular pueden emerger de procesos distribuidos en el espacio, la adaptividad enactiva no tiene que ser entendida en términos homunculares.

Por otro lado, desde los trabajos tempranos del enactivismo, se ha reconocido que la historia de las interacciones entre agente y ambiente juega un papel importante en el hacer-sentido. En palabras de Varela et al.: la cognición es el enactuar el mundo y una mente sobre la base de la

historia de las varias acciones que un ser en el mundo lleva a cabo (1991, p. 9). Inspirado en la teoría autopoietica, el enactivismo ha entendido el papel de la historia de manera causal. Esta concepción de la historia es capturada por el “axioma del determinismo estructural” que constriñe la extensión temporal de los procesos adaptivos de manera que los cambios de estado en un sistema siempre operan en el presente como resultado de su estructura actual y no están determinados por agentes externos o condiciones contextuales (Di Paolo, 2005, p. 434). Las interacciones pasadas son parte de las condiciones contextuales; así que, aunque estas determinan causalmente la estructura actual del sistema, no determinan la forma en que las interacciones actuales agente-ambiente ocurren.

Si el enfoque enactivo va por el camino correcto, podemos preguntarnos cómo exactamente el axioma del determinismo estructural se instancia en procesos adaptivos que se desarrollan en el tiempo. Los estudios psicológicos sobre cognición ultrarrápida permiten refinar esta pregunta y mostrar su pertinencia. En estos estudios se observa comportamiento con significado que ocurre tan rápido que no da tiempo suficiente para que el procesamiento de información en el cerebro juegue un papel en el control del desarrollo del comportamiento. El consenso es que el agente está listo para ser detonado por las circunstancias de manera que su respuesta converge automáticamente en el atractor comportamental correcto. El monitoreo y la regulación son propiedades que emergen de la interacción completa. Pero si esta explicación está en lo correcto, no queda claro cómo, dentro de la aproximación adaptiva, monitorear y regular pueden resultar de la estructura actual del sistema y no ser constituido normativamente por las condiciones históricas de la tarea completa; en otras palabras, no queda claro cómo el axioma del determinismo estructural puede sostenerse en la cognición ultrarrápida y, en ese sentido, la cognición ultrarrápida plantea un reto para el enactivismo. Veamos entonces, primero, las herramientas conceptuales que ha construido el enactivismo para entender las interacciones sensoriomotoras agente-ambiente, y segundo, en qué consiste el fenómeno de cognición ultrarrápida y la explicación enactiva de estos.

### 3.1.1. Esquemas y redes sensoriomotoras

En el capítulo anterior introduje los conceptos de autonomía y adaptividad, y expliqué en qué sentido son cruciales para la naturalización de la mente y la normatividad. Los sistemas autónomos son sistemas que (i) se auto-diferencian, es decir, que producen ellos mismos los procesos que los componen y los límites que lo diferencian del ambiente, y (ii) se auto-mantienen, es decir, que se abren a interactuar con el ambiente para obtener los recursos energéticos y materiales para continuar existiendo. De acuerdo con Di Paolo (2018), estas dos propiedades de los sistemas autónomos llevan a dos tendencias contrarias: cerrarse a la interacción con el ambiente para diferenciarse de este y abrirse para obtener los recursos necesarios para mantenerse. La tensión entre estas tendencias se resuelve por la capacidad adaptiva de los agentes, lo que significa que estos monitorean y regulan sus interacciones para abrirse a lo que necesitan para mantenerse o cerrarse a lo que amenace o sea irrelevante para su auto-mantenimiento. La normatividad básica surge precisamente en esta posibilidad de diferenciación que emerge de los sistemas autónomos y adaptivos.

Las definiciones de autonomía y adaptividad han sido ajustadas para dar cuenta de la complejidad de organismos con un sistema nervioso. Di Paolo, Buhrmann y Barandiaran (2017) proponen entenderlos como propiedades de la interacción entre un animal y su ambiente, y esta interacción es, a su vez, entendida en términos de esquemas sensoriomotores. Su propuesta captura las varias capas de normatividad que están en juego en la interacción entre humanos y su ambiente, y estas servirán para entender más adelante cómo las respuestas ultrarrápidas son posibles y las varias capas de normatividad que implican.

El concepto de esquema sensoriomotor describe los patrones de percepción-acción que surgen en la interacción entre un animal y su ambiente. Percepción y acción son co-dependientes porque, por un lado, la acción es perceptualmente guiada y, por otro, la percepción cambia y es guiada por la acción. Estas son entendidas como regularidades que ocurren en bucles sensoriomotores dinámicos, cerrados y auto-organizados. El concepto de esquema sensoriomotor de Di Paolo y colegas designa precisamente estos bucles e involucra todas las estructuras del cuerpo y el ambiente que sustentan la percepción y el hacer algo en el ambiente. Por ejemplo, bajo esta definición, agarrar una taza es un esquema sensoriomotor que implica la coordinación cuerpo-ambiente.

Los esquemas sensoriomotores se auto-organizan en redes autónomas y adaptivas que permiten al sistema cognitivo llevar a cabo ciertas actividades al transitar de un esquema a otro. Estas redes son sucesiones de esquemas organizados de manera que llevar a cabo un esquema permite al agente transitar al siguiente y llevarlo a cabo —ej., agarrar una taza permite al agente dar un sorbo a su bebida—, y toda la sucesión permite al agente llevar a cabo una actividad cabo —ej., tomar café con un amigo.

De acuerdo con la propuesta de Di Paolo y colegas, la normatividad bajo la que los esquemas se organizan en una red es impuesta por la organización interna misma de la red y por la situación particular en la que es llevada a cabo. Por un lado, la organización interna de una red está determinada por los objetivos propios de la actividad que la red le permite al agente llevar a cabo. Al beber café, la organización interna de la serie de esquemas de agarrar la taza, tomar un sorbo y poner la taza de vuelta sobre la mesa está determinada normativamente por si permiten o no al agente tomar café realmente, continuar interactuando con el amigo, no quemarse ni quemar a nadie más con el café caliente, etc. Por otro lado, la red sensoriomotora debe ser sensible al contexto particular en el que condiciones tanto ambientales como organísmicas juegan un papel. Por ejemplo, un agente puede agarrar un vaso porque alguien está a punto servirle agua y porque su cuerpo está deshidratado; llevar a cabo esta red de esquemas sensoriomotores no solo evita la descortesía de no recibir la bebida adecuadamente en esa situación, sino que permite eventualmente al agente llevar el agua a su boca y beber. En el próximo capítulo discutiré con más detalle cómo se constituyen estas normas. Por ahora, en pro del argumento, quisiera resaltar que seguir esta normatividad requiere que se monitoree y regule constantemente la manera en la que la red de esquemas se lleva a cabo en tales circunstancias: si sigue las normas internas de la actividad o no, y si va encaminada a conseguir los objetivos propios de la actividad en tales circunstancias. El monitoreo y la

regulación requeridas al llevar a cabo redes de esquemas sensoriomotores captura el requerimiento de regulación adaptiva de la agencia cognitiva.

### 3.1.2. Cognición ultrarrápida

Consideremos dos ejemplos de cognición ultrarrápida: tareas de perturbación y tareas go/no-go. En ambas tareas, las respuestas cognitivas ocurren más rápido que el tiempo fisiológicamente requerido para transmitir información al cerebro para ser procesada allí y posteriormente producir una respuesta controlada como resultado del procesamiento. A pesar de la rapidez, las respuestas de los sujetos son confiables, es decir, son correctas en una tasa significativa, y los sujetos se sienten agentes de estas respuestas cognitivas.

En las tareas go/no-go, a los participantes se les pide responder cuando un objetivo está presente y no responder cuando está ausente. Uno de los ejemplos más famosos es el experimento conducido por Macé et al. (2009) en el que se presentan a los participantes fotografías de escenas naturales durante 26 ms cada una. En una de estas tareas, se les pidió a los participantes apretar un botón continuamente y solo soltarlo cuando vieran un animal en la escena. Después de 1 segundo de mantener apretado el botón se consideraba que la respuesta era negativa, es decir, un no-go. La respuesta más rápida registrada fue de 260 ms para escenas en las que el animal presente era un perro, y el tiempo promedio de respuestas confiables fue de 390 ms. Otros experimentos han revelado que el significado general de una foto puede ser percibido en menos de 200 ms (Greene & Oliva, 2009; Grill-Spector & Kanwisher, 2005; Wallot & Van Orden, 2012, p. 4). Esto es significativamente más rápido que la escala de 1 s asociada a actos cognitivos completos, incluido el comportamiento perceptuo-motor, que son integradas a través de sincronía neuronal a gran escala en el cerebro (Varela, 1999).

En las tareas de perturbación, el comportamiento de los sujetos es perturbado para probar cómo estas pequeñas perturbaciones pueden ser superadas. Movimientos sacádicos (Altmann, 2011; Kirchner & Thorpe, 2006), el movimiento al caminar (Weerdesteyn et al., 2004), movimiento de brazos para indicar el reconocimiento de palabras (Moreno et al., 2011) y otro tipo de comportamientos (Kozhevnikov & Chistovich, 1965; Porter Jr & Castellanos, 1980) han sido examinados de esta manera. Una de las tareas de perturbación más conocidas son los experimentos de la perturbación de la quijada conducidos por Kelso, Tuller, Vatikiotis-Bateson y Fowler (1984). La mandíbula de los participantes fue ligeramente jalada hacia abajo mientras decían /baez/ o /baeb/ y los cambios en su lengua y labio superior e inferior fueron registradas. Los participantes no sabían cuándo el tirón iba a ocurrir. El experimento reveló que la compensación para cada palabra fue distinta: para /baez/, la compensación ocurrió primariamente en la lengua, mientras que para /baep/, ocurrió primariamente del labio superior. Ambas compensaciones ocurrieron entre 5 y 10 milisegundos, y las palabras fueron pronunciadas exitosamente.

Estos tiempos de respuesta presentan un reto para la explicación tradicional de respuestas cognitivas en términos de procesamiento de información. De acuerdo con Wallot and Van



Orden (2012), el tiempo restante para el procesamiento de información solo permite que demasiado pocas neuronas se activen, de manera que estas generarían una representación muy frágil. Además, incluso si la información fuera codificada por el orden de la activación neuronal y no su frecuencia, la complejidad de las fotografías de escenas naturales hace altamente improbable que la adecuación de las respuestas obtenidas pudiera ser explicada por “neuronas individuales conectadas en procesos *feedforward*” (2012, p. 6). Por lo tanto, estas dificultades exigen una aproximación diferente, una que apele a la organización dinámica del sistema sensoriomotor.

Tal aproximación ha sido propuesta por Kelso et al. (1984) y otros, y ha sido adoptada por Di Paolo et al. (2017). De acuerdo con Kelso y colegas, las diferencias observadas en la compensación para cada expresión apoyan la hipótesis de que son producidas por estructuras dirigidas a la producción estable de una acción significativa particular, y no por patrones de respuesta estandarizados como se seguiría de respuestas programadas o circuitos de input-output fijos (Kelso et al., 1984, p. 813). Las estructuras coordinativas coinciden con la noción de esquema sensoriomotor en la propuesta de Di Paolo et al. (2017), quienes coinciden con Kelso y colegas, e interpretan la rápida compensación como una “reacción adaptiva rápida a perturbaciones externas” (2017, p. 202).

De manera análoga, las respuestas rápidas observadas en las tareas go/no-go discutidas arriba son interpretadas por Di Paolo et al. (2017) como la ejecución de un esquema sensoriomotor. En ambos casos, un grupo de esquemas sensoriomotores apropiados para la tarea en el contexto es favorecido. Esto prepara al agente en un estado anticipativo lleno de inclinaciones igualmente válidas para responder, de las que todas comparten un fin común (2017, p. 201). En el caso del experimento de la perturbación de la mandíbula, el fin es producir las locuciones, y en el caso del experimento go/no-go, es soltar el botón cuando el objetivo esté presente en la imagen. Estar en tal estado anticipativo, de acuerdo con Di Paolo y colegas, es el aspecto intencional de la acción emergente. Un pequeño cambio en las circunstancias (internas o externas) puede entonces ser suficiente para romper la simetría del sistema y seleccionar una de las varias acciones posibles [i.e., esquemas sensoriomotores]. Esto es experimentado como iniciación de la acción (2017, p. 201). En ambos casos, el agente está preparado para producir un comportamiento significativo en las circunstancias particulares.

### 3.1.3. Determinismo estructural y anticipación: una tensión hasta ahora no reconocida

Los ejemplos de cognición ultrarrápida muestran que podemos tener respuestas con significado sin la necesidad de deliberación de alto nivel lo que favorece una explicación corporizada del hacer-sentido. Si aceptamos el axioma de determinismo estructural, las respuestas significativas, incluidas sus propiedades adaptivas y normativas, deben estar constituidas por la estructura actual del sistema y no por su contexto histórico. Esto debe aplicar a la cognición ultrarrápida también. Sin embargo, no es claro como monitorear y regular juegan un papel en lo que parece ser una respuesta meramente reflexiva que ocurre en menos de un segundo. Esto abre entonces la pregunta de cómo entender el concepto de

adaptividad en estas escalas cortas de tiempo. Como veremos, la explicación enactiva aceptada de la cognición ultrarrápida se sostiene en la historia de las interacciones agente-ambiente para explicar la dimensión normativa de estas respuestas rápidas, lo que revela una tensión entre la concepción enactiva clásica de la historia como un elemento causal y su papel aparentemente constitutivo en la cognición ultrarrápida.

Comprometerse con la ejecución de un esquema sensoriomotor no requiere que haya un procesamiento centralizado de información o deliberación explícita. Puesto que el agente ya está preparado en un estado anticipativo crítico apropiado para el contexto presente (2017, p. 202), el esquema seleccionado será una respuesta apropiada y significativa en las circunstancias presentes, estará dentro de una red de esquemas sensoriomotores que permiten al agente llevar a cabo una acción, y será experimentado de esa forma por el agente. Dentro de este marco dinámico, la cognición ultrarrápida puede ser explicada: incluso sin feedback, estas respuestas adaptivas son posibles, porque los agentes ya estaban preparados para dar una respuesta significativa y apropiada en tales circunstancias. En ese caso, el estímulo actúa como un detonante y no tanto como un contenedor de información que necesita ser procesada y decodificada (2017, p. 225). Estos estados anticipativos son interpretados no como un modelo predictivo, sino como una sincronización corporizada a la situación previamente establecida por la integración del feedback, los retrasos y las sinergias en su dinámica total (2017, p. 225).

La historia de interacciones agente-ambiente ha transformado al agente de manera que está preparado para que el ambiente detone la aparición de un atractor en el espacio de estados sensoriomotores posible en el que el agente simplemente caerá y de esa manera realizará la acción apropiada. De este modo, el feedback no es requerido durante las respuestas ultrarrápidas porque una respuesta significativa ya fue seleccionada en el estado anticipativo. Adicionalmente, la mayoría de las perturbaciones a la trayectoria de estados en esta respuesta ultrarrápida no serán suficientes para sacar al sistema sensoriomotor del atractor en el que se encuentra, lo que significa que compensará las desviaciones espontáneamente sin necesidad de feedback o procesamiento adicional. De acuerdo con Stepp y Turvey, la anticipación fuerte es puramente reactiva en algún nivel de análisis (2010, p. 156). Sin embargo, no es un reflejo sin sentido.

A pesar de la ausencia de feedback durante la respuesta adaptiva, los participantes hicieron sentido en sus respuestas ultrarrápidas. Sus respuestas ultrarrápidas siguen varios tipos de normatividad. Primero, los agentes deben distinguir correctamente si hay, por ejemplo, un perro en la imagen, lo que implica que tienen que hacer sentido de lo que ven. Segundo, deben adaptarse a las normas establecidas por el contexto experimental, e.g., deben soltar el botón, no gritar “¡perro!” o “¡guaf guaf!” ni saltar ni actuar de ninguna otra manera al ver un perro en la imagen. Tercero, los participantes están guiados por factores socioculturales que los hacen participar en el experimento, por ejemplo, al estar motivados por una recompensa monetaria, o por participar en un experimento científico, o por probar experimentalmente sus propias hipótesis en sí mismos.

La teoría enactiva del hacer-sentido predice que estas respuestas significativas son monitoreadas y reguladas de acuerdo con tales normas. Más específicamente, el que el agente pueda seleccionar o no su curso de acción, en este caso, mantener el botón oprimido o soltarlo, es resultado de sus procesos adaptivos. Por un lado, durante el curso de la interacción está monitoreando las trayectorias que esta toma para, ej., mantener la concentración, tener la mirada fija en la pantalla, mantener los músculos de la mano y el brazo contraídos, y entonces poder llevar a cabo la interacción exitosamente. Monitorear, en este caso, puede consistir en mantener la trayectoria en la que está el sistema, es decir, mantenerlo en el atractor en el que se encuentra. Por otro lado, el sistema regula su interacción cuando empieza a desviarse del curso de acción que le permite llevar a cabo la acción exitosamente responde con mecanismos para regresar a este, ej., si deja de oprimir el botón con la fuerza necesaria porque sus músculos se están cansando, reacomoda su posición para poder seguir apretando el botón. Ambos mecanismos de monitoreo y regulación permiten que el agente salga de su estado actual y suelte el botón cuando aparece la imagen a la que se supone que debe responder. En ese caso, el sistema es detonado para salir del estado actual y pasa al punto de equilibrio para el que estaba preparado el sistema sin que los mecanismos de regulación lo impidan porque en este caso se trata de la respuesta correcta.

Sin embargo, si la adaptividad está sujeta al axioma de determinismo estructural, el momento mismo de caer en un punto de equilibrio debe ser determinado por la estructura presente del sistema y no por su contexto histórico. El tipo de determinación en juego debe involucrar monitoreo y regulación para contar como una respuesta que hace-sentido. Sin embargo, no es claro cómo estos pueden emerger o ser operacionalizados en tiempos de respuesta tan rápidos, porque no hay suficiente tiempo para que el feedback o ningún otro mecanismo dedicado al monitoreo y la regulación jueguen algún papel. En esa medida, la cognición ultrarrápida revela una tensión entre la formulación clásica de la adaptividad que, al adoptar el axioma de determinismo estructural, entiende la historia como una cadena causal y el concepto operativo actual de adaptividad que enfatiza el papel constitutivo del contexto histórico.

### 3.1.4. Posibles soluciones

Para abordar este reto, puede sugerirse que los procesos adaptivos emergen de las interacciones organismo-ambiente en un lapso que incluye condiciones tanto presentes como pasadas. Por un lado, las intenciones y las sensibilidades que surgen en la historia de interacciones de un organismo establecen estados anticipatorios y ayudan a regular normativamente aspectos del involucramiento habilidoso con el mundo. Por otro lado, el contexto histórico de la interacción impone criterios de corrección que le dan sentido a la respuesta ultrarrápida. Por ello, sugerimos que el monitoreo y la regulación son propiedades emergentes de procesos que se extienden no solo en el espacio sino también en el tiempo. Esta idea es explícitamente rechazada por el axioma de determinismo estructural de la teoría autopoietica clásica. Sin embargo, el concepto de adaptividad ha evolucionado desde su primera formulación, de manera que ahora acepta el papel normativo de la historia de interacciones del organismo y con ello rechaza implícitamente el axioma de determinismo estructural. Al hacer este cambio explícito, se vuelve explícita también la necesidad del

enactivismo de considerar la mente como extensa temporalmente y no solo extensa espacialmente. Para llegar a esta conclusión, conviene considerar las diferentes soluciones posibles a este reto.

1. *Primera posibilidad: Las respuestas ultrarrápidas son hacer-sentido en las que hay monitoreo y regulación ultrarrápidos*

Podría argumentarse que el monitoreo y la regulación ocurren en la escala de los milisegundos mientras se desarrollan las respuestas ultrarrápidas y que están basados en los procesos que no requieren feedback de largo alcance. Es posible que sea feedback local lo que soporte estas respuestas y que presumiblemente involucre la médula espinal o bucles locales en el cuerpo. Este podría ser el caso con señales de control que pueden viajar en todas las direcciones y a través de niveles de organización sin ser canalizados primero a través de las capas ordenadas jerárquicamente (Flament-Fultot, 2016, p. 163) que podría ser suficiente para monitoreo y regulación ultrarrápidas. Esta posibilidad necesitaría ser desarrollada en más detalle, pero podría ser una manera de asegurar una base constitutiva de normatividad en el momento presente, y mantendría la consistencia entre la cognición ultrarrápida y el axioma del determinismo estructural. Bajo esta posibilidad, el papel de las interacciones agente-ambiente precedentes sería puramente causal, y tendrían el efecto de llevar al agente a estar preparado para responder adaptivamente de manera ultrarrápida.

Sin embargo, hay dos consideraciones que apuntan en la dirección contraria, es decir, a que la cognición ultrarrápida está potencialmente constituida por una escala temporal más amplia: i) La explicación estándar de cognición ultrarrápida en términos sensoriomotores es altamente dependiente del contexto histórico, y ii) tal extensión temporal de la explicación es más consistente con la necesidad de apelar a las interacciones precedentes para poder entender lo que el agente está haciendo en su respuesta ultrarrápida, ej., que decidió participar de un experimento psicológico que involucra ciertas expectativas. Aunque estas consideraciones no sean razones suficientes para descartar por completo esta primera posibilidad, sí apuntan a una alternativa prometedora.

2. *Segunda posibilidad: las respuestas ultrarrápidas no son actos de hacer-sentido, son parte de un acto de hacer-sentido que se desarrolla en una escala temporal más larga.*

Una persona que lleva a cabo una tarea go/no-go ciertamente está monitoreando y regulando su actividad durante todo el experimento, por ejemplo, al tratar de seguir las instrucciones, de hacer bien la tarea, de evitar distraerse, etc. Así que, quizás, estas respuestas ultrarrápidas adaptivas y normativas derivan su sentido de estar situadas en estos contextos y escalas temporales más amplias. Esta segunda posibilidad es consistente con la concepción de la dimensión temporal de la adaptividad que defiende Di Paolo:

Un evento adaptivo (o acto) puede estar formado por la concatenación y coordinación paralela de muchos otros eventos regulatorios, pero habrá un punto por debajo del cual no sea posible ninguna descomposición adicional sin perder la estructura de tiempo

del acto. En ese punto lo que queda son procesos crudos. Consecuentemente, hay una granularidad temporal mínima en la adaptividad (Di Paolo, 2005, p. 444)

Las respuestas ultrarrápidas pueden estar bajo el umbral de descomposición puesto que no dan tiempo a que eventos regulatorios paralelos ocurran y se coordinen. En cambio, estas respuestas pueden ser precisamente uno de los eventos regulatorios o procesos crudos, y su significado, si es que tienen alguno, puede ser solo derivativo de su pertenencia a un acto de hacer-sentido que se despliega en una escala temporal más amplia. Por ende, una respuesta ultrarrápida puede ser concebida como un elemento de la estructura temporal del acto de hacer-sentido en su totalidad, cuya duración se extiende más allá de la respuesta ultrarrápida.

Di Paolo caracteriza la estructura temporal mínima de un acto de hacer-sentido de la siguiente manera:

Toda regulación adaptiva es un evento estructurado con fases claramente definidas de inicio (la sensación de una tendencia negativa), aceleración (la activación del mecanismo adaptivo), consumación (la anulación de la tendencia negativa) y cadencia (la desactivación de la respuesta adaptiva) (Di Paolo, 2005, p. 444)

Las repuestas ultrarrápidas ocurren después de que el agente esté en un estado anticipatorio, lo que significa que está dispuesto a responder. Cuál de todos los esquemas sensoriomotores posibles será llevado a cabo sigue siendo indeterminado en el estado anticipatorio, pero la simetría entre tales posibilidades se rompe cuando hay un pequeño cambio en las circunstancias. El estado anticipatorio puede ser entendido dentro de la estructura temporal propuesta por Di Paolo como la activación del mecanismo adaptivo, esto es, la fase de aceleración, porque en ese estado el agente se vuelve sensible a cambios en las circunstancias y está listo para llevar a cabo el comportamiento adecuado dados tales cambios —en este estado ocurre el monitoreo. Por otra parte, una respuesta ultrarrápida, que se considera un comportamiento que tiene sentido una vez la simetría se rompe (Di Paolo et al., 2017, p. 201), puede ser entendido como el estado de consumación en el que una tendencia negativa es anulada —en este momento ocurre la regulación. La tendencia negativa en este caso consiste en no responder apropiadamente y no tener un buen desempeño en el experimento.

Pero si este es el caso, la aproximación enactiva aún debe explicar cómo actividades que en sí mismas no requieren ser monitoreadas ni reguladas pueden derivar su significado de procesos globales de hacer-sentido, dado que, en su corta temporalidad y en ellas mismas, no son instancias de hacer-sentido. Conviene recordar que las repuestas ultrarrápidas siguen la normatividad impuesta por la tarea, por ejemplo, de identificar si hay un perro en la imagen; por ende, difícilmente pueden verse como procesos crudos dentro de un proceso de hacer-sentido más amplio.

Aceptar esta opción implica, además, que la base de las relaciones significativas agente-ambiente en las que ocurren las respuestas ultrarrápidas no está enteramente en el aquí y el ahora en términos de su regulación y monitoreo, sino que fueron constituidas en el pasado. Pero esto requeriría renunciar al axioma de determinismo estructural. Si el significado de las

respuestas ultrarrápidas se deriva de su pertenencia a un acto significativo más amplio, entonces el cambio de estado en el sistema en un momento estaría contextualmente determinado por una condición anterior.

Finalmente, no es completamente claro qué tan lejos en el pasado o incluso en el futuro un acto mínimo de hacer-sentido debería abarcar. Como se indicó en la sección 3.1.3, hay al menos tres capas de normatividad que guían una respuesta ultrarrápida y que presumiblemente la hacen significativa: (i) las normas que guían la respuesta ultrarrápida, por ejemplo, responder a la presencia de la imagen de un perro, (ii) las normas establecidas en el estudio experimental por las instrucciones, y (iii) los factores socioculturales que hacen que participar en el experimento sea un acto con sentido. Es importante notar que el primer nivel de normatividad sería incomprensible sin los otros dos: soltar un botón cuando hay un perro visible en una pantalla solo tiene sentido cuando está situada en un contexto sociocultural y normativo más amplio. Cada uno de estos niveles de normatividad corresponde a diferentes escalas temporales en las que se desarrolla la acción y va de los milisegundos en (i) a años de socialización en (iii). Si la normatividad que guía la respuesta ultrarrápida en un momento dado viene de su pertenencia a actos de hacer-sentido más amplios, podría ser el caso que tal acto haya empezado tan temprano como el comienzo del proceso de socialización del participante.

3. *Tercera posibilidad: explicar la normatividad de las respuestas ultrarrápidas como constituidas exclusivamente por la historia de interacciones pasadas del agente.*

Esta posibilidad nos lleva a considerar la aproximación preferida por el enactivismo radical (Hutto & Myin, 2013, 2017). Bajo su perspectiva, la normatividad de las interacciones habilidosas debe ser explicada en términos de la historia de selección, sea en la escala evolucionaria, ontogenética o de aprendizaje. Por lo tanto, la actividad actual de los organismos es normativa solo respecto a lo que ha funcionado en el pasado. Y, de hecho, parte de la explicación en términos dinámicos de la anticipación implica la apelación a la historia de acoplamiento agente-ambiente que ha dejado al sistema preparado de la manera correcta para responder adaptivamente a escalas de tiempo muy rápidas. Así que argumentar que esta misma historia es también responsable del hecho de que estas respuestas son también formas de hacer sentido parece requerir solo un pequeño paso, que, en palabras de Hutto y Myin, equivaldría a la intencionalidad dirigida al mundo del organismo con sus propiedades fenoménicas. Como en la segunda posibilidad, sin embargo, esto requeriría que el axioma del determinismo estructural sea abandonado.

Sin embargo, la propuesta de Hutto y Myin de que la normatividad de actos presentes está solamente constituida por el pasado, a saber, por los actos similares que funcionaron o no en el pasado, tiene implicaciones problemáticas. Primero, hay un problema de novedad: una normatividad basada puramente en el pasado dejaría sin explicar la sensibilidad normativa de los agentes a la especificidad de la situación actual, esto es, su habilidad para producir respuestas ultrarrápidas a perturbaciones que nunca han sido experimentadas antes, como el que tiren de la mandíbula mientras se pronuncian sílabas. Segundo, esta explicación

presupone aquello que pretende explicar: asumir que nuevas instancias requieren un comportamiento normativamente guiado porque son suficientemente similares a otras instancias en el pasado deja sin explicar por qué, en primer lugar, el agente puede reconocer las similitudes relevantes en la nueva situación, que es en sí mismo un acto normativo. En ambos casos, las circunstancias actuales y los detonantes específicos del comportamiento no jugarían ningún papel en la determinación de las condiciones de éxito y fracaso porque el criterio normativo sería puramente histórico. Esto sugiere que no deberíamos excluir completamente el papel constitutivo del momento presente.

4. *Cuarta posibilidad: explicar la normatividad de las respuestas ultrarrápidas como constituida por interacciones tanto pasadas como presentes.*

Abandonar el axioma del determinismo estructural nos permite definir la cognición ultrarrápida de forma más holística y relacional, puesto que no hay razón para que el enfoque enactivo evite apelar al contexto histórico del agente. El monitoreo y la regulación se consiguen operacionalmente al estar en un estado crítico anticipatorio. Pero su carácter normativo no se deriva exclusivamente del estado actual del sistema sino de la interacción completa que se está desarrollando. Esto potencialmente implica que los mecanismos de monitoreo y control pueden ser un proceso que se desenvuelve en el presente pero que viene del pasado. De hecho, podríamos concebir la apelación del enactivismo radical al papel constitutivo del pasado como una referencia precisamente a estos procesos de temporalidad más extendida.

La normatividad entonces puede emerger respecto tanto a comportamientos exitosos en el pasado como a particularidades y detonantes de la situación actual, y así fuentes de normatividad tanto pasadas como presentes serían condensadas en la transición de un estado fuertemente anticipatorio a una interacción. La historia de selección pasada da forma a las posibles interacciones del individuo y al espacio de constreñimientos normativos del presente, mientras que las peculiaridades de la situación actual permitirían al agente llevar a cabo la respuesta adecuada y, cuando esta respuesta no sea posible para el agente, aprender nuevas respuestas. Si este es el caso, el enactivismo radical tendría que incluir las especificidades del momento presente como otro elemento constitutivo además de la historia pasada, lo que permitiría evitar el problema de presuponer la normatividad implicada en la selección de la respuesta pasada correcta en el momento presente.

Concepciones similares de procesos cognitivos han sido propuestas recientemente. Por ejemplo, Jurgens y Kirchhoff (2019) han argumentado que los procesos involucrados en la cognición social están constituidos diacrónicamente por procesos en la microescala y la macroescala que están en una relación de causación circular. Así que parece razonable adoptar como hipótesis de trabajo la idea de que hacer-sentido está constituido por interacciones tanto presentes como pasadas, y la idea de que los constreñimientos normativos pueden venir del pasado.

### 3.1.5. Expansión temporal y subdeterminismo

El fenómeno de la cognición ultrarrápida pone de manifiesto un reto clarificador para la propuesta enactiva respecto a cómo operacionalizar el monitoreo y la regulación adaptativas, y respecto a la escala temporal del hacer-sentido. Las respuestas ultrarrápidas ocurren demasiado rápido para que feedback de largo alcance ocurra, lo que el enactivismo explica en términos del papel constitutivo y normativo del contexto histórico de las respuestas. Sin embargo, esta explicación está en tensión con el axioma de determinismo estructural que captura la concepción original de la teoría autopoiética de la historia como una cadena de causas que llevan a la respuesta presente pero que no la constituyen. Esto abre la pregunta por cómo reconciliar el axioma del determinismo estructural, que es central en teoría de sistemas dinámicos, con la idea de que las respuestas ultrarrápidas sean una selección de trayectoria que depende de eventos del pasado, y en esa medida considerar las limitaciones de los sistemas dinámicos para modelar respuestas adaptativas cuyas condiciones de corrección dependen de su historia. Sin embargo, esta pregunta por la operacionalización no solo está fuera del propósito de esta tesis, sino que requiere un trabajo conceptual previo para precisar qué significa que la historia no solo tenga un papel causal sino también uno constitutivo.

Hay al menos dos maneras en que puede abordarse la pregunta por qué significa que eventos del pasado jueguen un papel constitutivo en las respuestas del presente, una ontológica y otra normativa. La primera posibilidad ha sido defendida, por ejemplo, por Kirchhoff (2014, 2015) quien captura la idea de una ontología extendida en el tiempo con el concepto de constitución diacrónica. Este concepto desafía la distinción entre, por un lado, la idea de causa como evento que precede y lleva a otro evento distinto, y, por otro lado, la idea de constituyente como evento o elemento que hace parte sincrónicamente de otro evento. Esta distinción no tiene sentido si se atiende a los fenómenos de la naturaleza que se despliegan en el tiempo:

No hay algo como un proceso en un instante o punto sincrónico en el tiempo. Por ejemplo, frecuentemente uno lee que el agua está constituida por H<sub>2</sub>O. Esta suposición es una suposición práctica que se hace en ciencia, así como en la vida cotidiana. Pero no debe tomarse como evidencia para la afirmación subsiguiente de que el agua está constituida por H<sub>2</sub>O en un punto sincrónico en el tiempo. En cambio, el agua está constituida por “oxígeno e hidrógeno en varias formas poliméricas, como (H<sub>2</sub>O)<sub>2</sub>, (H<sub>2</sub>O)<sub>3</sub>, etc., que continuamente se están formando, disipando y reformando en cortos periodos de tiempo de manera que dan lugar a las propiedades familiares del tipo macroscópico de agua” (Ladyman and Ross, 2007, p. 21) (Kirchhoff & Kiverstein, 2020)

Así como la existencia del agua es diacrónica, Kirchhoff y otros como Spivey (2008) han argumentado que la existencia de la mente y de los procesos mentales es también ontológicamente diacrónica.

Sin embargo, la dimensión normativa de la pregunta por la extensión temporal de la cognición es la que quisiera abordar. La teleosemántica hace una propuesta precisamente en esta dirección: propone derivar las condiciones normativas de los contenidos mentales de la



historia evolutiva, y examinar esta propuesta podría revelar herramientas útiles para entender el papel de la historia en la constitución de los fenómenos normativos en general. En la siguiente sección discuto esta propuesta, sus ventajas y limitaciones para naturalizar la normatividad y capturar su carácter temporalmente extendido.

## 3. 2. Teleosemántica

Las aproximaciones teleosemánticas buscan naturalizar las representaciones, en particular, su naturaleza normativa y lo hacen apelando a su historia de selección. La estrategia básica es derivar el contenido y las condiciones de corrección de las representaciones de las funciones biológicas, usando una concepción etiológica de función biológica originalmente propuesta por Larry Wright en los 70 y desarrollada y popularizada por varios otros autores en las décadas siguientes. Una de las propuestas dominantes es la de Ruth Millikan quien defiende que (i) un ítem tiene una función propia si resulta de una historia de selección, y que (ii) el término ‘representación’ es un término de función, lo que significa que las representaciones tienen una función propia que las define como representaciones, y esta es guiar el comportamiento del consumidor mediante isomorfismos entre representaciones y el ambiente. La historia de las representaciones y, en general, de cualquier ítem que tenga una función determinan el estándar respecto al que se evalúa su comportamiento. La concepción de cognición que subyace a las aproximaciones teleosemánticas es una concepción representacional. Esto significa que se asume que el rasgo definitorio de los comportamientos cognitivos es que están guiados por representaciones, y la razón es que estas permiten explicar la flexibilidad del comportamiento de los agentes cognitivos.

(i) El análisis del término ‘función’ de Millikan parte del uso ordinario de palabras que se refieren a objetos que tienen una función, por ejemplo, un abrelatas o una máquina de escribir (Millikan, 2020). Característicamente, estos objetos pueden desempeñarse bien o fallar, y el tener una función es crucial para ello:

Tener una función determina si la cosa tiene que ser evaluada como buena o excelente o correcta... determina un estándar respecto al cual se mide si puede ejecutar su función y qué tan bien lo hace. A menos que algo tenga una función, propósito, *telos* no hay un estándar para este tipo de evaluación (Millikan, 2020, p. 2).

Hay dos puntos que me interesa resaltar acá: primero, el tener una función es lo que establece criterios de normatividad. Segundo, la concepción de función se establece bajo el paradigma de los artefactos y cualquier otro uso de función y propósito es por analogía a estos (Millikan, 2021).

Los objetos que tienen una función propia vienen con normas respecto de las cuales se puede comparar su comportamiento y establecer si hicieron lo que se supone que deberían. Estas normas son producidas por su historia de selección. En el caso de los artefactos que Millikan considera paradigmático, la historia relevante es el proceso de selección en su diseño en el que formas y materiales se seleccionaron para permitir ciertos comportamientos.

Análogamente, la función biológica de un rasgo está determinada por su historia en la que los efectos causales de tal rasgo contribuyeron a la supervivencia de la especie y por ello fueron seleccionados. En estos procesos de selección se establece en qué consiste el funcionamiento normal de aquello que se selecciona: producir los efectos por los que ha sido seleccionado evolutivamente. Consideremos, por ejemplo, un órgano como el corazón:

La función de mi corazón es bombear sangre porque bombear sangre es lo que los corazones de mis ancestros hicieron que contribuyó a la supervivencia y reproducción de mis ancestros, y por ende contribuyeron a la persistencia de los corazones de ese tipo en la población, y que entonces explica mi posesión de tal corazón (Kingsbury, 2006, p. 2)

De acuerdo con la propuesta teleosemántica, la historia de selección determina que su función propia sea bombear sangre, a pesar de que también pueda servir también para determinar si alguien está teniendo un ataque de pánico.

Bajo esta concepción, el carácter normativo de la función biológica de un órgano o rasgo está determinado por el valor que tuvo para la supervivencia del organismo en su historia de selección. Esto implica que la función es un estándar o medida respecto de la que los comportamientos reales pueden diferir, pero no impone un deber o un mandato normativo (*ought*) (Millikan, 2005, pp. 83-84). En ese sentido, la historia de selección establece un criterio de normatividad.

(ii) Las representaciones, de acuerdo con Millikan, son términos de función. Esto significa que, así como corazones y riñones, las representaciones tienen una función propia que está determinada por su historia de selección natural. ¿Para qué son las representaciones entonces? La estrategia de Millikan es de nuevo recurrir al uso del lenguaje para encontrar características comunes y, de acuerdo con ella, el lenguaje mismo es el ejemplo más claro que tenemos de representación. La estructura general de las representaciones tiene tres elementos fundamentales: un productor, un intérprete y un propósito en la interpretación o consumo de las representaciones. Estas son producidas por algo para ser usadas por algo para algo. En particular, el propósito de las representaciones es guiar el comportamiento de los agentes cognitivos, controlar sus acciones o guiar a los intérpretes mientras llevan a cabo sus funciones (Millikan, 2020, p. 6).

La concepción de tres partes de las representaciones es frecuentemente ilustrada en la literatura considerando cómo una rana caza mosquitos para alimentarse. La rana se representa el mosquito, esta representación tiene la función de ser consumida por el sistema visual de la rana para que guíe su comportamiento y le permita cazar al mosquito y alimentarse. Los sistemas visuales y motor de la rana consumen representaciones y estas representaciones tienen la función de permitir a la rana cazar su alimento. Tales representaciones hacen posible que la rana tenga un comportamiento flexible y versátil, y que pueda enfrentarse a condiciones que nunca ha enfrentado antes.

Hay una forma característica en la que las representaciones llevan a cabo su función: una representación tiene una relación isomórfica con el ambiente. Esto significa que su estructura mapea estructuras externas en el ambiente y, en esa medida, se trata de una representación propiamente y no de cualquier otro elemento en la cadena causal que va desde el evento en el ambiente hasta la compleción de la respuesta de la criatura. Estos isomorfismos y la cadena causal de estímulo-respuesta a la que pertenecen, de acuerdo con Millikan (2004; citado en: Shea, 2013a), son seleccionados en la historia evolutiva y es esta historia la que establece los criterios para determinar si son correctas o si hay un error.

### 3.2.1. Dos críticas: intencionalidad e intensionalidad

Una de las críticas mejor conocidas a la teleosemántica es la crítica de Fodor, quien argumenta que la historia de selección no permite determinar los criterios normativos de las representaciones de manera suficientemente específica. El problema es que la historia de selección no permite diferenciar entre formas distintas de representarse el mundo, pero esta diferenciación es crucial para tener estados cognitivos. Veamos por qué esta diferenciación es crucial.

De acuerdo con la tradición en filosofía del lenguaje del siglo XX, los estados cognitivos están compuestos por dos elementos: un contenido, usualmente una proposición, y una actitud hacia ese contenido. El contenido especifica un estado de cosas en el mundo que puede ser verdadero o falso, por ejemplo, hay una mosca volando sobre mi cabeza, y la actitud proposicional es la forma en que el agente se dirige hacia esa proposición, por ejemplo, puede creer que es verdadera, puede saber que es el caso, puede dudar de ella o puede desear que sea el caso. La identidad de un estado cognitivo está determinada por estos dos elementos y esto implica que, si cambia el contenido, cambia también el estado cognitivo. No es lo mismo creer que hay una mosca volando sobre mi cabeza a creer que hay alimento para rana volando sobre mi cabeza a pesar de que ambas proposiciones se refieran el mismo estado de cosas en el mundo. Esta distinción no solo permite capturar que nuestras creencias dependen de la forma en que describimos el mundo, sino que también puede dar cuenta de los casos de construcción de conocimiento en los que se descubre que un objeto bajo alguna descripción, ej., Véspero, el lucero vespertino, es el mismo objeto que se conoce bajo otra descripción, Fósforo, el lucero matutino, que fue una preocupación muy presente en la tradición analítica del siglo XX (Frege, 1892 [1980]; Kripke, 1972).

Como vimos arriba, la explicación teleosemántica de Millikan de los estados cognitivos propone que estos pueden identificarse por su función propia: guían el comportamiento del sistema que los consume y llevan esta función a cabo mediante un isomorfismo entre ellos mismos y el ambiente. Esta es la función propia de las representaciones porque son los efectos para los que fueron seleccionadas a lo largo de nuestra historia evolutiva. Sin embargo, la mayoría de las instancias en las que, por ejemplo, una rana se ha representado una mosca podría también haberse representado un pequeño punto negro moviéndose, y el isomorfismo se mantendría. La razón es que en las circunstancias normales en las que viven y han vivido las ranas, ambas descripciones son co-extensionales. En general, no podemos esperar que la

historia evolutiva nos permita diferenciar entre las múltiples descripciones que podemos hacer de un conjunto de estados de cosas.

Millikan ofrece dos respuestas a esta crítica. La primera es que la objeción de Fodor surge de una interpretación que se concentra en las causas o sistemas productores de representaciones, e ignora el papel de sus consumidores (Millikan, 1990; Papineau, 1998). La historia evolutiva no selecciona solo los sistemas que puedan producir representaciones isomórficas con todos los estados de cosas co-extensionales pero descritos de forma distinta. En cambio, la historia evolutiva, selecciona las representaciones que permitan guiar exitosamente la acción del sistema consumidor: la selección natural es selección *para* algo. Por ello, no cualquier representación del mismo estado de cosas, ej., hay una mosca volando frente a mí, va a producir una representación que pueda guiar exitosamente a la rana para alimentarse a lo largo de su historia. De esta forma, es decir, reconociendo el papel tanto del productor como del consumidor de las representaciones, parte de la indeterminación de la representación bajo la teleosemántica de Millikan se resuelve.

Esta forma de selección, sin embargo, no determina completamente cuál es la descripción bajo la que el ambiente es representado, aun contemplando tanto el papel del sistema productor que debe generar una representación isomórfica con aspectos del ambiente como el papel del sistema consumidor cuyo comportamiento debe ser guiado exitosamente por la representación. La razón es que, dadas las condiciones normales del ambiente en la que la historia de selección ocurrió, aún es posible que tanto una representación que describe un estado de cosas en el mundo, ej., como pequeño punto negro comestible que se mueve, como una que lo describe, ej., como mosquito volando, frente a mí sean ambas exitosas. En ese sentido, Fodor afirma: A Darwin le importa cuántas moscas te comes, pero no bajo qué descripción te las comes (Fodor, 1990, p. 73). Pero si ese es el caso, la explicación teleosemántica no puede dar cuenta de cómo emerge el contenido específico de las representaciones y, en ese sentido, no puede dar cuenta ni de su significado ni sus condiciones de verdad.

Millikan ha respondido al menos de dos maneras a esta crítica. Por un lado, ha señalado que nuestras acciones intencionales y nuestro uso cotidiano del lenguaje no requieren descripciones tan específicas como las que Fodor demanda (Millikan, 2021). Millikan no desarrolla este punto extensamente en (2021), pero el corazón del argumento es presumiblemente que el significado de las representaciones, incluyendo las representaciones que le dan sentido al lenguaje, está determinado por su isomorfismo con el ambiente y la acción que estas deben permitir llevar a cabo al sistema que las consume. En ese sentido, el requerimiento de especificidad de Fodor es inapropiado para el contexto y uso de las representaciones. Por ejemplo, para nosotros cotidianamente no hace ninguna diferencia si nos representamos un estado de cosas bajo la descripción ‘el plátano está sobre la mesa’, o bajo la descripción, ‘el plátano de 102 gr está sobre la mesa’; y así mismo para las ranas en su historia evolutiva probablemente no hacía ninguna diferencia formarse una representación que describe la mosca como ‘pequeño punto negro comestible en movimiento’, y una que describe

el mismo estado de cosas como ‘cosita voladora comestible’ desde que lo lleven a cazar exitosamente su alimento en condiciones normales.

La segunda respuesta de Millikan es que su propuesta teleosemántica no busca dar cuenta de cómo se determina el contenido específico de las representaciones, ni los detalles que están en juego en este proceso. En cambio, el propósito es explicar el mecanismo general mediante el que surge la normatividad, y mostrar que si entendemos las representaciones como el tipo de cosas de la naturaleza que tienen una función propia, podemos dar cuenta de forma naturalista de su carácter normativo. En ese sentido, no es pertinente exigirle a la propuesta teleosemántica de Millikan que explique la finura descriptiva que la historia de selección, ej., de las ranas, permite. De manera que, si quisiéramos determinar específicamente el tipo de contenido de una clase de representaciones o el contenido específico de una representación, la teleosemántica nos indica hacia donde mirar, i.e., la historia de selección, pero no la respuesta que vamos a encontrar allí.

Desde una aproximación corporizada, Hutto y Myin proponen una respuesta alternativa a la crítica de la subdeterminación de Fodor y, en general, a este debate de los noventa. Hutto y Myin, defienden que, en efecto, la propuesta teleosemántica no determina el contenido de las representaciones, pero esto no es una razón para rechazarla sino para rechazar la idea de que las formas más básicas de mente en la naturaleza deban concebirse en términos representacionales. A pesar de que las objeciones y respuestas alrededor de este debate continuaron en las siguientes décadas y las teorías teleológicas de contenido mental han seguido desarrollándose hasta ahora, quisiera concentrarme en la aproximación enactiva mejor conocida a la teleosemántica que puede encontrarse en la literatura, a saber, el enactivismo radical de Hutto y Myin.

### Enactivismo radical

En *Radicalizing enactivism: basic minds without content* (2013), Hutto y Myin proponen el enactivismo radical. Esta propuesta tiene un pie en la tradición enactiva y el otro en la teleosemántica. Por un lado, el enactivismo radical comparte con la tradición enactiva la idea de que la cognición, la percepción en particular, está orientada a la acción y que es un error asumir entidades mentales internas como representaciones o contenidos para explicar la cognición. Por otro lado, Hutto y Myin toman de la teleosemántica la idea de que los criterios de corrección de la cognición están determinados por la historia de selección. Sin embargo, el enactivismo radical critica ambas posturas por asumir que el *quid* de la cognición radica en entidades o procesos internos. En el caso del enactivismo se trata de una consecuencia de defender el cuerpo vivo como un sistema que crea significados o interpretaciones del mundo. De acuerdo con Hutto y Myin, el enactivismo clásico o autopoietico no lleva su compromiso de volcar la cognición hacia la acción y rechazar el cognitivismo lo suficientemente lejos (Hutto & Myin, 2013, pp. 34, 35). En el caso de la teleosemántica porque derivan propiedades semánticas, es decir, descripciones que pueden ser verdaderas o falsas como las discutidas en la sección anterior, de mecanismos que solo pueden producir relaciones extensionales (Hutto & Myin, 2013, p. 79).

Hutto y Myin identifican al menos tres problemas con las explicaciones de la cognición basadas en representaciones. Primero, no es necesario apelar a representaciones o modelos mentales del mundo para explicar las formas básicas de cognición. Bajo la concepción corporizada de mente que defienden Hutto y Myin, las formas básicas de cognición son interacciones entre agente y ambiente en donde los agentes pueden ser entendidos como sistemas cuyas variables (estas son predominantemente los estados de los sistemas de percepción y movimiento) están acopladas a las variables del ambiente en una relación no lineal y en bucle. Por ejemplo, los estados del sistema de percepción visual varían con el comportamiento de la luz en el ambiente y la disponibilidad de los aspectos del ambiente para el agente varía constantemente con sus movimientos oculares y corporales. Hutto y Myin señalan que estos acoplamientos pueden ser explicados como relaciones de covariancia o como correlaciones, y esta relación es tan estrecha que es imposible identificar propiamente un dominio interno y otro externo, y mucho menos un dominio que represente al otro. Pero si este es el caso, difícilmente las representaciones, en tanto, elementos con propiedades semánticas, pueden contribuir realmente a este acoplamiento.

Esto nos lleva al segundo problema que identifican Hutto y Myin: aún si hubiera representaciones o modelos del mundo en un dominio interno, difícilmente estos podrían mantenerse actualizados y ser procesados con la rapidez suficiente para guiar eficientemente el comportamiento de los agentes. Esto es evidente no solo en casos de cognición ultrarrápida como los que discutí arriba, sino también en los desarrollos de la robótica en la década de los 90 (Beer, 1996; Brooks, 1991). Para construir robots con una capacidad similar de navegación en el espacio a la de los agentes cognitivos, se optó por prescindir de modelos internos y de procesamiento centralizado de la información que hacían sus reacciones excesivamente lentas e ineficientes, y se modeló esta capacidad a través de bucles locales de acoplamiento de manera que los robots pudieran interactuar directamente con el ambiente y no mediante ninguna representación.

El tercer problema es que no es del todo claro exactamente qué son las representaciones dentro de estas explicaciones ni tampoco qué significa exactamente que una parte de un organismo sea productor o consumidor de representaciones. Hutto y Myin señalan que frecuentemente se asume que las representaciones se refieren a la relación no semántica de correlación o covariancia entre parámetros en la interacción (Hutto & Myin, 2013, p. 41). Sin embargo, Hutto y Myin sostienen: añadir usuarios sensibles a la información o consumidores de información a la historia deja las relaciones de covariancia inalteradas; estas permanecen tal como son y donde están (2013, p. 72). Pero si este es el caso, no hay forma legítima dentro de una explicación naturalista de dar el paso de correlaciones puramente extensionales a propiedades semánticas o intensionales como las descripciones definidas o conjuntos de descripciones que discutí arriba.

Este último problema hace eco de la crítica de Fodor: tanto él como Hutto y Myin señalan la imposibilidad de derivar propiedades semánticas o descripciones intensionales de relaciones puramente extensionales. Pero, a diferencia de Fodor, en lugar de tratar de derivar los contenidos representacionales usando una estrategia distinta, Hutto y Myin optan por una

estrategia minimalista ontológicamente, y proponen abandonar por completo las apelaciones a representaciones en tanto entidades internas con propiedades semánticas para explicar las formas básicas de cognición, de ahí que se asuman como una propuesta radical. La forma de hacerlo es fundamentando dos aspectos esenciales de la cognición en la historia de los agentes cognitivos: la intencionalidad, es decir, la capacidad de los agentes cognitivos de dirigirse hacia algo y el carácter normativo de la cognición. Para ello recurren a las aproximaciones teleosemánticas y proponen adoptar una versión que no apele a representaciones.

Siguiendo la idea central de la teleosemántica, el punto de partida es que la selección natural produce organismos con ciertas estructuras y establece sus normas y funciones biológicas. Algunas de estas estructuras tienen la función de orientar al organismo hacia aspectos específicos del ambiente y responder a estos. En particular, los sistemas de percepción tienen la función de variar correlativamente con propiedades del ambiente, y estas correlaciones permiten guiar la acción del organismo para responder a tales propiedades, ej. cazando a la mosca o huyendo de la serpiente. Estas correlaciones y el bucle de percepción-acción en el que juegan un papel hacen que la respuesta del organismo sea intencional, es decir, que esté dirigida a aquello con lo que se supone que debe correlacionarse, y la historia evolutiva hace además que esta respuesta sea normativa porque las estructuras del organismo que participan han sido seleccionadas durante la historia evolutiva del organismo para producir tales correlaciones y ese tipo de respuestas. La diferencia crucial entre las aproximaciones teleosemánticas y el enactivismo radical de Hutto y Myin es que estos últimos no derivan ni pretenden derivar de esta explicación entidades con propiedades semánticas como las representaciones, es decir, entidades que pueden ser verdaderas o falsas. En ese sentido, proponen explicar la cognición en términos de comportamiento intencional y no entidades internas e intensionales.

### 3.2.2. Giro hacia la acción: criterios normativos extendidos en el tiempo

La sección anterior estuvo dedicada a la explicación teleosemántica de cómo surgen parámetros normativos en la naturaleza; esta sección busca incorporarla a la propuesta enactiva clásica para dar cuenta de la extensión temporal de la cognición. Si bien Hutto y Myin ya dieron un primer paso en esta dirección y defendieron ampliamente que las representaciones no juegan realmente un papel explicativo dentro de las aproximaciones teleosemánticas, hay un par de puntos más que deben ser discutidos para incorporarlas a una aproximación a la cognición fundamentalmente orientada a la acción como la que he defendido hasta ahora. Los dos puntos por discutir en más detalle son, por un lado, el establecimiento de criterios normativos en las propuestas teleosemánticas más allá de las representaciones, y segundo, el papel de los agentes individuales en tal normatividad.

Shea ofrece una explicación de las representaciones que no se limita a relaciones seleccionadas en la historia evolutiva ni tampoco a relaciones isomórficas. En *Representation in Cognitive Science* (2018), propone que las representaciones son producto de funciones que pertenecen a un sistema y cumplen dos condiciones. Primero, son ‘funciones de tarea’ que define como funciones robustas y estables cuyo input son sensaciones causadas por estados

del mundo y cuyo output son movimientos. El que sean robustas significa que son funciones que producen un mismo resultado en diferentes circunstancias y con diferente input. Mientras que el que sean estables significa que se han mantenido a lo largo del tiempo o bien porque contribuyeron al éxito evolutivo de los sistemas del mismo tipo o porque fueron favorecidas por los aprendizajes del sistema o porque contribuyen directamente a que el sistema continúe existiendo. La segunda condición para que una función produzca representaciones es que sea una relación explotable con aspectos del ambiente, esto es, que haya una relación de correspondencia estructural entre la función y aspectos del ambiente o de correlación, y, además, que esta relación sea procesada de acuerdo con algoritmos. Puesto que el interés de esta tesis son los fenómenos normativos, quisiera concentrarme en la idea de función de tarea como fuente de normatividad, y no en los elementos que la hacen producir representaciones.

La propuesta de Shea tiene varias virtudes. Primero, los requerimientos de estabilidad y robustez de la función de tarea hacen explícito qué en la historia de la especie o del individuo hace que se establezca una función y, entonces, un estándar normativo. Segundo, toma la evolución como solo una de las maneras de lograr que una función sea estable, y reconoce también la escala temporal de la vida individual del sistema como una fuente de estabilización en donde el organismo aprende a comportarse de ciertas formas y desarrolla patrones de comportamiento para mantenerse con vida. Estos dos puntos, los patrones compartidos establecidos en el tiempo y la preservación de la vida individual, son cruciales para dar cuenta del surgimiento de fenómenos normativos.

Modificar la teleosemántica de esta forma y caracterizar el tipo de comportamientos relevantes como robustos y estables en el tiempo nos acerca a una explicación naturalista de un elemento esencial para el surgimiento del significado bajo la concepción que defendí en el Capítulo 1: los patrones compartidos. Los patrones compartidos se establecen en la medida en que un comportamiento y sus variaciones se repite lo suficiente y empieza a ser reproducido de una generación a la otra, incluso cuando las circunstancias cambian. Tales patrones establecen la normalidad y es uno de los estándares respecto al que comportamientos individuales pueden evaluarse. Pero, al contrario de lo que parecen sugerir autores como Millikan al comparar los patrones establecidos por la selección natural con un promedio, estos patrones tienen el estatus de norma no meramente por ser regularidades nomológicas sino porque son regularidades que constituyen una forma de vida. Shea apunta a una idea similar al definir su concepto de función de tarea. En este, combina dos formas de teleología: por un lado, un comportamiento teleológico es uno que ocurre con regularidad y ocurre *para* algo, y por otro lado, el comportamiento de un organismo es causado por las consecuencias que este trae (Shea, 2018 p. 48). De acuerdo con la aproximación naturalista de Shea, el último tipo de comportamiento es posible si ese tipo de consecuencias han ocurrido con regularidad en el pasado. En ese sentido, una función robusta, es decir, una que tiende a producir los mismos resultados, hace que tal función se estabilice en el tiempo, y entonces, hacen posible que los organismos produzcan comportamientos para algo, o lo que es lo mismo, de acuerdo con Shea, estas son funciones de representación que son la base para determinar el contenido (Shea, 2018 p. 48).



El problema, es que el contenido y las funciones de representación no parecen ser una motivación propia de los seres vivos. Sugiero que aquello para lo que quisiera un organismo vivo actuar es, fundamentalmente, para mantener su forma de vida, reproducirse o expandir sus posibilidades de acción<sup>9</sup>. En ese sentido, sugiero que las funciones se estabilizan en tanto que permite a las criaturas vivas mantenerse con vida, reproducirse e interactuar con el ambiente y con otros seres vivos, lo que me lleva al segundo punto: el organismo individual tiene un papel fundamental en el surgimiento de la normatividad.

Para mostrar la importancia del organismo individual en la explicación de los fenómenos normativos quisiera regresar a la propuesta de Millikan y señalar un punto que ella deja sin resolver: qué hace que una característica adquiera una función propia, es decir, qué criterio permite reconocer cuando una característica es seleccionada o simplemente azarosa. Presumiblemente, algo más que un criterio temporal es necesario y se requiera un criterio que involucre cómo sus efectos causales contribuyen al mantenimiento de la vida del organismo individual y su capacidad para reproducirse en el aquí y el ahora.

Millikan y las aproximaciones que derivan la normatividad exclusivamente de los procesos evolutivos de selección parecen dar por sentado que existe precisamente esa otra forma de normatividad en el nivel individual y la escala temporal de su vida. Por ejemplo, Millikan afirma:

Un mero promedio, después de todo, es también una clase de norma. Las formas comportamentales que han tenido en el pasado valor de supervivencia son una medida de cuáles disposiciones comportamentales, tanto pasadas como presentes, pueden diferir, pero estas diferencias no están prescritas en ningún sentido. De hecho, las diferencias a veces resultan ventajosas (Millikan, 2005, pp. 83-84).

Si la normatividad solo emergiera de los procesos de selección natural, sería imposible, incluso un sinsentido, que salirse de las normas producto de la evolución pudiera ser ventajoso. De hecho, como señalan Christensen y Bickhard (2002), sin una normatividad que permita determinar la utilidad o conveniencia respecto a los individuos no hay nada para qué seleccionar porque el proceso de selección es un proceso que en sí mismo, así resulte en un estándar normativo, también responde a una normatividad.

La analogía con los artefactos desde donde Millikan entiende el concepto de función oscurece la necesidad de la normatividad individual. En la historia del diseño de un artefacto se seleccionan sus características para que cumplan un propósito, y es esta historia la que hace que, para usar uno de los ejemplos de Millikan misma, una máquina de escribir tenga una función propia así haya sido dañada. Pero esta historia no sería una historia de selección si no hubiera una diseñadora o un grupo de diseñadores con un propósito para el que llevan a cabo

---

<sup>9</sup> Frances Egan (2020), en una crítica parecida, ha defendido que la apelación a las representaciones en la propuesta de Shea tiene una función heurística y realmente no juega un papel en la explicación. Aunque estoy de acuerdo con su diagnóstico, no estoy de acuerdo con su suposición de que el sentido no puede ser parte de una explicación naturalista de la mente.

el proceso de selección. Por supuesto, esta analogía no implica que haya una diseñadora detrás de los procesos de selección natural, pero sí implica la existencia de organismos cuyo propósito es mantener su forma de vida individual. Estos individuos constituyen los puntos de vista que conforman el patrón de selección y respecto de quienes se trata de un proceso de selección y no meramente de una estadística indiferente. Por eso, reconocer el papel de los organismos individuales que buscan mantener una forma de vida es crucial.

Si bien Shea incluye explícitamente el mantenimiento de la vida individual como uno de los mecanismos que hacen estable y robusta una función, no hay una explicación, y tampoco tiene que haberla porque no es el propósito de su propuesta, de qué hace de la vida individual una fuente de criterios normativos. Entonces, la propuesta teleosemántica, tanto de Millikan como de Shea, deja la pregunta por el individuo sin resolver y esta es una de las motivaciones para recurrir a la aproximación enactiva. Ninguna de las propuestas explica qué hace que cierta clase de individuos en la naturaleza estén sujetos a procesos de selección o busquen mantenerse existiendo, i.e., los seres vivos. Responder esta pregunta permite reconocer los individuos para quienes las consecuencias de los procesos de selección importan, es decir, son ellos quienes se vuelven más o menos adaptados a un ambiente y quienes seguirán o no viviendo y reproduciéndose como resultado de lo que es seleccionado, quienes cometen o no errores cognitivos. El que haya procesos de selección natural o de aprendizaje presupone la existencia de organismos vivos cuya vida establece una norma fundamental: mantenerse con vida y reproducirse por generaciones. Sin responder cuál es el tipo de sistema que está sujeto a normatividad por sí mismo y no solo en virtud de la atribución de un observador, queda incompleta la pregunta por los fenómenos normativos en general y por la cognición en particular.

Hay dos consecuencias de estas observaciones que merecen ser notadas y que discutiré en el próximo capítulo. La primera es que los patrones de comportamiento compartidos por una forma de vida emergen de los organismos vivos individuales que la constituyen a través de generaciones, pero al mismo tiempo, la constitución y el comportamiento individual de los organismos está formado por los patrones compartidos constituidos en la historia evolutiva. Esto sugiere la hipótesis de que las especies o grupos de organismos que compartan una forma de vida constituyan un sistema autónomo a través de las generaciones. La segunda consecuencia es que, si bien es necesaria una agencia, una que surge de la constitución misma del organismo, esta no implica un compromiso con que el significado se localice en representaciones internas ni en ningún otro proceso o entidad interna y por definición invisible para otros. Esto se sigue tanto de la concepción de sentido como actividad pública que defendí en el capítulo 1, como de la caracterización de la propuesta enactiva del capítulo anterior. En el próximo capítulo propondré una definición de normatividad que captura estas dos consecuencias.

Para cerrar, quisiera regresar al problema de la cognición ultrarrápida que motivó el reconocimiento de la extensión temporal de la cognición al comienzo de este capítulo. Las respuestas ultrarrápidas son respuestas intrínsecamente sujetas a criterios de normatividad establecidos por la historia del agente y en esa medida son significativas. Tal historia se

refiere, por un lado, a la escala evolutiva en donde la historia de selección de la especie a la que pertenece el agente establece la función propia de, ej., los ojos para reconocer si hay o no un animal en el horizonte o la mandíbula para, entre otras cosas, producir sonidos sutiles de manera consistente. Por otro lado, también la historia de interacciones de la comunidad a la que pertenece el agente establece criterios normativos adicionales que, en este caso, delimitan aún más las posibilidades de respuestas correctas e incorrectas en la situación experimental. Tal historia de interacciones sociales humanas establece el espacio de posibilidades tanto de lo que cuenta como respuesta con sentido, correcta o incorrecta, o sin sentido en cuyo caso sería una respuesta incorrecta, como del comportamiento general durante el experimento que cuenta como apropiado e inapropiado. Es importante notar que la historia de evolutiva y de interacciones establece una diferenciación en el espacio de posibilidades de acción del agente, pero no determina su comportamiento ni su respuesta en el presente, de lo contrario, no habría posibilidad de una respuesta adaptiva genuina.

La historia evolutiva y de interacciones puede establecer la diferenciación entre regiones de posibles comportamientos y respuestas correctas e incorrectas porque en esa historia emergen patrones compartidos de comportamiento y respuestas, y tales patrones continúan desarrollándose en el presente. La idea de constitución diacrónica de Kirchhoff (2014) (cf. p. 104) captura la naturaleza de estos patrones compartidos como fenómenos que están constituidos por diferentes formas de interacciones y comportamientos que en conjunto y a lo largo del tiempo le dan al fenómeno sus características específicas. En el caso de la cognición ultrarrápida y la cognición en general, los fenómenos formados por diferentes formas de interacciones y comportamientos a lo largo del tiempo son formas de vida compartidas por una especie o una comunidad; y las condiciones de mantenimiento de tales formas de vida son precisamente los criterios normativos que le dan sentido a comportamientos, interacciones y respuestas específicas de los individuos que la constituyen. Este es el corazón de la concepción de normatividad que propondré y defenderé en el próximo capítulo.

### 3. 3. Conclusión

Comencé este capítulo ofreciendo un argumento a favor de la naturaleza temporalmente extendida de la cognición que desarrollé junto a Froese (Mojica & Froese, 2019). Para ello, discutimos los casos de cognición ultrarrápida en los que se exhiben respuestas no solo adaptivas sino cognitivas tan rápidas que no solo no dejan tiempo suficiente para el procesamiento interno de representaciones o información, sino que no dejan del todo claro cómo funcionan los mecanismos adaptivos que propone la explicación enactiva. Argumentamos que la manera más viable de enfrentar este reto del enactivismo es hacer explícito el carácter temporalmente extendido de la cognición de manera que se reconoce que la dimensión normativa de la respuesta ultrarrápida no se limita al instante en el que se da, sino que viene de interacciones y condiciones pasadas que continúan desarrollándose en el presente.

Bajo tal motivación, discutí las aproximaciones teleosemánticas a la naturalización de la normatividad que proponen precisamente que la dimensión normativa de las propiedades y

comportamientos del presente resulta del proceso de selección del pasado. Discutí la propuesta de Millikan en la que la cognición se concibe como un asunto representacional, las representaciones se conciben como funciones y las funciones están determinadas por la historia de selección. La función de un rasgo consiste en los efectos causales que tal rasgo tuvo en el pasado y en virtud de los cuales contribuyó a la sobrevivencia de la especie y fue seleccionado. Esta propuesta define las funciones propias de los rasgos mediante sus consecuencias prácticas. En consonancia con la noción de significado como consecuencias prácticas que presenté en el 1.3, el establecimiento de funciones puede ser entendido como el establecimiento de patrones que se automantienen en el tiempo, i.e., una forma de vida compartida por los miembros de una especie a lo largo de la historia evolutiva, y estos, a su vez fungen, como criterios de evaluación para el comportamiento de tales individuos.

Discutí dos limitaciones importantes en la propuesta de Millikan. Primero, para Millikan, las representaciones se caracterizan porque llevan a cabo su función mediante isomorfismos con el ambiente, pero mostré que, bajo la concepción misma de función de Millikan, este requerimiento es superfluo en tanto que las funciones propiamente cognitivas pueden haber sido seleccionadas sin necesidad de que haya un isomorfismo. Segundo, mostré que la concepción de función por analogía con los artefactos de Millikan obscurece un punto fundamental: la necesidad de individuos con puntos de vista para quienes tales consecuencias prácticas hagan la diferencia. Estos son necesarios para que surja el fenómeno de la normatividad intrínseca y motiva la integración de una explicación naturalista no solo del surgimiento de los criterios de normatividad como explica la teleosemántica, sino también del surgimiento de individuos con agencia para quienes las consecuencias prácticas puedan hacer la diferencia, como lo hace el enactivismo. En los capítulos que restan, desarrollo mi propuesta de concepción de normatividad en la que se integran ambos elementos y se reconoce la extensión temporal y el carácter situado de los fenómenos normativos y de la cognición en particular.

## Capítulo 4 Normatividad relacional y situada: una propuesta integracionista

Como mostré en el Capítulo 2, el enactivismo autopoietico ofrece una explicación naturalista de surgimiento de la agencia, es decir, el surgimiento de un sistema que tiene la capacidad de seleccionar cursos de acción que hacen la diferencia para su propia vida; a este tipo de interacciones se les conoce como respuestas adaptivas en el enactivismo. En el Capítulo 3, argumenté que las respuestas adaptivas están extendidas temporalmente porque parte esencial de su carácter normativo, a saber, el que están sujetas a criterios normativos, está constituida en una escala temporal que sobrepasa la impresión inmediata del agente que actúa. La propuesta teleosemántica ofrece una buena explicación de cómo los criterios normativos, los biológicos en particular, se constituyen a lo largo de la historia evolutiva de selección de las especies o en la historia personal del individuo. Sin embargo, mostré que su propuesta carece de una explicación del surgimiento de la agencia, pero este es un aspecto esencial para dar cuenta del fenómeno de la normatividad como argumenté en el primer capítulo. Por eso, en este capítulo regreso a la propuesta enactiva y muestro cómo esta puede modificarse para ofrecer una explicación de la agencia que reconoce la distinción entre la agencia de un sistema por un lado y los criterios normativos que lo rigen pero que sobrepasan sus impresiones presentes por el otro.

Comienzo discutiendo una pregunta que está abierta en el enactivismo todavía. El enactivismo autopoietico defiende una continuidad entre las normas biológicas, cognitivas y sociales fundamentada en su concepción de normatividad básica como automantenimiento. De acuerdo con el enactivismo, el automantenimiento de la identidad sensoriomotora de un agente establece las normas cognitivas que regulan su comportamiento, y el automantenimiento de su identidad social determina las normas sociales. Sin embargo, no hay una explicación clara de cómo los individuos, que por su propia constitución están movidos para interactuar con el mundo bajo la norma del automantenimiento, podrían interactuar con el mundo guiados por normas no individuales. Además, entender toda la normatividad como automantenimiento del agente que actúa hace que no sea claro cómo los agentes establecen interacciones sociales y adquieren hábitos siguiendo normas que no tienen ninguna implicación en su constitución como individuos o que incluso atentan contra ella. En este capítulo muestro estas limitaciones del enactivismo y, para hacer frente a estos retos, propongo una definición de normatividad que incorpora dos elementos que introduje en los capítulos anteriores, a saber, una noción de normatividad basada en una concepción del significado wittgensteiniana y orientada a la acción, y la distinción entre dos dimensiones del fenómeno de la normatividad: un agente y criterios normativos. Específicamente, propongo que, primero, los criterios normativos están siempre establecidos por el mantenimiento de un sistema autónomo y, segundo, los comportamientos normativos son producidos por un agente, es decir, un sistema autónomo y adaptivo, y tales comportamientos parte del sistema autónomo que establece los criterios normativos que lo rigen. O bien, el agente que actúa puede ser también el sistema autónomo que establece los criterios normativos, o bien los

criterios normativos pueden ser establecidos por un sistema distinto al agente, pero cuyos comportamientos pueden contribuir al mantenimiento de tal sistema.

El argumento central de este capítulo está basado y es en gran medida una traducción literal de mi artículo *The enactive naturalization of normativity: From self-maintenance to situated interactions* (Mojica, 2021). Hacia el final del capítulo, hay una explicación de la idea de causalidad circular propuesta por Thomas Fuchs; tal explicación fue tomada literalmente de mi reseña crítica a su libro publicada como parte de mi investigación de doctorado (Mojica, 2019).

## 4. 1. Variedades de normatividad: la continuidad vida-mente-agencia social

En esta sección trazo una distinción entre tres variedades de normatividad que rigen el comportamiento o constitución de los seres vivos pero que son cualitativamente distintas: la biológica, la sensoriomotora y la interacción entre agentes. Todas estas variedades son formas intrínsecas de normatividad. Un rasgo, un acontecimiento o una propiedad es intrínsecamente normativa si es producida por el mismo sistema que produce normas o propósitos respecto a las que se evalúan. Por ejemplo, la actividad metabólica de un organismo es intrínsecamente normativa ya que su existencia y su actividad son posibles porque pertenecen al sistema más amplio constituido por el organismo vivo, y su actividad es correcta o incorrecta, mejor o peor con respecto a su contribución para mantener al organismo con vida y, por tanto, para mantener el proceso metabólico en marcha. En contraste, algo es extrínsecamente normativo si su existencia no depende de lo que establece la norma o el propósito con respecto al cual se evalúan. Por ejemplo, las características y el comportamiento de un trozo de hierro no son normativos por sí mismos porque este no tiene ninguna finalidad por sí misma según la cual puedan evaluarse. Cuando el hierro se utiliza para hacer, ej., una olla o un pulmón de hierro, sus propiedades se evalúan normativamente como mejores o peores para hacer el artefacto y cumplir los propósitos que tienen sus diseñadores humanos, y en este caso, sus características y comportamiento son extrínsecamente normativos. Vale la pena notar que puede haber características y comportamientos extrínsecamente normativos solo en la medida en que exista un sistema distinto con normas y propósitos propios con respecto al que estas se puedan evaluar. Esto implica que la normatividad extrínseca no existiría sin sistemas intrínsecamente normativos, y por ello, la normatividad intrínseca es en la que me he concentrado hasta ahora y lo seguiré haciendo a lo largo de la tesis.

Este tipo de normatividad es la que está en juego cuando se busca explicar el fenómeno de la normatividad en la naturaleza: la emergencia de normas y propósitos respecto de los cuales eventos, rasgos o características pueden ser evaluadas como mejores o peores, convenientes o inconvenientes, correctas o incorrectas, etc. Tanto los parámetros normativos que rigen el comportamiento como el comportamiento en sí determinan el tipo de normatividad en juego, y con ese criterio, podemos identificar varios tipos de fenómenos normativos. En esta sección identifiqué tres tipos de normatividad: primero, la normatividad biológica cuya forma más

básica es el mantenimiento de la vida propia y adquiere más capas normativas con la historia selección natural; segundo, la normatividad que rige las interacciones sensoriomotoras, es decir, los comportamientos que un agente lleva a cabo guiado por su percepción, y que ha sido identificada como la forma más básica de cognición en el enactivismo; y tercero, la normatividad que guía la interacción entre agentes en donde surgen formas de normatividad que no dependen exclusivamente de ninguno de los individuos que interactúan. Muestro que la noción de normatividad sensoriomotora necesita ser refinada y sugiero que la normatividad de la interacción entre agentes puede llegar a ser constitutiva de la sensoriomotora.

#### 4.1.1. Normatividad biológica: enactivismo e interactivismo

Tanto los enfoques organísmicos como los evolutivos han identificado la normatividad biológica como el tipo de normatividad que se refiere a las normas impuestas por el propósito de mantener la vida de los individuos de una especie. En ese caso, el comportamiento de los organismos vivos, así como de sus sub-sistemas, procesos y órganos internos tienen una función, lo que significa que los organismos vivos se constituyen y se comportan de acuerdo con un propósito propio o una norma, y su corrección está determinada por las consecuencias prácticas que hacen una diferencia, ya sea respecto de su propia supervivencia o de la supervivencia de la especie a largo plazo. Como vimos en el capítulo anterior, una diferencia crucial entre los enfoques evolutivo y organísmico es en dónde identifican la fuente de los criterios normativos: en la escala temporal evolutiva o en las consecuencias a corto plazo para la vida de un individuo.

Vimos en el capítulo anterior que los enfoques teleológicos, en particular, la propuesta de Shea (2018), proponen que los patrones de comportamiento de los individuos o sus rasgos fisiológicos que se mantienen en diferentes circunstancias a lo largo de la historia evolutiva e individual establecen criterios normativos para su comportamiento y su actividad biológica. Estos patrones que Shea llama ‘funciones de tarea’ permiten determinar si, por ejemplo, los órganos, los procesos metabólicos o las funciones y comportamientos reproductivos son correctos o no. Pero, vimos arriba, queda abierta la pregunta por qué caracteriza a los seres vivos que les permite estar sujetos a estos criterios normativos.

Enactivismo e interactivismo ofrecen una explicación de la constitución individual de los organismos vivos que establece la base de la normatividad en la vida de un individuo, y consideran la historia evolutiva como derivativas, aunque cruciales para explicar las formas de vida como las conocemos. La fundamentación de la normatividad en el organismo individual se ha presentado como contrapuesta a las propuestas etiológicas de naturalización de la normatividad (ej., Bickhard, 2000), en donde los parámetros de constitución y comportamiento correctos están completamente determinados por la historia de selección.

Sin embargo, esta contraposición no es necesaria. Por ejemplo, Ruiz-Mirazo et al. (2004) defienden que para que un sistema sea un sistema vivo no es suficiente con que se autoindividúe, sino que también debe hacer parte de una red histórica de relaciones entre sistemas similares que evolucione en el tiempo. En esta misma línea sugiero más adelante que

explicaciones organísmicas y evolutivas pueden reconciliarse reconociendo que las interacciones en la vida individual están doblemente sujetas a normatividades establecidas por el individuo en tanto sistema y por la especie en tanto sistema colectivo de patrones compartidos en las escalas temporales evolutivas. Esta posibilidad de reconciliación es consecuencia de la concepción de normatividad que propondré, y permite reconocer los casos en los que estos los criterios individuales y evolutivos entran en conflicto -por ejemplo, las especies cuya forma de reproducción implica la muerte de los individuos.

Las propuestas basadas en el organismo individual tienen la ventaja de mostrar cómo su carácter viviente establece una norma, y ofrece potencialmente un criterio parcial para determinar por qué un rasgo llega a ser seleccionado. Tanto el enactivismo como el interactivismo conciben a un ser vivo como un sistema autónomo, lo que significa que está organizado de manera que los procesos que lo constituyen son mantenidos por el sistema mismo. Mientras que la explicación interactiva se basa en los sistemas termodinámicos y en una ontología de procesos, la enactiva tiene sus raíces teóricas en la teoría autopoietica como vimos ampliamente en el capítulo 2. A pesar de los antecedentes teóricos diferentes y de las diferencias en cuanto al papel de las representaciones, existe una amplia coincidencia entre sus respectivas concepciones sobre lo que fundamentalmente hace que un sistema genere sus propias normas.

De acuerdo con la explicación interactiva, un organismo vivo es un sistema autónomo que se auto-mantiene recursivamente. El interactivismo define la autonomía como una propiedad graduada de los sistemas termodinámicos que surge cuando los sistemas tienen tres propiedades básicas: el sistema forma una unidad que es cohesiva, está lejos del equilibrio termodinámico y contribuye al mantenimiento de sus propias condiciones de cohesión.

- (i) Que un sistema sea cohesivo significa que está organizado de manera que “interactúa con el entorno como un todo integrado causalmente” (Christensen & Bickhard, 2002, p. 17). Para que esto ocurra, el sistema debe estar compuesto por una organización que se mantiene constante a lo largo del tiempo. Una roca, por ejemplo, es una unidad cohesiva porque su estructura se mantiene constante en el tiempo e interactúa como una unidad con su entorno. Una roca, sin embargo, no es un sistema autónomo porque su organización está en equilibrio termodinámico, lo que significa que no necesita de energía para mantenerse, y por ello, puede ser aislada del ambiente y no desaparecer.
- (ii) La segunda condición, estar lejos del equilibrio termodinámico, significa que la organización del sistema necesita de influjo de energía constante para mantenerse, para ello el sistema debe estar abierto a la interacción para evitar llegar al equilibrio, lo que significaría perder su organización y dejar de existir. Esto implica que, si el sistema es aislado de su entorno, desaparece. Algunos de estos sistemas se mantienen porque tienen una entrada de energía desde el exterior, este es el caso de las células de convección de Rayleigh-Bénard que se mantienen gracias a una fuente de calor externa. Otros sistemas contribuyen a mantener su propia organización lejos del equilibrio, por ejemplo, la llama de una vela mantiene la temperatura por encima de su



umbral de combustión, vaporiza la cera de la vela para tener un suministro de energía, induce el movimiento del aire que introduce oxígeno al sistema y expulsa desechos de la combustión (Bickhard, 2000). Estos procesos son la forma en la que la llama contribuye al mantenimiento de su propia existencia: introducen activamente la energía que la organización del sistema necesita para mantenerse y, en cierta medida mantienen activamente las condiciones que permiten que la llama siga existiendo.

Esta segunda condición de estar lejos del equilibrio termodinámico es una forma en la que se realiza la condición de precariedad en la concepción enactiva de sistemas autónomos: los sistemas vivos deben mantener sus procesos constitutivos renovándose constantemente para seguir existiendo y están en riesgo constante de dejar de existir. El enactivismo define la precariedad como una propiedad de todos los procesos materiales en tanto que todos dejarán de existir eventualmente; pero en el contexto de los sistemas autónomos, la precariedad se refiere a que los procesos que constituyen tales sistemas desaparecen o decaen rápidamente si son aislados de la red a la que pertenecen, y cuando están dentro de tal red, deben ser renovados constantemente para que no decaigan (Di Paolo & Thompson, 2014, pp. 4-5). La segunda condición que propone el interactivismo de que los sistemas estén lejos del equilibrio termodinámico es la forma en la que los sistemas autónomos que se encuentran en la naturaleza son precarios. La formulación del interactivismo es, sin embargo, mucho más clara que la del enactivismo en tanto que hace explícita la diferencia entre sistemas cuya organización está lejos del equilibrio termodinámico y otros sistemas cuya organización está en equilibrio y por ello no debe ser renovada constantemente. Esta diferencia que el interactivismo logra capturar claramente en su definición es crucial porque el estar lejos del equilibrio termodinámico hace que los sistemas autónomos necesiten de la interacción constante con el ambiente.

Los sistemas que forman una unidad cohesiva, que están lejos del equilibrio termodinámico y que además contribuyen al mantenimiento de su propia existencia como las llamas de las velas o los tornados son un “interesante caso límite” y la diferencia entre sus formas de auto-producción y las formas de auto-producción de los agentes biológicos es aún una pregunta abierta dentro del enactivismo (Di Paolo et al., 2017, pp. 126-127). El interactivismo, sin embargo, ofrece una respuesta a esta pregunta. En primer lugar, los sistemas que forman una unidad cohesiva y que están lejos del equilibrio termodinámico son sistemas de donde surgen condiciones de normatividad básicas: independientemente de si un sistema lejos del equilibrio contribuye a su propio mantenimiento o no, la necesidad positiva de mantener condiciones energéticas para que el sistema se mantenga establece un parámetro normativo en donde aquello que permite mantener esas condiciones es conveniente y cumple una función para ese sistema. Si el sistema además tiene mecanismos para mantener tales condiciones como el fuego o los tornados, estos no solo son sistemas normativos sino también autónomos (Bickhard, 2016, p. 23). Esta, sin embargo, no es la misma forma de autonomía de los organismos vivos. Para estos, Bickhard propone una tercera condición, el auto-mantenimiento recursivo.

Antes de revisar esa tercera condición quisiera discutir la afirmación de Bickhard de que una normatividad básica surge de sistemas cohesivos y lejos del equilibrio. Como Bickhard mismo lo señala, su organización permite diferenciar entre condiciones y procesos activos que permiten que el sistema se mantenga y condiciones en las que decaería. Esta diferenciación del espacio de posibilidades establece los criterios de acuerdo con los que una interacción entre el sistema y su ambiente puede evaluarse. Sin embargo, tanto en los casos en los que los sistemas básicos como la llama de una vela o un tornado contribuyen a su propia existencia como los casos en los que las condiciones de mantenimiento vienen de fuera, no hay una agencia o un sistema individual respecto del cuál pueda decirse que cometió un error. La razón es que estos sistemas no pueden seleccionar entre distintos cursos de acción que tengan resultados diferentes respecto a su propio mantenimiento como sistema que forma una unidad cohesiva. Puesto que es imposible que estos sistemas respondan de formas distintas a cambios en el ambiente, no tiene sentido decir que estos sistemas cometieron un error. Como argumenté anteriormente, para que surja el fenómeno de la normatividad propiamente es necesario tanto un agente del que tenga sentido decir que puede errar como criterios normativos para determinar si su comportamiento es correcto. En el caso básico de sistemas cohesivos y lejos del equilibrio termodinámico, no hay agencia, y por lo tanto no hay normatividad genuina. Esta surge, en cambio, con la tercera condición que propone Bickhard.

(iii) La tercera condición que propone Bickhard es que el sistema contribuya de forma recursiva a mantener sus propias condiciones de cohesión. Esto significa que estos sistemas pueden desplegar diferentes procesos como respuesta al cambio en las circunstancias, así, por ejemplo, una bacteria no solo mantiene sus procesos metabólicos con la energía que obtiene de su alimento, el azúcar, sino que también puede moverse para cambiar de ambiente una vez los recursos energéticos se agotan. La bacteria puede seleccionar entre, al menos, tres tipos distintos de comportamiento: quedarse en donde está, nadar en una dirección o ir dando tumbos. Estas posibilidades permiten que la bacteria responda de maneras distintas a diferentes circunstancias del ambiente y, por ende, hace que esta pueda mantenerse a sí misma en varios rangos en los que las condiciones cambian (Bickhard, 2000). Esta capacidad de responder a cambios en las condiciones se contrapone a la limitada forma de automantenimiento de los sistemas no recursivos que desaparecen una vez los recursos del ambiente se agotan, como la llama de la vela que no puede buscar más fuentes de energía una vez se consume toda la cera de la vela. En ese sentido, de acuerdo con el interactivismo de Bickhard, esta forma de autonomía captura las condiciones en las que emerge el tipo de normatividad propiamente cognitiva, aunque en una versión mínima, en donde el agente debe seleccionar el curso de acción a partir de las indicaciones de posibilidades de interacción disponibles en el ambiente para el agente<sup>10</sup> (Bickhard, 2016).

---

<sup>10</sup> De acuerdo con Bickhard, estas son las condiciones de donde surge la normatividad representacional porque las indicaciones pueden ser verdaderas o falsas. Sin embargo, como argumenté en el capítulo 1 no es necesario reducir todas las formas de normatividad genuina a condiciones de verdad o falsedad ni siquiera cuando se trata de la normatividad que determina el sentido del lenguaje. Puesto que el propósito de esta sección no es discutir en detalle la propuesta interactiva sino introducir algunos de sus elementos para hacer más sólida la propuesta que he desarrollado hasta ahora, no discutiré el compromiso del interactivismo con las representaciones.

Los sistemas autónomos que pueden automantenerse recursivamente son cualitativamente distintos a otros sistemas autónomos en tanto que estos pueden seleccionar sus propias trayectorias de comportamiento. De acuerdo con Bickhard, esto es posible gracias a que el sistema tiene estructuras que pueden cambiar o ser activadas según las circunstancias; estas estructuras que normalmente están lejos del equilibrio termodinámico y ellas mismas deben ser mantenidas (Bickhard, 2000). Pero esta capacidad recursiva tiene una consecuencia más que me interesa resaltar: introduce las condiciones en las que surge la agencia mínima. La razón es que bajo tales condiciones el sistema puede seleccionar entre distintas posibilidades que hacen una diferencia práctica para su propia existencia. En la medida en que el sistema puede seleccionar, puede también errar, y en la medida en que tales selecciones tienen consecuencias para su propia existencia, este está intrínsecamente motivado a seleccionar activamente entre tales posibilidades. Por lo tanto, cuando las tres condiciones propuestas por Bickhard se cumplen no solamente emerge un sistema autónomo y vivo, sino también los dos elementos necesarios para que emerja la normatividad propiamente dicha.

El concepto de automantenimiento recursivo del interactivismo es simétrico con el concepto de adaptividad del enactivismo en tanto que ambos buscan capturar las condiciones en las que un sistema puede responder de forma flexible, intencional y normativa a los cambios en el entorno. Como vimos en el capítulo 2 (p. 65), la adaptividad es el requerimiento introducido por Di Paolo (2005) de que el sistema tenga mecanismos dedicados o emergentes para monitorear y regular sus propias tendencias y evitar acercarse a sus propios límites de viabilidad. La capacidad de automantenimiento recursivo es una forma de adaptividad en la que el sistema regula sus interacciones con el ambiente y sus propias estructuras ante cambios en el ambiente. El concepto de automantenimiento recursivo hace explícito desde la definición misma el carácter cambiante de la organización o procesos constitutivos del sistema, y que estas respuestas flexibles y normativas son además respuestas en las que el sistema modifica su interacción con el ambiente. Aunque estas dos características de los comportamientos normativos de los seres vivos han sido reconocidas en la literatura enactiva, estrictamente hablando, estas no se siguen del concepto de adaptividad como ha sido definido en la literatura. Por ello, propongo adoptar el concepto de automantenimiento recursivo del interactivismo como la forma en que los sistemas vivos realizan la adaptividad.

La caracterización de normatividad biológica que propongo incluye tanto procesos y rasgos del funcionamiento del organismo como sistema autónomo, ej., procesos metabólicos o de crecimiento, como patrones de comportamiento en donde el organismo interactúa con el ambiente guiado por su percepción. Por lo tanto, la clasificación que propongo implica que las normatividades biológica y sensoriomotora se superponen.

#### 4.1.2. Normatividad sensoriomotora

El enactivismo amplía la explicación de la normatividad mínima a formas más complejas de interacción en las que la normatividad rectora va más allá de las necesidades del organismo individual de mantenerse vivo y de mantener sus trayectorias lejos de los límites de la

viabilidad. En la sección 3.1.1, introduje brevemente los esquemas y redes sensoriomotoras que son los conceptos con los que el enactivismo explica estas formas más complejas de creación de sentido. En esta sección reviso con más detalle esta explicación.

Desde sus inicios, el enactivismo ha defendido que la cognición consiste en interacciones sensoriomotoras (Varela et al., 1991). Esta afirmación ha sido desarrollada posteriormente en la rama sensoriomotora del enactivismo dedicada a defender que la percepción y la acción consisten en bucles sensoriomotores, y también en la rama autopoiética del enactivismo. Esta última afirma no sólo que la percepción y la acción están constituidas por bucles sensoriomotores, sino también que la actividad sensoriomotora forma patrones que constituyen la identidad sensoriomotora de los agentes. Según el enactivismo autopoiético, estos patrones de actividad sensoriomotora forman redes autónomas precarias de las que surgen tanto el propósito de auto-mantenimiento como los criterios normativos. Por tanto, los agentes se ven movidos a interactuar con el entorno para mantenerse como organismos vivos, como agentes sensoriomotores y veremos más adelante, eventualmente como agentes sociales.

Como vimos anteriormente, la naturaleza autónoma y precaria de los organismos vivos da lugar a su comportamiento normativo: interactúan con el mundo para mantenerse vivos. Algunos enactivistas, como Di Paolo y Thompson (2014, pp. 71-72), consideran que esta forma básica de normatividad y comportamiento intencionado es ya una forma de cognición básica. Barandiaran (2017), sin embargo, ha propuesto reservar este término para el comportamiento normativo cuyas consecuencias no están ligadas a la preservación de la vida del organismo. Esta distinción está en consonancia con la concepción original de la cognición ofrecida en *The Embodied Mind* (Varela et al., 1991), pero lo que es más importante, hace justicia a la intuición de que la cognición no consiste en mantenerse vivo, como podría ser el caso de la función biológica, sino en interactuar con el entorno a través de la percepción y la acción propia de una manera intencionada, normativa, flexible y que se adapta a las circunstancias. Así, es posible diferenciar las funciones de mantenimiento de la vida, como el bombeo del corazón y la actividad digestiva, de las actividades cognitivas, como la elaboración de una herramienta o el mantenimiento de una conversación. Crucialmente, el comportamiento normativo desacoplado de las necesidades metabólicas del organismo es posible por el desacoplamiento parcial del sistema nervioso de las necesidades metabólicas del organismo. El sistema nervioso es entonces una condición de posibilidad para que surja la cognición.

Di Paolo et al. (2017) han propuesto el concepto de esquemas sensoriomotores para caracterizar las redes que constituyen las identidades sensoriomotoras de los agentes:

La idea de un esquema sensoriomotor describe una organización de patrones de coordinación sensoriomotora que son usados regularmente por el agente porque han sido evaluados como preferibles (junto con algún marco normativo relevante) para lograr una meta particular. El desarrollo o adquisición de un esquema sensoriomotor describe cómo un agente se sincroniza a una situación específica al seleccionar y modular patrones de coordinación sensoriomotores de acuerdo con normas relevantes (2017, p. 58).

Si la cognición consiste en interacciones sensoriomotoras, entonces el marco normativo, la meta particular y las normas específicas relevantes de los esquemas sensoriomotores a las que se refieren Di Paolo et al. son los que permitirán entender las normas propiamente cognitivas.

El enactivismo autopoiético defiende que estas normas son intrínsecas al mantenimiento de hábitos sensoriomotores, y estos se forman por la ejecución frecuente de esquemas sensoriomotores organizados de forma que permiten a los agentes llevar a cabo actividades particulares. Los hábitos forman una identidad que se incorpora a la identidad del agente quien está movido a mantener su propia identidad. Por tanto, es crucial entender qué son los hábitos para comprender cómo se explica la normatividad cognitiva desde el enactivismo autopoiético.

Cuando las redes de esquemas sensoriomotores son llevadas a cabo repetidamente por el agente y se convierten en un sistema autónomo, se convierten en hábitos. Esto ocurre cuando la coordinación frecuente entre el agente y el entorno hace que las estructuras del cuerpo y del entorno que intervienen en la realización de la red se organicen de manera que, con el tiempo, sea más fácil llevarlo a cabo. Por ejemplo, cuando un agente pinta con frecuencia, tanto su cuerpo como el entorno se acondicionan mejor a la actividad; por ejemplo, en cuanto a su cuerpo, los músculos implicados se fortalecen, su capacidad de percepción del color y su coordinación se vuelven más eficientes y precisas para dibujar, etc., en cuanto al entorno, el agente puede acondicionar un espacio con una superficie limpia y lisa para dibujar y pintar, y puede empezar a acumular diferentes tipos de materiales para diferentes propósitos específicos, etc. Estas estructuras suelen ser precarias, es decir, necesitan conservarse, de lo contrario, decaen; y como permiten la ejecución de los esquemas sensoriomotores, su precariedad significa que el hábito mismo también es precario. A medida que el agente y su entorno se habitúan a la ejecución de determinados esquemas sensoriomotores, el hábito sensoriomotor puede adquirir una “vida propia”: es a la vez condición y consecuencia de su propia promulgación (Di Paolo et al., 2017, p. 144). Esto significa que cuanto más ejecuta el agente los esquemas sensoriomotores pertinentes, más fuerte se vuelve el hábito y más probable es que lo ejecute en el futuro, es decir, el hábito se convierte en un sistema autónomo más robusto.

De acuerdo con el enactivismo, los hábitos también establecen de forma autónoma una identidad y una norma intrínseca de automantenimiento, así como lo hacen las redes autónomas y precarias de procesos fisicoquímicos que establecen una identidad y la norma de mantenerse a sí mismas frente a su propio deterioro. La identidad de un hábito consiste en el patrón que forma la ejecución repetida de acciones (esquemas sensoriomotores) organizadas jerárquicamente de manera que permiten que una actividad se lleve a cabo. Estos patrones incluyen tanto las estructuras del ambiente como las estructuras internas al organismo del agente que permiten que tales acciones se lleven a cabo. Vale la pena notar que la sucesión de acciones que permiten llevar a cabo una actividad y que constituyen el hábito no debe ser exactamente igual ni en tanto al orden ni en tanto a los elementos individuales. Esto es evidente en el ejemplo del hábito de pintar en el que no es necesario e incluso podría ser impertinente pintar lo mismo en cada iteración del hábito. Así que la normatividad que

emerge con un hábito puede entenderse como la región de posibilidades de acción que el hábito mismo impone para seguir existiendo que, a su vez, está determinada por si estas acciones pertenecen al patrón de acción situada que ha constituido al hábito en el pasado. Dentro de esta región de posibilidades hay grados de flexibilidad en la ejecución e implica el mantenimiento de estructuras internas y externas que hacen la acción posible.

Aunque en la literatura enactiva se sugiere que los agentes incorporan los hábitos a su propia identidad, veremos más adelante que esta propuesta no deja del todo claro cuál es exactamente la relación de un hábito con el agente individual que lo ejecuta; y por eso, no deja del todo claro si los hábitos constituyen o no un fenómeno normativo completo.

### 4.1.3. Interacción y normas sociales

El tercer tipo de normatividad surge de la interacción entre agentes que se reconocen como tales. Este caso coincide con los hábitos sensoriomotores en que los agentes tienen que coordinar sus movimientos con algo externo, pero esta vez no solo se trata de un entorno cambiante sino de otro agente, y coincide además en que tal coordinación está guiada por valores distintos a sus requerimientos metabólicos para mantenerse con vida. Esto supone una diferencia crucial: a diferencia de los hábitos sensoriomotores ejecutados individualmente que vimos en la sección anterior, las interacciones entre agentes no se guían exclusivamente por los objetivos e intereses individuales de ninguno de ellos, sino que cada uno aporta los suyos a la interacción y tiene su propio punto de vista que puede transformarse en la interacción. Así, los agentes coordinan no sólo su actividad sensoriomotora sino también los objetivos e intereses que guían la interacción. En la literatura enactiva, se han propuesto al menos tres conceptos para caracterizar y clasificar este tipo de interacción. Estos son: hacer-sentido participativo (De Jaegher & Di Paolo, 2007), interacciones socio-cognitivas (Froese & Di Paolo, 2011) e interacciones lingüísticas o *linguaging* (Di Paolo et al., 2018).

El concepto de hacer-sentido participativo es una de las primeras formulaciones del enactivismo para caracterizar la esfera de la cognición social (De Jaegher & Di Paolo, 2007). A diferencia de los enfoques tradicionales en la cognición, los encuentros entre personas se entienden primariamente como procesos interactivos en el que los agentes crean sentido juntos y no como si se tratara de descifrar los estados mentales escondidos en la mente del otro. En estas interacciones, los agentes se entienden como dos sistemas que se coordinan entre sí. Por supuesto, no todas las formas de acoplamiento entre sistemas constituyen interacciones sociales, aun cuando tales sistemas sean organismos humanos. De acuerdo con De Jaegher y Di Paolo, deben cumplirse dos condiciones. La primera es que el acoplamiento sea regulado [conjuntamente por los agentes que interactúan] para generar y mantener una identidad en el dominio relacional (De Jaegher & Di Paolo, 2007, p. 493). Esto se refiere a que en las interacciones sociales hay una doble influencia: lo que hace cada agente individual influye la forma en que se desenvuelve la interacción, y esta a su vez predispone a los agentes individuales a actuar de formas que llevan a que la interacción se mantenga. Esta doble influencia hace que la interacción misma sea un sistema autónomo distinto a cada sistema individual, y puesto que requiere de la participación de los agentes para mantenerse es

también precaria. Es importante resaltar que la contribución de cada agente al mantenimiento de la interacción puede ocurrir sin que estos tengan la intención de hacerlo; así como es el caso en el clásico encuentro en el pasillo en el que los agentes se coordinan para mantenerse delante del otro a pesar de su propia intención de apartarse y seguir caminando. La segunda condición es que los individuos involucrados sean y se mantengan siendo autónomos en su papel de interactuantes (De Jaegher & Di Paolo, 2007, p. 493).

De acuerdo con De Jaegher y Di Paolo, las interacciones que cumplen con estas dos condiciones son instancias de hacer-sentido participativo. El razonamiento es el siguiente: cada agente individual hace sentido de la interacción desde su propia perspectiva y, bajo la perspectiva enactiva, la forma en que los agentes hacen sentido es a través de sus interacciones con el ambiente, esto es, la forma en que sus movimientos se coordinan con su entorno. En las interacciones sociales ambos agentes se coordinan el uno con el otro lo que implica que se coordinan en sus respectivas actividades de hacer-sentido. Puesto que las interacciones sociales como las definen De Jaegher y Di Paolo constituyen un sistema autónomo y precario, estas establecen sus propias condiciones de mantenimiento, lo que significa que surge una distinción entre conveniente e inconveniente, adecuado e inadecuado, correcto e incorrecto respecto a la interacción misma. Este sistema autónomo no está disponible para ninguno de los individuos sin el otro, y forma entonces un espacio de significado al que cada uno puede acceder gracias a la coordinación de su hacer-sentido con el otro (De Jaegher & Di Paolo, 2007, p. 497).

Froese y Di Paolo (2011) identifican dos condiciones necesarias para que un agente pueda establecer interacciones no solo participativas, sino sociales: deben ser capaces de comportarse siguiendo valores no metabólicos, y deben ser capaces de percibir al otro como agente con sus propios propósitos. Vimos arriba que tener un sistema nervioso parcialmente desacoplado de las necesidades metabólicas del organismo permite que emerjan valores no-metabólicos que guían las interacciones. Entre agentes con estas capacidades pueden establecerse interacciones genuinamente socio-cognitivas. Froese y Di Paolo coinciden con la propuesta de De Jaegher y Di Paolo (2007) en tanto que un dominio autónomo debe emerger de la coordinación entre agentes y que su carácter de agentes debe mantenerse para que una interacción cuente como interacción social. Sin embargo, señalan que estas dos condiciones incluyen interacciones en las que o bien el otro no es reconocido como agente o en las que el carácter conjunto de la coordinación no hace ninguna diferencia para la manera en la que los agentes hacen sentido. Así que ofrecen una definición operacional de las interacciones sociales que añade una tercera condición a las ya propuestas por De Jaegher y Di Paolo con su concepto de hacer-sentido participativo, a saber, que la regulación que un agente cognitivo hace sobre su acoplamiento sensoriomotor es complementado por la regulación coordinada de al menos otro agente cognitivo (Froese & Di Paolo, 2011, p. 23). Esta condición hace explícita la intuición de que la interacción social es social cuando no puede ser llevado a cabo sin el reconocimiento del otro en tanto otro.

El dominio relacional que surge en las interacciones socio-cognitivas puede entenderse como el resultado de la negociación dialéctica entre agentes; la interacción se vuelve autónoma por

sí misma y establece sus propias normas de automantenimiento. Pensemos, por ejemplo, en el acicalamiento mutuo, el baile o una conversación. En todos los casos, cada agente tiene su propia posición, y regula sus propios movimientos y actos de habla según sus propios hábitos de movimiento y uso del lenguaje. Sin embargo, la interacción misma, es decir, el baile y la conversación, impone normas a cada agente sobre lo que tiene que hacer para que la interacción continúe. De hecho, la interacción puede continuar a pesar de las intenciones individuales de cada agente, por ejemplo, las discusiones que surgen y siguen adelante a pesar de las intenciones de los agentes de dejar de discutir.

Las interacciones socio-cognitivas entre agentes pueden convertirse en patrones a lo largo del tiempo y, como se ha señalado en la literatura enactiva, las comunidades pueden desarrollar guiones preestablecidos de interacción de manera que sus individuos saben cómo coordinarse entre sí de formas complejas incluso en su primera interacción. Esto ocurre a medida que los agentes aprenden a realizar actos complementarios parciales, que se completan en la interacción con otros (Di Paolo et al., 2017, p. 245). Este es el mecanismo central que nos permite desarrollar un lenguaje y utilizarlo para establecer interacciones complejas entre nosotros. Es importante destacar que las interacciones que se repiten en el tiempo se convierten en hábitos interactivos compartidos, y entonces, de acuerdo con el enactivismo, surge la tercera dimensión de la identidad en el individuo que incorpora estos actos parciales, es decir, la agencia social (Di Paolo et al., 2017, p. 216).

Di Paolo, Cuffari y De Jaegher (Cuffari et al., 2015; Di Paolo et al., 2018) ofrecen una explicación detallada de la emergencia de las interacciones lingüísticas (*linguaging*) a través de un modelo conceptual y uno del desarrollo. El modelo conceptual es un modelo dialéctico que parte del concepto de hacer-sentido participativo, en donde no se asume que los agentes se reconocen entre sí como agentes, y llega a interacciones lingüísticas mostrando que las diferentes formas de interacción tienen tensiones propias cuya resolución resulta en una forma de interacción distinta que, a su vez, tiene sus propias tensiones y que producen una tercera forma de interacción que, a su vez, tiene sus propias tensiones, etc. Por ejemplo, la tensión propia del hacer-sentido participativo ocurre entre las normas individuales de cada agente y la normatividad que impone la interacción. Esta tensión se resuelve mediante la emergencia de una agencia social que significa que cada agente incorpora las normas sociales y aprende a llevar a cabo actos parciales que no solo requieren de la regulación del otro, sino que disminuyen la posibilidad de conflicto (Cuffari et al., 2015, pp. 1098-1099). La emergencia social es un segundo de los 8 niveles distintos de interacción con sus respectivas tensiones que consideran Di Paolo, Cuffari y De Jaegher. Por supuesto, el último nivel, el nivel de las interacciones lingüísticas tiene sus propias tensiones, pero la emergencia de este nivel es el fenómeno que les interesa explicar y por ello el análisis para allí (Cuffari et al., 2015, p. 1110). El mecanismo fundamental en la resolución de las tensiones propias de cada nivel de interacción es la incorporación a cada individuo de normas sociales y de formas de actuar con otros (Cuffari et al., 2015, p. 1116). Este concepto de incorporación será central en la discusión de las siguientes secciones así que volveré a este en la siguiente sección.

\* \* \*



Hasta ahora, hemos identificado tres hitos en el continuo de la normatividad en la naturaleza que propone el enactivismo autopoietico:

1. La normatividad biológica está establecida por la identidad de un organismo vivo y por los patrones establecidos en la historia evolutiva. Este tipo de normatividad está constituida por procesos fisicoquímicos cuya organización les permite autodiferenciarse como sistema individual y automantenerse ante la precariedad.
2. La normatividad sensoriomotora o habitual se establece por la identidad sensoriomotora del agente y se constituye por el patrón de actividades que este realiza regularmente.
3. La normatividad social se establece, en primer lugar, por las condiciones de automantenimiento que establecen las interacciones autónomas entre agentes, y en segundo lugar, por la interiorización de normas sociales en los hábitos de interacción del agente con otros y su ambiente.

Todos estos hitos se basan en la normatividad autodirigida: los acontecimientos se evalúan normativamente siempre con respecto al interés del agente por preservar su identidad, bien sea su identidad biológica, su identidad sensoriomotora o su identidad social. En todos los casos, la normatividad surge como un principio de automantenimiento, la corrección está determinada por si se satisfacen las necesidades y se contrarrestan las amenazas a esa identidad. Esta explicación se compromete con un continuo entre vida, mente y agencia social en tanto que apela al mismo principio de auto-constitución y emergencia de la normatividad para explicar los tres fenómenos.

## 4. 2. Una brecha en el continuo

En esta sección, muestro que este tipo de normatividad no es suficiente para explicar la riqueza de las criaturas vivas y cognitivas. En particular, muestro que la naturalización enactiva de la normatividad propuesta llega a un callejón sin salida cuando las posibilidades sensoriomotoras no hacen ninguna diferencia con respecto a la identidad de los organismos, lo que hace imposible explicar la transición a la interacción social y a la cognición. Esta transición constituye una brecha importante, aunque no letal, para el continuo naturalista de vida-mente-agencia social que propone el enactivismo autopoietico.

La primera brecha que quisiera señalar está en el paso de la normatividad biológica a la normatividad sensoriomotora. En la explicación enactiva de la agencia sensoriomotora, la forma más básica de normatividad es la biológica en la que los agentes autónomos se automantienen y para ello diferencian y regulan su propia actividad respecto a lo que beneficia y lo que amenaza su propia existencia. Esto incluye sus propios procesos y condiciones internas, como la temperatura interna y la tasa de metabolización, así como sus interacciones con el entorno, de modo que incorporan los recursos necesarios y rechazan lo que es perjudicial para sí mismo. La autonomía fisicoquímica establece el tipo más básico de normatividad, es decir, la normatividad biológica, en la que los sistemas se convierten en agentes, y su comportamiento se somete a criterios normativos.

Ser un agente, en este caso, que el sistema diferencia sus interacciones con el entorno y sus potencialidades, así como sus propios procesos internos, en términos de lo que les permite mantener su propia vida y lo que la amenaza. El éxito o el fracaso de las respuestas internas o interactivas de los sistemas está determinado por criterios normativos que van más allá de sus impresiones actuales: mantener la viabilidad del sistema y evitar acercarse a los límites de lo que es viable para la existencia del sistema. Siguiendo el mismo impulso normativo de automantenimiento, podrían desarrollarse organizaciones más complejas de seres vivos: la capacidad de percibir más matices para moverse mejor y de moverse en función de lo que se percibe continuamente puede ser cada vez más sofisticada, ya que los organismos necesitan controlar su movimiento hacia los recursos y alejarse de las amenazas. De este modo, surgen patrones de comportamiento sensoriomotor.

Sin embargo, aunque surjan patrones sensoriomotores, no ha surgido una motivación normativa diferente ni un criterio normativo distinto al biológico: el agente sigue siendo movido a interactuar con el entorno para preservarse como organismo vivo, y puede diferenciar los acontecimientos del entorno sólo en esos términos. La agencia en este caso sigue siendo predominantemente biológica, lo que significa que dentro del ámbito de todas las interacciones sensoriomotoras que en principio podrían ser posibles para el agente en su entorno, el agente sólo puede diferenciar las que tienen consecuencias para su vida, puede estar motivado sólo para ejecutar acciones con este tipo de consecuencias y, por tanto, sólo éstas pueden convertirse en hábitos. Los esquemas sensoriomotores que se convierten en hábitos tienen como objetivo preservar la vida del agente; entonces, la única fuente de normatividad externa para evaluar las interacciones sensoriomotoras son sus consecuencias para la preservación de la vida del agente. En otras palabras, el hábito se constituye para sostener la identidad biológica del organismo y, si la norma general es la autosustentación, no hay razón para hacer más distinciones normativas más allá de la biológica.

Entre los hábitos sensoriomotores que sostienen la vida de un agente, se pueden diferenciar los que necesitan ser aprendidos de otros, –por ejemplo, la caza en algunos mamíferos–, de los que pueden surgir únicamente de las posibilidades fisiológicas del organismo y su ambiente, –por ejemplo, el agarre en todos los primates–. Independientemente de que el hábito se forme en la interacción con otros o a través de la interacción individual repetida con el entorno, los agentes individuales están motivados a ejecutar el esquema sensoriomotor con suficiente frecuencia para que se convierta en un hábito por la necesidad de preservar sus vidas individuales. Por lo tanto, hasta ahora no hay ninguna razón para que el organismo desarrolle una agencia desvinculada de sus necesidades de automantenimiento. ¿Cómo pueden entonces surgir en la naturaleza hábitos sensoriomotores y comportamientos sociales irrelevantes o incluso dañinos para la vida? Esta pregunta es crucial porque, como vimos anteriormente, la idea de sentido relevante para la cognición es un sentido que va más allá de lo necesario para mantener la vida. Esta pregunta se resume en la Figura 5.

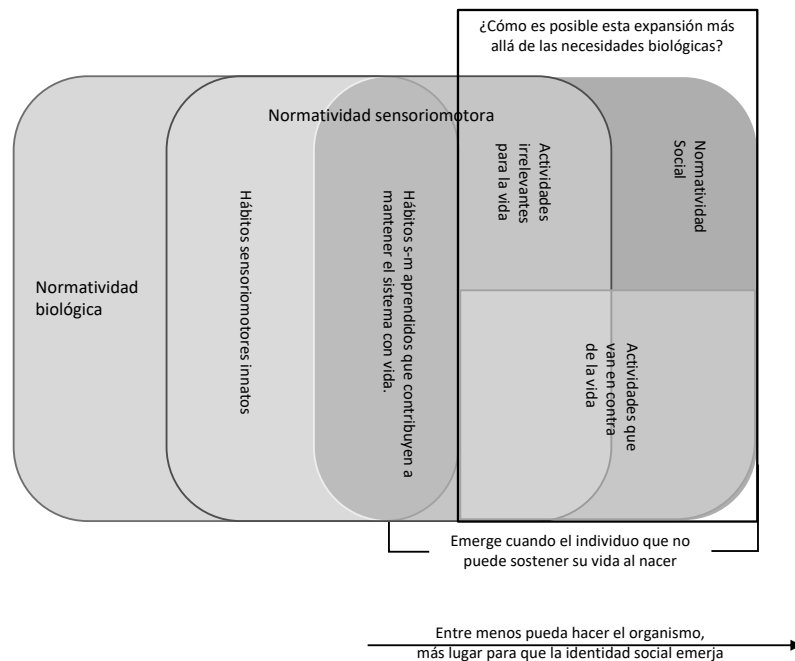


Figura 5

El principio de automantenimiento permite explicar la emergencia de los primeros tres espacios normativos de la izquierda. Sin embargo, no es suficiente para explicar la emergencia de espacios normativos en donde las diferentes posibilidades no hacen una diferencia para el mantenimiento de la vida de los organismos. Estos últimos espacios normativos son los dos últimos de la derecha.

Froese y Di Paolo (2009) apuntan a una pregunta similar. Recuperando la crítica de Clark (2000), señalan que en las explicaciones enactivas que se habían ofrecido hasta entonces, había una brecha entre formas mínimas de cognición y formas más complejas. No parece ser el caso que los mismos tipos de organización y mecanismos que permiten dar cuenta de formas de vida simples, y en particular, de puntos de vista regidos por la normatividad biológica, permitan explicar también problemas más abstractos o de más alto nivel cognitivo. Froese y Di Paolo coinciden con Clark en que la complejidad de las interacciones es crucial para cerrar esta brecha, pero señalan que estas no son interacciones con objetos como suele ser el caso en los ejemplos y consideraciones de la mente extendida –ej., una calculadora, locuciones o marcas en un papel que hacen parte del lenguaje–, sino que frecuente y crucialmente, estas son interacciones con otros agentes. Así que proponen recurrir a las interacciones entre agentes para cerrar la brecha que señalaron entonces.

Si bien el concepto de hacer-sentido participativo ya había sido propuesto para ese entonces, Froese y Di Paolo argumentan que el enfoque en las dinámicas del sistema que emerge en la interacción debe ser complementada con el análisis del sentido que cada agente puede construir en tal interacción. De esta manera sería posible explicar cómo surgen formas de

cognición complejas en los individuos y no solo cómo surgen las interacciones entre ellos. Para abordar la pregunta por el sentido, Froese y Di Paolo adoptan la perspectiva predilecta del enactivismo, aunque ciertamente no la única posible: la fenomenología; y delimitan el problema a una pregunta más específica, a saber, la pregunta por cómo es posible percibir que un objeto es tridimensional si solo podemos ver una de sus caras a la vez.

La solución que ofrecen tiene dos pasos. El primero es reconocer que fenomenológicamente la percepción está extendida temporalmente, lo que significa que no percibimos una imagen en un instante de tiempo, sino que percibimos eventos que se desenvuelven en el tiempo, y así podemos percibir las distintas caras de un mismo objeto. Este primer paso no requiere de una interacción intersubjetiva y es común tanto a la percepción humana como la percepción de otras formas de vida que interactúan con el mundo sin reconocer que lo que perciben puede ser distinto para otros. El segundo paso es reconocer que las interacciones intersubjetivas permiten al sujeto pasar de una comprensión de aspectos del ambiente como posibilidades para sí a una comprensión de estos como objetos que pueden ser entendidos de forma distinta por otros. Cuando un agente interactúa con el otro y lo reconoce como un agente autónomo con sus propios propósitos, el significado que el mundo adquiere para este deja de estar centrado exclusivamente en sus necesidades individuales y se convierte en un algo externo que puede tener un sentido distinto para otro. En otras palabras, en la intersubjetividad, el sentido de los aspectos del mundo deja de ser exclusivamente presentes para mí, sino que se reconocen como algo que es o puede estar presente para otros. En ese sentido, Froese y Di Paolo caracterizan el paso de formas mínimas de cognición a formas más complejas como el paso de la relación organismo-*umwelt* a yo-mundo-otro.

Por supuesto, el enactivismo autopoietico se ha desarrollado desde entonces. Como vimos en la sección anterior, en la última década ha ofrecido explicaciones del surgimiento de la identidad sensoriomotora y la identidad social y lingüística como peldaños importantes en la constitución de formas más complejas de cognición. Vimos que la sugerencia de incorporar las interacciones intersubjetivas como constitutivas de formas complejas de cognición fue escuchada en la explicación del surgimiento de la identidad social y lingüística, pero pudo prescindirse de ella en la constitución de la identidad sensoriomotora. Sin embargo, ni en el artículo de Froese y Di Paolo ni en los desarrollos más recientes queda claro cómo la expansión de la identidad y la agencia es posible bajo la concepción enactiva de normatividad como automantenimiento del agente; esta es, en otras palabras, la pregunta por cómo puede surgir una normatividad indiferente al mantenimiento de la vida de los organismos.

Preguntar por las condiciones de posibilidad de una normatividad desvinculada de las necesidades del organismo vivo tiene al menos dos dimensiones. En primer lugar, apunta a la cuestión de cómo un agente puede llegar a estar motivado para comportarse según normas irrelevantes para su automantenimiento básico y hacerlo con la frecuencia necesaria para que se conviertan en hábitos. Esto es, en otras palabras, cómo un agente podría adoptar un punto de vista normativo en el que asuntos irrelevantes para su mantenimiento biológico pueden hacer una diferencia para sí mismo. En segundo lugar, parte de las condiciones de posibilidad de una normatividad desvinculada de la vida es que existan criterios normativos que permitan

evaluar una interacción más allá de la impresión inmediata del agente. Estos criterios normativos establecen la diferencia que el comportamiento supone para los agentes, es decir, sus consecuencias prácticas, y permite diferenciarlo entre correcto e incorrecto, adecuado o inadecuado, etc., a pesar de que este no suponga una diferencia práctica o inmediata para su identidad biológica o a pesar de que incluso atente contra esta.

Con respecto al primer punto, a saber, las condiciones de posibilidad de una agencia desvinculada de las necesidades organísmicas, se pueden encontrar algunas pistas en la literatura enactiva.

### 1ra pista: emergencia sensoriomotora

Di Paolo, Buhrmann y Barandiaran ofrecen una explicación detallada en términos de sistemas dinámicos de cómo se constituyen las identidades sensoriomotoras durante el proceso de desarrollo (2017, pp. 76-108, 141-181). A partir de las convergencias de la perspectiva enactiva y el trabajo de Piaget sobre el desarrollo sensoriomotor de infantes, Di Paolo et al. hacen un análisis de la configuración de redes sensoriomotoras desacopladas de las necesidades metabólicas del individuo en términos dinámicos. Para entender los principios de organización propios de las redes sensoriomotoras, consideran una situación simplificada pero que tiene los elementos esenciales para su surgimiento: un agente individual en interacción con un ambiente genérico, más concretamente, la interacción entre un agente y su biberón (Di Paolo et al., 2017, p. 77). Esta suposición metodológica no implica que ignoren la importancia de factores socioculturales en el desarrollo sensoriomotor como ocurre en la realidad. De hecho, Di Paolo et al. reconocen que nuestras capacidades perceptuales están tremendamente influenciadas por nuestra historia de interacciones y nuestra situación dentro de un entorno sociocultural particular (2017, p. 77), y reconocen que virtualmente no hay nada en nuestro mundo ni en nuestra percepción que no haya sido tocado por nuestras normas socioculturales. Sin embargo, considerar un individuo en interacción con un ambiente genérico tiene, para ellos, la ventaja de mostrar los elementos estructurales de la constitución de redes sensoriomotoras y, por ende, de la normatividad propia del nivel sensoriomotor que emerge de estas redes (2017, p. 155). Esta suposición tiene sentido si, como vimos arriba, se asume que la dimensión sensoriomotora y la social de un agente son niveles o dimensiones distintas entre sí.

Di Paolo et al., muestran que en la interacción agente-biberón surge una red sensoriomotora que le permite tomar el biberón. Esta está constituida por la serie de esquemas: succionar-tragar-respirar-succionar-tragar-respirar-succionar-tragar-respirar-etc. Esta serie constituye una red en sí misma que se autoperpetúa en el sentido en que succionar lleva a tragar, y tragar lleva a respirar y al haber respirado permite al agente succionar otra vez, etc. En ese sentido, se forman relaciones estructurales entre esquemas sensoriomotores y en donde la ejecución de unos tiene la función de permitir la ejecución de otros esquemas subsiguientes. Así, Di Paolo et al. afirman como principio general de las redes sensoriomotoras:

El conjunto de dependencias estructurales y funcionales entre esquemas define las condiciones de viabilidad del mantenimiento continuo de la red sensoriomotora, así

como la autoindividuación molecular define cómo la concentración de comida, presión o temperatura afectan la viabilidad de la vida celular (Di Paolo et al., 2017, p. 155)

Entonces la red que forman los esquemas sensoriomotores, que involucra tanto estructuras del ambiente como del organismo, es una red que se autoindividua y se automantiene. Por lo tanto, de estas redes emerge una normatividad intrínseca: procurar las trayectorias del sistema que permitan la repetición de la red y evitar la no repetición.

En suma, la idea es que de la interacción de un agente con su entorno pueden emerger redes de esquemas sensoriomotores que les permiten realizar actividades y de las que emerge una normatividad propiamente sensoriomotora. Esta explicación tiene la virtud de que no apela a entidades internas y amplía la definición de agencia cognitiva como sistema autónomo interactivo a interacciones más complejas que las guiadas por la normatividad biológica. Sin embargo, todavía no queda claro un elemento crucial: cómo ocurre el paso normativo y motivacional del propósito de mantener la propia identidad biológica al propósito de actuar por actuar y actuar con la suficiente frecuencia para que se convierta en un hábito. En otras palabras, entre todas las posibilidades de interacción que existen para un agente en un ambiente antes de que este forme hábitos sensoriomotores, ¿cómo decide con cuáles involucrarse y cuáles llevar a cabo repetidamente?

## 2da pista: incorporación

Una segunda pista importante proviene de descripciones fenomenológicas de cómo normatividades que no surgen del individuo mismo pueden regir su comportamiento autónomo. Por ejemplo, Barandiaran, Di Paolo, y Rohde, en el contexto de la discusión de las condiciones necesarias de la agencia, afirman:

Incluso si el origen de algunas normas no recae completamente dentro del individuo (ej. normas sociales), siempre es el individuo quien las internaliza, actúa de acuerdo con ellas, tiene éxito o falla al hacerlo y, podríamos añadir, el fracaso o el éxito tiene algún efecto sobre este (ej. sobre su identidad socialmente construida) (2009, p. 6)

La manera en que el enactivismo explica el papel de las normas externas en la constitución de las formas de actuar de un individuo es en términos de internalización o incorporación. En ese mismo sentido, en el contexto de la explicación en términos de desarrollo de cómo los humanos pueden aprender el pensamiento abstracto bajo el enfoque enactivo, Di Paolo et al. sostienen:

Los procesos sociales pueden intervenir pero no de una manera determinante en la identidad autoconstituida del agente autónomo (de otra forma, dejaría de ser autónomo) (Di Paolo et al., 2017, p. 244)

Esta observación puede aplicarse a todas las formas de desarrollo de todas las formas de vida gracias a la concepción enactiva de normatividad como automantenimiento del agente

individual. Como vimos arriba, de acuerdo con el enactivismo las normas intrínsecas emergen de una identidad autónoma que se mantiene a sí misma, sean estas biológicas, sensoriomotoras o interactivas. El razonamiento detrás es que, si las normas que rigen el monitoreo y regulación de las interacciones del agente no emergen de su necesidad intrínseca de mantener su propia identidad, la interacción no es realmente adaptativa, no proviene de la naturaleza autónoma del agente y, por tanto, no es una instancia en la que el agente haga sentido en su interacción con el ambiente. Así que para poder establecer interacciones en las que se hace sentido estas deben estar dirigidas a mantener alguna forma de identidad del agente, en este caso, una identidad sensoriomotora o social. Esto, como vimos, implica establecer interacciones con el entorno o con los demás que constituyen dicha identidad. Por ende, ninguna norma que no sea autoproducida por el individuo puede desempeñar un papel constitutivo en las interacciones cognitivas, a menos que sea interiorizada y entonces automantenida por el agente. Si este es el caso, es crucial explicar qué significa interiorizar una norma y cómo esto es posible. Sin embargo, este punto parece contraintuitivo para un enfoque cuyos propósitos incluyen una cuenta no internalista de la mente. Por eso y para resaltar la naturaleza corporizada de la agencia, Cuffari, De Jaegher y Di Paolo prefieren usar el término ‘incorporación’ y no internalización (2018, p. 224).

Como señalé en la sección anterior, en la explicación de la emergencia de las interacciones lingüísticas que proponen Cuffari, De Jaegher y Di Paolo, la incorporación de normas de interacción juega un papel crucial en la constitución de la agencia sociolingüística (Cuffari et al., 2015, p. 1120). La definición de incorporación que ofrecen es la siguiente:

Esto es lo que llamamos incorporación: el punto en el que puede decirse que el cuerpo se ha apropiado un hábito, una prótesis, una herramienta, un estilo, etc. No solo el enactuar el patrón habitual depende del cuerpo autónomo como su fuente de valor y de medios para la acción efectiva, sino que el cuerpo ha confiado algunas de sus condiciones internas de clausura operacional a las regularidades del hábito. De esta forma, la lógica autosustentada de los patrones habituales también se incorpora en la forma de valores no metabólicos (Di Paolo 2005). Vale la pena notar que la incorporación no es simplemente un proceso de aprendizaje de una habilidad donde el éxito sigue una norma pre-existente, sino que es un proceso de transformación corporizada del agente al punto de que la nueva forma de vida incorpora normas que inicialmente solo tenían sentido al nivel de patrones de comportamiento sostenidos (Cuffari et al., 2015, p. 1116).

En *Linguistic bodies* (2018), Di Paolo, Cuffari y de Jaegher usan el concepto de incorporación para explicar el proceso de transformación del agente cuando aprende habilidades sensoriomotoras y cuando es socializado para usar el lenguaje. Por un lado, el proceso de incorporación de habilidades sensoriomotoras ocurre a través de la asimilación y acomodación de situaciones nuevas a patrones de interacción existentes. Aunque muchas de las habilidades sensoriomotoras que aprendemos los humanos las aprendemos en nuestra interacción con otros humanos, estas no requieren de la interacción con otros para ser incorporados. En cambio, el proceso de incorporación de habilidades lingüísticas implica

necesariamente la incorporación de interacciones con otros y, en esa medida, las habilidades lingüísticas de un agente reflejan su historia de interacciones y las formas de vida compartida con otros a las que ha pertenecido (2018, p. 192). Sin embargo, este concepto tiene limitaciones que motivan una definición más precisa de normatividad que la que se ha usado hasta ahora en la literatura enactiva.

## Limitaciones

Considerar las interacciones entre agentes nos permite ver por qué la idea de incorporar normas que no se encuentran totalmente dentro del individuo no puede tratarse de una incorporación total. Como vimos arriba, las interacciones sociales son sistemas autónomos constituidos por la coordinación entre dos agentes. Estas interacciones son sociales cuando no están guiadas exclusivamente por ninguno de los interactuantes, por lo que sus objetivos, intereses y preocupaciones se moldean, negocian y transforman en la propia interacción. La interacción social es tal porque ninguno de los interactuantes hace suyas las normas que guían el comportamiento del otro agente o que guían la interacción como sistema autónomo. Como se reconoce acertadamente en las explicaciones enactivas de la interacción social, incluso en las interacciones altamente convencionalizadas, como las que se establecen a través del lenguaje, las habilidades de los agentes han de ser usadas en la interacción con otro agente con sus propios objetivos, propósitos y preocupaciones (Cuffari et al., 2015). Los interactuantes se coordinan entre sí subsumidos a la normatividad de la propia interacción; así, interiorizar las normas que guían la interacción significaría que ésta deja de ser social y empieza a estar guiada por los objetivos e intereses de uno de los agentes. La esencia de la interacción social es que en ella surgen normas que habrían sido imposibles para los individuos por sí solos, estas normas permiten a los agentes dar sentido al mundo, a los demás o incluso a sí mismos, pero estas normas no son ni pueden ser interiorizadas sin acabar con la interacción social misma.

Las interacciones con otros agentes son la forma paradigmática en que los agentes se incorporan a las prácticas sociales; esto significa que, a través de las interacciones con otros, los agentes aprenden a ajustarse a patrones de actividad compartidos. Como argumenté en la sección 1.3, las consecuencias prácticas son el criterio normativo por excelencia; en el caso del lenguaje, estas consecuencias se establecen mediante patrones compartidos, y en el caso de la función biológica, estas consecuencias son la vida y la muerte, la salud y la enfermedad, la reproducción, la interacción y la similitud con quienes comparten la misma forma de vida dentro de la especie, etc. Esto nos lleva a la segunda dimensión de la pregunta por las condiciones de posibilidad de la normatividad no biológica, a saber, cómo surgen los criterios que establecen consecuencias prácticas para los agentes más allá de sus necesidades biológicas y sus impresiones inmediatas. Si aceptamos la noción wittgensteiniana de significado que defendí, los patrones de actividad socialmente compartidos constituyen tales criterios normativos. Esta respuesta es coherente con la concepción enactiva de interacción y las normas sociales. Sin embargo, esta respuesta también implica que los agentes no pueden incorporar las normas sociales a pesar de que sean constitutivas de la manera en que hacen



sentido en sus interacciones: por definición, se constituyen más allá de las impresiones de un solo individuo.

He considerado hasta ahora las limitaciones de la idea de incorporación de normas cuando se trata de interacciones y normas sociales. ¿Qué pasa con los hábitos sensoriomotores? La suposición de que un agente individual que explora las posibilidades de acción que le ofrece su entorno es suficiente para explicar la emergencia de redes sensoriomotoras tiene una limitación importante: no hay ninguna motivación, razón normativa o explicación causal de por qué un agente seleccionaría un subconjunto de posibilidades de acción sobre todos los demás antes de que se forme un hábito sensoriomotor cuando no supone ninguna diferencia para su mantenimiento biológico y, por ende, no supone ninguna diferencia en absoluto. La explicación enactiva de una red de esquemas sensoriomotores que se automantiene no ofrece una respuesta en este punto porque no hay ninguna red, nada que mantener, no existe todavía ningún hábito que pueda establecer un criterio de normatividad o una motivación para el agente. Las interacciones y prácticas sociales proporcionan precisamente la fuerza normativa externa que hace la diferencia entre las posibilidades de acción como mejores o peores cuando la identidad biológica del agente no está en juego. Estos criterios normativos ejercen una fuerza selectiva sobre el individuo porque hacen una diferencia para sus posibilidades de pertenecer a una forma de vida compartida. Más adelante desarrollaré este punto, pero por ahora basta con notar que esto significa que la interacción con agentes que pertenecen a formas de vida compartidas fundamenta normativamente el desarrollo cognitivo y lo hace posible.

En suma, el problema de la idea de incorporación de normas dentro del enfoque enactivo es el siguiente: las interacciones normativas de los agentes que no están guiadas por sus necesidades biológicas están conformadas por criterios normativos cuya existencia no depende de la existencia de un individuo. Estos criterios normativos, que defiendo en el siguiente apartado, están constituidos por patrones de interacción compartidos que se extienden en el tiempo y, por tanto, no son ni pueden ser incorporados por un agente. Por ende, la agencia no coincide con las normas externas y entonces la incorporación es necesariamente parcial; pero, a su vez, esto implica que el sentido de las interacciones entre un agente cognitivo y su ambiente no están completamente determinadas por su identidad ni por las normas que implican su automantenimiento como agente.

Podemos volver ahora a la afirmación con la que abrí esta sección: la propuesta enactiva de naturalización de la normatividad tiene una brecha en la transición de la normatividad biológica centrada en el mantenimiento del individuo a las formas de normatividad orientadas a la acción misma, es decir, los hábitos sensoriomotores, y orientadas a las interacciones con los demás y la vida social. Esta brecha hace eco del vacío en el continuo entre vida y formas complejas de mente que señalaron Froese y Di Paolo ya en el 2009. La brecha se refiere a dos vacíos. Primero, no hay una explicación de lo que hace que, entre las diversas posibilidades de acción que tiene un agente en su entorno, determinados esquemas sensoriomotores se conviertan en un patrón de interacción en primer lugar y no otros. No hay tal explicación porque estos esquemas no suponen una diferencia para la identidad biológica del agente, que

es la única identidad que éste tiene antes de desarrollar sus propios hábitos sensoriomotores. Segundo, no es claro cómo los agentes pueden participar de forma autónoma en interacciones con el mundo y los demás y darles sentido siguiendo una normatividad cuyos criterios son, por definición, imposibles de incorporar completamente. Este es el caso cuando los agentes se involucran en interacciones sociales con otros agentes que siguen normas sociales.

### 4. 3.Cerrar la brecha: normatividad relacional y situada

Afortunadamente, esta brecha no es insuperable. Un paso en esa dirección es refinar la noción de normatividad adoptada por el enactivismo autopoiético. Una comprensión explícita de la normatividad como fenómeno relacional nos permite entender cómo es posible que un agente interactúe normativamente con los demás y con el mundo sin incorporar a su identidad normas generadas externamente. Además, una definición más clara de normatividad nos permitiría tender un puente entre la forma biológica de la normatividad que pretende preservar la vida del individuo o la forma de vida de una especie y las formas de normatividad que no tienen impacto en la vida biológica de los individuos o las especies. En esta sección propongo una definición de normatividad que busca recoger los elementos que he discutido hasta ahora y superar las limitaciones que he señalado.

La definición de normatividad debe cumplir varios requerimientos. Primero, debe capturar la idea, que defendí en el capítulo 1, de que la normatividad requiere criterios normativos y una agencia mínima. Segundo, tal definición debe capturar la idea de que el fenómeno de la normatividad requiere la posibilidad de que el agente pueda seleccionar diferentes cursos de acción que hagan una diferencia práctica. Tercero, debe prestarse para una explicación naturalista de los fenómenos normativos en general y eventualmente de la cognición. Esto incluye, por supuesto, la normatividad biológica, es decir, la determinación de la función de los procesos constitutivos de un organismo, sus interacciones con el ambiente para mantenerse con vida y el papel de la historia evolutiva de la especie a la que pertenece el organismo.

Para ello, propongo una concepción de normatividad relacional y situada. La idea central que defiende tiene dos partes. Primero, los criterios normativos siempre son establecidos por un sistema autónomo y la norma que establecen es la de su automantenimiento —estos sistemas son, por ejemplo, un organismo vivo, una especie en la escala evolutiva, un hábito sensoriomotor o una práctica social—. Segundo, los comportamientos correctos o incorrectos, mejores o peores, etc. son seleccionados por un agente, es decir, un sistema autónomo y adaptivo, y tales comportamientos son correctos o incorrectos en función de si contribuyen o no al mantenimiento del sistema autónomo al que pertenecen —se evalúan respecto a si contribuyen al mantenimiento del agente mismo, de la especie a la que pertenece o de los hábitos y prácticas de las que participa—. Propongo operacionalizar esta idea en la siguiente definición de fenómeno normativo:

Hay una interacción normativa entre un sistema  $p$  y  $x$  cuando  $p$  constituye un agente, y su relación con  $x$  pertenece a un sistema de automantenimiento  $S$ . La interacción  $p$ - $x$

es correcta si contribuye al mantenimiento de  $S$  y es incorrecta si va en contra de su mantenimiento.

Donde

- $p$  es un agente que, según la definición del enactivismo, es un sistema precario, autodiferenciado y automantenido que puede controlar y regular sus interacciones en función de lo que necesita para preservarse.
- $x$  es un objeto de interacción, esto es, un proceso interno, un aspecto material del entorno u otros agentes.
- $S$  es un sistema de automantenimiento al que pertenece la interacción  $p-x$ , y puede ser un patrón de formas compartidas de hacer las cosas o la propia  $p$ .

Hay varios aspectos de esta definición que quisiera resaltar. Primero, el sistema autónomo que establece los criterios normativos, es decir, la norma de su automantenimiento, no es necesariamente el mismo que el sistema autónomo (y adaptivo) que selecciona el comportamiento o la relación correcta o incorrecta y esta distinción es crucial porque, veremos más adelante, permite dar cuenta de comportamientos correctos que son indiferentes a o van en contra del mantenimiento del agente. Segundo, ni el comportamiento del agente ni aquello con lo que se relaciona pueden ser normativos por sí mismos, tampoco un órgano o un proceso pueden ser normativos en sí mismos, sino solo en relación con un agente individual completo. Así, los fenómenos normativos surgen sólo en la interacción de un agente autónomo y adaptivo con su entorno o en su relación con sus propios componentes. Tercero, los agentes que establecen las interacciones normativas son agentes capaces de seleccionar el curso de sus interacciones y esto ha sido operacionalizado como sistemas autónomos y adaptivos que son, hasta ahora, solamente organismos vivos. En lo que sigue, muestro cómo funciona la definición de normatividad que propongo para dar cuenta de diferentes tipos de normatividad y cómo explica la posibilidad de que se establezcan normas que van en contra del automantenimiento del agente.

### Normatividades biológicas

Consideremos la función de los órganos, por ejemplo, el corazón. Según la definición, existe una interacción normativa entre el organismo ( $p$ ) y su propio corazón ( $x$ ) porque el organismo  $p$  constituye un agente que monitorea y regula sus interacciones respecto a lo que parece amenazar su identidad en un lapso dado, y su relación con su propio corazón  $x$  pertenece al sistema de automantenimiento,  $p$  mismo, de modo que  $S=p$  en este caso. Hay que tener en cuenta que dicho sistema de automantenimiento es el organismo en su conjunto, no ninguno de sus órganos constitutivos, procesos, subsistemas, etc. En este caso, los criterios normativos son los límites de viabilidad del organismo en la escala temporal de su vida individual, el agente es el organismo en su conjunto, y la interacción normativamente evaluable es la contribución y las respuestas del corazón al resto del cuerpo en el aquí y el ahora. Esta interacción es correcta si el corazón bombea suficiente sangre y a un ritmo constante al resto

del sistema, contribuyendo así al mantenimiento del organismo al que pertenece, y es incorrecta en caso contrario.

La relación entre un sistema vivo autónomo ( $p$ ) y la red que forman sus órganos y procesos constitutivos ( $x$ ) es un caso especial porque capta la emergencia de la normatividad en la naturaleza. Lo que lo hace especial es la relación circular que se establece entre  $p$  y  $x$ : la existencia de  $p$  como sistema autónomo a lo largo del tiempo depende, entre otras cosas, de la función adecuada de  $x$ , y la existencia y la función adecuada de  $x$  en el presente dependen de la capacidad de  $p$  para autosostenerse. Así, la relación que el sistema completo  $p$  tiene con  $x$  hace una diferencia con respecto a la propia existencia de  $p$  como entidad individual que se mantiene en el tiempo. Esta relación circular captura la normatividad propia de los sistemas cuya identidad depende de su propia actividad y que son adaptativos, es decir, así como los define el enactivismo autopoietico, que pueden monitorear y regular sus propias tendencias para mantenerse continuamente existiendo como individuos. Estos sistemas establecen interacciones que son intrínsecamente normativas e incluyen la clase de sistemas adaptativos autónomos. Esto es cierto para todos los sistemas vivos y sólo para ellos.

Existe un sistema autónomo más amplio  $S$  al que pertenece la interacción organismo-corazón: los patrones de interacción de los organismos de la misma especie y sus propios corazones. Estos patrones se extienden a lo largo del tiempo evolutivo y son moldeados y mantenidos por la vida de los organismos de la especie correspondiente a lo largo de las generaciones. Así, la relación de un organismo individual con su propio corazón, es decir, el comportamiento del corazón en un organismo, puede evaluarse en función de la continuidad de su comportamiento con el de las generaciones pasadas, esto equivale a un criterio de normatividad evolutivo, así como en función de su capacidad para contribuir a la continuación futura de la especie. En suma, el funcionamiento del organismo está sujeto a la normatividad impuesta por el automantenimiento de la especie como sistema autónomo.

Una de las motivaciones para proponer esta definición de normatividad para el enactivismo autopoietico fue que la definición actual en términos de automantenimiento del agente hace difícil entender los casos de comportamientos exitosos que van en contra del mantenimiento de la agencia individual. La definición de normatividad que propongo permite aclarar estos puntos. La razón es que permite que una determinada interacción  $p-x$  sea normativa con respecto a las condiciones de mantenimiento de un sistema  $S$  cuyas condiciones de existencia no coinciden con las de  $p$  y que incluso podrían ser independientes de la interacción particular  $p-x$ . Por tanto, las condiciones externas de normatividad impuestas por  $S$  pueden hacer que una interacción sea normativa sin que estas normas sean incorporadas por el agente o incluso vayan contra este.

Para ilustrar el conflicto entre normas biológicas evolutivas e individuales, consideremos el caso de la reproducción de los pulpos cuyo apareamiento detona cambios hormonales regulados por la glándula óptica que los llevan a la muerte. Por un lado, la hembra deja de comer después de poner los huevos para cuidar de ellos hasta que esta muere. El dejar de comer y dedicarse por completo al cuidado de los huevos va directamente en contra de sus necesidades metabólicas para mantenerse con vida; aun así, es correcto que lo haga porque de

lo contrario los huevos no se desarrollarían y morirían. Por otro lado, el apareamiento dispara el proceso de envejecimiento del macho y muere porque deja de comer y porque pierde la habilidad de escapar de sus depredadores. En este caso, lo que lo lleva a la muerte no tiene una función de cuidado de la siguiente generación, pero es un proceso normal, es decir, si no ocurriera iría en contra de lo que se supone que debe pasar en la vida de los pulpos. En ambos casos, los pulpos seleccionan sus trayectorias de comportamiento, incluso las que los llevan a la muerte después del apareamiento. La definición de normatividad que propongo permite explicar el carácter doble de este proceso: es exitoso con respecto a los criterios normativos impuestos por la especie en tanto sistema autónomo en la escala evolutiva, pero es fallido respecto a los criterios normativos impuestos por el organismo individual de cada pulpo en tanto sistema autónomo. Esto es una ventaja respecto a la aproximación enactiva actual en donde la explicación del fenómeno de la reproducción no es clara y menos aún de procesos reproductivos que llevan a la muerte como el que acabamos de considerar: no es evidente como procesos tan costosos y con ningún beneficio para el organismo individual podrían ocurrir en la naturaleza si mantuviéramos la idea de enactiva de normatividad como un asunto esencialmente de autopreservación del agente individual.

### Normatividades sensoriomotoras

Dos puntos: hábitos sensoriomotores y brecha entre formas básicas de normatividad y cognición sofisticada

Consecuentemente con la noción de normatividad como automantenimiento, el enactivismo explica la constitución de hábitos sensoriomotores como el mantenimiento de una identidad. Esta explicación descansa en la idea de la identidad individual como un entretejido de microidentidades sin un yo inaugurada por Varela (1991) que más recientemente, se han conceptualizado como constituidas por los hábitos sensoriomotores del agente (Di Paolo et al., 2017). La concepción de normatividad que propongo tiene el potencial de refinar esta idea: podría argumentarse que la identidad de un individuo coincide con su agencia, y en esa medida se sitúa en el cuerpo vivo que es capaz de seleccionar entre diferentes cursos de acción, mientras que los hábitos son criterios normativos que rigen las interacciones que los agentes establecen y son tales criterios los que le dan las múltiples capas a la identidad de un mismo agente. Veamos primero cómo las ideas de hábito y su relación con la identidad han sido caracterizadas en el enactivismo.

Di Paolo, Buhrmann y Barandiaran (2017), por ejemplo, caracterizan los hábitos como redes autónomas porque están constituidos por comportamientos que se repiten y tal repetición favorece su repetición futura y, por ende, su mantenimiento en el tiempo. De tal carácter autónomo se ha derivado la idea de que los hábitos pueden constituir una suerte de identidad por ellos mismos; Barandiaran, por ejemplo, ha sugerido tal idea:

Los hábitos exhiben formas de autoproducción circular análogas a otras formas autónomas como la autopoiesis. Un hábito individual, defiende Barandiaran (2017), provee una primera analogía con la vida y una primera aproximación a la concepción sensoriomotora de identidad y normatividad en la que, a través de la repetición, un

hábito puede tomar vida por sí mismo: es al tiempo causa y consecuencia de su propia enacción (James, 2020, p. 5).

Los hábitos están constituidos no solo por los movimientos iniciados por el agente sino por todas las estructuras del ambiente que sostienen y hacen posible tal repetición en el tiempo (Di Paolo et al., 2017). Así, por ejemplo, el hábito de pintar está sostenido no solo por los movimientos del agente de dibujar, mezclar y aplicar pintura, sino también por los lienzos, los pigmentos, aceites y solventes, los lápices y demás herramientas involucradas en el hacer pinturas. En ese sentido, se ha reconocido que las identidades sensoriomotoras del agente no terminan en los límites de su cuerpo biológico, sino que se extienden también al ambiente que posibilita su ejecución (Di Paolo et al., 2017).

Sin embargo, de acuerdo con la definición de normatividad que defiende, un hábito por sí mismo no constituye un sistema con agencia y, en esa medida, al contrario de lo que el término microidentidad sugiere, tampoco constituye propiamente una identidad. Cuando un agente  $p$  interactúa con un determinado aspecto del entorno  $x$ , dicha interacción forma parte de un sistema autónomo más amplio, a saber, el patrón compartido de hacer las cosas de su comunidad, su especie biológica o, en general, su forma de vida, sin que el agente tenga que incorporarlo de alguna manera y mucho menos saberlo explícitamente. Dicha interacción está normativamente sometida a los criterios impuestos por el patrón compartido de hacer las cosas de la forma de vida a la que pertenece, es decir, a una norma externa compartida con otros que está constituida por una forma de vida  $S$ .

En el caso de los humanos y otras especies cuya forma de vida es social, la razón por la que estos patrones son recogidos por agentes y repetidos hasta convertirse en hábitos sensoriomotores es que son reforzados por otros, ya sea como condicionamiento de la satisfacción de las necesidades biológicas básicas del agente a la adopción de estos patrones o mediante el reconocimiento por parte de otros agentes, normalmente los cuidadores, e interacciones con ellos. Por ello, para entender las condiciones de posibilidad del surgimiento de las redes sensoriomotoras, en particular, sus condiciones normativas, es fundamental considerar los patrones de actividad compartidos por los individuos con quienes se desarrolla y vive el agente. Así, permitir que las normas constituidas externamente sean constitutivas del carácter normativo de las interacciones y hábitos sensoriomotores sin que sean o tengan que ser incorporadas por el agente nos permite ver cómo es posible que surjan hábitos sensoriomotores irrelevantes para la viabilidad de la identidad biológica. Muestra, además, la naturaleza intersubjetivamente constituida de nuestra agencia humana que, si bien resuena con la convicción no-individualista y situada de la aproximación enactiva, no había sido capturada hasta ahora sino solo para las formas de vida más complejas y más explícitamente sociales como las humanas.

Para ilustrar cómo mi definición de normatividad da cuenta del surgimiento de hábitos sensoriomotores irrelevantes o incluso dañinos para el automantenimiento del agente quisiera considerar de nuevo el caso de la pintora. Arriba vimos que su hábito se configura porque lleva a cabo constantemente los esquemas sensoriomotores que le permiten dibujar y pintar, y esto ha llevado a que tanto sus músculos y posibilidades fisiológicas como la disposición

material de su ambiente hagan más fácil que esta actividad se siga repitiendo. La consolidación de este hábito no solo es irrelevante para su automantenimiento como agente, sino que puede ser dañina en el corto y el largo plazo: los movimientos repetitivos del dibujo y la pintura le generan una lesión en la muñeca que le provoca dolor y limita su movilidad progresivamente; además, en todas sus sesiones fuma cigarrillos que le dan la sensación de concentración pero que, en el corto plazo, disminuyen la cantidad de oxígeno que llega a su cerebro, su capacidad pulmonar y su sentido del olfato y le provocan congestión, y al largo plazo, dañan progresivamente su capacidad pulmonar, la capacidad general de su cuerpo de curarse a sí mismo y de combatir agentes dañinos externos, etc. y hacen mucho más probable que desarrolle cáncer.

El enactivismo explica el mantenimiento de estos hábitos como constitutivos de la identidad del agente como pintora y fumadora, pero es difícil entender por qué un hábito como este, que claramente atenta contra su agencia más fundamental, es decir, su agencia biológica, puede consolidarse en primer lugar. Si la definición de normatividad que propongo está en lo correcto, las rutinas de pintura de la pintora hacen parte de uno o varios sistemas autónomos más amplios que establecen los criterios normativos bajo los que es correcto que ejecute movimientos repetitivos con su mano y que fume mientras pinta. Este sistema autónomo más amplio es la práctica social de la pintura de la que ella participa y en la que fue socializada cuando aprendió a pintar o bien interactuando con otros en una escuela o bien siguiendo manuales y consumiendo productos culturales. Puede que el fumar también sea parte de la forma en que se vive como pintor en la práctica en la que participa y que el fumar sea para ella solo parte de ese aspecto de su vida; pero puede que fume también sea parte de las formas culturalmente mediadas en las que aprendió a expresar su ansiedad o su aburrimiento. Parte del proceso de socialización en la que la pintora adquiere sus habilidades como pintora y fumadora es que aprenda a usar su capacidad sensoriomotora de respuestas adaptivas para contribuir al mantenimiento de estas prácticas. Esto significa que aprenda a reconocer las situaciones en las que es pertinente llevar a cabo la actividad para mantener el hábito, atender a los aspectos del ambiente relevantes para hacerlo y saber cómo ejecutar la actividad correctamente en esas circunstancias.

Arriba vimos que Froese y Di Paolo (2009) señalan una brecha en el continuo que defiende el enactivismo entre formas básicas de cognición que emergen de las mismas estructuras que soportan la vida y formas de cognición más complejas, en particular, la cognición humana. Vimos que proponen una solución desde la perspectiva fenomenológica en donde se reconoce, primero, que la perspectiva del agente en su interacción con el ambiente es extendida temporalmente y, segundo, que las interacciones con otro agente permiten al individuo descentrar su comprensión de aspectos del ambiente y verlos como objetos que significan diferentes posibilidades de acción para otros agentes. La razón es que las interacciones intersubjetivas abren dominios de hacer-sentido que serían imposibles para cada agente individualmente. La concepción de normatividad que propongo recoge estos dos elementos y los integra a una concepción más general de normatividad como fenómeno extendido temporalmente y que en muchos casos no está completamente centrado en un individuo. Veamos cómo.

Por un lado, tanto las interacciones que establece un individuo que son el objeto de la normatividad ( $p-x$ ) como los sistemas autónomos que constituyen sus criterios de corrección ( $S$ ) son procesos que se desarrollan en el tiempo. La expansión temporal de las interacciones es evidente en interacciones entre dos o más agentes en las que la coordinación misma al, por ejemplo, bailar o tener una conversación, requiere tiempo para emerger y mantenerse. Análogamente, la interacción entre un agente y un aspecto material del ambiente requiere tiempo para que el agente pueda explorarlo, es decir, coordinarse con este y dominar las regularidades de la interacción. Esto es evidente en la interacción sensoriomotora que un agente establece con una esponja para determinar su suavidad, de ahí que sea uno de los ejemplos clásicos para mostrar que la percepción del ambiente consiste en la coordinación sensoriomotora de un agente con el ambiente material; pero también es el caso en otras situaciones menos evidentes como en los movimientos sacádicos de los ojos que se coordinan con los patrones de luz disponibles en el ambiente de los que depende la percepción visual (O'regan & Noë, 2001a) o en el movimiento de las células ciliadas del oído que se mueven para modular los sonidos del ambiente de los que depende la audición (Froese & González-Grandón, 2019). Finalmente, incluso las relaciones entre el agente con sus partes constitutivas, como sus propios órganos y procesos metabólicos, también se desarrollan y mantienen en el tiempo en tanto que los elementos constitutivos del sistema deben ser renovados constantemente por la actividad misma del sistema y, así mismo, el sistema se mantiene en el tiempo en la medida en que sus elementos constitutivos, que son procesos, se mantengan activos.

Sin embargo, el carácter extendido temporalmente de la normatividad no se limita a que las interacciones normativas se desenvuelven en el tiempo. Estas interacciones, sean agente-agente, agente-ambiente o agente-procesos constitutivos, hacen parte de un sistema autónomo que establece los criterios de corrección y es una red de procesos que se extiende temporalmente. Este último sistema autónomo existe en una escala temporal más amplia que la escala temporal de las interacciones particulares y en ese sentido va más allá de las interacciones específicas ( $p-x$ ) que lo constituyen y a las que impone criterios normativos. Esto es evidente cuando el sistema  $S$  al que pertenece cualquiera de estas interacciones consiste en la forma de vida compartida a la que pertenece el individuo, por ejemplo, una especie o una comunidad. Pero también es el caso cuando el sistema autónomo que establece los criterios de normatividad es una interacción con otro agente o el agente mismo en tanto organismo individual. La razón es que estos sistemas autónomos solo pueden servir como criterios de corrección para las interacciones individuales que los constituyen, y por ende son siempre temporalmente más amplios que las interacciones específicas. En ese sentido, también los criterios de normatividad son extendidos temporalmente.

### Normatividades interactivas

En la interacción social, cada uno de los interactuantes es un agente individual autónomo, llamemos a estos agentes  $p_1$  y  $p_2$ . La relación que se evalúa normativamente en este caso es la coordinación entre ellos,  $p_1-p_2$ , que contribuye al mantenimiento del sistema autosostenible  $S$  que la interacción entre estos agentes constituye a lo largo del tiempo (es decir,  $S =$  patrón



constituido por  $(p_1-p_2)$  a lo largo del tiempo). Nótese que las normas impuestas por la interacción  $S$ , es decir, lo que es bueno para la continuación de la interacción, no coinciden con las normas de la existencia individual de cada uno de los agentes interactuantes ni implican necesariamente que los objetivos y propósitos de los agentes coincidan. Sin embargo, el espacio de posibilidades de acción de cada agente debe permitir tanto la coordinación con el otro para contribuir al sostenimiento de su interacción, como el ver al otro como un agente cuyos propósitos y cuya perspectiva son irreducibles a la propia. Así, bajo la definición de normatividad que propongo, los agentes pueden participar y sostener una interacción social y hacer sentido juntos y del otro sin tener que incorporar las normas de la interacción o las normas de automantenimiento del otro agente.

Si bien una interacción entre dos agentes o más puede convertirse en un sistema autónomo por sí misma e imponer criterios normativos a los interactuantes, las interacciones humanas suelen hacer parte de un sistema autónomo más amplio: nuestras prácticas sociales. Trataré este punto en el capítulo siguiente.

Para terminar esta sección, quisiera señalar que queda una pregunta importante abierta, pero que tendrá que esperar para el próximo capítulo. La pregunta por en qué consiste la agencia bajo la concepción de normatividad que propongo. Un primer paso para responder es reconocer que, aunque los agentes no incorporan completamente normas generadas externamente cuando aprenden esquemas sensoriomotores o a coordinarse con otros y los convierten en sus hábitos, sí adquieren habilidades para actuar que cambian sus posibilidades y su identidad como agentes, y que les permiten buscar y regular sus interacciones de formas que antes les eran imposibles. Este punto de vista ampliado podría caracterizarse como el impulso de mantener la fluidez de las interacciones, de pertenecer a los demás y de tener sentido con y para los demás en el contexto específico que cohabita con otros. Como señalé ya desde el Capítulo 2, la constitución de la agencia implica la constitución de un individuo capaz de tener un comportamiento intrínsecamente normativo. Esto implica que el individuo tiene una motivación intrínseca para actuar y que no es indiferente a la forma en que transcurran sus interacciones.

### Tipos de normatividad y tensiones

Hemos visto hasta ahora que el comportamiento normativo es un fenómeno que surge de la interacción entre un agente autónomo (esto es, un sistema que se automantiene de forma recursiva) y alguna otra cosa (un órgano o proceso constitutivo, un aspecto material del ambiente u otro agente), y tal interacción se sitúa en el contexto más amplio de un sistema autónomo (esto es, el agente mismo o una forma de vida). Vimos que hay varios tipos de fenómenos normativos. Vimos una primera clasificación propuesta por el enactivismo en la que los comportamientos normativos se diferencian según si siguen una normatividad biológica, una sensoriomotora o una social/interactiva. Ahora que hemos diferenciado dos dimensiones distintas de todos los fenómenos normativos, es posible hacer una clasificación un poco más fina tomando cada una de estas dimensiones como criterio. En ese sentido tenemos la siguiente clasificación:

1er criterio: aquello con lo que interactúa el agente ( $p-x$ ):

- $x$  es un proceso constitutivo del organismo mismo, i.e., órgano o proceso interno. La interacción consiste en una relación circular de constitución.
- $x$  es un aspecto material del ambiente. La interacción consiste en la coordinación sensoriomotora agente-ambiente.
- $x$  es otro agente vivo. La interacción consiste en la coordinación sensoriomotora agente-agente.

2do criterio: sistema autónomo que hace posible la interacción, en la que esta se inserta y respecto del cual se evalúa su corrección ( $S$ ):

- $S$  es el organismo mismo en tanto sistema autónomo. Escala temporal de la vida individual.
- $S$  es la forma de vida compartida por una especie en la escala evolutiva. Escala temporal de la evolución, selección natural.
- $S$  es la forma de vida compartida por una comunidad de agentes. Escala temporal de la historia.

Por supuesto, los límites entre estas diferentes formas de fenómenos normativos son borrosos, es decir, hay interacciones que están sujetas a más de una clase de criterios de normatividad que incluso pueden estar en conflicto. Este es el caso, por ejemplo, cuando aquello con lo que el individuo establece una relación son sus propios órganos o procesos constitutivos. Tales interacciones están siempre sujetas a los criterios de normatividad establecidos por el organismo mismo como sistema autónomo; pero también, en una vasta mayoría de casos, están sujetos a los criterios establecidos por la forma de vida compartida por la misma especie, es decir, las funciones establecidas por la selección evolutiva. Los procesos biológicos están sujetos a criterios de normatividad que, en ocasiones, se oponen de manera que, ej., lo que es bueno para el organismo como sistema autónomo no es correcto para la forma de vida compartida por una especie –este es el caso en la forma de reproducción de la mantis religiosa en la que la hembra, en ocasiones, se come la cabeza del macho y lo mata, o es también el caso en la reproducción de los pulpos quienes mueren cuando nacen sus crías.

Una ventaja de distinguir criterios normativos de agencia es que permite diferenciar varios criterios de normatividad que pueden estar en juego en las acciones humanas sin equipararlos a una sola motivación de mantenerse a sí mismo como sistema individual. El objeto de la normatividad son las interacciones que un agente selecciona, pero una misma interacción puede estar sujeta a varios criterios de normatividad constituidos por diferentes sistemas autónomos de los que hace parte. Por ejemplo, cuando una persona consume alimentos, este acto individual está sujeto a criterios normativos biológicos que determinan si el alimento contribuye o no energía y nutrientes necesarios, si contribuye o no a mantener ciclos estables de producción de insulina, si está echada a perder y debe ser expulsada inmediatamente o si por el contrario es inofensiva para el organismo, si trae una sensación corporal de energía o somnolencia etc. Sin embargo, estos no son los únicos criterios que determinan si un evento

de consumo de alimentos es correcto, adecuado, aceptable, etc. Podemos distinguir al menos otros dos criterios normativos más en juego.

Por un lado, lo que comemos ahora hace parte de una amplia historia evolutiva que resultó en la configuración contemporánea de nuestra dieta omnívora. Esta forma de comer se mantiene en la medida en que la especie humana ha continuado consumiendo frutas, verduras y productos animales; sin embargo, estos patrones pueden cambiar en la medida en que los patrones de alimentación de la humanidad también lo hagan y entonces los criterios normativos que rigen nuestras formas de alimentación también cambien: así como ya transitamos una vez en la historia a una dieta de alimentos procesados y ultraprocesados, podríamos imaginar un cambio futuro a una dieta vegana. Por otro lado, la corrección y el significado de lo que comemos y cómo lo hacemos está altamente determinado por nuestras prácticas sociales; estas nos permiten hacer un sinnúmero de distinciones normativas sobre lo que comemos y cómo lo hacemos: consideramos aceptable y de buen gusto comer carne de vacas y cerdos pero consumir caballos, gatos y perros es repugnante y hasta éticamente reprochable, debemos usar cubiertos para comer pasta pero no para comer pizza, hay quienes consideran pecaminoso o prohibido comer ciertos alimentos ultraprocesados como donas o ramen pero no otros como batidos o barras de proteína, etc.

Todos estos criterios normativos convergen en nuestras interacciones particulares alrededor de la comida y constituyen sus múltiples capas normativas y de significado, y esto es posible, aún sin que el agente las tenga presentes en su interacción. Bajo la concepción de normatividad que propongo, esto es posible porque la interacción contribuye al mantenimiento de diferentes sistemas autónomos que, a su vez, constituyen las diferentes capas de normatividad que le dan sentido a la interacción y la hacen correcta o incorrecta en distintos niveles. En esa medida, la concepción de normatividad que propongo logra integrar y aclarar las diferentes fuentes de normatividad a las que está sujeto un agente individual en sus interacciones, a saber, su automantenimiento como organismo individual, como miembro de una especie y como miembro de un grupo social, y que pueden entrar en conflicto sin necesidad de que el agente las internalice.

La concepción de sentido y cognición que defendí en el Capítulo 1 implica que, dentro de esta variedad de fenómenos normativos, los cognitivos son paradigmáticamente los que están sujetos a criterios de normatividad impuestos por la forma de vida compartida por una comunidad. En el próximo capítulo regresaré al concepto de forma de vida. Por ahora, quisiera cerrar esta sección discutiendo la relación circular entre sistemas autónomos que establecen criterios de normatividad (S), y las interacciones que los constituyen en tanto sistemas y sobre las que tales criterios ejercen una fuerza normativa. Esta relación se asemeja a la causalidad circular, en particular, a la concepción de causalidad circular vertical que propone Thomas Fuchs en *Ecología del cerebro*.

### Constitución de los sistemas autónomos

La causalidad circular vertical ocurre dentro del organismo y entre los diferentes niveles de elementos que lo componen: átomos, células, órganos y organismo como un todo. Esta

causalidad no ocurre entre dos entidades distintas sino entre un todo y sus partes, y ocurre en ambas direcciones. Primero, cada nivel, es decir, célula, órgano, organismo completo, etc., hace posibles los procesos del nivel inferior porque abre y cierra posibilidades a los elementos de este nivel en función del papel que cumplen dentro del nivel superior al que pertenecen. Esta relación es de arriba a abajo o top-down. Por otro lado, cada nivel existe y es capaz de llevar a cabo procesos porque los elementos y sub-procesos que lo componen se organizan de manera que en conjunto constituyen tal nivel superior. Los procesos mentales, en tanto actos de un organismo vivo, y la neurofisiología del agente se relacionan precisamente de esta forma. Fuchs ejemplifica este punto esbozando una explicación de un acto de habla: la causa más próxima del movimiento de los músculos de la lengua y la laringe es la liberación de acetilcolina, pero estos se organizan en patrones de movimiento porque el agente está diciendo ciertas palabras y está intencionalmente orientado hacia su contenido (para una explicación operacional y detallada de cómo las estructuras coordinativas se organizan en función de la intencionalidad en la producción del habla ver (Kelso et al., 1984)). En este caso, la intencionalidad del agente actúa como causa de orden superior que selecciona cuál de los posibles movimientos de los músculos serán ejecutados y forma una dinámica del conjunto total de los músculos, esto es, causalidad top-down. Así mismo, el acto intencional del agente selecciona y da forma a los patrones de activación neuronal, que en conjunto con las demás dinámicas fisiológicas conforman el acto mismo de habla, esto es causalidad bottom-up. Puesto que la relación entre todo y partes es bidireccional, Fuchs la etiqueta como causalidad circular vertical.<sup>11</sup>

En el mismo sentido de la causalidad circular vertical de Fuchs, la relación de constitución circular de mi propuesta de normatividad busca capturar el que componentes y organismo como un todo se posibilitan entre sí y abren trayectorias de comportamiento que de otra forma serían imposibles. Sin embargo, hay una diferencia crucial entre la relación de constitución circular que propongo y la causalidad circular de Fuchs. La noción de causalidad sugiere una relación de determinación entre los términos de la relación, es decir, una en la que las trayectorias del sistema están determinadas por las condiciones iniciales del sistema. Sin embargo, es fundamental que tanto el organismo completo como sus partes puedan tomar diferentes trayectorias de manera que exista la posibilidad del error y por ende la posibilidad de agencia. Como vimos, esta flexibilidad permite que se puedan establecer diferentes formas de interacción en diferentes circunstancias y esto abre la posibilidad de seleccionar un comportamiento u otro, y así el comportamiento normativo de sistemas autónomos emerge. Esto significa que debe existir la posibilidad de selección tanto para las trayectorias de la interacción del organismo y sus órganos y procesos constituyentes como en la interacción del organismo el ambiente.

La relación de constitución circular no solo existe ente el organismo y sus componentes, sino que es también la relación entre una especie y sus organismos individuales, y entre una forma

---

<sup>11</sup> Este párrafo fue tomado literalmente de mi reseña crítica *Mentes en cuerpos de afecto: nota crítica sobre Ecology of the Brain, de Thomas Fuchs* publicada en: Mojica, L. (2019). Mentes en cuerpos de afecto: nota crítica sobre Ecology of the Brain, de Thomas Fuchs. *Crítica*, 51(153), 83-97. .

de vida y los humanos que la constituyen. Regresaré a esta última forma de constitución entre formas de vida sociales y humanos individuales en el próximo capítulo. Sin embargo, desarrollar la idea de especies biológicas como sistemas autónomos o de entender relaciones entre diferentes especies también como sistemas autónomos está fuera del alcance de esta tesis y tendrán que esperar para desarrollos conceptuales futuros.

## 4. 4.Conclusión

En este capítulo defendí que los fenómenos normativos surgen cuando simultáneamente hay un sistema autónomo del que surge la norma de su propio mantenimiento (ej. un organismo vivo, una especie en la escala evolutiva, un hábito o una práctica social) y un agente cuyo comportamiento contribuye o no al mantenimiento de tal sistema autónomo (ej. el mantenimiento de sí mismo, la especie a la que pertenece o los hábitos y prácticas de las que participa). Este último comportamiento es correcto o incorrecto según si contribuye o no al mantenimiento del sistema al que pertenece. La definición de normatividad que propongo tiene varias ventajas:

1. Permite resolver la brecha entre la normatividad autodeterminada de la vida individual que propone el enactivismo y las formas de normatividad que no hacen ninguna diferencia para la existencia de organismos en tanto seres vivos individuales. La solución es reconocer que los comportamientos de un agente hacen parte de sistemas autónomos más amplios que el agente mismo, por ejemplo, aquellos constituidos por la especie o por la comunidad a la que pertenece, por ello sus comportamientos están regidos normativamente también por el mantenimiento de esos otros sistemas y es posible que lo que es correcto respecto a tales sistemas sea irrelevante o incorrecto respecto al mantenimiento del agente mismo. Esta propuesta permite además hacer explícito el carácter situado de la agencia en la propuesta enactiva que hasta ahora ha mantenido una definición de agencia fundamentalmente centrada en las necesidades del individuo a pesar de sus propios intentos de alejarse del individualismo.

2. Logra integrar y aclarar las diferentes fuentes de normatividad a las que está sujeto un agente individual en sus interacciones, a saber, su automantenimiento como organismo individual, como miembro de una especie y como miembro de un grupo social, y que pueden entrar en conflicto. Esto es posible porque en la definición que propongo un mismo comportamiento puede pertenecer a diferentes sistemas autónomos simultáneamente, y por eso, puede ser, ej., bueno para la vida individual pero malo para el mantenimiento de la especie. Las diferentes dimensiones en las que se constituye la normatividad de los comportamientos del agente han sido ampliamente reconocidas en la literatura enactiva; mi propuesta permite especificar en dónde radica la distinción entre estas dimensiones.

En el próximo capítulo muestro que esta propuesta también contribuye a la integración de las dos aproximaciones radicalmente corporizadas más populares en las ciencias cognitivas, el enactivismo y la psicología ecológica, y permite aclarar la disputa sobre el carácter normativo de las affordances porque permite aclarar diferentes tipos de normatividad que hasta ahora no habían sido claramente diferenciados y por ello han motivado tal disputa.

## Capítulo 5

### Normatividad y psicología ecológica

Hasta ahora he afirmado que la concepción de normatividad que defiendo es situada, relacional y pública. Este capítulo está dedicado a profundizar esta afirmación, y a mostrar las implicaciones de la concepción de normatividad que propongo para otro de los enfoques corporizados a la cognición disponibles en la literatura, a saber, la psicología ecológica: argumento que las affordances, en tanto relaciones ambiente-agentes, existen en virtud de que está situadas en formas de vida compartidas en prácticas sociales o en especies biológicas. Me concentro en las prácticas sociales en tanto que en estas se constituye el comportamiento propiamente cognitivo, es decir, comportamiento cuyas motivaciones y criterios normativos sobrepasa las normas biológicas. Argumento que las prácticas sociales son sistemas autónomos que constituyen los criterios de normatividad de los involucramientos particulares de un agente con las affordances del ambiente. Este argumento se cimienta en la definición de normatividad que propuse en el Capítulo 4 en tanto que diferencia los criterios normativos y agente, y diferencia la escala temporal presente del agente en interacción de la escala temporal más amplia de constitución de las prácticas sociales.

Empiezo discutiendo en detalle un par de posibles objeciones a la atribución de normatividad a las affordances y muestro que, a pesar de que las razones que estas objeciones ofrecen apuntan a distinciones importantes entre diferentes fenómenos de la naturaleza, no son suficientes para rechazar su carácter normativo. Esta discusión es la traducción literal de mi respuesta *Reclaiming meaning, reclaiming normativity* publicada en (Mojica, 2020) al artículo *Précis of The Philosophy of Affordances* de Manuel Heras-Escribano (2020b). Una vez motivado el carácter normativo de las affordances, discuto dos conceptos cruciales para entender la relación cognitiva entre agente y ambiente y su carácter situado: forma de vida y prácticas sociales. Argumento que las interacciones cognitivas agente-ambiente están siempre situadas dentro de formas de vida compartidas que constituyen sistemas autónomos porque en estas se constituyen las habilidades de involucramiento con affordances y las consecuencias prácticas de su ejecución. Esto significa que tales formas de vida establecen los criterios de corrección para tales interacciones y establecen sus condiciones de posibilidad.

#### 5. 1. Affordances y su carácter normativo: la visión recibida

La psicología ecológica surge con el influyente trabajo de Gibson (2014 [1979]) en la segunda mitad del siglo XX como una alternativa tanto a las aproximaciones cognitivistas que explican la percepción y la acción en términos de representaciones internas como a las aproximaciones conductistas que las explican como una relación lineal de input-output. Según la perspectiva ecológica, los agentes perciben su entorno inmediatamente como un ambiente con sentido, y lo hacen en términos de las posibilidades de acción que este les ofrece, esto es, en términos de

affordances. El concepto de 'affordance' implica que las posibilidades de acción no dependen exclusivamente de un agente, de la especie o de la comunidad a la que pertenece; pero tampoco de la materialidad del entorno. En cambio, se refieren a la relación entre ambos que depende tanto de las habilidades de los agentes como de la materialidad del entorno (Chemero, 2003). De manera que la unidad mínima de análisis para entender la percepción y acción es el sistema que forman agente-ambiente, y las affordances son propiedades de este sistema.

De acuerdo con la aproximación ecológica, la percepción consiste en la detección de información ecológica disponible en el ambiente. Para hacerlo, los agentes se mueven en el ambiente y detectan los patrones de luz y sombra, de cambios en el sonido o de cambios en la estructura y la resistencia que se forman en relación con su movimiento. Por ejemplo, el agente detecta que el rebote de la luz se mantiene constante, ej., en las esquinas de las mesas y de la habitación, pero que algunas caras se hacen visibles y otras dejan de reflejar luz hacia el agente cuando este se mueve, ej., una taza o un sofá. Estos patrones forman una estructura de distribución de energía (Heras-Escribano, 2019, p. 36) que cuando el agente logra dominar, puede discriminar superficies, sustancias, objetos, etc. en su ambiente. Esto implica que la información ecológica no es lo mismo que la estimulación sensorial. La razón es que la estimulación sensorial no es suficiente para que el agente pueda especificar la estructura de distribución de energía; para usar el famoso ejemplo de Gibson, si la habitación estuviera llena de niebla, el agente recibiría el estímulo sensorial de la luz rebotando en la niebla, pero no podría detectar las estructuras de distribución de energía y, por ende, no podría percibir el ambiente (Gibson, 2014 [1979], pp. 47-49; Heras-Escribano, 2019, p. 36). La información ecológica entonces es detectable a través de patrones de percepción y movimiento en los que el agente se coordina con el ambiente, y se refiere a la relación entre las capacidades de percepción y movimiento de un agente en relación con el ambiente.

Detectar la información ecológica disponible para un agente en el ambiente le permite percibir las posibilidades de acción disponibles, esto es, las affordances disponibles. Por ejemplo, percibir una abertura que permite el paso de la luz de cierta manera es percibir la posibilidad de pasar por esa abertura. Sin embargo, tanto la información ecológica como las affordances del ambiente cambian según las posibilidades sensoriomotoras de los agentes y además existen más allá de la percepción y acción de un individuo particular. Por ejemplo, una superficie plana es un affordance para caminar para un humano adulto y para gatear para un bebé; una taza es un affordance para ser agarrada y llenada de agua para un humano adulto, pero es una posibilidad para ser escalada por una hormiga. Estas posibilidades existen a pesar de que no haya ni hormiga ni humano ni bebé que lleven a cabo estas. Las affordances se refieren a la relación que existe entre habilidades disponibles para una especie o para un animal potencial y la estructura o posibilidades que ofrece su ambiente. En la sección 5.3 veremos que esta relación entre ambiente y animal potencial es normativa respecto al mantenimiento de una forma de vida, y además establece criterios normativos para las interacciones entre agentes individuales y su ambiente.

El carácter normativo de esta relación no solo ha sido defendido por otros autores como Chemero (2009), Dings (2020) o Rietveld y Kiverstein (2014), también ha sido criticado

(Heras-Escribano, 2020b). A grandes rasgos, el carácter normativo suele defenderse argumentando que las habilidades para percibir e involucrarse activamente con los aspectos del ambiente como oportunidades para la acción tienen condiciones de corrección. Por ejemplo, uno puede acertar o equivocarse al percibir que puede pasar por una abertura, o puede fracasar o tener éxito al escalar una montaña. Así, las affordances parecen implicar que el organismo establece una forma de relacionarse con un aspecto del entorno que puede ser correcta o incorrecta. Defender el carácter normativo de las affordances exige explicar cómo se establecen las condiciones de corrección.

Hay al menos tres formas de abordar esta explicación. Los defensores del enactivismo autopoietico, como Di Paolo et al. (2017), sostendrían que las affordances, es decir, la relación que los agentes establecen con el ambiente en términos de posibilidades de acción, son normativas porque los agentes establecen y siguen su propia regla intrínseca cuando actúan, es decir, para preservar su forma de vida. Por otro lado, Chemero (2009) defiende que las affordances son normativas porque cumplen una función, a saber, permiten al agente actuar, pero el agente puede equivocarse o fallar al ejecutar la acción. Finalmente, Rietveld y Kiverstein (2014) defienden que las affordances son normativas en tanto que su existencia es relativa a una forma de vida compartida y tal forma de vida establece criterios normativos de comportamiento para los individuos que pertenecen a ella. Más adelante, en la sección 5.3, regresaré a este punto. Por ahora, es importante notar que en todos los casos la noción de normatividad está orientada a determinar el curso de acción correcto más que a formar afirmaciones verdaderas o a representar el entorno con precisión. Por lo tanto, afirmar que las affordances son normativas en este contexto significa, a grandes rasgos, que establecen cursos de acción correctos e incorrectos y que la forma en que los agentes se involucran con estas está sujetas a evaluaciones de corrección.

## 5.2. Reivindicando la normatividad de las affordances

Manuel Heras-Escribano (2019, 2020b) ofrece varios argumentos contra los enfoques que defienden que las affordances son normativas. En esta sección discuto dos de ellos. El primero apela al carácter externo e inherentemente social de la normatividad que deriva de las consideraciones de seguimiento de reglas y del argumento del lenguaje privado de Ludwig Wittgenstein, y el segundo se basa en la distinción entre los aspectos nomológicos y normativos en la explicación del comportamiento. Su posición tiene la virtud de distinguir claramente entre el error biológico y el cognitivo, y de destacar el papel constitutivo de la socialidad en la mente. Pero a pesar de estos méritos, muestro que sus argumentos no son suficientes para concluir que las affordances no son normativas.

### 5.2.1. El carácter público y social de la normatividad

El primer argumento de Heras-Escribano se basa en la idea de que la normatividad no puede separarse de la socialidad (2020b, §26). Argumenta que un fenómeno que puede ser explicado en términos puramente funcionales sin apelar a una dimensión social no es normativo y afirma que la percepción de affordances y el involucramiento con ellas pueden ser explicados



en términos puramente funcionales. Además, si la normatividad es intrínsecamente social, también es un error explicar la cognición, como lo hace el enactivismo autopoiético, en términos de agentes que establecen y siguen sus propias reglas intrínsecas cuando perciben y se relacionan con las affordances. La razón por la que la normatividad es social se deriva, para Heras-Escribano, de las consideraciones de Wittgenstein sobre el seguimiento de reglas y el argumento del lenguaje privado.

Vimos en el primer capítulo que las consideraciones de Wittgenstein sobre el seguimiento de reglas muestran que la interpretación de las reglas no puede determinar por sí misma si una acción es correcta o no (Wittgenstein, 2009 [1953], §201). La razón es que una regla siempre puede ser interpretada de tal manera que la misma acción puede ser considerada tanto incorrecta como correcta con respecto a la misma regla, por lo que no permite a los agentes distinguir si están en lo correcto de la mera impresión de que lo están. Por lo tanto, la interpretación de la regla no puede ser la forma en que se determina la corrección de una acción, y esto significa que apelar a la interpretación de una regla general no puede explicar lo que hace que una acción sea mala o correcta (Wittgenstein, 2009 [1953], §§185-201). ¿Qué se necesita entonces?

La normatividad, como bien señala Heras-Escribano, debe estar abierta a la posibilidad del error, y debe proporcionar criterios para distinguir entre actuar correctamente y la mera impresión de hacerlo. El argumento del lenguaje privado de Wittgenstein muestra que todo lo que sirve como criterio de corrección debe ser público y no puede ser una referencia a fenómenos privados (Wittgenstein, 2009 [1953], §270). Esto significa que los criterios de corrección deben estar a disposición de todos y, en el caso particular del significado del lenguaje, sólo las prácticas sociales del uso del lenguaje son lo que es común a todos, ni las reglas abstractas ni los pensamientos y sensaciones interiores (Wittgenstein, 2009 [1953], §23, §43, §260).

Si bien las prácticas sociales se establecen como las fuentes de normatividad del lenguaje porque son perfectamente públicas y lo único potencialmente disponible para todos los hablantes, esto no significa que la socialidad sea la única fuente posible de normatividad en general. Aún más, aunque la normatividad de los comportamientos lingüísticos y no lingüísticos sea intrínsecamente pública, esto no implica que todas las formas de normatividad sean intrínsecamente sociales: ¿por qué un organismo necesitaría criterios sociales para determinar cuál el curso correcto de acción? En efecto, hay formas públicas, pero no sociales, de determinar si un curso de acción es correcto: pasar por una abertura, cazar una presa sin salir gravemente herido en el intento, comer bayas no venenosas, etc., tienen condiciones de corrección perfectamente públicas. El carácter público significa que actuar correcta o incorrectamente, tener éxito o no, etc., tiene consecuencias prácticas que no sólo implican una diferencia para el curso de acción del agente, sino que serían evidentes para un observador. Es importante notar que estipular un observador potencial para una acción es un criterio demasiado débil para llamarlo social, aunque sí sea suficiente para considerarlo un criterio disponible públicamente. Esta noción de lo público incluye tanto el comportamiento socialmente constituido, por ejemplo, el uso del lenguaje, como el comportamiento biológico

no socialmente constituido, por ejemplo, un pulpo cazando comida. Este último tipo de comportamiento es, sin embargo, normativo, no sólo porque tiene una finalidad con respecto a la cual el pulpo puede fallar, sino también porque su éxito o fracaso supone una clara diferencia práctica para su supervivencia.

En este sentido, percibir el entorno e interactuar teniendo un propósito práctico son formas perfectamente públicas de interacción entre un agente y los aspectos del entorno. En el despliegue de una acción, los agentes se orientan a un aspecto concreto del entorno, esto significa que realizan una de las muchas posibilidades de acción disponibles para ellos en el ambiente, y en esa medida, sus propósitos se hacen presentes en el ambiente. Sus acciones tienen consecuencias prácticas que van más allá de sus intenciones presentes y hacen que su corrección o incorrección sea evidente para los propios agentes, aunque nunca sean observados por otros, por ejemplo, tener comida para alimentarse, poder pasar por una abertura para encontrar refugio. Esto significa que la percepción y el involucramiento con affordances son formas de relación perfectamente públicas que un agente establece con el entorno.

Aquí se pueden plantear dos posibles objeciones. En primer lugar, se podría argumentar que la percepción de affordances es sustancialmente diferente al involucramiento con affordances con respecto a su carácter público y normativo, porque mientras que involucrarse con un affordance implica un ejercicio manifiesto de habilidad, la percepción de affordances no. Sin embargo, establecer esta distinción entre acción y percepción para permitir que la primera sea normativa y no la segunda va en contra de uno de los compromisos centrales de la psicología ecológica: la inseparabilidad percepción-acción (Lobo et al., 2018). Como vimos arriba, la percepción no puede entenderse separada de la acción porque la percepción implica un bucle sensoriomotor en el que lo que el agente hace y siente están en una retroalimentación constante e inextricable. Aunque este bucle puede ser descrito causalmente, sigue teniendo condiciones normativas, es decir, que debe guiar al agente en su acción y flujo de experiencia. Recordemos que los criterios normativos en juego no se basan en la correspondencia o la preservación de la verdad, sino que son criterios pragmáticos cuya corrección, etc., se determina a medida que se desarrollan las acciones. Como afirman Michaels y Carello (1981), la percepción debería revelar aspectos útiles del entorno, pero a diferencia de ellas, yo sostengo que la utilidad implica una dimensión normativa.

Esto lleva a la segunda objeción: las condiciones de corrección que están en juego en la percepción de affordances y el involucramiento con ellas son funcionales. Como tales, pueden explicarse completamente en términos causales y, por tanto, no es necesario apelar a la normatividad para caracterizar las affordances. Este es precisamente el segundo argumento de Heras-Escribano.

### 5.2.2. Propósitos y la distinción nomológica-normativa

Heras-Escribano argumenta que a diferencia del comportamiento normativo, que tiene una dimensión social y que no puede reducirse a términos plenamente causales, podemos dar

pleno sentido a los comportamientos de los organismos en términos completamente causales, es decir, apelando a regularidades de las leyes de la naturaleza o a funciones evolutivas (2020b, §25). Por lo tanto, no hay necesidad de introducir el vocabulario normativo para explicar la respuesta a affordances y la percepción de estas (2020b, §29).

Si bien es cierto, como sostiene Heras-Escribano (2020b, §31), que la fuente de normatividad debe proporcionar criterios normativos y que un comportamiento normativo debe estar abierto a la posibilidad del error, los comportamientos de los organismos no pueden explicarse completamente en términos causales. El éxito o el fracaso de un determinado comportamiento es una evaluación normativa que depende de algún propósito o valor. Los propósitos y valores no pueden explicarse en términos puramente causales, por ejemplo, una descripción de las bacterias que nadan hacia un gradiente creciente de sacarosa estará incompleta y no tendrá sentido si no se reconoce que la sacarosa es valiosa como alimento para las bacterias. En general, cualquier descripción del comportamiento de un organismo vivo carece de sentido y está incompleta si no capta que hay una forma de selección y un propósito y un valor en la acción del organismo, a saber, el valor de preservar su forma de vida; de lo contrario, no estaría motivado a seleccionar y llevar a cabo una de las distintas posibilidades de acción disponibles para él.

Como vimos en la discusión de la teleología natural de Kant del Capítulo 2 (p. 55), la descripción en términos de causas y efectos debe subsumirse a una descripción de medios y fines cuando está en juego el comportamiento de un ser vivo. La finalidad de los organismos de mantenerse con vida es un criterio claro para evaluar su percepción y compromiso con las affordances. Por un lado, este criterio de evaluación es un criterio que va más allá de las impresiones inmediatas del agente. Los propósitos del agente se hacen presentes en el entorno mediante el despliegue de sus acciones, es decir, involucramiento con affordances, y sus consecuencias prácticas, por lo que su éxito o fracaso no sólo es observable para los demás, sino que también es experimentado por el agente a pesar de sus intenciones originales. Por otra parte, los organismos pueden fracasar en su percepción de affordances y en su compromiso con respecto a sus propios propósitos, lo que abre la posibilidad de error.

Aunque Heras-Escribano simpatiza con el reconocimiento de la agencia biológica y cognitiva en los organismos no humanos (2020b, p. 110), se abstiene de reconocer su carácter normativo porque busca establecer una distinción clara entre el componente social y el componente biofisiológico en el comportamiento con sentido. Sin duda, esta distinción es importante y no ha sido profundizada suficientemente en la literatura enactiva y corporizadas. Sin embargo, aceptar esta distinción no implica negar la naturaleza normativa de los comportamientos no sociales de los seres vivos; e incluso, negar la normatividad en el nivel básico de las posibilidades tiene un costo demasiado alto: oscurece la continuidad entre el ser humano y otras formas de vida y, por tanto, impide que los organismos no sociales puedan ser tomados como agentes en las explicaciones científicas. Ser un agente cognitivo requiere ser capaz de interactuar con el entorno de acuerdo con los propósitos propios del agente. Si la normatividad se excluye de la forma más básica de cognición, es decir, la percepción y el involucramiento con affordances, entonces la relación que los organismos no sociales

establecen con el entorno no puede explicarse en términos de sus propósitos. Sin embargo, esto implica que la agencia queda descartada para estas formas de vida. Por lo tanto, si Heras-Escribano quiere preservar la capacidad de acción de los organismos no sociales, tendría que aceptar una forma básica de normatividad no social como la que defendí en la sección anterior, o bien insistir en que una explicación causal es la explicación completa del comportamiento vivo no social y, por lo tanto, adoptar una posición eliminativista.

### 5.2.3. Objeciones y respuestas

Heras-Escribano ofrece dos respuestas a este argumento. Una respecto a la distinción entre lo público y lo social, y la otra respecto a la consecuencia eliminativista que señalo que se sigue de rechazar la normatividad en criaturas no sociales.

Primero, Heras-Escribano pone en duda la distinción entre lo público y lo socialmente constituido que uso para mostrar que es posible que existan fenómenos normativos que no sean sociales y sugiere que esto no es realmente posible. Una acción puede hacer una diferencia obvia para un observador cuando ciertos criterios de corrección pueden ser atribuidos, como el propósito de cazar una presa y alimentarse; y señala: Sin embargo, esta atribución puede ser hecha por miembros de una comunidad basados en prácticas sociales, porque solo para ellos puede ser obvio el despliegue de prácticas normativas por otros agentes (puesto que estas personas han desarrollado este sentido de normatividad que enfatizo siguiendo a Wittgenstein) (Heras-Escribano, 2020a, p. 232).

Por supuesto, la atribución de propósitos a otras criaturas por parte de una persona humana está constituida socialmente dentro de las prácticas compartidas a las que pertenece, crucialmente, a las prácticas de lenguaje a las que pertenece. Sin embargo, la noción de lo público que he defendido hasta ahora no solamente se refiere a lo público para un ser humano, sino que incluye también la posibilidad de que el comportamiento de un ser vivo sea observable y significativo para otras criaturas vivas no humanas. La posibilidad de reconocimiento entre seres vivos no humanos es evidente en las relaciones entre individuos de la misma especie, sobre todo en especies sociales, y en relaciones entre individuos de diferentes especies, ej., los animales de carroña que reconocen cuando otros están a punto de morir o los animales que fingen su propia muerte para evitar ser cazados, comportamiento conocido como tanatosis. En general, la vida y la muerte, y frecuentemente la salud y la enfermedad, son fenómenos observables por otros aún si su significado no está constituido dentro de prácticas sociales.

Heras-Escribano propone cinco condiciones adicionales para que un fenómeno sea normativo. (i) Debe ser situado (2019, p. 99), (ii) socialmente constituido (2019, pp. 100-102), (iii) la normatividad debe involucrar criterios públicos de corrección-incorrección y no solo condiciones de éxito o fracaso para capturar no solo que se logre un propósito sino también cómo se logra (2019, p. 92), (iv) debe haber un agente consciente de que está actuando y cómo lo está haciendo (2019, p. 94), y (v) el agente debe tener racionalidad práctica (2019, pp. 95-99). Estos últimos dos criterios apuntan a dos condiciones más generales necesarias para la

normatividad: primero, que el agente tenga la posibilidad de seleccionar cómo se desenvuelve su interacción con el ambiente y, segundo, que lo pueda hacer de acuerdo con sus propios propósitos. Estas dos condiciones están condensadas en mi definición de normatividad: Primero, la posibilidad de seleccionar entre diferentes cursos de acción se sigue de su capacidad de automantenerse recursivamente. Segundo, de la naturaleza autónoma de los agentes emergen propósitos intrínsecos, y en esa medida los pueden actuar de acuerdo con sus propios propósitos.

Las condiciones (i) y (iii) que propone Heras-Escribano son compatibles con la concepción de normatividad que he defendido hasta ahora. Por un lado, las interacciones están siempre situadas en un contexto material y hacen parte de un patrón de comportamientos compartidos que se reproduce a sí mismo en el tiempo. Por otro lado, los criterios de corrección incluyen potencialmente distinciones normativas entre diferentes formas de conseguir un propósito. Sin embargo, sostener que es condición necesaria de la normatividad que sea constituida socialmente es el punto de desacuerdo.

Como argumenté en la sección anterior, el argumento del lenguaje privado de Wittgenstein no es razón suficiente para concluir que todas las formas de normatividad son sociales, como lo hace Heras-Escribano. Para Wittgenstein, un lenguaje genuinamente privado es imposible, es un sinsentido, y este es uno cuyas expresiones “solo puedan ser conocidas para el hablante, [que se refieran] a sus sensaciones inmediatas y privadas. De manera que otro no pueda entender el lenguaje” (Wittgenstein, 2009 [1953], § 243) en: (Heras-Escribano, 2019, p. 101). El punto crucial para que el lenguaje tenga sentido es que exista la posibilidad de retroalimentación al comportamiento del hablante para guiar su comportamiento de manera correcta, y en el caso del lenguaje, esta retroalimentación viene siempre de otros hablantes de la comunidad. Argumenté que, a diferencia de los comportamientos lingüísticos, los comportamientos biológicos tienen consecuencias prácticas que guían y retroalimentan el comportamiento del agente que no están constituidas socialmente; estas, en cambio, están constituidas por la organización y necesidades de mantenimiento del agente y su relación con el ambiente, y tienen consecuencias prácticas para la vida y la posibilidad de establecer interacciones del agente. En ese sentido la constitución social no parece ser condición necesaria de la normatividad, aunque sea el caso en la normatividad lingüística.

Vale la pena preguntarnos qué buscamos tanto Heras-Escribano como yo al requerir criterios de corrección públicos o sociales. Esta pregunta no solo abre la posibilidad de llegar a un acuerdo, sino que hace explícito aquello esencial a la normatividad a lo que apuntan nuestros respectivos requerimientos. Ambos buscamos que los fenómenos normativos no sean fenómenos privados, es decir, que no sean fenómenos o entidades en la mente de los agentes y en principio inaccesibles por otros como frecuentemente se ha postulado en filosofía de la mente y el lenguaje; buscamos que exista la posibilidad del error para el comportamiento de los agentes y que puedan errar incluso cuando ellos mismos no se den cuenta en el momento de actuar. La definición de normatividad que propongo permite mostrar las diferentes formas en que este requerimiento puede cumplirse y que implican a su vez diferentes tipos de normatividad.

En la definición de normatividad que propongo, aquello que es sujeto de ser normativo es una interacción, y uno de los requerimientos que propuse es que al menos uno de los términos de tal interacción sea un agente. Como hemos visto en los capítulos anteriores, la forma mínima de agencia son sistemas autónomos y adaptivos que se mantienen a sí mismos recursivamente y que establecen su propia normatividad en la escala temporal de su vida. Esta forma mínima de normatividad es también la forma mínima en que surgen fenómenos no privados y susceptibles de error.

Por un lado, vimos arriba que los sistemas que se mantienen a sí mismos recursivamente tienen la capacidad de seleccionar sus formas de interacción para continuar manteniéndose a pesar de cambios drásticos en el ambiente (ver p. 122). Cuando el comportamiento del sistema resulta de tal capacidad de selección entonces es susceptible de ser correcto o incorrecto respecto a sus propios propósitos. En esa medida, estos comportamientos son susceptibles de error.

Por otro lado, las trayectorias de comportamiento que seleccionan las formas mínimas de agencia no son fenómenos privados porque tienen consecuencias que pasan por el ambiente. El comportamiento seleccionado puede ser o bien una interacción que el agente como sistema completo establece con el ambiente o bien la relación circular que existe entre el agente como sistema completo y sus propios órganos y procesos constitutivos. En el primer caso, es evidente que la interacción entre el agente y el ambiente es un fenómeno no privado en el sentido en que es potencialmente observable por otros, y además puede resultar incorrecto respecto al propósito del agente de mantenerse con vida incluso cuando a este le dé la impresión de que está en una trayectoria de comportamiento correcta. El segundo caso, sin embargo, no es tan evidente. Aunque la interacción entre los procesos internos constitutivos y el sistema completo no es inmediatamente observable por otros, sus consecuencias prácticas sí lo son; por ejemplo, si el corazón no bombea sangre con suficiente frecuencia y regularidad o si no produce suficiente insulina para controlar los niveles de glucosa en la sangre, eventualmente el agente como sistema completo perderá capacidades de interacción con el ambiente, por ejemplo, no poder ingerir ciertos alimentos, desmayarse o incluso morir. Estas consecuencias prácticas afectan al agente independientemente de si el sistema está bajo la impresión de que seleccionó la trayectoria correcta o no como ocurre, por ejemplo, cuando produce menos insulina de la necesaria para mantener los niveles de glucosa estables o cuando sigue bombeando sangre a células cancerígenas. En suma, la relación circular entre agente y sus procesos constitutivos puede ser correcta o incorrecta respecto al propósito del agente de mantenerse con vida, su corrección va más allá de las impresiones presentes del agente y tiene implicaciones públicas en tanto que eventualmente hacen una diferencia práctica para las posibilidades de interacción del agente con su ambiente que le permiten mantenerse con vida.

Esta forma mínima de normatividad no es suficiente para atribuir capacidades cognitivas a estos sistemas, y probablemente esta distinción, i.e., la distinción entre fenómenos biológicos y cognitivos, es la que Heras-Escribano busca mantener al negar el carácter normativo de los fenómenos biológicos:

La estrategia de subsumir el comportamiento de agentes tanto no sociales como sociales bajo la etiqueta de ‘normatividad’ sin explicar los detalles de cada tipo de comportamiento no es apropiada para explicar las diferencias entre todos los tipos de agencia en detalle. Debemos buscar explicar las sutilezas, en vez de difuminarlas en una aproximación general y sin refinar (Heras-Escribano, 2020a, p. 232)

Defender una diferencia cualitativa entre fenómenos biológicos y fenómenos cognitivos se contrapone a las posturas enactivas que sostienen que esta forma mínima de normatividad es una forma de cognición (ej., De Haan, 2020, pp. 53-59; Thompson, 2007, pp. 148-149), y muestra que son fenómenos lo suficientemente distintos para requerir distintas aproximaciones científicas y filosóficas. Por ejemplo, la percepción, que hace parte de los fenómenos biológicos, puede explicarse en términos de patrones que forma la reverberación de la luz o el sonido en relación con la posición del agente, su capacidad sensorial y sus movimientos. En cambio, para explicar la cognición es necesario apelar a procesos de desarrollo dentro de prácticas socioculturales constituidas por el uso compartido del lenguaje. Aunque el enactivismo no niega que haya diferencias importantes entre las formas mínimas de comportamiento normativo y cognición humana, el énfasis en la continuidad vida-mente puede oscurecer diferencias importantes entre diferentes formas de vida y de interactuar con el ambiente.

La distinción entre tipos de normatividad que se sigue de la concepción de normatividad que propuse en el capítulo anterior las clasifica según el tipo de interacción y la naturaleza de los criterios de normatividad bajo los que son evaluables. Esta clasificación evita el peligro de pasar por alto diferencias importantes entre fenómenos normativos distintos y en eso coincido con una de las motivaciones de Heras-Escribano.

Sin embargo, si bien es valioso hacer distinciones claras entre fenómenos tan distintos como el comportamiento para preservar la vida y los comportamientos cognitivos típicamente humanos, hay una distinción más fundamental que quiero preservar con la concepción de normatividad que propuse: la distinción entre sistemas que pueden seleccionar sus propias trayectorias de interacción con el ambiente y los demás, incluidos aquellos que son auto-constituidos y están lejos del equilibrio termodinámico pero que no tienen la posibilidad de cambiar sus propias trayectorias de interacción. El primer tipo de sistemas pueden seleccionar entre diferentes formas de interacción con implicaciones para su propia existencia, y de esta forma emerge la agencia: el sistema puede distinguir entre diferentes posibilidades en el entorno y percibir las como mejores o peores porque puede hacer una diferencia práctica al respecto. En otras palabras, sin la posibilidad de hacer una diferencia práctica seleccionando una u otra trayectoria, no hay posibilidad tampoco de percibir algo como significativo –esta es presumiblemente la convicción general que fundamenta y motiva la aproximación a la cognición de la psicología ecológica–. Es importante notar que esta forma básica de percepción implica una aproximación afectiva en tanto que el agente no es indiferente al mantenimiento de su propia existencia. En suma, los fenómenos normativos implican agentes a quienes les concierne su propia existencia, y la distinción entre este tipo de existencia y la

existencia de sistemas inertes e indiferentes es la que quiero preservar con la definición de normatividad que defiendo.

El salto de lo no-normativo a lo normativo es el salto de la no-agencia a la agencia, de lo inerte a lo vivo, de lo no-sintiente a lo sintiente. En este punto podemos regresar al párrafo de las *Investigaciones Filosóficas* con las que abrí el capítulo 1:

Mira a una piedra e imagina que tiene sensaciones.—Uno se dice a uno mismo: ¿Cómo puede uno si quiera hacerse la idea de adscribir una sensación a una cosa? ¿uno podría también adscribírsele a un número!—Ahora mira a la mosca retorciéndose y enseguida estas dificultades desaparecen y el dolor parece poder agarrarse de aquí, sonde antes todo era, por así decirlo, demasiado liso para ello.

Y así también un cadáver nos parece bastante inaccesible al dolor.—Nuestra actitud hacia lo que está vivo y hacia lo que está muerto no es la misma. Todas nuestras reacciones son diferentes.—Si alguien dice: “esto no puede venir simplemente del hecho de que una cosa viva va y viene de tal o cual forma y una cosa muerta no”, entonces quiero sugerir a esta persona que este es el caso de una transición ‘de cantidad a cualidad’. (Wittgenstein, 2009 [1953], p. § 284)

Heras-Escribano coincide con la atribución de agencia a otras formas de vida en tanto que son capaces de comportamiento con propósitos, incluso las no sociales (2020a, p. 232). Sin embargo, a diferencia de lo que he defendido hasta ahora, para Heras-Escribano, “una explicación nomológica es suficiente para comprender su comportamiento complejo”. Este es el punto central de su respuesta al riesgo eliminativista que señalo en su posición, pero para terminar esta sección, quisiera mostrar que esta no es la forma más adecuada de explicar un comportamiento con propósitos.

Por supuesto es posible explicar comportamientos complejos como los patrones de percepción-acción en términos nomológicos, es decir, en términos de leyes que gobiernan los fenómenos naturales. Sin embargo, esta explicación no es suficiente para capturar al menos dos rasgos cruciales de la agencia: la capacidad de comportarse de acuerdo con propósitos y la posibilidad de selección y de error.

Idealmente, una explicación nomológica deduce los comportamientos particulares de leyes generales que los rigen. La psicología ecológica y, en particular Heras-Escribano, defienden que tales leyes son los patrones de reverberación de energía en el ambiente, junto con los patrones de percepción y movimiento del agente, y estos patrones pueden ser capturados por ecuaciones que formalicen esas tres variables. Dentro de esa descripción, sin embargo, los propósitos del agente son completamente irrelevantes, después de todo tales patrones se mantienen independientemente de las necesidades o propósitos de un agente. Pero incluso si los propósitos fueran a ser incluidos como variables dentro de tal descripción, una descripción nomológica no da lugar a que haya posibilidad de selección una vez las circunstancias o variables están determinadas; pero si no hay posibilidad de selección, no hay posibilidad del error y por ende no hay realmente agencia. En ese caso, sería imposible, por ejemplo,



deslizarse y caer o no darse cuenta de hay un predador a pesar de que las circunstancias sean favorables; pero por supuesto, esto es posible y es un error porque el agente tiene la posibilidad de encauzar sus trayectorias de acción de una u otra forma y entonces establecer diferentes cursos de interacción sensoriomotora. Así que, si no hay posibilidad de selección y de error, entonces no hay agencia genuinamente. Por ende, una explicación nomológica no puede capturar todos los aspectos esenciales del comportamiento complejo de agentes capaces de actuar con propósito.

Aceptar la naturaleza normativa de formas de vida no sociales permite dar cuenta de la posibilidad del error necesaria para la agencia genuina y el comportamiento con propósitos. Además, como discutí respecto a la primera respuesta de Heras-Escribano, esta naturaleza normativa no es intrínsecamente social, sino que surge con sistemas que pueden seleccionar sus trayectorias de interacción y hacer la diferencia respecto a su propia existencia. Esas dos razones motivan la naturaleza normativa de las affordances que es el tema de la siguiente sección.

### 5. 3. Criterios de normatividad y formas de vida

La discusión con Heras-Escribano me permitió justificar el que se conciban las affordances como normativas; en esta sección muestro en qué sentido lo son: definiendo que las affordances son la relación entre aspectos del ambiente y habilidades disponibles en una forma de vida, en donde una forma de vida es un sistema autónomo que establece criterios de normatividad que regulan las interacciones particulares entre individuos y ambiente. Empiezo discutiendo la distinción entre dos tipos de información ecológica, una relativa al agente individual y la otra relativa a la especie, y las nociones de ambiente asociadas que proponen Baggs y Chemero (2018), y muestro cómo estas capturan la distinción fundamental entre agencia y criterios normativos que está a la base de la noción de normatividad que propuse en el capítulo anterior. Luego discuto la propuesta de Rietveld y Kiverstein (2014) quienes defienden una normatividad relativa a la forma de vida. Su propuesta tiene dos ventajas. La primera es que logra hacer una distinción entre diferentes niveles de análisis del ambiente en donde se sitúa normativamente el comportamiento tanto humano como no humano. La segunda ventaja es que su concepto de forma de vida me permite argumentar que las formas de vida son sistemas autónomos que establecen criterios de normatividad como propuse en mi definición de normatividad en el capítulo anterior.

#### 5.3.1. Ambiente, umwelt y hábitat

El concepto de affordance se refiere tanto a la percepción de un individuo de su ambiente en términos de sus posibilidades de acción, como a las posibilidades de acción que están disponibles en el ambiente para un individuo genérico de una especie. Después de todo, tales posibilidades de acción están disponibles en el ambiente independientemente de si un agente individual las percibe o no. Esto es evidente también cuando se consideran las affordances en los términos en los que se han operacionalizado, es decir, en términos de invariancias y patrones de información. Por supuesto, el ambiente tiene una estructura que existe

independientemente de si es percibida por un individuo en un momento y lugar específicos, así mismo la luz rebota sobre esas superficies independientemente de si llega a ser detectada por un agente, y la luz formaría los mismos patrones respecto de cualquier individuo con las mismas capacidades para sentir y moverse por el ambiente independientemente de si estos patrones están de hecho siendo percibidos por algún agente o no. A esta relación potencial entre ambiente y animal en la que se formarían patrones entre la estructura del ambiente y la percepción-acción del animal se le conoce como información ecológica. Siguiendo a Claire Michaels y Claudia Carello (1981), Edward Baggs y Anthony Chemero (2018) denominan a este tipo de información ‘información-sobre’ que caracterizan como la relación que va desde la estructura del ambiente hacia un animal potencial. Sin embargo, el concepto de ‘información-sobre’ no captura una instancia importante de la explicación, a saber, cuando un agente de hecho establece una interacción sensoriomotora con el ambiente, detecta información y esta información es ‘información-para’ el individuo, esto significa que percibe el mundo de cierta manera y puede actuar en él (Baggs & Chemero, 2018, p. s2185).

Baggs y Chemero (2018) muestran que la distinción entre estas dos formas de información ecológica, es decir, información-sobre e información-para, coincide con la distinción entre dos conceptos de ambiente. Por un lado, la información-sobre el ambiente, al ser relativa a un animal potencial, captura la idea de que el ambiente es un *hábitat* para individuos que comparten capacidades sensoriomotoras similares que son, paradigmáticamente, individuos de una misma especie. Un subconjunto de este hábitat se refiere a los aspectos del ambiente con los que efectivamente se involucra un agente, es decir, los que percibe y a los que dirige su acción. Baggs y Chemero (2018) llaman ‘*umwelt*’ a este subconjunto que se refiere al mundo que habita un agente y que tiene sentido para este dadas sus capacidades particulares. Identifican un tercer sentido de ambiente, a saber, el mundo físico, que se refiere a la existencia del ambiente independientemente de cualquier criatura viva que pudiera percibirlo y actuar en él pero que alberga todos los hábitats y *umwelten* posibles. Puesto que el mundo físico no es relativo a la percepción y acción de un ser vivo, este no da lugar a ningún tipo de información ecológica.

La información ecológica existe solo en la medida en que sea información para alguien, es decir, en la medida en que sea detectada y usada por agentes actual o potencialmente. De acuerdo con Chemero (2009), las habilidades de los agentes son el segundo término de la relación que permite a los agentes detectar la estructura del ambiente y usarla para actuar. Chemero define una habilidad en términos de funciones biológicas cuya existencia depende de la historia evolutiva y de la historia de desarrollo del individuo (2009, pp. 145-146). Chemero caracteriza una función biológica como una relación entre los organismos de una especie y el ambiente, que se establece porque en algún momento del pasado ayudó al animal o a sus ancestros a sobrevivir, reproducirse o prosperar en el ambiente (2009, p. 145). En el establecimiento de esta relación, el agente desarrolla rasgos fisiológicos y comportamientos que le permiten involucrarse con el ambiente y, a su vez, el ambiente es modificado por el agente para poder involucrarse con este. Esto significa que, con la suficiente repetición de las modificaciones, el ambiente es adaptado a los agentes y se convierte en su nicho; y, con la suficiente repetición de rasgos y comportamientos, los agentes de la especie desarrollan

habilidades que se adaptan a las posibilidades que ofrece el ambiente. En ese sentido, las habilidades consisten en las posibilidades de involucramiento con el ambiente que surgen por la relación entre los rasgos fisiológicos y comportamientos establecidos en la historia evolutiva y personal del agente, y las posibilidades del ambiente que habita.

Chemero contrapone su concepción de habilidades en términos de funciones a la concepción de habilidades como propiedades disposicionales que defendieron autores como Turvey et al. (1981). Esta contraposición apunta a al menos una diferencia importante: las propiedades disposicionales como la solubilidad se actualizan cuando están en las circunstancias adecuadas, ej., estar inmerso en agua, mientras que las funciones no necesariamente lo hacen. Esto implica que, a diferencia de las disposiciones, las funciones dan lugar a que el agente no responda de la manera adecuada a las posibilidades de acción que le ofrece el ambiente o que no responda en absoluto a pesar de que las posibilidades de acción estén disponibles sin ningún obstáculo de por medio. Esta distinción le permite a Chemero señalar que:

hay algo inherentemente normativo en las habilidades. Se supone que los individuos con habilidades se comporten de formas particulares, y puede fallar al hacerlo. Las disposiciones, de otro lado, nunca fallan; simplemente se manifiestan o no en las circunstancias apropiadas. (Chemero, 2009, p. 145)

Ambiente	↔	Agente
Umwelt	→información-para→	un individuo
Hábitat	←información-sobre←	animal potencial de una especie
Mundo físico	No	No

Tabla 1  
Inspirada en Fig. 1 en (Baggs & Chemero, 2018, p. S2185)

Esta caracterización del carácter normativo de las affordances implica que la normatividad es un rasgo de la interacción entre individuo y agente. Dada la definición de habilidad como función, podemos concluir que la historia de selección es aquello respecto a lo que se evalúa tal interacción. En otras palabras, para usar los términos que he propuesto, el criterio normativo que permite diferenciar las formas particulares en las que se supone que el agente debería actuar de las formas de comportamiento incorrectas es su historia evolutiva y de desarrollo. Concebir los criterios normativos de esta forma hace eco de las propuestas teleosemántica y del enactivismo radical (cf. p. 105) que defienden que la normatividad emerge en la historia de selección, pero, a diferencia de estas, no ignora el papel de la agencia. Sin embargo, no es del todo claro qué significa que las habilidades sean funciones y que estas

establezcan criterios de normatividad. En este punto, la aproximación a la teoría de affordances defendida por Erik Rietveld y Julian Kiverstein ofrece una respuesta.

### 5.3.2. Formas de vida como sistemas autónomos

En el capítulo anterior, propuse que los criterios normativos se establecen por un sistema autónomo; en el capítulo 2 discutí ampliamente la explicación que ofrece el enactivismo autopoietico de cómo un organismo individual en tanto sistema autónomo establece criterios de corrección para sus propias interacciones y procesos constitutivos (cf. p. 68) y, en el capítulo 4, propuse entender estos criterios como las consecuencias prácticas para el mantenimiento de la existencia del organismo que se desenvuelven en una escala temporal mayor a la de la interacción, esto es, la escala temporal de su vida. Sin embargo, no he discutido en qué sentido otros criterios de corrección son establecidos por otros sistemas autónomos ni cuál es la relación entre estos y las interacciones particulares para las que establecen criterios de normatividad. En esta sección quisiera abordar este punto y para ello propongo subsumir el concepto de hábitat de Baggs y Chemero como parte de un sistema autónomo que establece criterios de corrección y que pueden concebirse como un nivel de análisis del concepto de forma de vida que proponen Rietveld y Kiverstein.

Rietveld y Kiverstein (2014) introdujeron el concepto de forma de vida en el contexto de teoría de affordances. La manera en que lo entienden está inspirada tanto en el trabajo de Tim Ingold para quien las variaciones culturales o variaciones en forma de vida son variaciones en la habilidad (Ingold, 2002), como en la filosofía del Wittgenstein tardío quien famosamente afirmó que hablar un lenguaje es pertenecer a una forma de vida (Wittgenstein, 2009 [1953]). Rietveld y Kiverstein (2014) caracterizan su propuesta como una expansión de la concepción de affordances de Chemero quien las define como relaciones entre aspectos del ambiente y habilidades de los organismos, y argumentan que estas últimas están constituidas en las formas de vida compartidas. Ofrecen la siguiente caracterización general de la idea de forma de vida:

Lo que es común a los seres humanos no es solo la biología que compartimos sino también nuestro estar embebidos en prácticas socioculturales: nuestro compartir formas estables de vivir con otros, nuestras formas de continuar relativamente estables (cf. “feste Lebensformen”; Wittgenstein, 1993, p. 397). (Rietveld & Kiverstein, 2014, p. 329)

Rietveld y Kiverstein identifican al menos tres niveles de análisis que la noción de forma de vida permite capturar: el primero, es la forma de vida compartida por todos los humanos en tanto especie, el segundo se refiere a las prácticas socioculturales y el tercero se refiere a la forma de vida particular de cada individuo (Rietveld & Kiverstein, 2014, p. 330). Por supuesto, no solo los humanos tenemos formas regulares de interactuar entre nosotros, sino que otras especies sociales también las tienen entre ellos y entre especies (gatos y perros son un ejemplo clarísimo de este fenómeno). Aún más, todas las especies tienen formas relativamente estables de hacer las cosas –de ahí que tenga sentido estudiar el comportamiento característico de las especies y no solo su constitución físico-química. Sin

embargo, Rietveld y Kiverstein se centran en formas de vida como prácticas sociales humanas y desde ahí construyen la concepción de normatividad que defienden.

De acuerdo con Rietveld y Kiverstein, la noción de normatividad se aplica a las interacciones que un individuo con habilidades establece con las affordances (2014, p. 332). Identifican dos elementos en esta noción. Primero, un individuo con habilidades es capaz de distinguir “actividades correctas de incorrectas, mejores de peores, óptimas de subóptimas o adecuadas de inadecuadas” (2014, p. 332) –en los términos que he usado, esto se refiere a la agencia. Segundo, las interacciones están situadas en un escenario material que determina si la interacción del agente con affordances es adecuada o no, y en ese sentido, han denominado a esta normatividad, *normatividad situada* (Rietveld, 2008). En ese escenario hay dos aspectos que determinan la corrección de la interacción: por un lado, las condiciones materiales específicas, y por otro lado, la concordancia entre la interacción específica y lo que hacen los miembros de la práctica. Esta situación material se refiere entonces a los criterios de corrección.

Las propuestas de Chemero y de Rietveld y Kiverstein han contribuido significativamente a aclarar qué significa que las affordances sean normativas. Pero todavía dejan una pregunta abierta: ¿qué hace que un patrón de comportamiento compartido por varias personas pueda establecer criterios normativos para interacciones individuales? Al respecto, mi propuesta es apelar a la intuición fundamental del enactivismo de que la normatividad surge en sistemas autónomos y argumentar que tales patrones de comportamiento son sistemas autónomos. Precisamente esta respuesta está contenida en la concepción de normatividad que propuse en el capítulo anterior en el que una de las dos condiciones de posibilidad de los fenómenos normativos es que haya un sistema autónomo respecto del cual se establece la corrección de una interacción específica.

La característica crucial para la emergencia de la normatividad de los sistemas autónomos es que son sistemas autoconstituidos, esto significa que están constituidos por partes o procesos que son, a su vez, producidos por el sistema mismo. De manera que un sistema autónomo genera las condiciones de su propia existencia. Vimos que, no solo todas las formas de vida, incluso las más básicas como las células, cumplen con esta condición, sino que otros sistemas como los tornados o las llamas también la cumplen. Vimos además que, en la literatura enactiva, se ha defendido ya el carácter autónomo de diferentes niveles de análisis de las formas de vida; por ejemplo, con el concepto de *hacer-sentido participativo* se ha defendido la autonomía de las interacciones entre agentes que se desenvuelven en el tiempo (De Jaegher & Di Paolo, 2007), la autonomía de las interacciones del agente con su ambiente que se repiten en el tiempo se ha defendido con el concepto de *hábito sensoriomotor* (Di Paolo et al., 2017) y la autonomía del patrón de interacciones lingüísticas de los hablantes de una lengua ha sido capturada por el concepto de *linguaging* (Di Paolo et al., 2018). En todos estos casos emergen sistemas autónomos porque las interacciones se repiten en el tiempo y crean paulatinamente las condiciones de su propio mantenimiento; tales interacciones además deben ser recurrentes para que el sistema se mantenga, sea este un hábito, interacciones entre agentes o actividades que surgen con el uso del lenguaje.

## Prácticas sociales como sistemas autónomos

Hay al menos dos antecedentes importantes de la idea de que las interacciones forman sistemas autónomos: por un lado, Luhmann (2007 [1997]) había defendido ya que la sociedad forma un sistema autopoiético, por otro lado, desde la década de los 80, la teoría de la práctica en sociología propuso una aproximación que reconcilia el estructuralismo y el interaccionismo y, para ello, defiende que las estructuras sociales son sistemas que se automantienen y están constituidos por la organización de interacciones, acciones, etc.

Luhmann sitúa su concepción de sociedad en oposición a concepciones esencialistas en las que, por ejemplo, la solidaridad o el consenso es la esencia de la sociedad (2007 [1997], p. 48), y en oposición a concepciones interaccionistas (1998, p. 23). Su propósito es identificar de manera abstracta y universal qué constituye un sistema social, para lo que lo concibe como un sistema autopoiético formado por operaciones de comunicación. Sin embargo, su concepción de autopoiesis está excesivamente centrada en la autoproducción e ignora la naturaleza precaria y materialmente realizada de los elementos que constituyen al sistema, a saber, acciones llevadas a cabo por seres vivos, en particular, humanos. Este énfasis formal y no material tiene dos consecuencias, primero, ignora el papel constitutivo de las interacciones entre agentes autónomos corporizados, y segundo, ignora que las operaciones autónomas se transforman en la medida en que se realiza el sistema social. Por esa razón, la propuesta de Luhmann es incompatible en elementos clave con la concepción de mente y cognición que he defendido, y, a partir del enfoque de la teoría de la práctica quisiera defender una concepción incipiente de sistema social como sistema autónomo que se automantiene mediante las acciones de agentes corporizados y socializados y sus interacciones en las situaciones materiales o nichos que han sido habitados por quienes constituyen tales sistemas sociales. En esta propuesta, la corporización del individuo es crucial porque incorpora habilidades de coordinación sensoriomotora que son socialmente constituidas y compartidas, y que permiten mantener la interacción social como operación que tiene sentido solo dentro de una práctica social.

La aproximación de la teoría de la práctica surge con Bourdieu (2016 [1979]) y su concepto de *habitus*, y Giddens (1995 [1984]) y su teoría de estructuración. Estas aproximaciones surgen como respuesta al debate sobre si los fenómenos y patrones sociales deben concebirse como resultado de la sumatoria de las acciones individuales o, en cambio, como resultado de estructuras que le dan orden a la sociedad y que existen más allá de los individuos. La teoría de la práctica propone una tercera vía en donde (i) los fenómenos sociales son entendidos como prácticas sociales, (ii) las acciones individuales e interacciones entre individuos reproducen las prácticas sociales y las mantienen existiendo y (iii) las prácticas sociales hacen posibles tales interacciones como acciones sociales. Más recientemente, Butler (2002) ha defendido que el género es performativo, es decir, que sucede en la repetición de gestos, acciones, interacciones, etc. que reproducen y mantienen estructuras sociales y culturales que establecen lo que significa ser hombre o mujer en una sociedad. Particularmente relevante para la concepción de mente, cognición y normatividad que he desarrollado en esta tesis,

Schatzki (1996) se basa en la filosofía de Wittgenstein y desarrolla una concepción de inteligibilidad y de mente como constituidas en prácticas sociales.

En lo que sigue, me centraré en la concepción de práctica de Schatzki no solo por la afinidad teórica con la aproximación a la cognición que he defendido hasta ahora, sino porque la claridad de su formulación permite desarrollar un argumento a favor de la concepción de forma de vida humana como sistema autónomo y generalizable a otras formas de vida. Al usar la teoría de la práctica como la propone Schatzki no busco defender que todo el reino de lo social se pueda caracterizar como prácticas sociales, sino busco mostrar cómo fenómenos sociales pueden constituir criterios normativos para el comportamiento de los agentes. Asumiendo que la forma de vida humana está constituida por prácticas sociales, muestro que las prácticas sociales son sistemas autónomos. Para ello muestro, primero, que la existencia de una práctica social depende de la continuación del patrón organizado de acciones individuales, interacciones entre individuos y del mantenimiento de las estructuras materiales que las permiten; en otras palabras, muestro que la existencia de una práctica en tanto sistema como un todo depende de que sus componentes continúen existiendo. Esta es la parte relativamente fácil del argumento, pero no es suficiente para mostrar que las prácticas son sistemas autónomos porque esta característica es verdad de sistemas no-autónomos como una licuadora o una máquina de escribir. Así que en la segunda parte del argumento muestro que las acciones individuales, interacciones entre agentes y los espacios materiales en los que las llevan a cabo son posibles porque pertenecen a una práctica; esto significa que la existencia de los componentes de la práctica, esto es, las acciones e interacciones de los individuos y los nichos, son posibles solo dentro de la práctica misma. Si este es el caso, entonces las prácticas son un sistema autoconstituido.

1. De acuerdo con Schatzki (1996) la inteligibilidad de las acciones humanas se configura dentro de prácticas sociales, y éstas a su vez están constituidas esencialmente por (i) la habilidad de actuar de formas específicas de cada uno de sus miembros, (ii) de reconocer cuando los otros lo hacen y (iii) de responder a las acciones de otros. Estas tres habilidades compartidas existen en la medida en que individuos lleven a cabo tal acción regularmente, por ejemplo, usar platos y cubiertos para comer, no escupir la comida si no nos gusta, comer con la boca cerrada, etc., que los otros reconozcan regularmente tales acciones como la acción de comer, en oposición, por ejemplo, a jugar con la comida, y que reaccionen regularmente a tales acciones de manera que se mantenga la posibilidad de interacción con el otro, en el ejemplo, esto consistiría en acciones como decir “en este momento está comiendo, por favor, llámalo más tarde”, en oposición a no poder nombrar ni saber reaccionar a las acciones del otro. La iteración de tales acciones, y de tales formas de reconocimiento e interacción hacen que se conserven los componentes materiales que permiten que estas acciones e interacciones se lleven a cabo, ej., cubiertos, platos, ollas, manteles, comedores, etc., y que sean reemplazados si ya no pueden conservarse o si dejan de ser apropiados. Así que acciones e interacciones hacen que se mantenga el nicho propio de la práctica, y al mismo tiempo, la existencia del nicho permite que las acciones e interacciones propias de nuestras formas de comer

sigan llevándose a cabo. En suma, acciones, interacciones y nicho forman una estructura social en el sentido en que estos patrones estructuran las diferentes posibilidades de hacer, en el ejemplo, las posibilidades de alimentarse, de manera que se organizan las acciones temporalmente y las situaciones en las que se suelen llevar a cabo. Estas estructuras existen solo en la medida en que los individuos de una comunidad las lleven a cabo frecuentemente.

2. Para ver por qué estas acciones y nicho específicos solo son posibles dentro de una práctica, quisiera empezar recordando qué hace que una acción tenga significado. He defendido que una de las características cruciales del sentido, en particular, el del lenguaje es que de las acciones o usos del lenguaje significativos tienen criterios públicos de normatividad que permiten determinar su corrección más allá de las impresiones inmediatas del individuo que las ejecuta. Mostré que hay dos formas en las que los criterios de corrección de los fenómenos normativos pueden establecerse: una es el mantenimiento de la vida propia en cada organismo individual y otra está fundamentada en los patrones de acción e interacción compartidos. Cuando una interacción con el ambiente o con otro agente no hace la diferencia para el mantenimiento de la vida orgánica del agente, o incluso cuando la manera en que una interacción fundamental para la vida se lleva a cabo tampoco hace tal diferencia, los criterios de corrección de estas interacciones están determinados por los patrones de acción e interacción compartidos con otros con quienes el individuo comparte su cotidianidad. Los criterios de corrección son parte fundamental del sentido en tanto que establecen las consecuencias prácticas de la interacción. Las prácticas sociales establecen tales consecuencias prácticas y los criterios de corrección de las actividades humanas: lo correcto es mantener el patrón de hacer y decir compartidos, y lo incorrecto, para ponerlo en términos extremos, es hacer o decir algo tan lejano a tal patrón que los otros no puedan responder y entablar o mantener una interacción.

El que existan estas prácticas sociales hace que una acción tan irrelevante para la vida como usar cubiertos para comer tenga sentido. El que tenga sentido hace que mantengamos la producción y el uso de estos artefactos y hace que los individuos sigan llevando tal acción a cabo, de lo contrario se tacharía al individuo de 'loquito', maleducado o incivilizado, esto es, alguien que no pertenece a nosotros; el que tenga sentido permite también que los individuos interactúen frecuentemente a través de tales acciones, –ej. ofrecerle cubiertos a los invitados–, y hace, además, que otras personas sean entrenadas para pertenecer a la práctica –los niños y los recién llegados de lugares en los que no se usan cubiertos para comer son entrenados para usar cubiertos<sup>12</sup>. Sin la existencia de tales patrones compartidos, no llevaríamos a cabo tales acciones, –por ejemplo, la acción de llevarnos a la boca la comida con un

---

<sup>12</sup> Schatzki ofrece un argumento similar en el que muestra que una acción socialmente significativa Schatzki, T. R. (1996). *Social practices: A Wittgensteinian approach to human activity and the social*. Cambridge University Press.



minitridente— ni someteríamos a nadie al proceso de aprender a usar tales herramientas para llevar a cabo una actividad que podría ejecutar perfectamente con las manos. En esa medida, la existencia de una práctica es condición de posibilidad para que interacciones que no tienen implicaciones para la vida de los organismos ocurran.

Las prácticas sociales están constituidas por acciones, interacciones individuales y espacios y artefactos materiales que forman un patrón estructurado, y al mismo tiempo estos son haceres y objetos con sentido en tanto que pertenecen a tales prácticas sociales. Esto significa que las prácticas sociales y las interacciones individuales tienen una relación circular lo que implica que las prácticas en tanto sistemas se autoconstituyen y, por lo tanto, son sistemas mínimamente autónomos.

Hay, sin embargo, una acotación importante que debe hacerse al carácter autónomo de las prácticas sociales y es que su autonomía no implica necesariamente que tengan agencia. Como discutí en el capítulo II, la autonomía no es condición suficiente para atribuirle agencia al sistema, aunque sí es una condición necesaria. Vimos que hay al menos dos condiciones adicionales, la primera, es que el sistema como un todo tenga la posibilidad de monitorear y regular sus propias interacciones en función de sus propias necesidades de automantenimiento, —a esta capacidad se le conoce en la literatura enactiva como adaptividad—. Una instancia de adaptividad es la capacidad de automantenimiento recursivo que propone Bickhard (2009) como característica de los sistemas que tienen agencia. Esta se refiere a la capacidad de un sistema de cambiar los términos de interacción y sus respuestas cuando las condiciones del ambiente cambian. Mientras todos los seres vivos individuales tienen esa capacidad adaptiva de automantenerse recursivamente en un rango de situaciones, otros sistemas autónomos como los tornados no las tienen. Autores como Stapleton y Froese (2015) han argumentado a favor del carácter adaptivo y, por ende, de cierto nivel de agencia de sistemas sociales como por ejemplo las colonias eusociales de insectos que forman hormigas, abejas, o termitas (2015, p. 225). Aunque estas colonias sean sistemas autónomos, no son agentes en un sentido inequívoco puesto que no cumplen con la segunda condición necesaria de la agencia, a saber, que el sistema tenga un límite físico autoconstituido que lo separe del ambiente, lo delimite como unidad, pero que permita la interacción, mínimamente, el intercambio de materia y energía.

Los sistemas en la naturaleza que inequívocamente exhiben agencia en su comportamiento son de hecho sistemas con límites físicos que los diferencian del ambiente, por eso, esta condición es presumiblemente esencial para que surja la agencia y una perspectiva de primera persona. El límite físico permite que el sistema tenga una mayor capacidad de regular lo que entra al sistema y lo que repele y, en esa medida, crea las condiciones para que procesos y organizaciones complejas puedan emerger, como los órganos y subsistemas como el sistema nervioso. Sin embargo, como señalan Stapleton y Froese (2015), no es claro qué hace que esta condición sea necesaria para que el sistema sea un agente o tenga una perspectiva de primera persona. De hecho, sugieren que quizás serían posibles en circunstancias particulares apelando no solo a colonias eusociales, sino también a resultados experimentales de

interacciones diádicas entre humanos en donde presumiblemente se exhibe una intencionalidad compartida con otro<sup>13</sup>.

El punto que quiero mostrar con esta discusión es el mismo que defienden Stapleton y Froese (2015): la propiedad de ser agente es una propiedad gradual y no binaria en tanto que, en ocasiones, sistemas autónomos sin límites autoconstituidos exhiben rasgos característicos de la agencia como el comportamiento flexible y sensible al cambio en las circunstancias para lograr algo. En ese sentido, el que un sistema autónomo sin límites autoconstituidos tenga agencia o sea capaz de tener experiencia subjetiva es una pregunta potencialmente abierta. Sin embargo, puesto que hasta ahora no se han encontrado instancias de sistemas sin límites materiales que exhiban agencia inequívocamente, es altamente probable que tales límites sean necesarios y que la agencia surja, de acuerdo con Stapleton y Froese, de sistemas cuyos elementos constitutivos estén materialmente integrados en un cuerpo viviente (2015, p. 234).

La integración material del sistema autónomo en un cuerpo viviente presumiblemente es necesaria para que surjan dos características esenciales de la agencia individual. La primera es que un agente es un sistema que forma unidad causal, es decir, interactúa con el ambiente como una unidad y es capaz de monitorear y regular tales interacciones. El que los agentes formen una unidad causal captura el hecho de que sus partes no interactúan con el ambiente de forma completamente autónoma, sino que siempre lo hacen como parte del sistema, por ejemplo, ni los ojos ni el cerebro ven autónomamente, sino que interactúan con otras partes del sistema, como los músculos, y esa interacción resulta en el comportamiento del sistema como unidad. La segunda característica es que los agentes, en su comportamiento como unidad, tienen un comportamiento expresivo. Como argumenta Schatzki (1996) su comportamiento expresa sus condiciones de vida, es decir, el curso de sus trayectorias como sistema, ej., si necesita alimento o descanso, o si está satisfecho, etc.

Podemos regresar ahora a la acotación sobre la agencia de las prácticas sociales en tanto sistemas autónomos. Evidentemente, no cumplen con el segundo requerimiento, es decir, las prácticas sociales no están delimitadas por un límite físico que las separe de otras prácticas y por eso difícilmente les atribuiríamos agencia. Sin embargo, a diferencia de un tornado, las prácticas sociales están constituidas por interacciones entre agentes individuales cuyo

---

<sup>13</sup> Stapleton y Froese se refieren a los resultados de experimentos psicológicos conocidos como el paradigma del cruce perceptual ('perceptual crossing' paradigm) conducido por Froese et al. (2014): usando la configuración minimalista de realidad virtual propuesta por Auvray y colegas (2009), que se conoce como el paradigma de 'cruce perceptual' (Auvray y Rohde 2012), exploraron las condiciones bajo las cuales la gente reconoce la presencia de otra persona sobre la base de las diferencias en la dinámica de la interacción únicamente. La interacción entre un par de participantes espacialmente separados estuvo mediada por una interfaz humano-computadora que reducía el alcance de su interacción a mover un avatar en un espacio virtual lineal e invisible y a recibir retroalimentación táctil cuando coincidían con un objeto virtual. Se les pidió a los participantes que colaboraran para encontrarse en el espacio virtual y evadir objetos distractores. Uno de estos objetos era estático y el otro era una reproducción instantánea, pero que no reacciona, de los movimientos del avatar del compañero. Se les dio 15 intentos en los que se les pedía que hicieran clic cuando sintieran que habían encontrado al otro participante, y después de cada intento se les pidió que calificaran la claridad de la experiencia de su compañero si habían hecho clic. No se les daba retroalimentación respecto a la corrección del clic durante el experimento. A pesar de la interacción social reducida y de la presencia de objetos distractores, la mayoría de las parejas de participantes lograban resolver la tarea en la mayoría de los intentos. (Stapleton y Froese, 2015, p. 231)

comportamiento es normativo, es decir, individuos que pueden seleccionar las trayectorias de sus interacciones respecto a sus intereses y errar o acertar respecto a una norma. Esto puede implicar que el comportamiento de una práctica social como sistema puede ser más flexible y resiliente que el de un tornado, aunque no necesariamente implica que la práctica misma también tenga agencia. Por otro lado, hay una diferencia crucial entre prácticas sociales y las colonias eusociales: mientras las prácticas están constituidas por clases de acciones e interacciones, las colonias están constituidas por individuos completos. Por ejemplo, las prácticas alimentarias están constituidas por familias de acciones, ej. usar cubiertos y platos, salir a comer en ocasiones especiales, comprar los alimentos para cocinarlos en vez de sembrarlos nosotros mismos, etc. Un mismo individuo puede pertenecer a tal práctica alimentaria porque lleva a cabo tales acciones cotidianamente y pertenecer a otras prácticas simultáneamente, como a las prácticas académicas o las prácticas de ser mujer. En contraste, las colonias eusociales están constituidas por individuos completos, su identidad individual está completamente formada por su lugar dentro de la colonia, ej., los olores y huellas químicas que producen las abejas u hormigas individuales depende de la colonia a la que pertenecen, y por eso, un mismo individuo no puede pertenecer simultáneamente a dos colonias. Las prácticas sociales se conectan entre sí en al menos dos sentidos, uno individual y uno colectivo.

### 5.3.3. Individuo y umwelt

En el nivel individual, las acciones de una práctica pueden llevar al individuo a otras prácticas, por ejemplo, la restricción alimentaria que hace parte de la práctica de hacer dieta frecuentemente lleva a los individuos a involucrarse en prácticas de alimentación desordenada o trastornos de la conducta alimentaria caracterizados por, ej., patrones de atracones de comida seguidos de purga o compensación. Adicionalmente, algunas prácticas se organizan rutinariamente en la misma sucesión temporal, ej., ir a trabajar, regresar a casa y cenar con la familia, ver televisión, dormir, ir a trabajar, etc., de forma que unas llevan al individuo a las otras sin esfuerzo. Estas dos observaciones apuntan a un hecho más general y es que un mismo individuo pertenece a múltiples prácticas distintas simultáneamente, por ejemplo, ser madre, ser académica, ser fan de las cumbias, etc. Dentro de tales prácticas se constituyen los roles sociales de la persona porque en estas se establecen sus capacidades para interactuar con el mundo que se expresan y ejecutan en su cuerpo, y porque estas establecen los criterios de corrección de su comportamiento. Tales prácticas, además, constituyen la forma de vida del individuo.

Las habilidades de interacción de un agente humano con otros y con el ambiente material se forman durante la socialización o, para usar la terminología wittgensteiniana que introduce en el Capítulo 1, en el entrenamiento de los sujetos para introducirlos a una práctica. Como lo ha defendido la literatura enactiva ampliamente, estas habilidades, es decir, el producto del entrenamiento es la sedimentación de patrones de interacción en el cuerpo de los agentes individuales. James y Loaiza (2020) proponen el concepto de inter-identidades interpersonales (IIs) para capturar tanto la sedimentación de patrones de interacción en una diada o en un grupo, como a la forma en que las normas establecidas por las prácticas sociales

a las que pertenecen son reproducidas y transformadas en sus interacciones concretas (2020, p. 4). Las inter-identidades interpersonales se manifiestan en los intereses y preocupaciones de cada individuo en la interacción, en sus acciones concretas y sus formas de ejecutarlas, y están animadas por la práctica a la que pertenecen. Para continuar con el ejemplo que James y Loaiza usan, la interacción cotidiana de una pareja romántica reproduce y transforma las prácticas a la que pertenecen ambos, por ejemplo, si suelen comer juntos, reproducen las prácticas de alimentación al seguir usando cubiertos o al designar la responsabilidad de comprar y preparar los alimentos a la mujer, o las transforman al incorporar palillos a sus formas de comer cotidianas después de su viaje a Japón o al decidir un horario compartido en donde ambas partes de la pareja comprenden y preparan los alimentos por igual.

Como discutí arriba, la psicología ecológica propone concebir la percepción como inseparable de la acción tanto porque el agente debe explorar activamente el ambiente para poder descubrir los patrones de reverberación de la luz y el sonido, como porque estos patrones hacen sentido para el agente en tanto que ofrecen posibilidades de acción. Mostré además que uno de los elementos constitutivos de las prácticas sociales son las habilidades para actuar. Como señalan Rietveld y Kiverstein (2014), estas dos ideas implican que cuando un agente pertenece a una práctica social, percibe el ambiente en términos de las posibilidades de acción establecidas en tal práctica social; por ejemplo, pertenecer a las prácticas alimenticias compartidas en Japón permite a quienes pertenecen a ella percibir los palillos como una herramienta para comer.

Rietveld y Kiverstein (2014) defienden que los individuos perciben el ambiente como un paisaje de affordances; se pueden diferenciar al menos dos sentidos de esa afirmación. Primero, pertenecer a una práctica implica normalmente tener la habilidad de ejecutar varias acciones que se conectan entre sí –ej., comer normalmente con palillos, abrir adecuadamente el empaque de un onigiri y saber comerlo, etc., o comer normalmente con cubiertos, saber la cantidad de comida indicada para poner en una tortilla y cómo sostener el taco para comerlo, etc.–, en donde los aspectos materiales en los que se ejecuta esta serie de acciones se presentan como posibilidades de acción interconectadas. En ese sentido, un individuo que pertenece a una práctica percibe el ambiente como un paisaje de affordances. Segundo, puesto que un individuo pertenece simultáneamente a varias prácticas, su percepción del ambiente no se limita a solo una familia de posibilidades dentro de una práctica. De hecho, puede percibir diferentes capas de significado en un mismo ambiente en tanto que aspectos de este le permiten enactuar diferentes prácticas. Por ejemplo, una mesa puede ser una posibilidad para sentarse a comer, o para sentarse a escribir artículos o para jugar Monopoly. Por lo tanto, los individuos perciben el ambiente como un paisaje de affordances conectadas por la misma práctica y como affordances que se constituyen en diferentes prácticas.

Este paisaje de affordances es el *umwelt* que, vimos arriba se refiere al aspecto del ambiente que ofrece información para el individuo y es relativo a sus habilidades particulares. Después de la caracterización de prácticas sociales que he hecho hasta ahora podemos refinar el concepto de *umwelt* y entenderlo como relativo al conjunto de prácticas sociales a las que pertenece el individuo. Esta especificación tiene una consecuencia interesante: incluso el

*umwelt* en tanto el subconjunto del ambiente que habita un individuo tiene condiciones de corrección que no están completamente determinadas por el individuo, y en el caso humano que estamos considerando, están determinados por los patrones de acción e interacción compartidos con otros. La razón es que, como argumenté arriba, pertenecer a una práctica requiere ser reconocido por otros, interactuar y mantener abiertas las posibilidades de interacción con ellos, y por ende, esto permite determinar si la forma en la que el individuo interactúa con su propio ambiente es correcta o no.

Hay, por supuesto, varios sentidos en que los criterios de normatividad pueden ser incorporados parcialmente por un individuo que interactúa. Por ejemplo, puede incorporar al horizonte de sus posibilidades de acción aspectos del ambiente que no son perceptibles inmediatamente. En el sentido más básico, esto significa que aspectos invisibles de un objeto, como las caras escondidas de un objeto tridimensional son incorporadas como posibilidades de acción en el presente, y como proponen Froese y Di Paolo (2009) en su explicación de cómo emerge la percepción de objetos que existen en sí mismos más allá de las posibilidades presentes del agente, se vuelven parte de la percepción presente como parte del horizonte de posibilidades.

Un segundo elemento que los individuos pueden incorporar a su agencia son posibles eventos en la interacción presente o en interacciones futuras provocados por otro agente que implicarían cambios en las trayectorias de la interacción. Por ejemplo, como argumentan Cuffari et al. (2015), los individuos pueden incorporar actos parciales a su repertorio de posibilidades acción en los que se anticipa la respuesta del otro con su propia agencia, como por ejemplo en el acto de dar o de saludar.

Sin embargo, como señalé en el capítulo anterior, esta incorporación no significa que el individuo pueda hacer por completo suyas todas las condiciones de normatividad que le dan sentido a sus interacciones como acciones sociales. La razón es que las consecuencias prácticas de sus interacciones y de la forma en que las lleva a cabo dependen de si hacen o no parte del patrón de interacciones compartidos con otros, y lo que hace que una interacción particular cuente como parte del patrón compartido es que otros mantengan sus interacciones con el agente como un individuo que hace sentido. Así que mantener la interacción con otros es crucial para hacer sentido más allá de la normatividad impuesta por las necesidades biológicas. Esto implica, a su vez, que la identidad de la persona es también establecida por la manera en que otros dentro de las prácticas a las que pertenece la reconocen. Esto explica, por ejemplo, que una persona pueda tener características con las que ella no se identifica, como ser racista, por el significado de sus acciones en el contexto de las prácticas sociales a las que pertenece.

#### 5.3.4. Formas de vida compartidas y hábitat

Por supuesto, no todos los agentes de una misma comunidad participan de las mismas prácticas, ej., tienen trabajos, pasatiempos, identidades sexuales y edades diversas. A pesar de la diversidad, comparten un piso de formas comunes de interactuar con el ambiente e

interactuar entre sí que les permite reconocerse unos a otros como individuos que hacen sentido y que pertenecen a una misma forma de vida. Siguiendo los diferentes niveles de análisis que proponen Rietveld y Kiverstein (2014), en el nivel más general, compartimos una forma de vida humana: nos reconocemos unos a otros como miembros de la misma especie, mínimamente reconocemos al otro como un cuerpo vivo y expresivo que tiene necesidades de, ej., alimento, descanso y socialización, similares a las propias, le reconocemos como alguien que pertenece a prácticas sociales construidas con el lenguaje aunque no compartamos ninguna, etc.<sup>14</sup> En un análisis más específico, podemos identificar como formas de vida compartidas a los patrones de interacción que forman los hablantes de una misma lengua o, en un análisis más específico, el que forman los académicos dedicados a estudiar la función de ciertas construcciones lingüísticas comunes en el alemán hablado o los fotógrafos que se dedican al fotolibro en iberoamérica o las mujeres blancas de clase alta de Estados Unidos o las mujeres transgénero que viven en la Ciudad de México.

Este nivel de análisis relativo a lo colectivo al que da lugar el concepto de forma de vida como lo proponen Rietveld y Kiverstein (2014), permite profundizar la definición de hábitat que proponen Baggs y Chemero (2018). Como vimos arriba el concepto de hábitat busca capturar los aspectos materiales que ofrecen posibilidades de acción independientes de las habilidades específicas de un organismo, pero relativas a las habilidades compartidas de una especie. Para ello, Baggs y Chemero (2018) conceptualizan el hábitat como los aspectos del ambiente relativos a un animal genérico de una especie. Sin embargo, la idea de un animal genérico es problemática por dos razones. Primero, no es claro qué criterios deberíamos usar para decidir qué habilidades o características cuentan como genéricas y cuáles cuentan como particularidades. Segundo, la idea misma de un animal genérico prescinde de la diversidad de habilidades entre organismos de una misma especie y en esa medida, excluye todos los cuerpos y sus habilidades considerados fuera de lo estándar. Precisamente este fenómeno ha ocurrido en la medicina en donde hasta hace muy poco tiempo se ha considerado que el humano genérico es joven adulto (Cherubini et al., 2011), blanco (Hussain-Gambles et al., 2004) y su sexo biológico es masculino. Bajo esta idea de humano genérico, algunas formas de diagnóstico han ocultado padecimientos cuya manifestación es distinta en humanos que no coinciden con el estándar, por ejemplo, las fallas cardíacas en mujeres (Heiat et al., 2002) o el trastorno por déficit de atención en mujeres (Quinn & Madhoo, 2014) tienen síntomas distintos a las de los hombres y por eso han sido más difíciles de diagnosticar. Para evitar este riesgo de invisibilización asociado con la idea de un animal genérico propongo entender el hábitat como relativo a formas de vida en el nivel colectivo de análisis.

---

<sup>14</sup> Por supuesto, el reconocimiento del otro como miembro de la misma forma de vida en el nivel de generalidad o especificidad que sea no implica un trato ético hacia el otro. En Pescador Canales, C., & Mojica, L. (2022). Making us Autonomous: The Enactive Normativity of Morality. *Topoi*, 41(2), 257-274. Cassandra Pescador y yo argumentamos que este reconocimiento ocurre en la interacción y que implica el imperativo moral de cuidar las autonomías tanto del otro como la propia. La forma en que estos cuidados se llevan a cabo y algunos de los espacios y situaciones que implican tal necesidad de cuidado varían de acuerdo con las prácticas sociales a las que pertenecen los interactuantes. Sin embargo, el imperativo moral de cuidar la autonomía del otro en tanto miembro de la forma de vida compartida por la humanidad es invariable y rige moralmente tanto en la interacción con el otro como en el diseño de nuestros espacios.

Las formas de vida colectivas, como he insistido, se constituyen en los patrones de interacción compartidos. Estos patrones incluyen la interacción dirigida a aspectos materiales del ambiente y la interacción con otros a través de aspectos materiales como teléfonos, computadoras o marcas de tinta en un papel. Tales aspectos materiales son significativos para los miembros de una misma forma de vida en tanto que les ofrecen posibilidades de acción a sus miembros o a un subconjunto de ellos y estos constituyen el hábitat relativo a diferentes formas de vida. Según el nivel de análisis que la situación específica requiera, puede elegirse el nivel de generalidad de la forma de vida que puede ir desde el nivel más amplio de ser parte de una misma especie hasta particularidades específicas de la forma de vida que forma una pareja de personas casadas durante años.

Para terminar, quisiera regresar a la afirmación que defendí en la primera parte del capítulo, a saber, que las affordances son normativas. Después de haber discutido los conceptos de forma de vida y prácticas sociales, ahora es posible explicar más precisamente qué significa esta afirmación. El concepto de affordance captura la relación entre el hábitat y las habilidades disponibles en una forma de vida en el nivel colectivo. Esta relación consiste en los patrones de interacción entre agentes y una familia de aspectos materiales del hábitat con los que han establecido o podrían establecer una interacción similar.

Por un lado, el concepto de affordance se refiere a las interacciones entre habilidades de los agentes y aspectos materiales del ambiente. Esta relación no se limita a aspectos materiales específicos del hábitat con los que ya han ocurrido interacciones de hecho, sino que incluye aspectos con los que potencialmente podrían interactuar los agentes dadas las habilidades disponibles en su forma de vida. Por eso propongo definir el concepto de affordance en términos de una familia de aspectos materiales. Esto se refiere a un conjunto de aspectos materiales similares entre sí en algún respecto, pero no necesariamente el mismo, como es el caso en una familia. Lo que une a esta familia de aspectos materiales es que todos permiten que algún agente ejecute cierta habilidad disponible en su forma de vida compartida, es decir, es un conjunto definido en función de una habilidad. Definir el aspecto material de un affordance de esta manera captura el hecho de que hay aspectos materiales que ofrecen posibilidades de acción a individuos de una forma de vida a pesar de que nadie haya interactuado con estos nunca, por ejemplo, una rama de un árbol ofrece a un mono la posibilidad de colgarse a pesar de que nunca haya sido usada por ningún mono para ello. El que exista un aspecto material que sea un affordance significa entonces que hay una habilidad que podría ser ejecutada en la interacción con este, y por lo tanto, implica la existencia de criterios de corrección para las interacciones que un agente establezca con dicho aspecto material.

Por otro lado, los patrones de interacción se establecen en la medida en que los individuos llevan a cabo repeticiones de la misma acción, y como defendí arriba, estas se llevan a cabo porque contribuyen al mantenimiento de su identidad biológica, o porque se insertan en el patrón de interacciones compartidas en la forma de vida a la que pertenece, o ambas. Entonces, el patrón de interacciones se establece como consecuencia normativa o bien de las necesidades de automantenimiento de una forma de vida colectiva en la que se sitúa la

interacción específica, o bien de las necesidades de automantenimiento de la vida individual de cada ser vivo. En cualquiera de los dos casos, la interacción está sujeta a los criterios de normatividad establecidos por el mantenimiento de una forma de vida. En caso de que se trate de la forma de vida de un individuo, se trata no solo de un sistema autónomo que, por ende, establece criterios de corrección, sino que también es un sistema adaptivo y por ello implica también la posibilidad de selección de su comportamiento como sistema completo y entonces agencia y un comportamiento sujeto a criterios de normatividad. En caso de que se trate de una forma de vida compartida por un conjunto de individuos se trata también de un sistema autónomo que establece criterios de diferenciación normativa; sin embargo, no es adaptivo porque no tiene la capacidad de seleccionar sus propias trayectorias y además no tiene límites materiales autoconstituidos por lo que su comportamiento en tanto sistema completo no está sujeto a criterios de normatividad y entonces no constituye un agente. Por lo tanto, sea al nivel individual o colectivo de una forma de vida, los patrones de interacción que constituyen un affordance establecen criterios de corrección para las interacciones particulares de los individuos que pertenecen a tal forma de vida, y por ello, las affordances son normativas.

## 5. 4. Conclusión

En este capítulo mostré que las affordances, entendidas como la relación entre habilidades disponibles en una forma de vida y aspectos materiales del ambiente, son normativas. Empecé refutando las razones que ofrece Heras-Escribano para negar la normatividad de las affordances. Argumenté que, si bien la normatividad requiere de criterios públicos, esto no implica que deban ser sociales como él defiende, y por ello no es necesario limitar la normatividad a fenómenos cuyos criterios estén constituidos en prácticas sociales. Argumenté además que las formas de interacción de los seres vivos con el ambiente requieren de una explicación en términos normativos porque exhiben la capacidad de selección y un comportamiento flexible de manera que pueden errar o acertar respecto al mantenimiento de su propia forma de vida. Concluí esta discusión mostrando que, a pesar de que el comportamiento de todos los seres vivos es normativo, no todos son capaces de establecer interacciones cognitivas. Estas últimas están reservadas, como correctamente sugiere Heras-Escribano, a las interacciones cuyos criterios de normatividad están constituidos socialmente.

En la segunda mitad del capítulo, argumenté que las affordances son normativas en tanto que se refieren a la relación entre habilidades propias de una forma de vida y aspectos del ambiente. Había mostrado ya en capítulos anteriores que los organismos vivos forman sistemas autónomos, y en ese nivel biológico son una forma de vida individual que establece los criterios normativos para sus propias interacciones en la escala temporal de su vida. En este capítulo me centré en la discusión de las formas de vida compartidas como sistemas autónomos que establecen criterios de corrección sociales y cognitivos en contraste con tal normatividad biológica. Empecé conceptualizando formas de vida como prácticas sociales y mostré que estas son sistemas autónomos, pero no sistemas con agencia, en tanto que constituyen interacciones con sentido y están constituidas por el patrón organizado de tales interacciones. Las prácticas sociales, en virtud de su naturaleza autónoma, establecen los criterios de corrección de interacciones particulares.



## Recapitulación y conclusiones

Defendí que la normatividad ocurre cuando dos condiciones se cumplen: cuando un individuo es capaz de seleccionar entre diferentes formas de interacción o respuesta, llamé a esta capacidad agencia mínima, y cuando existen criterios normativos para determinar la corrección de tales interacciones y respuestas que sobrepasan la interacción particular y las impresiones que el agente pueda tener al respecto. Esto implica, por un lado, que, si un individuo tiene alguna capacidad de selección, pero no hay criterios normativos para determinar su corrección más allá de la interacción misma, entonces las interacciones producto de esta selección no son propiamente normativas. Por otro lado, esta concepción de normatividad implica también que, si hay criterios normativos para determinar interacciones o respuestas correctas, pero no hay individuos que de hecho puedan seleccionar sus cursos de interacción o sus respuestas, entonces tales criterios tampoco constituyen por sí mismos un fenómeno propiamente normativo.

La forma más básica en que la normatividad ocurre en la naturaleza es cuando un individuo es capaz de seleccionar cursos de interacción en el presente que hacen la diferencia sobre su propia existencia en la escala temporal de su vida. Este es el caso en sistemas cuya organización es autónoma, es decir, cuando son sistemas que se autodiferencian y se automantienen en el tiempo, pero que además tienen la capacidad de modificar sus formas de automantenimiento para seguir existiendo cuando las condiciones cambian —a esta última capacidad se le conoce como adaptividad en la literatura enactiva o como capacidad de mantenimiento recursivo en el interactivismo. En estos sistemas básicos surgen simultáneamente agencia y criterios normativos. Por un lado, la organización autónoma del sistema hace que este tenga que mantener activamente su existencia lo que establece los criterios de corrección porque hace posible diferenciar sus propias respuestas según las implicaciones prácticas que estas hacen para la existencia del sistema: o permiten mantener su organización o contribuyen a su disipación. Es crucial notar que las implicaciones prácticas de los eventos que se diferencian normativamente son las respuestas del sistema o falta de ellas e impactan la existencia del sistema mismo, y por ello son criterios normativos propios de su existencia. Esto lleva al segundo punto, a la agencia mínima: cuando los sistemas autónomos son además adaptivos, tienen la capacidad de seleccionar respuestas diferentes para mantener su existencia. La adaptividad implica que el sistema puede hacer realmente una diferencia para su propia existencia y por ende que puede responder correcta o incorrectamente respecto a los criterios de corrección que este mismo establece.

Aunque en esta forma básica de normatividad, los criterios de corrección y la agencia mínima surgen como dimensiones de un mismo sistema que se automantiene, estas constituyen dimensiones conceptual y fácticamente distintas. Por un lado, los comportamientos que selecciona el agente están en una escala temporal más corta a la escala temporal en la que se constituyen los criterios normativos. Esto hace posible que los sistemas seleccionen cursos de acción de manera ultrarrápida, pero que esta selección sea evaluada en una escala temporal más amplia. Por otro lado, un agente puede actuar correcta o incorrectamente respecto a criterios normativos que no son relevantes o que incluso pueden ser dañinos para su propia

agencia. Esto es evidente tanto en casos de función biológica en los que hacer lo correcto implica la muerte del organismo (ej. reproducción de los pulpos) y en casos de actividades humanas que atentan contra la viabilidad del organismo (ej. fumar).

La noción de normatividad del enactivismo se ha fundamentado en el automantenimiento del organismo, pero hasta ahora no se ha diferenciado explícitamente en la literatura la dimensión inmediata de la experiencia del organismo individual en interacción de sus consecuencias en una escala temporal más amplia que la regulación inmediata de tal interacción. El no distinguir explícitamente estas dos dimensiones de la normatividad ha hecho que en ocasiones se equipare el presente de la experiencia de primera persona del organismo con la naturaleza temporalmente extendida de su existencia, y que se equipare además la agencia con los criterios normativos. El no separar estas dimensiones ha llevado a que la propuesta enactiva dé la impresión de defender inadvertidamente una posición internalista en el que los agentes calculan internamente las consecuencias prácticas de sus trayectorias —esta es precisamente la objeción que hacen Hutto y Myin al enactivismo autopoietico (2013)—, o bien de defender mecanismos de monitoreo y control sin que sea clara su escala temporal y que entonces hace que no sea evidente cómo es posible que hayan respuestas ultrarrápidas normativas cognitivamente—esta es el punto que Froese y yo planteamos en (2019)—, o incluso de no reconocer la complejidad de los fenómenos cognitivos y buscar reducirlos a la normatividad básica del mantenimiento de la vida en donde agencia y criterios normativos están fundamentados en la autonomía y adaptividad de un mismo organismo lo que hace difícil, si no imposible, explicar cómo es posible que existan consecuencias prácticas y criterios normativos para diferenciar interacciones que son irrelevantes para nuestro mantenimiento como seres vivos.

Propuse la siguiente definición de normatividad para capturar esta distinción: un fenómeno normativo consiste en (i) interacciones que agentes autónomos y adaptivos establecen, y (ii) que están sujetas a criterios normativos; tales criterios son siempre la norma de automantenimiento de un sistema autónomo que puede ser igual al agente que establece la interacción o puede contenerlo. Un agente es, por ejemplo, un óvulo y su comportamiento está sujeto a los criterios normativos de su propio automantenimiento, el mantenimiento del ser humano en el que se encuentra y el mantenimiento de la especie humana. Esta definición de normatividad tiene al menos dos motivaciones. Por un lado, busca capturar la idea de que el sentido de una acción o del uso del lenguaje radica en sus consecuencias prácticas y públicas, y por ende, sobrepasa las impresiones individuales del agente que actúa; y busca generalizar esta idea a otros fenómenos normativos. Por otro lado, busca refinar dos propuestas del enactivismo respecto al fenómeno de la normatividad: primero, que surge de sistemas autónomos capaces de hacer una diferencia para su propia existencia y, segundo, que las formas básicas de normatividad forman un continuo con formas más complejas de normatividad que pueden encontrarse en la naturaleza y que todas están fundamentadas en la norma del automantenimiento.

Diferenciar los criterios normativos de la agencia permite separar las consecuencias prácticas de una interacción de la perspectiva del agente, y esto hace que no sea necesario que el agente internalice ni que sea la única fuente que determina las consecuencias de sus acciones en el

aquí y el ahora para que estas sean normativas, tengan sentido y sean propiamente cognitivas. Esto tiene varias ventajas:

La primera ventaja es que esta distinción da lugar a una explicación de la cognición consistente con el propósito del enactivismo de no recurrir a representaciones mentales ni cálculos internos, sino que apela a los mecanismos y procesos emergentes de los sistemas dinámicos, lo que responde a las acusaciones de internalismo de Hutto y Myin (2013). Esto es una ventaja porque este tipo de explicación no postula entidades o procesos internos como criterios para determinar la corrección de interacciones cognitivas; si este fuera el caso, tales criterios no estarían disponibles más allá de las impresiones propias del agente y, por lo tanto, no habría una posibilidad real u objetiva de errar o estar en lo correcto. Los criterios normativos, en cambio, están situados en la constitución temporalmente extendida de un sistema autónomo.

La segunda ventaja de esta distinción es que permite dar cuenta del surgimiento y consolidación de comportamientos correctos que atentan contra el automantenimiento del agente. Cuando un agente está situado o participa de un sistema autónomo más amplio, sus comportamientos están sujetos no solo a sus propias normas de automantenimiento, sino también a la norma del automantenimiento del sistema más amplio. En muchas formas de vida, es común que esta última norma sea irrelevante para el mantenimiento del agente o que incluso atente contra este. Así que el agente puede estar actuando correctamente con respecto a las normas impuestas por el sistema autónomo al que pertenece, aunque este mismo comportamiento sea inadecuado para mantenerse con vida o lo lleve a la muerte.

La tercera ventaja de esta distinción es que abre la posibilidad de que las interacciones normativas puedan ocurrir en una escala temporal ultrarrápida mientras que los procesos de evaluación en los que se determina su corrección se desenvuelven en escalas temporales más amplias, lo que responde directamente la pregunta que planteamos Froese y yo en (2019) a partir de los casos de cognición ultrarrápida. La razón es que la diferenciación del papel del agente individual, por un lado, de las condiciones históricas y situacionales en las que se debe desenvolver su interacción para ser un fenómeno normativo, por el otro, permite que no se asuma intencional o inadvertidamente que el individuo es el único que constituye la normatividad y el sentido de sus interacciones ni que debe hacerlo en cada instante de su interacción. En esa medida, la selección ultrarrápida de un curso de interacción es un fenómeno cognitivamente normativo en tanto que se inserta en los patrones de acción compartidos por nuestras comunidades, y quizás nuestra especie, y que establecen los criterios para determinar la corrección de tales interacciones individuales, además de que la interacción misma es de hecho la selección de un curso de acción entre las varias posibilidades de acción del agente.

Una última consecuencia que quisiera resaltar de la caracterización de normatividad que he defendido es que el sentido de nuestras acciones, incluido nuestro uso del lenguaje, no está determinado completamente por nuestras motivaciones o impresiones individuales, sino que está ampliamente determinado por sus consecuencias prácticas en el ambiente sociomaterial que habitamos y respecto a otros. Esto se sigue de que los criterios normativos determinan la

corrección de una acción y existen más allá de la interacción individual y de las impresiones que el agente pueda tener sobre esta. El significado social de nuestras acciones es un punto que ya ha sido ampliamente reconocido por las ciencias sociales, y es evidente en el hecho de que podemos hacer y decir cosas de cuyo significado realmente no somos conscientes. Pero si el sentido de las interacciones que establecemos está determinado por prácticas sociales, es decir, por las consecuencias prácticas para el ambiente y para otros que permite mantener o no una práctica como sistema autónomo, entonces quienes constituyen la comunidad a la que el agente pertenece son fundamentales para darle sentido a su propia vida.

Esto hace explícita una dimensión afectiva no solo de los humanos sino de cualquier criatura cuya vida cotidiana transcurra en interacción constante con otros, a saber, mantener la forma de vida propia implica mantener la forma de vida de otros con las que se entretiene la cotidianidad propia y contribuye a darle sentido. Es importante resaltar que por 'afectivo' me refiero a la motivación sentida a actuar, interactuar con otro o con algún aspecto del ambiente, y esta puede tomar formas distintas y valencias tanto positivas como negativas. Por lo tanto, la motivación afectiva del automantenimiento no está únicamente dirigida a uno mismo, sino que contempla inmediatamente al otro y a comunidad a la que uno pertenece como dimensión constitutiva de lo que hace sentido y de lo que hace una diferencia en la propia vida más allá de mantener el metabolismo corriendo. Podría pensarse en esta motivación afectiva como la motivación para actuar y atender aquello que está en el ambiente que uno habita; y esto es evidente, por ejemplo, en las relaciones de cooperación entre especies y de cuidado de los miembros de una misma manada o descendencia.

La motivación afectiva hacia el otro y la comunidad en la que uno habita tiene además una dimensión moral en el caso de los humanos. Tal motivación afectiva hacia el otro es una capacidad de interactuar de ciertas maneras y hace parte de lo que podemos seleccionar o no en nuestras interacciones, en esa medida, puede estar sujeta a criterios normativos. Por su parte, tales criterios normativos están constituidos en nuestras prácticas sociales y le dan significado a la forma en que nos acercamos afectivamente a otros. Quizás, una de las formas de motivación afectiva primarias es actuar como aquellos que nos cuidan o con quienes convivimos cotidianamente y pertenecer a esa comunidad, habitar con ellos; pero esta motivación no agota el significado ni todas las dimensiones normativas de las acciones llevadas a cabo por esta motivación, sino que está sujeta a los criterios normativos más amplios que forman las prácticas sociales a las que pertenece y reproduce esa comunidad aun cuando el agente no sea consciente de ellas.

Una instancia relevante de este fenómeno es cuando nuestras acciones reproducen prácticas de exclusión, por ejemplo, prácticas clasistas o racistas, sin darnos cuenta o incluso creyendo explícitamente que no lo somos; es posible incluso nunca darse cuenta del significado que tales acciones o dichos tienen para otros y continuar reproduciéndolas solo bajo la motivación afectiva de pertenecer junto con los nuestros. Como defendemos Canales y yo (artículo en proceso de revisión), tales prácticas son excluyentes porque disminuyen la capacidad de otros de actuar y elegir cursos de acción, y precisamente cuando la capacidad de actuar de otro o de uno mismo está en juego, las acciones e interacciones están sujetas a criterios normativos morales. Hay entonces, por un lado, una responsabilidad moral

individual de descubrir el significado de nuestras acciones cotidianas más allá de la impresión afectiva inmediata que tienen en nosotros de ser lo que se hace en nuestra comunidad y transformar tales interacciones cuando disminuyan la autonomía de otros o de uno mismo. Hay, por otro lado, una responsabilidad moral colectiva de eliminar estructuras materiales y desfavorecer mecanismos institucionales que perpetúen prácticas en donde se disminuya sistemáticamente la capacidad de seleccionar cursos de acción de ciertas poblaciones; tomar medidas en este sentido disminuiría la posibilidad de que las acciones individuales disminuyan la autonomía de otros inadvertidamente. Por supuesto, la pregunta más acuciante para una investigación futura al respecto es cómo hacerlo.

# Bibliografía

- Agmon, E., Gates, A. J., & Beer, R. D. (2015). Ontogeny and adaptivity in a model protocell. *Artificial Life Conference Proceedings* 13,
- Allen, C., & Neal, J. (2020). Teleological Notions in Biology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 ed.): Metaphysics Research Lab, Stanford University.
- Altmann, G. T. (2011). Language can mediate eye movement control within 100 milliseconds, regardless of whether there is anything to move the eyes to. *Acta psychologica*, 137(2), 190-200.
- Baggs, E., & Chemero, A. (2018). Radical embodiment in two directions. *Synthese*. <https://doi.org/10.1007/s11229-018-02020-9>
- Barandiaran, X. E. (2017). Autonomy and enactivism: Towards a theory of sensorimotor autonomous agency. *Topoi*, 36(3), 409-430.
- Barandiaran, X. E., Di Paolo, E., & Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5), 367-386.
- Barandiaran, X. E., & Egbert, M. D. (2014). Norm-establishing and norm-following in autonomous agency. *Artificial Life*, 20(1), 5-28.
- Barbaras, R. (2008). Life, movement, and desire. *Research in Phenomenology*, 38(1), 3-17.
- Barrett, N. (2017). The normative turn in enactive theory: An examination of its roots and implications. *Topoi*, 36(3), 431-443.
- Barrett, N. (2019). On the nature and origins of cognition as a form of motivated activity. *Adaptive Behavior*, 28(2), 89-103.
- Barrett, N. (2020). Extremal properties and self-preserving behavior. *Adaptive Behavior*, 28(2), 113-118.
- Beck, J. (2012). The Generality Constraint and the Structure of Thought”, *Mind*, 121(483): 563–600. doi:10.1093/mind/fzs077
- Beer, R. D. (1996). Toward the evolution of dynamical neural networks for minimally cognitive behavior. *From animals to animats*, 4, 421-429.
- Bennett, M., & Hacker, P. (2003). *Philosophical Foundations of Neuroscience*. Blackwell Publishing.
- Bickhard, M. H. (2000). Autonomy, function, and representation. *Communication and Cognition—Artificial Intelligence*, 17(3-4), 111-131.
- Bickhard, M. H. (2009). The interactivist model. *Synthese*, 166(3), 547-591.
- Bickhard, M. H. (2016). Inter-and En-activism: Some thoughts and comparisons. *New Ideas in Psychology*, 41, 23-32.
- Bloor, D. (1983). *Wittgenstein: A Social Theory of Knowledge*. Columbia University Press.
- Bourdieu, P. (2016 [1979]). *La distinción: criterio y bases sociales del gusto*. Taurus.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial intelligence*, 47(1-3), 139-159.
- Butler, J. (2002). *Gender Trouble: Feminism and the Subversion of Identity*. Routledge.
- Canguilhem, G. (1966 [1978]). *On the Normal and the Pathological* (C. R. Fawcett, Trans.). D. Reidel Publishing Compny.
- Cely Avila, F. E., & Mojica, L. (2019). Afecto y sentido. *Pensamiento. Revista de Investigación e Información Filosófica*, 75(285), 913-929.
- Chemero, A. (2003). An outline of a theory of affordances. *Ecological psychology*, 15(2), 181-195.
- Chemero, A. (2009). *Radical embodied cognitive science*. MIT Press.
- Cherubini, A., Oristrell, J., Pla, X., Ruggiero, C., Ferretti, R., Diestre, G., Clarfield, A. M., Crome, P., Hertogh, C., Lesauskaite, V., Prada, G.-I., Szczerbinska, K., Topinkova, E., Sinclair-Cohen, J., Edbrooke, D., & Mills, G. H. (2011). The Persistent Exclusion of Older Patients From Ongoing Clinical Trials Regarding Heart Failure. *Archives of Internal Medicine*, 171(6), 550-556. <https://doi.org/10.1001/archinternmed.2011.31>
- Christensen, W. D., & Bickhard, M. H. (2002). The process dynamics of normative function. *The Monist*, 85(1), 3-28.
- Clark, A. (2000). *Mindware: An introduction to the philosophy of cognitive science*. Oxford University Press.
- Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.
- Cornejo, G. V. (2009). Fisiología de la reproducción humana. *Revista Mexicana de Medicina de la Reproducción*, 1(4), 115-130.
- Cuffari, E. C., Di Paolo, E., & De Jaegher, H. (2015). From participatory sense-making to language: there and back again. *Phenomenology and the Cognitive Sciences*, 14(4), 1089-1125.
- Cussins, A. (2002). Experience, thought and activity. *Essays on nonconceptual content*, 147-163.
- Davidson, D. (1982). Rational animals. *dialectica*, 36(4), 317-327.

- Davidson, D. (1985). Reply to Quine on Events. In LePore E. and McLaughlin B. (Eds.), *Actions and Events: Perspectives on the Philosophy of Donald Davidson* (pp. 172- 176). Blackwell.
- De Haan, S. (2020). *Enactive psychiatry*. Cambridge University Press.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the cognitive sciences*, 6(4), 485-507.
- Dennett, D. C. (1981). True believers: The intentional stance and why it works. *Mind Design II: Philosophy, Psychology, and Artificial Intelligence*, 57-79.
- Di Paolo, E. (2001). Artificial life and historical processes. European Conference on Artificial Life,
- Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the cognitive sciences*, 4(4), 429-452.
- Di Paolo, E. (2018). The enactive conception of life. In A. Newen, L. D. Bruin, & S. Gallagher (Eds.), *The Oxford handbook of cognition: Embodied, embedded, enactive and extended* (pp. 71-94). Oxford University Press.
- Di Paolo, E. (2019, 18 November, 2019). *Response to questions about the book Sensorimotor Life* YouTube.
- Di Paolo, E., Buhrmann, T., & Barandiaran, X. E. (2017). *Sensorimotor life: an enactive proposal*. Oxford University Press.
- Di Paolo, E., Cuffari, E. C., & De Jaegher, H. (2018). *Linguistic bodies: The continuity between life and language*. MIT Press.
- Di Paolo, E., & Thompson, E. (2014). The enactive approach. In *The Routledge handbook of embodied cognition* (pp. 86-96). Routledge.
- Dings, R. (2020). Meaningful affordances. *Synthese*, 1-21.
- Dromm, K. (2003). Imaginary naturalism: The natural and primitive in Wittgenstein's later thought. *British journal for the History of Philosophy*, 11(4), 673-690.
- Flament-Fultot, M. (2016). Counterfactuals versus constraints: towards an implementation theory of sensorimotor mastery. *Journal of Consciousness Studies*.
- Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard university press.
- Fodor, J. A. (1981). The mind-body problem. *Scientific american*, 244(1), 114-123.
- Fodor, J. A. (1982). Cognitive science and the twin-Earth problem. *Notre Dame Journal of Formal Logic*, 23(2), 98-118.
- Fodor, J. A. (1990). *A theory of content and other essays*. The MIT press.
- Frege, G. (1892 [1980]). *On sense and reference*. De Gruyter Mouton.
- Froese, T. (2011). From second-order cybernetics to enactive cognitive science: Varela's turn from epistemology to phenomenology. *Systems Research and Behavioral Science*, 28(6), 631-645.
- Froese, T. (2017). Life is Precious Because it is Precarious: Individuality, Mortality and the Problem of Meaning. In G. Dodig-Crnkovic & R. Giovagnoli (Eds.), *Representation and Reality in Humans, Other Living Organisms and Intelligent Machines* (pp. 33-50). Springer International Publishing. [https://doi.org/10.1007/978-3-319-43784-2\\_3](https://doi.org/10.1007/978-3-319-43784-2_3)
- Froese, T. [Neural basis of Consciousness] (2021). Consciousness Talks 6: Phenomenology & Embodied mind: Irruption theory of consciousness. [Archivo de video]. YouTube. <https://youtu.be/f1DYNBvhsoQ>
- Froese, T., & Di Paolo, E. (2011). The enactive approach: Theoretical sketches from cell to society. *Pragmatics & Cognition*, 19(1), 1-36.
- Froese, T., & Di Paolo, E. A. (2009). Sociality and the life–mind continuity thesis. *Phenomenology and the Cognitive Sciences*, 8(4), 439-463.
- Froese, T., & González-Grandón, X. (2019). How passive is passive listening? Toward a sensorimotor theory of auditory perception. *Phenomenology and the Cognitive Sciences*, 1-33.
- Gibson, J. J. (2014 [1979]). *The ecological approach to visual perception*. Psychology Press.
- Giddens, A. (1995 [1984]). *La constitución de la sociedad: Bases para la teoría de la estructuración*. Amorrortu Editores.
- Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, 20(4), 464-472.
- Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science*, 16(2), 152-160.
- Grzankowski, A. (2015). Not all attitudes are propositional. *European Journal of Philosophy*, 23(3), 374-391.
- Heiat, A., Gross, C. P., & Krumholz, H. M. (2002). Representation of the Elderly, Women, and Minorities in Heart Failure Clinical Trials. *Archives of Internal Medicine*, 162(15). <https://doi.org/10.1001/archinte.162.15.1682>
- Heras-Escribano, M. (2019). *The Philosophy of Affordances*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-98830-6>
- Heras-Escribano, M. (2020a). Author's Response: Affordances as a Basis for a Post-Cognitivist Approach to the Mind. *Constructivist Foundations*, 15(3), 231-237.

- Heras-Escribano, M. (2020b). Précis of The Philosophy of Affordances. *Constructivist Foundations*, 15(3), 199-213.
- Hussain-Gambles, M., Atkin, K., & Leese, B. (2004). Why ethnic minority groups are under-represented in clinical trials: a review of the literature. *Health & Social Care in the Community*, 12(5), 382-388. <https://doi.org/https://doi.org/10.1111/j.1365-2524.2004.00507.x>
- Hutto, D. (2011). Philosophy of Mind's New Lease on Life Autopoietic Enactivism meets Teleosemantics. *Journal of Consciousness Studies*, 18, 44-64.
- Hutto, D., & Myin, E. (2013). *Radicalizing enactivism: basic minds without content*. MIT Press.
- Hutto, D., & Myin, E. (2017). *Evolving enactivism: basic minds meet content*. MIT Press.
- Ingold, T. (2002). *Culture and the perception of the environment*. Routledge.
- James, M. M. (2020). Bringing Forth Within: Enhabiting at the Intersection Between Enaction and Ecological Psychology [Hypothesis and Theory]. *Frontiers in Psychology*, 11(1348). <https://doi.org/10.3389/fpsyg.2020.01348>
- James, M. M., & Loaiza, J. M. (2020). Coenhabiting Interpersonal Inter-Identities in Recurrent Social Interaction [Hypothesis and Theory]. *Frontiers in Psychology*, 11(577). <https://doi.org/10.3389/fpsyg.2020.00577>
- Jonas, H. (2001 [1966]). *The phenomenon of life: Toward a philosophical biology*. Northwestern University Press.
- Jurgens, A., & Kirchhoff, M. D. (2019). Enactive social cognition: Diachronic constitution & coupled anticipation. *Consciousness and Cognition*, 70, 1-10.
- Kant, I. (1992 [1790]). *Crítica de la facultad de juzgar* (P. Oyarzún, Trans.). Monte Ávila. (1790)
- Kant, I. (2010 [1787]). *Crítica de la razón pura* (P. Ribas, Trans.). Gredos.
- Kelso, J. S., Tuller, B., Vatikiotis-Bateson, E., & Fowler, C. A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance*, 10(6), 812.
- Kingsbury, J. (2006). A proper understanding of Millikan. *Acta Analytica*, 21(3), 23-40.
- Kirchhoff, M. D. (2014). In search of ontological emergence: diachronic, but non-supervenient. *Axiomathes*, 24(1), 89-116.
- Kirchhoff, M. D. (2015). Extended cognition & the causal-constitutive fallacy: In search for a diachronic and dynamical conception of constitution. *Philosophy and Phenomenological Research*, 90(2), 320-360.
- Kirchhoff, M. D., & Kiverstein, J. (2020). Attuning to the World: The Diachronic Constitution of the Extended Conscious Mind [Hypothesis and Theory]. *Frontiers in Psychology*, 11(1966). <https://doi.org/10.3389/fpsyg.2020.01966>
- Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision research*, 46(11), 1762-1776.
- Kiverstein, J. D., & Rietveld, E. (2018). Reconceiving representation-hungry cognition: an ecological-enactive proposal. *Adaptive Behavior*, 26(4), 147-163.
- Kozhevnikov, V. A., & Chistovich, L. u. d. A. (1965). Speech: Articulation and perception.
- Kripke, S. A. (1972). Naming and necessity. In *Semantics of natural language* (pp. 253-355). Springer.
- Kripke, S. A. (1982). *Wittgenstein on rules and private language: An elementary exposition*. Harvard University Press.
- Lobo, L., Heras-Escribano, M., & Travieso, D. (2018). The history and philosophy of ecological psychology. *Frontiers in Psychology*, 9, 2228.
- Luhmann, N. (1998). *Complejidad y modernidad: De la unidad a la diferencia*. Trotta.
- Luhmann, N. (2007 [1997]). *La sociedad de la sociedad*. Herder, UIA.
- Macé, M. J.-M., Joubert, O. R., Nespoulous, J.-L., & Fabre-Thorpe, M. (2009). The time-course of visual categorizations: you spot the animal faster than the bird. *PloS one*, 4(6), e5927.
- Maturana, H. (2002). Autopoiesis, Structural Coupling and Cognition: A history of these and other notions in the biology of cognition. *Cybernetics & Human Knowing*, 9, 5-34.
- Maturana, H. (2016). Confusion of Reflective Domains? *Constructivist Foundations*, 11(2), 213-214. <http://constructivist.info/11/2/213.maturana>
- Maturana, H., & Varela, F. (1980). *Autopoiesis and cognition : the realization of the living*. D. Reidel Pub. Co. Publisher description <http://www.loc.gov/catdir/enhancements/fy0814/79024724-d.html>
- Table of contents only <http://www.loc.gov/catdir/enhancements/fy0814/79024724-t.html>
- Michaels, C. F., & Carello, C. (1981). *Direct perception*. Prentice-Hall Englewood Cliffs, NJ.
- Millikan, R. G. (1989). In defense of proper functions. *Philosophy of science*, 56(2), 288-302.
- Millikan, R. G. (1990). Compare and contrast Dretske, Fodor, and Millikan on teleosemantics. *Philosophical Topics*, 18(2), 151-161.
- Millikan, R. G. (2004). *Varieties of meaning: the 2002 Jean Nicod lectures*. MIT press.
- Millikan, R. G. (2005). *Language: A biological model*. Oxford University Press on Demand.



- Millikan, R. G. (2017). *Beyond concepts: Unicepts, language, and natural information*. Oxford University Press.
- Millikan, R. G. (2020). Neuroscience and teleosemantics. *Synthese*, 1-9.
- Millikan, R. G. (2021). Interactivism in Perspective Preconference Reading Group and Talks: Ruth Millikan presents her work. In.
- Mojica, L. (2014). *Anxiety: a grammatical investigation* [Universiteit van Amsterdam]. Amsterdam.
- Mojica, L. (2019). Mentes en cuerpos de afecto: nota crítica sobre Ecology of the Brain, de Thomas Fuchs. *Crítica*, 51(153), 83-97.
- Mojica, L. (2020). Reclaiming Meaning, Reclaiming Normativity. *Constructivist Foundations*, 15(3), 216-218. <https://constructivist.info/15/3/216.mojica>
- Mojica, L. (2021). The enactive naturalization of normativity: from self-maintenance to situated interactions. *History and Philosophy of the Life Sciences*.
- Mojica, L., & Froese, T. (2019). On the spatiotemporal extensiveness of sense-making: ultrafast cognition and the historicity of normativity. *Synthese*, 1-14.
- Moreno, A., & Barandiaran, X. E. (2004). A Naturalized Account of the Inside-Outside Dichotomy. *Philosophica*, 73.
- Moreno, A., & Mossio, M. (2015). *Biological autonomy: A philosophical and theoretical enquiry* (Vol. 12). Springer.
- Moreno, M. A., Stepp, N., & Turvey, M. (2011). Whole body lexical decision. *Neuroscience letters*, 490(2), 126-129.
- Noë, A. (2004). *Action in perception*. MIT press.
- Nowakowski, R. M., & Dziembowski, W. A. (2003). Nonlinearity of Nonradial Modes in Evolved Stars. *Astrophysics and Space Science*, 284(1), 273-276. <https://doi.org/10.1023/A:1023260223343>
- O'regan, J. K., & Noë, A. (2001a). A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, 24(5), 939-973.
- O'Regan, J. K., & Noë, A. (2001b). A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, 24(5), 939.
- Papineau, D. (1998). Teleosemantics and indeterminacy. *Australasian Journal of Philosophy*, 76(1), 1-14. <https://doi.org/10.1080/00048409812348151>
- Papineau, D. (2021). Naturalism. *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2021/entries/naturalism/>>.
- Pescador, C., & Mojica, L. (en proceso de publicación). Making us autonomous: The enactive normativity of morality. *Topoi*.
- Pitt, D. (2020). Mental Representation. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 ed.). <<https://plato.stanford.edu/archives/spr2020/entries/mental-representation/>>.
- Port, R. F., & Van Gelder, T. (1995). *Mind as motion : explorations in the dynamics of cognition*. MIT Press.
- Porter Jr, R. J., & Castellanos, F. X. (1980). Speech-production measures of speech perception: Rapid shadowing of VCV syllables. *the Journal of the Acoustical Society of America*, 67(4), 1349-1356.
- Quinn, P. O., & Madhoo, M. (2014). A review of attention-deficit/hyperactivity disorder in women and girls: uncovering this hidden diagnosis. *The primary care companion for CNS disorders*, 16(3).
- Richardson, M. J., & Chemero, A. (2014). Complex dynamical systems and embodiment. *The Routledge handbook of embodied cognition*, 39-50.
- Rietveld, E. (2008). Situated normativity: The normative aspect of embodied cognition in unreflective action. *Mind*, 117(468), 973-1001.
- Rietveld, E., & Kiverstein, J. (2014). A Rich Landscape of Affordances. *Ecological Psychology*, 26(4), 325-352. <https://doi.org/10.1080/10407413.2014.958035>
- Ruiz-Mirazo, K., Peretó, J., & Moreno, A. (2004). A universal definition of life: autonomy and open-ended evolution. *Origins of life and evolution of the biosphere*, 34(3), 323-346.
- Schatzki, T. R. (1996). *Social practices: A Wittgensteinian approach to human activity and the social*. Cambridge University Press.
- Shea, N. (2013a). Millikan's isomorphism requirement. In *Millikan and her critics* (pp. 63-80). Wiley-Blackwell.
- Shea, N. (2013b). Naturalising representational content. *Philosophy Compass*, 8(5), 496-509.
- Shea, N. (2018). *Representation in cognitive science*. Oxford University Press.
- Spivey, M. (2008). *The continuity of mind*. Oxford University Press.
- Stapleton, M., & Froese, T. (2015). Is Collective Agency a Coherent Idea? Considerations from the Enactive Theory of Agency. In C. Misselhorn (Ed.), *Collective Agency and Cooperation in Natural and Artificial Systems: Explanation, Implementation and Simulation* (pp. 219-236). Springer International Publishing. [https://doi.org/10.1007/978-3-319-15515-9\\_12](https://doi.org/10.1007/978-3-319-15515-9_12)
- Stepp, N., & Turvey, M. T. (2010). On strong anticipation. *Cognitive systems research*, 11(2), 148-164.
- Taguchi, S. (2016). Can the lived experience of living beings be approached through inference? *Constructivists Foundation*, 11(2), 216-216.

- Thompson, E. (2007). *Mind in life: biology, phenomenology, and the sciences of mind*. Belknap Press of Harvard University Press.
- Turvey, M. T., Shaw, R. E., Reed, E. S., & Mace, W. M. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn (1981). *Cognition*, 9(3), 237-304.
- Varela, F. (1981). Describing the logic of the living: The adequacy and limitations of the idea of autopoiesis. *Autopoiesis: A theory of living organization*, 36-48.
- Varela, F. (1991). Organism: A meshwork of selfless selves. In *Organism and the Origins of Self* (pp. 79-107). Springer.
- Varela, F. (1999). The specious present: A neurophenomenology of time consciousness.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: cognitive science and human experience*. MIT Press.
- Villalobos, M. (2013). Enactive cognitive science: revisionism or revolution? *Adaptive Behavior*, 21(3), 159-167. <https://doi.org/10.1177/1059712313482953>
- Villalobos, M. (2020). Living beings as autopoietic bodies. *Adaptive Behavior*, 28(1), 51-58. <https://doi.org/10.1177/1059712319879747>
- Villalobos, M., & Razeto-Barry, P. (2020). Are living beings extended autopoietic systems? An embodied reply. *Adaptive Behavior*, 28(1), 3-13. <https://doi.org/10.1177/1059712318823723>
- Villalobos, M., & Ward, D. (2015). Living systems: Autonomy, autopoiesis and enaction. *Philosophy & Technology*, 28(2), 225-239.
- Villalobos, M., & Ward, D. (2016). Lived experience and cognitive science: Reappraising enactivism's Jonasian turn. *Constructivist Foundations*, 11(2).
- von Uexküll, J. (1957). A stroll through the worlds of animals and men, Instinctive Behavior. *Int. Univ. Press, New York*, 5-80.
- Wallot, S., & Van Orden, G. (2012). Ultrafast cognition. *Journal of Consciousness Studies*, 19(5-6), 141-160.
- Weber, A., & Varela, F. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1(2), 97-125. <https://doi.org/10.1023/A:1020368120174>
- Weerdesteyn, V., Nienhuis, B., Hampsink, B., & Duysens, J. (2004). Gait adjustments in response to an obstacle are faster than voluntary reactions. *Human movement science*, 23(3-4), 351-363.
- Wittgenstein, L. (1980). *Remarks on the philosophy of psychology* (Vol. 2). Basil Blackwell.
- Wittgenstein, L. (2003 [1953]). *Investigaciones filosóficas*. Instituto de Investigaciones Filosóficas.
- Wittgenstein, L. (2009 [1953]). *Philosophical investigations*. John Wiley & Sons.